CONFERENCE PROCEEDINGS

ALLIANCE UNIVERSITY

Private University established in Karnataka State by Act No.34 of year 2010
Recognized by the University Grants Commission (UGC), New Delhi

Alliance College of Engineering and Design

BENGALURU

ALLIANCE INTERNATIONAL CONFERENCE ON ARTIFICIAL INTELLIGENCE AND MACHINE LEARNING (AICAAM)

26 - 27 APRIL, 2019

ISBN:978-93-5361-299-3

ALLIANCE UNIVERSITY

Private University established in Karnataka State by Act No.34 of year 2010
Recognized by the University Grants Commission (UGC), New Delhi

Alliance College of Engineering and Design

Conference Proceedings

ALLIANCE INTERNATIONAL CONFERENCE ON ARTIFICIAL INTELLIGENCE AND MACHINE LEARNING

(AICAAM)

26-27 APRIL 2019

ISBN:978-93-5361-299-3

Editors

Dr. R.Selvarani

Dr. Reeba Korah

Prof. Sudaroli

Alliance College of Engineering and Design
Alliance University - Bangalore

Proceedings of Alliance International Conference on
Artificial Intelligence and Machine Learning
26th and 27th of April 2019

ISBN:978-93-5361-299-3

Preface

It is a great privilege to present the conference proceedings of 'Alliance International Conference on Artificial Intelligence and Machine Learning, (AICAAM 2019)'.

The theme Artificial Intelligence (AI), a cutting-edge technology with impacts on a wide range of businesses and activities, is rapidly becoming part of our everyday lives, changing how we work, travel and interact with each other. Yet we are only at the cusp of discovering the manner in which AI will have an impact on business, society and culture.

AICAAM 2019 aims to observe the innovative practices and novel ideas in diverse domains of science, technology, business and other interdisciplinary fields. The focus is on addressing societal challenges globally and supporting interdisciplinary and cross-disciplinary scientific problems through innovative approaches in the field of machine learning and artificial intelligence.

The ever-changing scope and rapid development of artificial intelligence and machine learning is driving important developments such as autonomous vehicles, medical diagnosis, advanced manufacturing and many others that create new problems and questions, resulting in the real need for sharing brilliant ideas and stimulating good awareness of this significant research field.

As AI moves from the theoretical realm to the global marketplace, AICAAM 2019 produces a bright picture and fascinating landscape for artificial intelligence and machine learning. The support received and the enthusiasm witnessed in this conference have truly been remarkable. The responses to the call-for-papers had been overwhelming – both from India and overseas.

We hope that the conference proceedings will prove an invaluable resource for researchers, businesses, and policymakers in the field, as well as for those general readers who want to find out about contemporary research on artificial intelligence and machine learning.

We would like to express our gratitude and appreciation for the reviewers who helped in maintaining high quality of manuscripts included in the proceedings published by Alliance University. We would also like to extend our thanks to the members of the organizing team for their meticulous support in bringing out the proceedings and conducting this event.

Dr. Anubha Singh
Patron – AICAAM 2019
Pro Vice-Chancellor, Alliance University

Table of Contents

Business Ventures Modelling using Chemical Bonding Concepts

Tejeswarsai Sasipalli*, Venkata SanyasiRao Sasipalli, Sashibhushana Rao Majji[+], R Venkataraman[++]**

**Kan-on Senior High School, Hiroshima, Japan*
***ICT Wing, Center for Excellence in Computer Technology, Hiroshima, Japan*
[+]Satya Institute of Technology and Management, Vizianagaram, AP, India
[++]Rinanu Semiconductors, Bangalore, Karnataka, India

Abstract- Modelling successful business ventures can be a challenging process, which needs innovative concepts or methodologies. We propose an innovative methodology that is to apply the chemical bonding concepts in business ventures, where chemical elements bonding can be mapped to people association. This paper introduces implementation concepts of how the business ventures can be realized utilizing chemical bonding process and describes the model in detail with numerical examples. A business matching table is also introduced to identify suitable matches (read partners) to enter into business ventures based on the valences. The identification of a suitable partner is done through a search algorithm, and as the data gets larger and larger, the time complexity also gets bigger and the best algorithm takes O(n log(n)) time.

Index Terms- Business Ventures, Chemical Bonding, Valency, Normalization, Business Matching Table, Search Pattern

I. INTRODUCTION AND FORMATION

A) Business Ventures

Capital accumulation occurs mostly through promotion of business ventures, where investments flow into new business domains, strengthen or expand the existing domains. Business venture is a kind of partnership association, where seeker may not have sufficient knowledge or strength in a certain business domain but want to enter into that business domain, where some others have good amount of strengths, and vice versa [1]. Two types of Ventures can happen a) between domestic and domestic or b) domestic – foreign investors. Most of the present ventures in developing countries are of the type domestic – foreign in the form of FDIs [2] or through technology transfer [3][4]. In whatever form the association of two partners may be, it can be viewed as

conglomeration of Strengths, Weaknesses and Needs; and promotion and realization of a concrete association that leads to business venture in a mutually beneficial way.

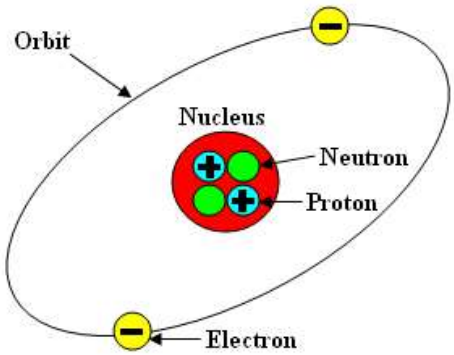
But how concrete the association can be realized? who does this job? and how it can be done? Is there any scientific methodology? These are all questions which have many answers and varied solutions based on situational needs. We consider chemical bonding concepts where atomic structure itself can be mapped to organization setup, and atomic number as strength of business domain, identifying the corresponding chemical element, etc., details of which are described in the next sections. Organization and company are interchangeably used to mean the same.

B) Chemical Concepts

Chemical elements are combined to form compounds. In the combination of the elements we see bonding among all the participating elements. This is due to valency or the electronic charge in the outermost shell of atom. Similarly, individual organizations combine their common business interests to form venture by sharing the strengths and weaknesses [5].

Atomic structure

Atom is a smallest constituent unit of ordinary matter that has the properties of a chemical element. Every atom is composed of Nucleus and Electron orbits bound to the nucleus. Nucleus is composed of protons (positive charges) and neutrons (no charge), while the Electrons (negative charges) orbit around it. The number of protons of an atom defines its chemical element. We don't go in details of quantum mechanics of atoms, but for simplicity we stick to the Bohr model of atom, focus on atomic number and the valency (number of electrons in outermost shell), who can give or take electrons to form chemical bonding.

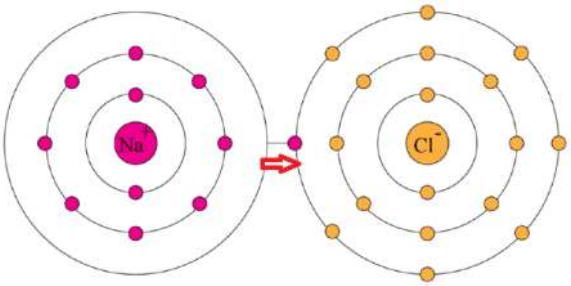


Valency

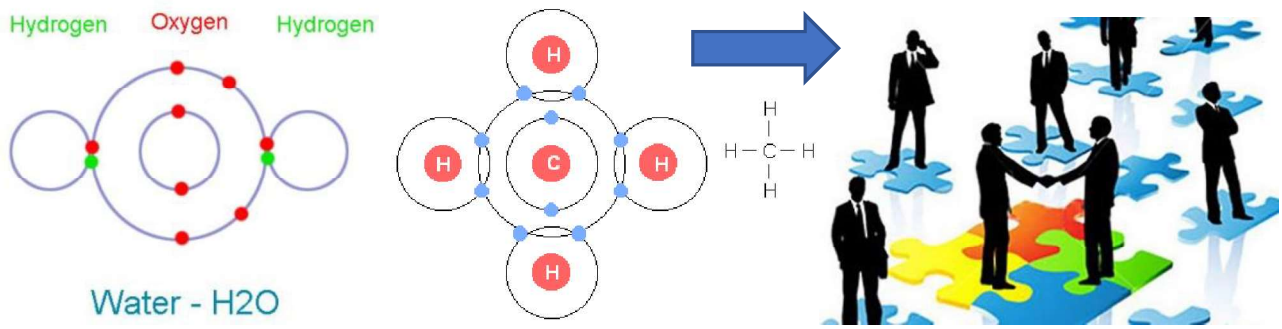
Electrons in an atom are arranged in orbitals or shells and those electrons that are present in the outermost orbit are called valence electrons, which are excess in the configuration [6]. If the outermost shell gets filled, there will be very little chemical activity and the combining capacity turns to zero eventually. Noble gases have no valence electrons. All the other elements look for bonding. Bohr model says that the outermost shell consists of a maximum of 8 valence electrons [7] in the electron configuration of $2n^2$, where n represents (K, L, M, N, …) shell. Eg. 12 electrons = $2 * 2^0 + 2 * 2^2 + 2$, which implies K, L shells are full and M shell contains 2 electrons. Therefore, valence is 2.

Chemical Bonding

Chemical bonding involves transfer of an electron, so if one atom gains an electron while the other atom loses an electron. Because opposite charges attract, the atoms bond together to form a molecule.



Transforming this concept to business venture that an organization wants to enhance in a certain domain but has no strength and there exists one or more organizations which have this required strength which can be shared to form a new venture (molecule). In the depicted image, Sodium Na has 1 electron excess and Chlorine has 1 electron deficient, hence these two elements can combine to form Sodium chloride [8] (ionic bonding), like a new venture. Some more examples are shown in the following pictures. These pictures are anonymous to only describe the concepts.

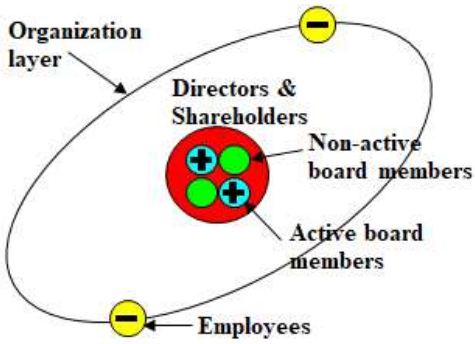


Oxygen forms two bonds with two hydrogens to form water molecule (covalent bonding). Carbon form 4 bonds with hydrogen to form Methane (covalent bonding). As such, companies form partnerships with other companies to form new business ventures. As Garry describes the science that results from the integration of a constellation of underlying disciplines such as chemistry, biology and physics to evolve biotechnology, nanotechnology, in his science-based business paper [9], the proposed methodology can also be a result of proper implementation of chemistry principles to evolve successful business ventures. Ionic bonding is stronger in some compounds and Covalent is stronger on other compounds. We use both types of bonding as with the situation needs.

II. ATOMIC AND ORGANIZATIONAL CONCEPTS AND CALCULATIONS

Atom is a structured configuration of particles while company is a structured configuration of people. Although atomic structure and company structure are two different perspectives, a commonality can be identified. We identify the commonalities in two structures and map them to observe principal character that can be utilized in forming business ventures.

A) Atomic Organization



Every combination of people or elements or for that matter anything in nature whether physical or abstract, is a perfect organization of charges that move around in search of a mate or matching for evolution. It is nature. Business is a need of people who are charged with positive or negative and coexist in an organizational setup.

If we look at the basic structures, one can see commonness between an atomic structure and a company structure as described in below Table 1. Let us consider the atomic number of an element, which provides valency information for that element. And consider the business domain strength of a company as an atomic number of that business domain, where we can calculate the valency, which describe the capability to donate or acquire at employee (knowledge) level.

	Atom	Company
Nucleus	Nucleus	Board of Directors & shareholders
Protons	Particles with positive electric charge	Active directors & shareholders
Neutrons	Particles with no net electric Charge	Non-active directors & shareholders
Orbit	Path in which electrons move	Organization layer
Electron	Particle with negative electric charge, characterize the element	Employees and others responsible for the performance of organization
Valence electrons	Excess or Deficient electrons in the outermost orbit	Strengths and weaknesses

Table 1: Atomic Organization Mapping

B) Normalized Domain Strength Calculation

All companies create their customer bases for each domain of their business (for example, mobiles, cars, toys, computers, sports equipment, etc) in different regions (North, South, East and West). In order to find out atomic number of a business domain, we summarize the strengths information in Table 2, in terms of customer bases in the four as given below. Also, we summarize the existing ventures or collaborations information in the same table.

n	N	e
W	A#	E
w	S	s

Table 2: Strength-identify for each business domain

A#: Normalized domain strength of a particular business domain
N, S, E, W: Existing customers in North, South, East, West regions
n, s, e, w: Existing ventures and collaborations in these regions respectively.

Normalized domain strength of a company can be calculated using below formula.

Normalized domain strength: $nC_i = C_i - \frac{\sum_i C_i}{\sum_j V_j}, \forall i, j$

where,

$$C_i = \{\text{existing customers}\}, i \in [N, S\ E, W],$$
$$V_j = \{\text{existing ventures}\}, j \in [n, s, e, w]$$

Special cases:

1) $V_j = 0, \forall j$: means that there is no at least one venture in any region. Without having venture, no organization can survive. Business truth, so $\sum_j V_j \neq 0$
2) $\sum_j V_j = 4, \forall j$: means that there is at least one venture in each region, which cannot help determining the strength, so $\sum_j V_j > 4$, no need to be an integer value, it can be a fraction value.
3) $\sum_i C_i = 8, \forall i$: then nC_i becomes < 1. We cannot work with chemical periodic table, so $\sum_i C_i > 9$.

C) Numerical Examples

Numerical examples can help understand the calculation of normalized domain strength (A#), we provide two example, general case and extreme case:

Example1 (General case): Let the existing customers in four regions be: $C_i = \{20, 30, 15, 50\}$ where $i \in [N, S, E, W]$ and the existing ventures in four regions be: $V_j = \{3, 2, 0, 5\}$ where $j \in [n, s, e, w]$

3	20	0
50		15
5	30	2

=> Normalized customers =>

	8.5	
39		3.5
	19	

=> A# =>

	69	

This 69 can be taken as an atomic number of this particular business domain (not the whole organization). We can now look into the chemical periodic table which chemical element it represents. In this particular example, it represents Thulium whose valency is 3+. Place this valency in the Business Matching Table which is described in next section, to find out matching partner for a perfect venture.

Another example which calculates with bottom line values.

Example 2 (Boundary case) Let the existing customers in four regions be: $C_i = \{1, 2, 2, 4\}$ where $i \in [N, S, E, W]$ and the existing ventures in four regions be: $V_j = \{1, 1, 1, 1.5\}$ where $j \in [n, s, e, w]$

1	1	1
4		2
1	2	1.5

=> Normalized customers =>

	-1	
2		0
	0	

=> A# =>

	1	

The atomic number of this particular business domain is 1, which represents Hydrogen in the chemical periodic table and whose valency is 1+. The negative value in normalized customers table signals the organization need to mobilize the strengths in the corresponding region if interested.

III. BUSINESS MATCHING TABLE CONCEPT

Business matching table (BMT) is a table that helps business units to identify themselves in a position, based on their strengths and weaknesses (valences), also can find a right match to form a perfect chemical bonding or business venture to fulfil their business interests. A company which has less than 5 employees is termed as Micro and can be categorized as Sole Proprietor organization, similarly, a company which has more than 2500 employees can be termed as a corporation and can be categorized as large corporation. These are common terminologies in the business world.

Business Matching Table

		Weakness									Strength							
		-8	-7	-6	-5	-4	-3	-2	-1	0	1	2	3	4	5	6	7	8
		A8	A7	A6	A5	A4	A3	A2	A1	AB	B1	B2	B3	B4	B5	B6	B7	B8
Corporation	> 2500 Corp.																	
Small Corp	< 2500 XXLarge																	
	< 1000 Xlarge																	
LLC	< 500 Large																	
	< 250 Medium																	
Partnership	< 50 Small																	
	< 25 Mini																	
Sole Proprietor	< 5 Micro																	

Table 3: Business Matching Table

A1-A8: Negative valency numbers from -1 to -8
AB: No valency, inert state
B1-B10: Positive valency numbers from 1 to 8

Strength / Weakness: Every company will have a kind of strength / weakness in certain business domain. Such as a service provider company can have weakness in mobile technology or security technology, etc. Weakness can be overcome by making bonds with suitable companies which are stronger in those business domains.

Each cell in business matching table shall represent a set of characteristics which will be reflected in all those organizations. In this BMT definition, organizations will have continuous opportunity to move from one cell to another based on their changing strength in customer base. The relative gap in characteristics of cells will help guide organizations to determine the required resources to move in to desired cell by calculating the valency.

Initially, most of the companies shall be established on their own strengths, however, in the course of business running, their strengths change. Strengths and weakness change continuously. If each organization documents the valences of their business domains in this table, it is easy to find venture partner from this BMT table.

Business matching table though looks simple, but if we do simulation of strengths / weaknesses (valences) and business aspiration (needs) in all geographical regions and in all the business domains, the input size (number of parameters) increase, which leads to big-data analytics, we skip the details of this direction of research in this paper.

IV. RESULTS AND ANALYSIS

Thulium valency is 3+, it can have strong bonding with elements whose
valency = 1 to form 3 bonds
valency = 1 to form 1 bond + valency = 2 form 1 double bond give total 3 bonds. Thulium has chemical bonding solutions with Oxygen and Ferrous to from Thulium Oxide (Tm_2O_3) and Thulium Fluoride (TmF_3), these are common. Here, we are not interested in the core chemistry but interested in the patterns of chemical bonding which can reflect in forming partnership towards ventures. In this particular example, Thulium valency can be positioned in our BMT at (Xlarge, B3) cell as shown in below figure. From the BMT table (Xlarge, B3) can form 3 bonds with any of A1 cells. In business terms it can form 3 single partnership associations with the companies whose valency is in A1 column. Or, (Xlarge, B3) can form 1 bond with any cell in A1 column and one double with any cell in A2. This can be seen in below figure.

Business Matching Table

		Weakness									Strength							
		-8	-7	-6	-5	-4	-3	-2	-1	0	1	2	3	4	5	6	7	8
		A8	A7	A6	A5	A4	A3	A2	A1	AB	B1	B2	B3	B4	B5	B6	B7	B8
Corporation	> 2500 Corp.																	
Small Corp	< 2500 XXLarge								1									
	< 1000 Xlarge								1				3					
LLC	< 500 Large							2										
	< 250 Medium								1									
Partnership	< 50 Small																	
	< 25 Mini								1									
Sole Proprietor	< 5 Micro																	

In business ventures it always recommended to have minimum number of bindings, which necessitates to a search problem in order to find closer valency to 3. Here the valency is small number, we do not worry much but if the valency gets larger and positioned in B7, we will have to search all columns A1-A6 for suitable pairs. We can use "Find a pair with given difference" algorithm to achieve this search. In this particular search the difference number can be half of valency number so that we can minimize the number of associations. The time complexity of this search algorithm can be reduced to O(nlog(n)) time by writing best logic algorithm, where big-O is order and n is the valency number.

V. CONCLUSION

Capital accumulation strategies are tools in economy growth. However, the ultimate players in this process are organizations. These organizations in the globalization trend scenarios need to go into ventures across the world. Forming ventures is not easy, which may end up in mutual losses. To avoid any losses, it is recommended to adopt suitable concepts or methodologies. One of these concepts described in this paper is chemical bonding concepts. The paper described what the chemical elements are and how the bonding concept could be applied in business. Then creating a mapping with business organization structure where the chemical bonding process is implemented. The business matching table introduced here provides a way to find the partners to form stronger ventures. The process of finding venture partners is described with one example, but if we fill all the columns of the BMT, we will get large sets of data. As an extension of this research we create datasets and implement search algorithms to find the best match for a given valency.

ACKNOWLEDGMENT

We acknowledge the support of RAMTEJ Technologies Corporation provided in the form of Research Grant No. RG012019. This paper content reflects only authors' views and the funding organization is not responsible for any use of the information it contains. We wish to thank the anonymous reviewers.

REFERENCES

[1] Richard Florida and Donald F. Smith, Jr., "Venture Capital Formation, Investment, and Regional Industrialization", *Annals of the Association of American Geographers,* Vol. 83, No.3 (1993), pp. 434-451

[2] Dirk Willem te Velde, "Policies towards foreign direct investment in developing countries: emerging best-practices and outstanding issues", *Overseas Development Institute, London* (March 2001), pp. 1-62

[3] Manas Chatterji, "Technology Transfer in the Developing Countries", *Springer publications*, ISBN 978-1-349-20560-8 (1990)

[4] Dominique Redor et Mohamed Saadi, "International Technology Transfer to Developing Countries: when is it immiserizing?" *Dans Revue D'Economie Politique*, Vol. 121, (2011), pp. 409-433

[5] Richard Langley, Stephen F., "Strengths of Ionic and Covalent Bonds" *Chemical Bonding and Molecular Geometry*, (2012), Ch.7

[6] Chris Hren, Peter J. Mikulecky, "Why are Valence Electrons Important?" *Chemistry Workbook for Dummies, 3rd Edition*, (2017), ISBN 978-1-119-35745-2

[7] Anne Marie Helmenstine, "Bohr Model of the Atom Explained" *https://www.thoughtco.com/bohr-model-of-the-atom-603815* (Jan 2019)

[8] "Ionic bonding in sodium chloride" https://www.bbc.com/bitesize/guides/ztc6w6f/revision/2

[9] Garry P. Pisano, "The Evolution of Science-Based Business: Innovating How We Innovate", Industrial and Corporate Change, Special Issue in Honor of Alfred D. Chandler, Harvard Business School, Jan 2010.

Evaluation of Machine Learning Methods for Organic Apple Authentication by Diffraction Grating and Image Processing

Hui Wang*, Weiran Song*, Nanfeng Jiang**

** Ulster University, United Kingdom*
*** Fujian Normal University, China*

Abstract- In this work we present an extensive evaluation of a low-cost, non-destructive sensor system on a task to differentiate between organic and conventional apples. This system includes a diffraction grating sheet, a smart phone, and a flashlight. The flashlight is used as a light source to illuminate an apple, the diffraction grating sheet splits the reflected light from the apple into different colors to create a rainbow, and the smartphone records the rainbow image. We acquired 150 apples from a local supermarket, 75 being organic and 75 being conventional. We used the sensor system to collect rainbow images under room temperature and dark condition, and then used image processing tools to convert images into spectra. We ran twelve machine learning algorithms on the spectral data created this way and conducted further experiments using a commercial NIR spectrometer on the same set of apples. The locally weighted partial least squares classification (LW-PLSC) reached the highest accuracy of 93% on rainbow image data, while the partial least squares discriminant analysis (PLS-DA) correctly classified all spectral data. The sensor system is convenient and low-cost, which provides a variable solution for in-line food authentication.

A Statistical Approach to Quantify Promotion Corrected Measure of Item Loyalty

Diptarka Saha*

* WalmartLabs, Bengaluru

Abstract - Customers' loyalty towards an item is an indispensable component of the item's performance, yet this is an unobservable quantity and may be studied from many angles, as have been done by various authors. Another major goal of any item performance metric would be to eliminate the effects of promotion while calculating loyalty. This study provides a general framework for calculating item loyalty as a function of the exponentially weighted average of the customers' purchases along with penalties and rewards that reflect the effect of promotion. The methods developed have been applied on Walmart's panel data to generate meaningful and robust promotion corrected loyalty values for several products.

Index Terms- *Item Loyalty, Promotion Corrected Loyalty, Panel data, Walmart Canada*

I. INTRODUCTION

As a core objective of marketing activities, maintaining and enhancing brand loyalty have been widely acknowledged [1]. The value of brand loyalty is self-evident for a company [2]. To have consumers that are loyal towards your brand is something that every brand prioritizes, since those customers are a great source of possible revenue in the form of lesser price elasticity, word of mouth promotion etc. [3].

As a result, few topics have captured as much attention among practitioners and academics alike as brand loyalty. Despite the abundance of the studies performed, no universal model for measuring customer loyalty has been accepted [4]. Heuristically, the word loyalty is used to describe the customer's tendency to repurchase the same brand [5]. A *conceptual definition* provided by Jacoby and Kyner [6]. The definition is expressed by a set of six necessary and collectively sufficient condition. These are that brand loyalty is

1. Biased or non-random
2. Behavioural response (i.e. from purchase)
3. Expressed over time
4. By some decision-making unit (e.g. households)
5. With respect to a set of alternatives
6. Function of some psychological process

In particular, any measure of brand loyalty should differentiate itself from market-share, there may be items, typically high-priced items that don't generate high volume of sales compared to other (typically low-priced) items, but have a customer base that is immensely loyal toward the brand due to some unique selling point of the brand. A so-called '*niche*' brand [7]. This study provides a definition of loyalty that is able to detect such items and thus provide important insights that is not apparent on the surface.

Another important question to keep in mind while measuring loyalty is what other factors can collude sales other than loyalty – one very common occurrence we see in real world data is that sales is driven

by promotion; which to a casual eye may look like a result of customer's loyalty towards the item. One of the major goals of this study is to eliminate effects of promotion while calculating loyalty.

II. Literature Review

Majority of the *operational definitions* of loyalty maybe categorized into two major classes: behavioral or attitudinal, based on their emphasis on purchasing or cognitive component [8]. Both approaches have their respective strength and weaknesses. Behavioral loyalty focuses on actual purchases especially giving emphasis on repeat purchases and purchase volumes [9]. Since they are based on long-term purchase data, data for behavioural loyalty measures are easy to collect and are unlikely to be incidental. Biggest limitation of this approach is its unidimensional nature, especially the fact that the distinction between repeat purchases and loyalty is non-existent; as many authors have pointed out that a customers' inertia alone may result in her repeat purchases – that do not necessarily imply her unrivalled affection towards that product – the so called '*spurious loyalty*' [10]. In fact, nearly three quarters of customers purchases are based on their emotion and attitude [11].

On the other hand, attitudinal measures rely on stated preferences, attitudes, intentions; these measures are typically based on surveys. For example, Anselmsson [12] presents an interesting case study of how customers are willing to pay a price premium for their preferred brands. Although attitudinal loyalty measures do provide insights into customer mindsets that are not captured by purchase history they still suffer from a number of disadvantages. First and foremost, would be the availability of data, behavioral measures require point of sale data, which is easily available due to automatic capturing, however, attitudinal measures require ground surveys – which can be time consuming, expensive in terms of money and human resources and overall hard to come by. Along with this, attitudinal measures may prove to be inaccurate as it depends on the mood of the person being surveyed when he's being surveyed; moreover, intentions may not always imply loyalty or even a probability of future purchase, one may consider example of expensive cars in this regard. And finally, the survey results typically represent one single time point and hence may not always convey robust information.

Due to these reasons, more recent measures pioneered by Dick and Basu [13] have incorporated behavioural and attitudinal measures simultaneously. They proposed attitude as a cause of repeat purchases. In fact, purchases induce a favourable outlook towards the brand in the mind of the customer which, in turn, generates more repeat purchases and hence loyalty is seen as a function of *attitude-manifested behaviour*. [14]

III. Methodology

A. Data

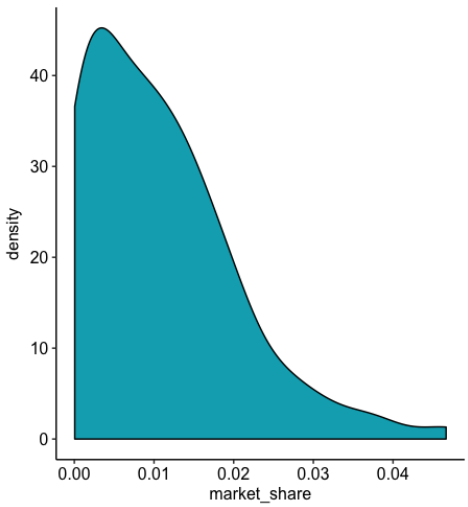
This work is based on point of sale data provided by Walmart Canada. As new product contenders fight for the shelf space in Walmart, identifying suitable items for a modular drop becomes all the more crucial and hence the need for an accurate yet interpretable loyalty scores. We will be demonstrating our work using products under the category *Mouthwash*. The motivation behind choosing mouthwash for this task is our belief that customers are likely to be unwilling to change their choices of mouthwash and tend to stick to their preferred brand/product. However, since this is also a relatively expensive product and vanity product in the sphere of oral care products, there would be a segment of the population who will be inclined to give in to promotional frenzies and yet another segment of the population who will display loyalty to their preferred product in face of extensive promotional bonanza – this creates an ideal situation to isolate effects of promotion on loyalty. The

data contains all purchases made in Walmart Canada from June 2017 to June 2018. Each entry in this data (sample: table 1) contains one household ID, the product ID (UPC), purchase date, the price, and whether or not that product was on promotion at the time of the purchase. As can be seen, this data doesn't contain any attitudinal features, hence we will be proposing a behavioural measure of loyalty. Although this doesn't encompass every aspect of customer loyalty this fits well into our framework and is sufficiently informative, objective, and reasonable regarding the aspect of available data.

UPC	***Purchase Date***	***Household ID***	***Price***	***Promotion***
6379370730	02/01/2018	101	5.77	1

Table 1: Example of the Data

In the given period, the data consists around 17 million purchases of 94 products by over 3 million households. For calculating the loyalty measures, we discard purchases made by light users, households with less than 3 purchases in the past 52 weeks. In the figures below, we summarise basic properties of the data.



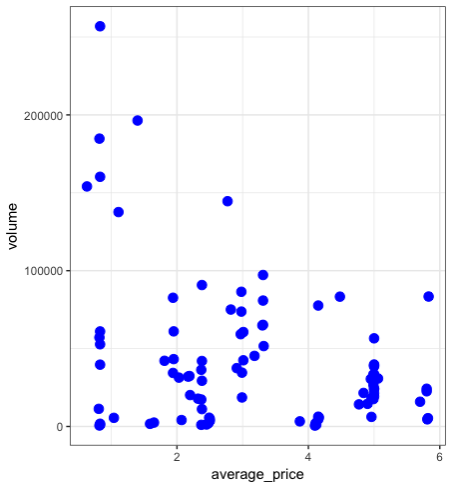


Figure 1(a): Empirical Density Plot of Market Share of The Products

Figure 1(b): Scatterplot of Volume vs Price

From these plots the data provides some insights, figure 1 shows positively skewed distribution which is expected as most of the items have low market-share and very few have (relatively) large share. Similarly, in figure 2 we have an overall negative correlation between price and volume of products, with a typical funnel like shape.

B. Definitions

With these specifications in mind, we begin describing our proposed measure. The following definitions will be used while describing the measure.

First, to define the premise, we have H households $\mathbb{H} = \{1, 2, \dots, H\}$. Every household makes t_h many visits to buy a mouthwash, in the given 52 weeks – each time they have n products $\mathbb{J} = \{1, 2, \dots, n\}$ to choose from. They buy exactly one product at each visit.

In what follows suffixes h, j, t will always denote household ID, product ID, and visit number respectively.

1. **Transaction (T):** The random variable transaction is defined as
$$T_h(t) = j \text{ if HH h bought the product } j \text{ at visit number } t$$

2. **Promotion (P):** This is an indicator variable that defines whether a transaction was made under promotion or not. So, for a transaction $T_h(t) = j$
$$P(T) = \begin{cases} 1, & j \text{ was on promotion duringt visit } t \text{ of HH } h \\ 0, & j \text{ wasn't on promotion during visit } t \text{ of HH } h \end{cases}$$

3. **Promotion Delta (ΔP):** This is another categorical variable that takes value 1 if the HH's last transaction was not under promotion but the current one is and vice versa. Formally,
$$\Delta P(T) = P\big(T_h(t)\big) - P\big(T_h(t-1)\big)$$

4. **Switch (S):** Finally, this variable denotes if a transaction is different from the one before it.
$$S(T) = \begin{cases} 1, & T_h(t) \neq T_h(t-1) \\ 0, & T_h(t) = T_h(t-1) \end{cases}$$

5. **Loyal HH:** A household is said to be a loyalty HH if they buy the same product in every visit.
$$T_h(t) = j \; \forall t \; for \; some \; j \in \mathbb{P}$$

   Let $\mathbb{L} \subseteq \mathbb{H}$ denote the set of all the loyal HHs

   Along with these also define, $\mathbb{H}_j \subseteq \mathbb{H}$ as
$$\mathbb{H}_j = \{h \in \mathbb{H} \mid j \in \bigcup_t T_h(t)\}$$

   Having these definitions in hand, we attempt to find out how much sale of a product was due to promotion. One simple measure would be

6. **Beta ($\boldsymbol{\beta}$)**: This measures the proportion of sales of a particular product under promotion. More formally,
$$\beta(j) = \frac{\sum \mathbf{I}(T = j \cap P(T) = 1)}{\sum \mathbf{I}(T = j)} \; \#(1)$$
   Where, **I** is the indicator function defined by: $\mathbf{I}(A) = \begin{cases} 1 & A \; true \\ 0 & else \end{cases}$

   However, this is a very aggregated view and doesn't give us any information about the propensity to change preference of any HH; hence, it doesn't answer the question if a HH will switch preference if j was under promotion. To facilitate more intricate insights, we define the following statistics.

7. **Gamma ($\boldsymbol{\gamma}$)**: The γ of a product j measures if households switch from other products to j due to promotion.
$$\gamma(j) = \frac{\sum \mathbf{I}(T = j \cap S(T) = 1 \cap \Delta P(T) = 1)}{\sum \mathbf{I}(T = i \cap S(T) = 1)} \#(2)$$

   This may be understood as the proportion of switches to product j in which $\Delta P = 1$, i.e. among all the transactions where a HH switched to product j, what proportion was "*due to*" introduction of promotion, as one should understand *if this is high then the sale of product j is*

primarily driven through promotion and all such transactions should be penalized while calculating loyalty.

8. **Omega (ω)**: In a similar tone we also define ω of a product j, which measures if households don't switch from product j once promotion is withdrawn.

$$\omega(i) = \frac{\sum \mathbf{I}(T = i \cap S(T) = 0 \ \cap \Delta P(T) = -1)}{\sum \mathbf{I}(T = i \cap \ S(T) = 0)} \#(3)$$

This may be understood as the proportion of transactions where HH did not switch from product j in which $\Delta P = -1$, i.e. among all the transactions where a HH did not switch from product j, what was the proportion where promotion was withdrawn, *as one should understand if this is high then even though the HH may have started buying the product j due to the promotion they didn't switch back when the promotion was withdrawn developing some kind loyalty to the product and all such transactions should be rewarded while calculating loyalty.*

We will use these three statistics for each product to measure the compound effect of promotion on sale and loyalty - business decisions can also be taken by having a quick look at these numbers only, however in the next section we will use similar concepts to extend and adjust Loyalty for promotion.

C. L-prob

Using definitions used above, we will define two measures of loyalty; but before that we will digress a little and introduce a simple and quick behavioural measure of loyalty we call '*L-prob*'. It measures 100% commitment of a household towards a product. This can be difficult to find in a category such as say juice, as people tend to switch around quickly and this switching is more often than not idiosyncratic. However, this may be very useful in categories such as electronics. That being said let's introduce the measure

$$L(j) = \frac{|\mathbb{L} \cap \mathbb{H}_j|}{|\mathbb{H}_j|} \#(4)$$

In informal terms, this is the probability that a HH never buys anything but the product j. The measure is really simple and easy to compute and interpret, however – it lacks the rigor to capture hidden patterns in the purchase data, we will primarily use this as a baseline to judge the performance of our loyalty measures.

C. Loyalty Measures

Now we are ready to introduce two measures of loyalty $\{\psi_1, \psi_2\}$. Both are similar in anatomy; however, $\psi_1$is based only on past purchases – whereas, ψ_2 also utilises information on promotion during those purchases. Comparison between ψ_1, ψ_2 in the light of statistics β, γ, ω defined above will provide accurate insights regarding the effect of promotion and sale and loyalty of the products.

To give a brief overview, we create '*Loyalty Proportion*' for each household as a function of an *exponentially weighted average of its past purchases* (for ψ_1) along with penalties and rewards that reflect the effect of promotion (for ψ_2). Once we have this, we aggregate results for each household to get an overall value for loyalty that reflects how much the market is loyal to that product after correcting for promotion.

To define *Loyalty Proportion* for a HH we have to first create the following indicator variables:

- $Y_{hj}(t) = \begin{cases} 1, & T_h(t) = j \\ 0, & else \end{cases}$
- $Z_{hj}(t) = \begin{cases} 1, & T_h(t) = j \text{ and } P(T) = 1 \\ 0, & else \end{cases}$
- $U_{hj}(t) = \begin{cases} 1, & T_h(t) = j \text{ and } (T) = 1 \text{ and } \Delta P(T) = 1 \\ 0, & else \end{cases}$
- $W_{hj}(t) = \begin{cases} 1, & T_h(t) = j \text{ and } S(T) = 0 \text{ and } \Delta P(T) = -1 \\ 0, & else \end{cases}$

Now define for some $\alpha \in (0,1)$,

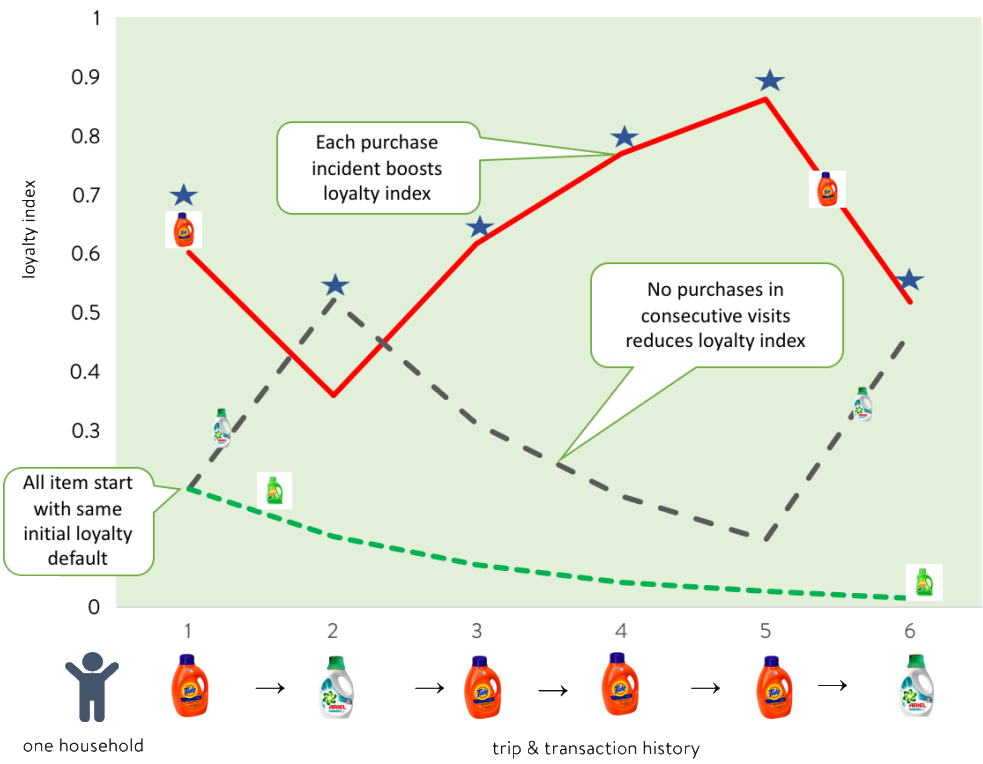
$$Loy_{jh}^{1}(t) = \alpha\, Loy_{jh}^{1}(t-1) + (1-\alpha)\, Y_{ij}(t) \quad \#(5)$$

And a more general version of this using the promotional information as

$$Loy_{jh}^{2}(t) = \alpha\, Loy_{jh}^{2}(t-1) + (1-\alpha)\, Y_{ij}(t) - \beta\, Z_{ij}(t) - \gamma\, U_{ij}(t) + \omega\, W_{ij}(t) \quad \#(6)$$

Where, $\alpha \in (0,1), \beta = \frac{1}{n}\sum_j \beta(j),\ \gamma = \frac{1}{n}\sum_j \gamma(j),\ \omega = \frac{1}{n}\sum_j \omega(j)$

Note, the variable $Loy_{jh}^{1}(t)$, henceforth called *loyalty index*, was introduced in [5] to be used as a predictor variable in predicting customer purchase probability using an *MnL* model. However, for our purposes we will deviate from that and use this as a descriptive statistic that will capture the bias of each households towards a particular product against its alternatives. This can be viewed in the following toy example involving 6 purchases of a household involving 3 items



★ *denotes the item with highest loyalty index after that visit, e.g. if the red line has* ★, *then* *had highest index at that time*

Figure 2: Toy Example 1

The variable $Loy_{jh}^{2}(t)$ adds to this index by penalising purchases made under promotion and rewarding purchases made without promotion and hence will be able to eliminate the effect of promotion while calculating the aforementioned bias. Loy^{1} and Loy^{2} will act as backbones of ψ_1, ψ_2 respectively.

Once we have the *loyalty indices* for each household/product at each time point we proceed to define *loyalty proportion* for the household.

Define,

$$M_h^1(t) = argmax_j \left(Loy_{jh}^1(t)\right) \#(7)$$

$$M_h^2(t) = argmax_j \left(Loy_{jh}^2(t)\right) \#(8)$$

We understand, this value denotes the product which the household h is most biased towards at time t. Since these values are dependent for a household; over time, the long run frequency of these values will be a good proxy for that household's overall bias towards a particular product against its alternatives.

Finally, define:

$$p_h^1(j) = \frac{1}{t_h}\sum_t \mathbf{I}\left(M_h^1(t) = j\right) \#(9)$$

$$p_h^2(j) = \frac{1}{t_h}\sum_t \mathbf{I}\left(M_h^2(t) = j\right) \#(10)$$

$p_h^1(j), p_h^2(j)$ will serve the *loyalty proportion* for the household, in the sense if the loyalty of a household is a constant number this is the share for each product.

Finally, we summarise $p_h^1(j), p_h^2(j)$ across households as follows:

$$\psi_1(j) = \frac{\sum_j p_h^1(j)\mathbf{I}\left(p_h^1(j) > 0\right)}{\sum_h \mathbf{I}(p_h^1(j) > 0)} \#(11)$$

$$\psi_2(j) = \frac{\sum_j p_h^2(j)\mathbf{I}\left(p_h^2(j) > 0\right)}{\sum_h \mathbf{I}(p_h^2(j) > 0)} \#(12)$$

Note, if $p_h(j) = 0$ for some j, h, we don't consider household h for calculating loyalty for the product j. This is due to the fact that generally $|\mathbb{J}| = n \gg t_h$ and hence most of the entries in p_h will be zero – so it will be ambiguous to consider these "disloyal" households while calculating loyalty as they will change the measure by huge margin and the measures will have no meaning.

This completes our definition of loyalty measures for each product under a category – both with and without adjustment for promotion. Before we dive into results from real life data, let us illustrate how

the results actually differ using this simulated data of one household's purchases in a two – product scenario.

Product bought	A	A	B	B	B	A
Promotion index	0	0	1	1	0	0
Δ P	-	0	1	0	-1	0
Z	0	0	1	1	0	0
U	-	0	1	0	0	0
W	-	0	0	0	1	0

Table 1(a): Toy Example 2 - Data

This shows most of B's purchases were a result of promotional activity, however – it also shows some new-born loyalty toward B at visit number 5 where the HH continued with B even when the promotion was taken away. The loyalty numbers for this HH would be

	Basket Share	$\boldsymbol{\psi_1}$	$\boldsymbol{\psi_2}$
A	0.5	0.5	0.67
B	0.5	0.5	0.33

Table 1(b): Toy Example 2 – Results

Expectedly, the fact that B's sales are not organic in nature but are rather induced by promotion is completely ignored by $\psi_1$reflected in the value of ψ_2 and hence although both products are bought exactly three times, our promotion corrected measure of loyalty is able to differentiate between them.

IV. RESULTS

We now present our findings for products in mouthwash category in Walmart Canada. Due to privacy reasons, the authors may not always be able to produce exact names and information in a tabular form and will hence focus on figures and summary statistics to judge validity of our results.

First, a brief overview of the statistics β, γ, ω for the products is shown here:

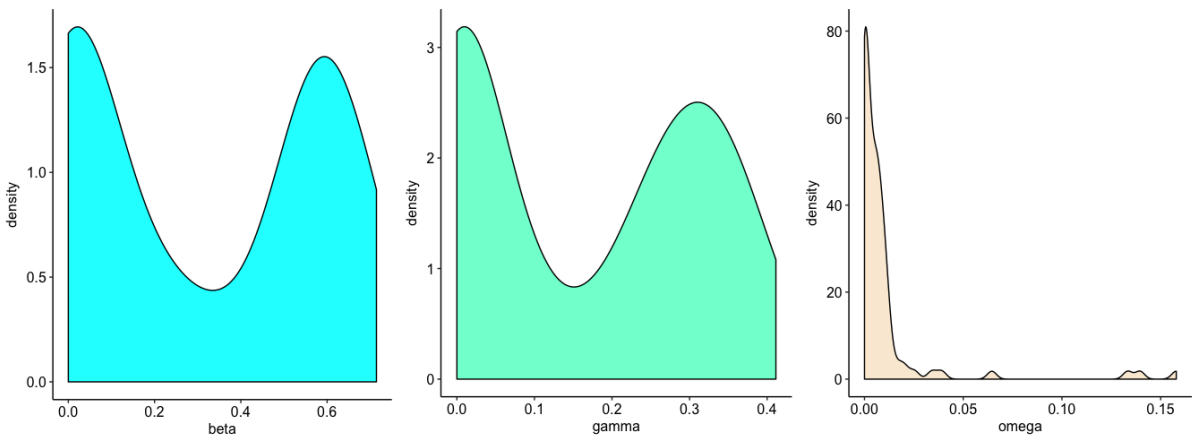


Figure 3: Density of statistics β, γ, ω for Walmart Mouthwash products

So, interestingly β, γ have a similar bimodal distribution – this indicates most of the products were not on promotion for long in the given time, however there are a bunch of products who were on promotion on almost more than 50% of the time, this will create a great contrast in the product universe, we should be able to see vastly differing ψ_2 in products having similar ψ_1. On the other hand, the behaviour of ω is much more routine, most of the mass is concentrated towards zero – which indicates that for most products, very few households were willing to continue with the product once promotion was withdrawn, a passive overview that the promotions were not really effective in building long – term item loyalty.

Proceeding to actual loyalty values – first, we would like to compare the three loyalty measures $\{L_{prob}, \psi_1, \psi_2\}$ of the products with their respective market shares. As we specified earlier our loyalty measure must be able to differentiate itself from market shares. Moreover, as we expect high degree of correlation between L_{prob} & ψ_1. We also add the following figures along with associated correlation matrix between these item performance metrics for comparison.

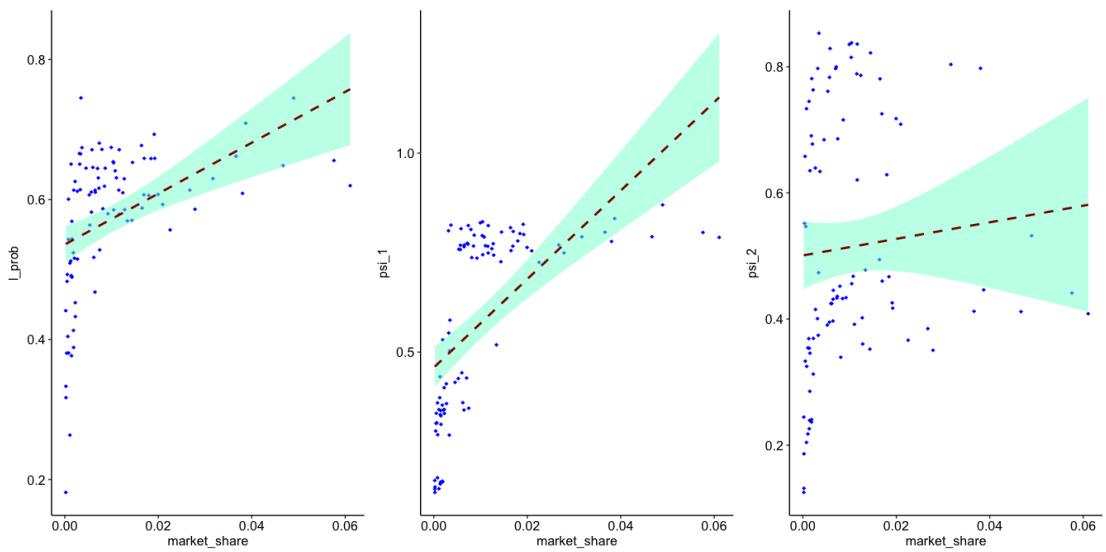


Figure 4: Scatterplot of L_{prob}, ψ_1, ψ_2 with Market share for Walmart Mouthwash products

	market_share	*l_prob*	*psi_1*	*psi_2*
market_share	1.00	0.44	0.60	0.08
l_prob	0.44	1.00	0.78	0.53
psi_1	0.60	0.78	1.00	0.43
psi_2	0.08	0.53	0.43	1.00

Table 2: Correlation between Market Share and Loyalty measures

Let us deep dive into the results,

- First, we see ψ_1 is fairly correlated with market share, however the correlation is not very high, in fact, the plot identifies products which might have low market share but relatively high loyalty, few such UPCs are

UPC	*Average price*	$\boldsymbol{\psi_1}$	*Market Share*
69702932417	$ 11.97	0.76	0.005
6081503995	$ 9.97	0.82	0.003

Table 3: Low-Performance-High-Loyalty Items

- Next, there is a strong correlation between L_{prob} & ψ_1 – this is also expected as described earlier.
- However, by far the most interesting outcome would be the relationship of ψ_2 with other metrics, we have seen the influence of promotion is quite high from the results of figure 3 – As a result, $\psi_2$is virtually uncorrelated with market-share as most of the purchases are done due to promotion and item loyalty is penalised for such instances. We also, saw around 30 - 40 % products have a high β, γ value this suggests large deviations between ψ_2 & ψ_1. This is observed in the data as correlation between them is only 0.43 – were promotions not causing the preponderance of sales we would expect this to be higher since ψ_2 & $\psi_1$will become identical functions.

Finally, we compare some individual products to see exactly where we can see effects of promotion and how that is being adjusted in the definition of loyalty. Consider the products

UPC	**market_share**	$\boldsymbol{\beta}$	$\boldsymbol{\gamma}$	$\boldsymbol{\omega}$	$\boldsymbol{\psi_1}$	$\boldsymbol{\psi_2}$
6260095395	3.66 %	0.6272	0.3827	0.0109	0.8013	0.4122
7906801318	3.17%	0.0004	0.0000	0.0000	0.7896	0.8037

Table 4: Difference between ψ_1, ψ_2 due to promotion: effect of β, γ

These two products have almost similar market share and similar $\psi_1$values, but if we add the promotion information we can see product in the first row (call it P_1) has very high β, γ values and low ω values – customers are only buying this due to promotion, on the other hand product in the second (call it P_2) row has virtually non-existent β, γ values – suggesting the sales are organic; we can see this difference mirrored in the ψ_2 values of the products. We see, $\psi_2(P_2) \approx \psi_1(P_2)$. For P_1, ψ_2 drops drastically. Thus, this measure provides us with a unique view of loyalty where it has successfully eliminated any 'spurious' loyalty created due to promotional activities.

Also consider, this following pair of products (row 1: P_3; row 2: P_4)

UPC	**market_share**	$\boldsymbol{\beta}$	$\boldsymbol{\gamma}$	$\boldsymbol{\omega}$	$\boldsymbol{\psi_1}$	$\boldsymbol{\psi_2}$
5800031086	0.08%	0.5372	0.2877	0.1579	0.1844	0.2043
6260095976	3.87%	0.5935	0.3821	0.0096	0.8356	0.4462

Table 5: Difference between ψ_1, ψ_2 due to promotion: effect of ω

Now, P_4 seems like a typical example from above, high market-share and ψ_1 but sales are actually driven by promotion and hence reduced $\psi_2$values. However, the case of P_3 is actually very interesting. It has similar β, γ values as P_4 – so most of its sales are also initiated via promotion, however we notice it has (relatively) high ω, this means this product has been somewhat successful in retaining customers even after end of promotional activities and hence doesn't see a huge difference between ψ_1, ψ_2; almost as if counter – acting forces balancing each other.

V. CONCLUSIONS

This study presents novel statistical measures for product loyalty based on past purchases of the customers along with promotional information. We attempt at deriving a theoretically robust loyalty measures that aren't only dependent on sequential purchase and are able to successfully ascertain and eliminate effects of promotions, as a result, we can distinguish between products that have high sells and products that actually have high loyalty – this holistic approach helps while making optimal assortment decisions such as deletion, expansion. Apart from these, we have presented simple, interpretable statistics, β, γ, ω and L_{prob} which along with market-share can provide important aspects of promotion and loyalty in a glance due to their intuitive nature.

The future scopes of this study include but is not limited to –

- Enhance this process by introducing some cognitive properties thus making it multidimensional.
- Evaluate loyalties at several time points to test for a seasonality component.
- Evaluate demand transfer dynamics within category/ substitutable products, e.g. what happens when a high / low loyalty product is dropped?

REFERENCES

[1] D. A. Aaker, Managing Brand Equity: Capitalizing on the Value of a Brand Name, New York: The Free Press, 1991.

[2] F. Reichheld, The Loyalty Effect, Boston: Harvard Business School Press, 1996.

[3] S. Ganguli and S. K. Roy, "Generic technology- based service quality dimensions in banking: Impact on customer satisfaction and loyalty," *International Journal of Bank Marketing,* vol. 29, no. 2, pp. 168-189, 2011.

[4] I. Skačkauskienė, N. Vilkaitė-Vaitonė and S. Vojtovic, "Model for measuring customer loyalty towards a service provider," *Journal of Business Economics and Management,* vol. 16, no. 6, pp. 1185-1200, 2015.

[5] J. L. Peter M. Guadagni, "A Logit Model of Brand Choice Calibrated on Scanner Data," *INFORMS,* 1983.

[6] J. Jacoby and D. B. Kyner, "Brand Loyalty vs. Repeat Purchasing Behavior," *Journal of Marketing Research,* vol. 10, no. 1, pp. 1-9, 1978.

[7] K. Mandarino, "Niche Brands: Understanding how niche fashion startups connect with Millennials," *Joseph Wharton Scholars,* 2016.

[8] M. Mellens, M. Dekimpe and J.B. Steenkamp, "A Review of Brand-Loyalty Measures in Marketing," *Tijdschrift voor Economie en Management,* vol. 41, pp. 507-533, 1996.

[9] R. L. Oliver, "Whence consumer loyalty," *Journal of Marketing,,* vol. 63, pp. 33-44, 1999.

[10] G. S. Day, "A Two - Dimensional Concept to Brand Loyalty," in *Mathematical Models in Marketing,* Springer Berlin Heidelberg, 1976.

[11] D. D. Gremler and S. Brown, "Service loyalty: Its nature, importance and implications," *Advancing Service Quality — A Global Perspective,* vol. 5, pp. 171 - 181, 1996.

[12] J. Anselmsson, "Brand image and customers' willingness to pay a price premium for food brands," *Journal of Product & Brand Management,* vol. 23, no. 2, pp. 90 - 102, 2013.

[13] A. Dick and K. Basu, "Customer loyalty: Toward an integrated conceptual framework," *Journal of the Academy of Marketing Science,* vol. 22, no. 2, pp. 99-113, 1994.

[14] S. Hart, A. Smith, L. Sparks and N. Tzokas, "Are Loyalty Schemes a Manifestation of Relationship Marketing?," *Journal of Marketing Management,* vol. 15, no. 6, pp. 541 - 562, 1999.

Analysis of Voting Techniques in Ensemble based Learning

Meenakshi A. Thalor*

* Department of Computer Engineering, AISSMS Institute of Information Technology, Pune, Maharashtra, India
thalor.meenakshi@gmail.com

Abstract- Now a day's data analytics and data science is at boom which works on big data and come up with some hidden patterns after data mining and analysis. Classification process on big data is a challenge as single classifier is unable to give an unbiased decision because of plasticity–stability problem. This paper gives the introduction about the ensemble based learning where multiple base classifiers participate in predicting the class of unlabeled data. Different voting techniques to combine the outcome of different base classifiers of ensemble are also explained. At the end a strict comparison is drawn to show the impact of voting technique in decision of ensemble.

Index Terms- Classification, Base Classifiers, Ensemble, Voting

I. ENSEMBLE BASED LEARNING

Sometimes the data on which one wants to perform learning is so huge that it's not possible to process or analyze such vast amount of data by one classifier. This huge amount of data can be analyze by partitioning data into smaller chunks and pass smaller chunks to multiple classifiers(ensemble) rather than one classifier as shown in fig. 1. Once the learning over different chunk of data is done or model is ready the unseen data can be passes to different classifiers for classification.

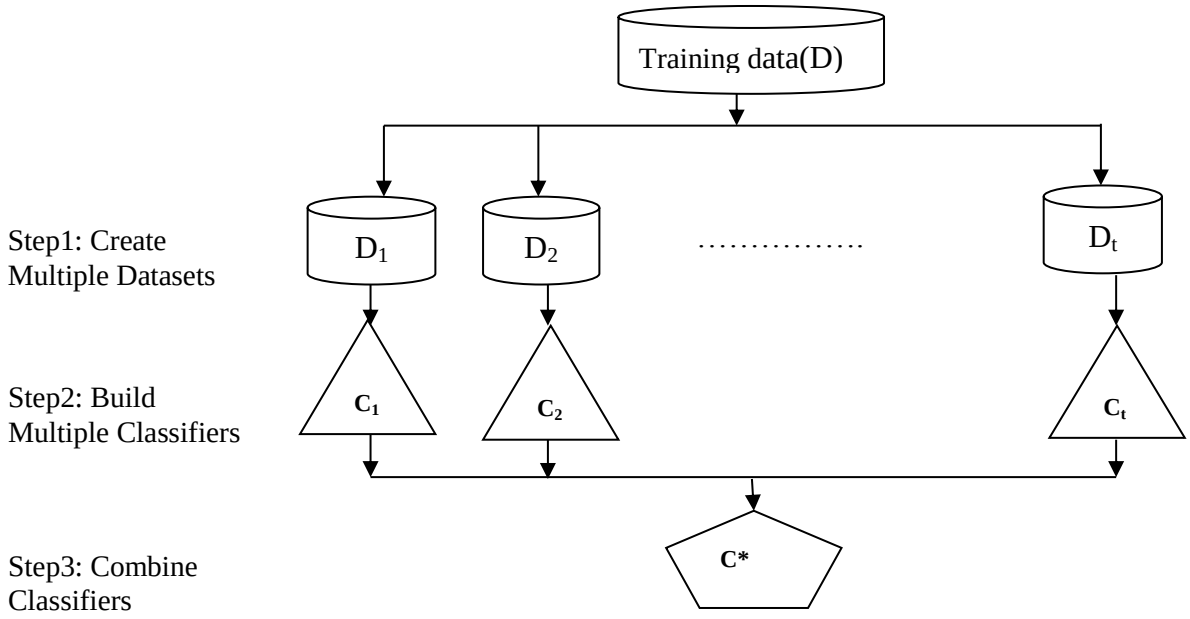


Fig. 1: Ensemble based Learning

Each classifier will predict the class of unseen data then by using different voting approaches one can combine the outcome and come to a final class conclusion. The Ensemble based classification [1,2] achieved by grouping the classifiers by some way then individual predictions of each classifier is combined to classify unseen data.

II. COMPONENTS OF ENSEMBLE

Ensemble is combination of set of diverse classifiers, which collectively do decision making and predict the label for unseen data. Generation of classifiers in ensemble can take place in different ways. Ensemble [3,4,5] can be generated in sequence, in parallel or in layered form. In order to do the classification using an ensemble approach, first one has to consider following elements of ensemble.

A. Training Set

For supervised learning, the basic requirement is a training dataset. A labeled training dataset consists of instances and their corresponding class variable. An instance is feature –value vector which can be represented in a variety of languages. Generally, set of instances are represented by X which consist of n attributes X= $\{x_1, ..., x_i , ..., x_n\}$ and Y to represent the class label or the target attribute.

B. Base Classifier

The base classifier is a classification algorithm that attains a training data and makes a model that specifies the association between the input variables and the class variables.

C. Diversity Generator

The main element which is responsible for creating the different classifiers is called diversity generator. In order to create diverse classifiers in ensemble various techniques are proposed in literature [6,7,8] as mentioned below:

a. *Using diverse training sets:*
   In order to create district classifier, diversity can be introduced in training set by partitioning it into N subsets to train individual classifier. Each member of ensemble will get a different part of training set and unique learning takes place.
b. *Using diverse feature subsets:*
   In order to create distinct classifier, diversity can be introduced by providing separate set of features to train individual classifier. Rather than dividing a training set, each member of ensemble will get a different part of feature/attribute set and unique learning takes place.
c. *Using diverse classifier models:*
   In order to create distinct classifier, diversity can be introduced by combining different types of individual base classifier. Here diversity in ensemble is introduced by keeping same training set and same set of features but using different base classifiers like naïve bayes, support vector machine and decision trees etc.
d. *Using diverse combination schemes:*
   In order to create district classifier, diversity can be introduced by using different types of combination rules. The way you combine the prediction of each base classifier is also introduce diversity in ensemble even though training data, feature set and base classifiers are same.

D. Combiner

The purpose of combiner is to join the prediction outcome of the various classifiers. The outcome of an ensemble depends on the selection of appropriate rule to unite the decision of each classifier. Voting rule is applied at the final step of ensemble system. In literature [9-14], various voting rules are presented which are explained in next section.

III. VOTING TECHNIQUES IN ENSEMBLE

Let x_t is testing instance, each base classifier(h_j) makes an prediction of class label for testing instance x_t which is given by as where Y_i=1….C where C represent no. of class labels.

A. Geometric Average Rule

Geometric Average Rule gives the prediction outcome of the testing instant x_t to that class which maximizes the product of as specified in eq. 1.

$$x_t \rightarrow Y_i \; satisfy \quad \max_{Y_i} \prod_{j=1}^{L} P_j(Y_i|x_t) \tag{1}$$

B. Arithmetic Average Rule

The Arithmetic Average rule is defined as discovering the maximal value of the arithmetic average of $P_j(Y_i|x_t)$ as specified in eq. 2.

$$x_t \rightarrow Y_i \; satisfy \quad \max_{Y_i} \frac{1}{L} \sum_{j=1}^{L} P_j(Y_i|x_t) \tag{2}$$

C. Median Rule

The Median Value rule will select the target class label whose median value is maximum using eq. 3.

$$x_t \rightarrow Y_i \; satisfy \quad \max_{Y_i} \{median(P_j(Y_i|x_t))\} \tag{3}$$

D. Majority Rule

In ensemble, each individual classifier provides the prediction of class label for testing sample and then majority voting rule select the final prediction as the one that takes the majority of votes from each individual classifier in the ensemble as specified in eq. 4. When there is tie among the class labels then a random class label is selected among all tie members.

$$x_t \rightarrow Y_i \; satisfy \quad \max_{Y_i} \sum_{j=1}^{L} \Delta_j(Y_i|x_t) \tag{4}$$

Where

$$\Delta_j(Y_i|x_t) = \begin{Bmatrix} 1; if \; h_j(x_t) = Y_i \\ 0; otherwise \end{Bmatrix}$$

E. Max Voting Rule

Max rule selects the target predicted class label depending on information provided by the maximal value of $P_j(Y_i|x_t)$ across all potential class labels as specified in eq. 5.

$$x_t \rightarrow Y_i \; satisfy \quad \max_{Y_i}\{\max_j(P_j(Y_i|x_t))\} \tag{5}$$

F. Min Voting Rule

Min rule selects the target predicted class label depending on the maximal of the minimal values of $P_j(Y_i|x_t)$ across all potential class labels as specified in eq. 6.

$$x_t \rightarrow Y_i \; satisfy \quad \max_{Y_i}\{\min_j(P_j(Y_i|x_t))\} \tag{6}$$

G. BC Rule

The Borda Count rule uses positional-scoring procedure where individual $Pj(Y_i|x_t)$ provides rank to each class label as specified in eq. 7. Based on the prediction given by each classifier, each classifier will rank all the candidate class labels. For a multi class problem, each class label gets 0 rank least vote received, 1 rank point for each next least vote, etc., up to C-1 points for majority vote received. Finally, the class label that receives largest votes will win the election.

$$x_t \rightarrow Y_i \; satisfy \quad \max_{Y_i} \sum_{j=1}^{L} \Omega_j(Y_i|x_t) \tag{7}$$

Where if classifier h_j ranked x_t in the kth position for class label Y_i , and C is the number of possible classes in multiclass problem.

H. Weighted Majority Voting Rule

A weight coefficient is consigned to each classifier in ensemble and after each prediction the weight of individual classifier is updated based on its performance. Here selects the target predicted class label depending on the weight of each individual classifier as specified in eq. 8.

$$x_t \rightarrow Y_i \; satisfy \quad \max_{Y_i} \sum_{j=1}^{L} w_j . \Delta_j(Y_i|x_t) \tag{8}$$

Where w_j is a weight coefficient for classifier h_j: w_j>=0 and $\sum_{j=1}^{L} w_j = 1$

IV. RESULTS AND ANALYSIS

The main aim of this paper is to determine which voting methods performed best in ensemble. Fig.2 depicts the performance of different voting methods where using diverse base classifier diversity is introduced in ensemble and accuracy is evaluated using different ensemble size. The “Pima Indian diabetes dataset” downloaded from “UCI Machine Learning Repository” is considered here for experimentation purpose. On small ensemble size, the product, min and max voting rules perform well but as ensemble size grow number of correctly classified instances decrease. On large ensemble size Avg., Weighted majority and borda count perform better as they consider the overall voting and less demanding than the confidence method. Lastly, the majority-vote performance is low as it rejects a large

number of samples as they are not majority candidate which results the actual errors are much lower than the recognition performance suggests.

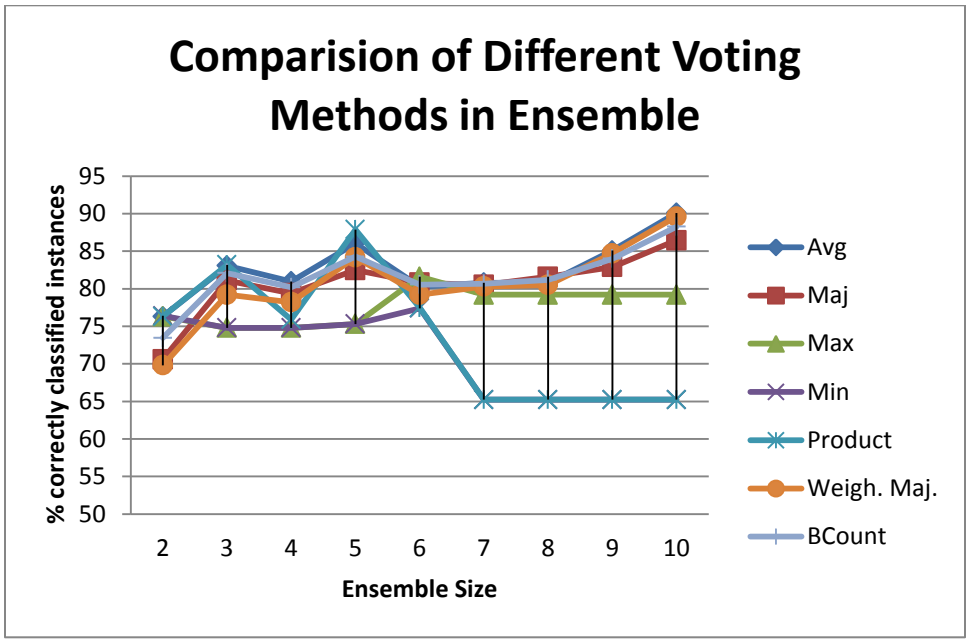


Fig. 2: Results of different voting methods

V. CONCLUSION

This paper gives a detail description of ensemble learning, basic components of ensemble, approaches to combine the decision of each classifier in ensemble and comparison of voting methods with respect to ensemble size. Based on review, conclusion is that the average voting and weighted majority voting are effective approaches as compare to others. The product and sum rule performed best on the smaller sized ensembles, while the Borda count gave good recognition results on larger ensembles. The weighted majority voting significantly improves the prediction accuracy on big data as consider the weight of each classifier.

REFERENCES

[1] L. Rokach, "*Ensemble-based Classifiers*," Artificial Intelligence Review, vol. 33, pp. 1-39, 2010.

[2] Robi Polikar, "*Bootstrap-Inspired Techniques in Computational Intelligence*", IEEE Signal Processing Magazine, 2007

[3] Meenakshi A. Thalor, Dr. S. T. Patil, "*Review of Ensemble Based Classification Algorithms for Nonstationary and Imbalanced Data*", IOSR Journal of Computer Engineering, Impact Factor 1.213, e-ISSN: 2278-0661, Vol. 16, pp. 103-107, February 2014.

[4] Meenakshi A.Thalor ,Dr. S.T.Patil, "*Comparison of Ensemble Based Classification Algorithms*", International Journal of Advanced Research in Computer Science and Software Engineering, Vol. 4 Issue 4, Impact Factor 2.080, pp. 1126-1132 ,ISSN: 2277 128X, April 2014

[5] Meenakshi A.Thalor and S.T. Patil, "*Learning on High Frequency Stock Market Data Using Misclassified Instances in Ensemble*" International Journal of Advanced Computer Science and Applications (IJACSA), May 2016, Vol. 7, Issue 5, pp. 283-288.

[6] Meenakshi A. Thalor and S. T. Patil, "*Incremental Learning on Non-stationary Data Stream using Ensemble Approach*", International Journal of Electrical and Computer Engineering, CiteScore-0.37, SJR-0.164, August 2016,Vol. 6, No. 4, pp. 1811-1817.

[7] Meenakshi A. Thalor and S. T. Patil, "*Learning Framework for Non-stationary and Imbalanced Data Stream*", International Journal of Engineering and Technology (IJET), CiteScore-0.42, SJR-0.17, October – November 2016, Vol. 8, Issue 5, pp. 1942-1945.

[8] Meenakshi A. Thalor, Dr. S. T. Patil, "*Ensemble for Non Stationary Data Stream: Performance Improvement over Learn++.NSE*", IEEE International conference on Information Processing, December 2015.

[9] A.Salini, U. Jeyapriya A, "*Majority Vote Based Ensemble Classifier for Predicting Students Academic Performance*," International Journal of Pure and Applied Mathematics, Vol. 118, no. 24, 2018.

[10] K.Madasamy, Dr. M. Ramaswami, "*Parallel Heterogeneous Voting Ensemble for Effective Classification of Imbalanced Data*", International Journal of Innovative Research in Science,Engineering and Technology, Vol. 7, Issue 8, August 2018.

[11] N. Littlestone and M. Warmuth, "*Weighted majority algorithm*," Information and Computation, vol. 108, pp. 212–261, 1994.

[12] R. Battiti and A.M. Colla, "*Democracy in neural nets: Voting schemes for classification*", Neural Networks, vol. 7, no. 4, pp. 691–707, 1994.

[13] R.E. Schapire, Y. Freund, P. Bartlett, and W. S. Lee, "*Boosting the Margin: A New Explanation for the Effectiveness of Voting Methods*," Annals of Statistics, vol. 26, no. 5, pp. 1651–1686, 1998.

[14] E. Bauer and R. Kohavi, "*An empiricial comparison of voting classification algorithms: bagging, boosting and variants*," Machine Learning,vol. 36, pp. 105–142, 1999.

EEDCR: Energy Efficient Delay Compensation Routing in Distributed Networking System

Dr R. Selvarani*, Dr Vinodha.K**

*Program Director – Doctoral Program, Alliance College of Engineering & Design, Bangalore, India
*Selvarani.riic@gmail.com
**Dept. of ISE, The Oxford College of Engineering, Bangalore, India.
**mohan.vinodha@gmail.com

***Abstract*—** Delay minimization plays a crucial role in ensuring Quality of Service (QoS) and Quality of Experience (QoE) in future internet architecture, which is still in infancy stage of development. In the distributed networking system with heterogeneous communication protocols, it is difficult to reduce delay. As the nodes executing applications on future internet architecture are wireless with mobility features, an adverse effect of delay leads to high energy consumption. In this paper an Energy Efficient Delay Compensation Routing technique, (EEDCR) is presented which is specially designed to sustain the peak traffic condition of the distributed networking system. Supported by simple and cost effective analytical model with a three level frame work viz, domain, gateway node and operator node, EEDCR has proved to be quite efficient in controlling the delay, minimizing energy consumption and improving the throughput. It is also capable of processing more traffic load compared to the existing conventional techniques.

Index Terms-: Distributed Networking System, Delay Minimization, Energy Efficiency, Future Internet Architecture.

I. INTRODUCTION

The existing architecture of internet renders services of different types and also provides the platform for executing number of applications from various domains. The objective of the future internet architecture is to develop a distributed platform for the purpose of enabling wide range of services without any significant impact on the ongoing services [1]. The future internet architecture is actually meant for developing a highly sophisticated and collaborative platform that can develop a well-integrated framework for harnessing different forms of technologies together. A slow pace of development of such architecture is the evolution of Internet-of-Things that connects cloud with sensors [2][3]. The characteristics of future internet architecture are i) provide seamless connectivity between nodes in the distributed networking system, ii) offer faster and reliable accessibility to complex services, iii) leverage parallel processing and task scheduling iv) offer robust security [4][5][6]. In this paper we discuss the problem of delay and energy consumption which are considered as the main components of the heterogeneous devices. The upcoming applications that utilizes the services of future internet architecture are highly distributed in nature, hence these devices require an effective communication protocol [7][8]. This leads to the problem of compute time, which is the time required to process routing message from one network device (e.g. switch/gateways) to other network device. The computation of processing time is highly detrimental for real-time applications and it results in delayed services. Hence, delay is one of the critical factor that play a significant role in both QoS (Quality-of-Service) and QoE (Quality-of-Experience) in upcoming distributed applications [9]. It is observed that, there is only a limited focus on the analysis on delay and the energy consumption in a situation like highly Distributed

networking system. Hence, future internet architecture (FIA) is quite a complex structure that calls for a closer investigation on different networking factors and routing mechanisms.

Here, we consider the energy associated with every node involved in communication at that instant. The energy optimization depends on routing and data processing policy of a particular communication technique. Hence, it is noted that if a node suffers undue delay, the node energy will be reduced.

In this paper, we present a novel technique that emphasizes on delay minimization followed by energy optimization exclusively for distributed network system. We discuss some of the research works related to delay minimization and energy optimization in Section II followed by brief discussion of problem identification in Section III. The proposed system and its significant contribution are presented in Section IV. Section V discusses the research methodology followed by Algorithm implementation in Section VI. The accomplished simulated outcomes are discussed in Section VII and finally the summary of the paper is presented in Section VIII as the conclusion.

II. RELATED WORK

The techniques introduced by various researchers for addressing the problems caused due to delay in a distributed network system are discussed in this section.

Bie et al. [10] have presented a technique for modeling the delay in vehicular traffic system. The author have discussed about the queuing model for minimizing the delay in the communication process. The outcome of the work was assessed using prediction error percentage with respect to time interval. A study towards vehicular communication system is being carried out by Tiwari and Kumar [11] recently. Using the simulation-based approach, the authors have discussed a scheme for authenticating the cache at the gateways that significantly reduces the broadcasting operations; by this the delay is minimized. The outcome of the study was tested with response time of a query.

Gancet et al. [12] presented a technique that attempts to minimize the communication latencies for underwater communication. They have designed and developed a remotely operated vehicle with an aid of satellite and communication interfaces to link onshore and offshore communication modules. Study on reducing the delay in planetary networks was carried out by Zhang et al. [13]. The authors have presented an algorithm that can significantly control topology and hence minimize the delay.

Byun et al. [14] have presented a technique to address the problems of delay and energy efficiency in a communication system. The author has used slotted ALOHA protocol [15] for the purpose of minimizing the delay and enhance the rate of transmission. Similar studies on delay and energy minimization was also carried out by Park et al. [16]. However, the authors use controlling mechanism to control voltage and frequency in microprocessor. The authors considered case study of DC-DC converter of PID controller.

Raza et al. [17] have introduced a system that can compute the cumulative delay of the links over multiple networks. Here the study emphasized more on extracting the correlation among the delay with other associated network metrics. Saleh et al. [18] have discussed about minimizing energy dissipation by scheduling the transmission rate for a cognitive network with multiple accesses. The outcome of the work was tested with respect to throughput and delay.

Li et al. [19] have used Markovian chain for analyzing the tradeoff among delay, data quality, and energy for sensor networks using distributed algorithm. Motoyoshi et al. [20] have designed a system that enhances the quality of communication for future mobile networks considering OpenFlow clusters. The outcome of their work was found to minimize the cost significantly.

Hence, it can be seen that there are adequate studies that focus on delay minimization. The next section describes the problems in the existing system.

III. PROBLEM IDENTIFICATION

Future Internet Architecture uses complex distributed networking system where the nodes of different domains communicate with each other through a gateway node. The existing border gateway protocol is applicable for communication in the existing internet architecture but not for future internet architecture because of its dynamic nature. The similar problem exists with energy issues too. The following are the various issues, to be answered for the future internet architecture growth and development:

- How to calculate delay in heterogeneous network domain?
- Even if we calculate delay in heterogeneous network domain, how it is going to be helpful in designing delay compensation technique? (as it may lead to lot of stale routing information)
- Does delay affect energy utilization in distributed network?
- How to develop a simulation model considering heterogeneous distributed network and then address delay problem?
- Should we go for delay calculation at the end (i.e. delay efficient approach) or should be keep on estimating delay in progress of transmission (i.e. delay aware approach)

The distributed network possesses many network domains where the nodes in each domain may be mobile. Mobility introduces dynamicity in the network domain, where there is a higher probability of delay (due to data packet conversion, searching for definite route, existing job processing etc). We have also illustrated in our prior work that futuristic internet-based architecture will use more heterogeneous network as compared to homogeneous network. Developing such type of network topology will obviously invite various challenging scenarios in distributed networking system. Even if our prior framework uses stochastic rate control, cross layer control, Markov modeling (for congestion control), it doesn't address delay compensation. If delay compensation is not carried out in distributed network, it could lead to serious violation of SLA (as future architecture needs cloud environment potentially). Hence, the problem statement of the proposed study can be stated as "*It is computationally challenging to formulate a unified algorithm for delay minimization and energy efficiency for distributed networking system*" The following section presents the proposed system in order to address this issue.

IV. PROPOSED SYSTEM

The proposed system aims at minimizing the communication delay and optimizing the energy consumption in distributed networking system. Our prior study has already laid an emphasis on the need for addressing the complexities in distributed network system [21][22]. Our recent works have also introduced a Rate Control

Metric (RCM) that is responsible for directly controlling the traffic flow rate to minimize the traffic congestion in future internet architecture [23]. This paper is a continuation of the prior work, where we lay emphasis on the delay factor which was not addressed in our earlier studies. We name the proposed technique as Energy Efficient Delay Compensation Routing that is meant for smoothening and normalizing the traffic situation in peak hours especially built for distributed networking system. Fig.1 shows the architecture of the proposed system EEDCR. The proposed work develops a simulation study with distributed nodes, different network domains (to implement heterogeneous distributed network) with one test-destination point for assessing the time duration of receiving the data. A tree-based approach is used as it can ensure that data dissemination time will not be increased beyond the controlled level. Finally, the proposed study develops a model that can calculate the energy consumed by the transceiver in a distributed network. In the entire process of modeling, the prime focus is on computing the time required for performing communication on different scenarios of distributed computing. The significant novelty of the presented work is that it is simple and a potential technique for energy saving. It can jointly conserve energy (apart from calculating it) along with selection of the communication channel with less amount of delay prior to formulate decision of routing in highly challenging dynamic architecture. The objectives are as under:

- To develop a framework which can provide a balanced routing process between two major parameters Viz. energy efficiency and delay factor in the distributed networking system.
- To develop an analytical model that can conceptualize the routing mechanism for future internet architecture by incorporating new modules like domain, gateway node and operator node. Here, domain represents a group of nodes.
- Finally, framing a novel energy efficient routing architecture with delay compensation for heterogeneous environment (EEDCR).

The next section discusses about the research methodology followed by the implementation of the algorithms.

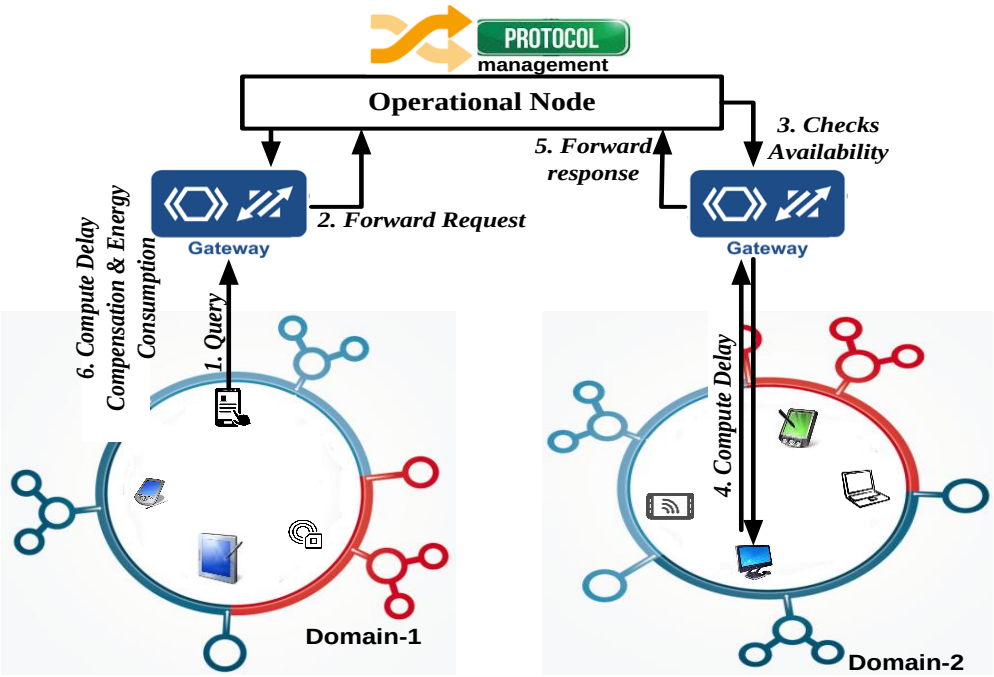


Figure:1 Proposed Architecture of EEDCR

V. RESEARCH METHODOLOGY

The EEDCR is analyzed using empirical and mathematical models. Through these methods the communication delay is minimized. The formulation of EEDCR is discussed below.

A. *Core Modules Involved:*

Usually, the future internet architecture will consist of number of nodes, which are obviously quite heterogeneous in nature. We introduced two types of nodes namely *gateway node* and *operator node* along with conventional concept of *domain*. A *domain* is formulated by a group wise communication of homogeneous networks of similar applications or with different applications if they follow similar routing protocols in distributed networking system. The member nodes present inside the domain can communicate with each other but not with members of different domains. This is because of the incompatible compliance of routing protocol on different domains. This problem is addressed by considering a *gateway node* as shown in fig 2 that can process messages from different domains and also assist in routing process. The energy is managed at an optimal level by the use of a third party node called as *operator node*. A gateway node is responsible for gathering the data from its domain and forwards it to *operator node*

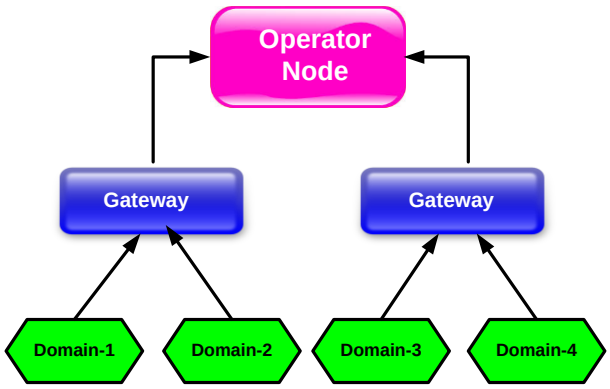


Figure 2 Relationships among Domain, Gateway and Operator Node

The gateway nodes maintain the details of other gateway nodes in-order to process the request generated by the member nodes of one domain to member nodes of other domains. The information gathered by the gateway nodes is forwarded to the operator node which is an intermediated node in the communication network that identifies the gateway node closer to the destination node.

Routing Technique Applied

The routing technique for communication among nodes in heterogeneous network is shown in (Fig.3). The routing process is initiated by forwarding a query generated by a node located in a particular domain (say D_1). The density of nodes in a domain is uneven, just to assume real time situation. The query forwarded by the node is received by its respective gateway node. Normally, a gateway has the list of the entire registered operator nodes, which must get connected in order to deliver the message to another domain. In our proposed system, we do not establish a direct contact among the domains in order to minimize the energy dissipated due to the data processed by the gateway. As the spatial distance and traffic intensity of the operator nodes may differ from each other, the requesting gateway node has to select an appropriate operator node that can ensure less delay and high energy efficiency.

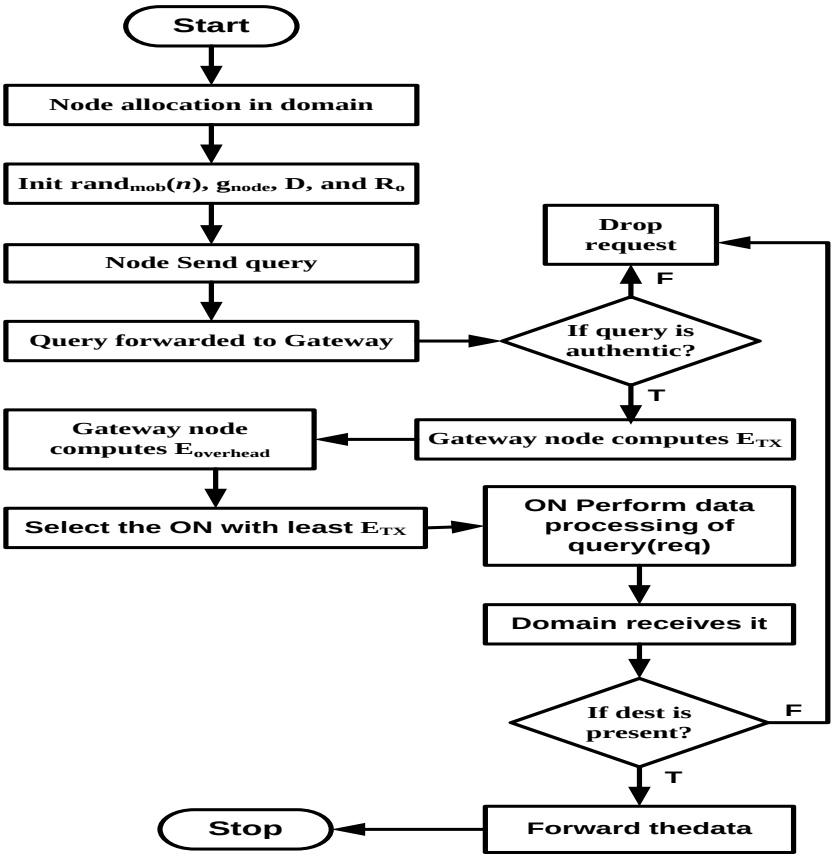


Figure 3 Routing Technique of EEDCR

Upon receiving a request from the gateway (say Gateway-1), the Operator Node (ON) searches for an appropriate gateway and instantly responds to the requesting gateway with total time required to process and transmit the data to the end node. The requesting gateway node broadcasts the request to all the available ON and in turn receives the responses that have time duration details in terms of latency. The requesting gateway node then computes delay and also calculates how much delay will be required to compensate for all the response along the different routes. Based on the route with lowest energy depletion and least delay, the requesting gateway node selects the final ON forwards the same and processes the data to its destination gateway node (say gateway-2) reaching the final destination node.

Delay Compensation

In order to compute the delay, the proposed EEDCR determines the time duration required for performing specific task with respect to given number of modules involved. The process flow of delay compensation technique is highlighted in Fig.4, which shows that initially, the system starts allocating different number of nodes in all the domains that computes time slots for all the essential routing operation during peak traffic condition in the distributed network. The computation of timeslots potentially depends on size of network and number of domains involved in the process.

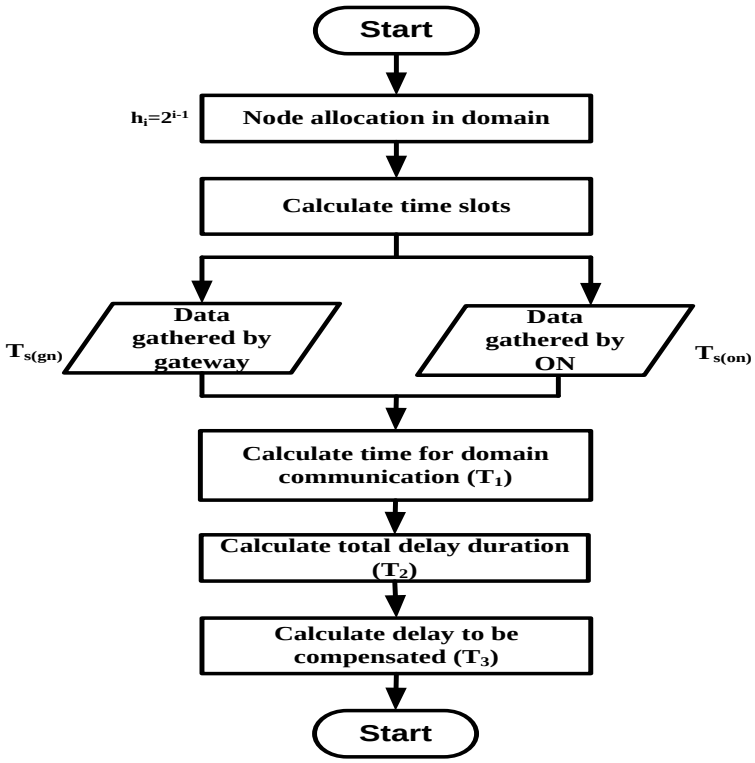


Figure 4 Delay Compensation Technique of EEDCR

The system computes the duration of the time required for data gathering by the gateway node $T_{s(gn)}$ and the operator node $T_{s(on)}$. The system then computes the total time T_1 required for domain-based communication which is then followed by the computation of total delay (T_2) using probability theory. Finally, based upon the response from the ON, the requesting gateway node calculates the exact amount of delay to be compensated. This technique is similar to routing as it complements the routing during the situation of congestion in future internet architecture.

Energy Optimization

The energy efficiency of the EECDR is based on observing the amount of energy being dissipated during routing process using threshold-based scheme in order to control avoidable energy depletion. The schema of energy efficiency and control is highlighted as in Fig.5.

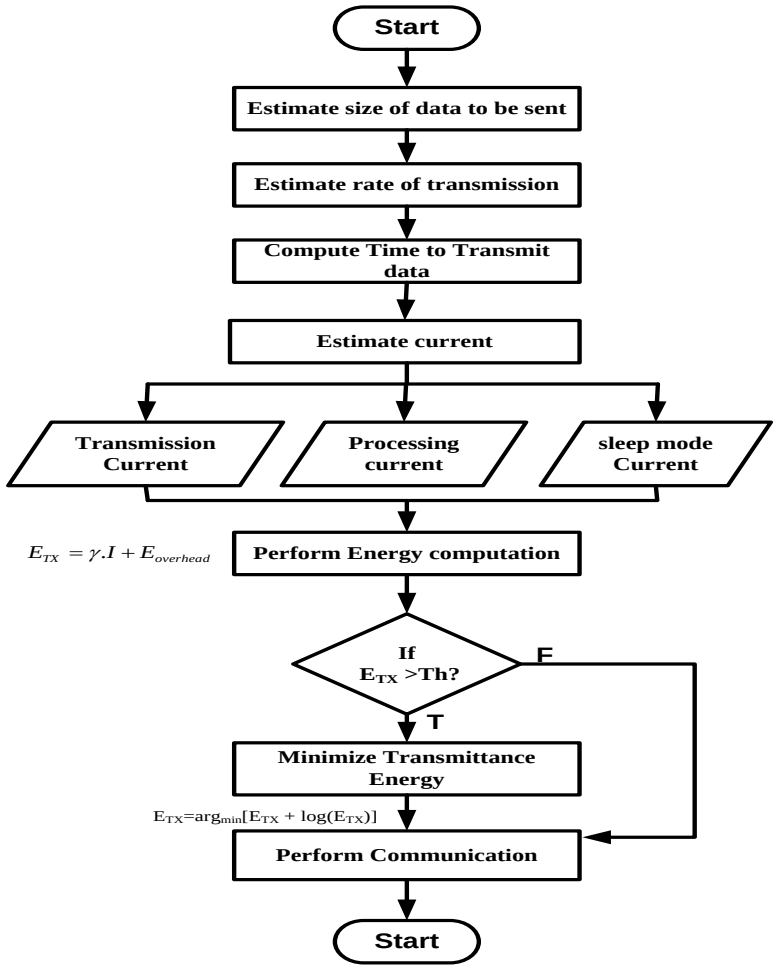


Figure 5 Energy Efficiency Technique of EEDCR

VI. ALGORITHM IMPLEMENTATION

The development of EEDCR is carried out using mathematical modeling and implemented through MATLAB. The complete design of the proposed algorithm is focused on developing and incorporating delay calculation and compensation techniques during the peak traffic condition in distributed networking scenario. The important simulation attributes used for the proposed study is as shown in Table1.

Table 1 Simulation Attributes used

Simulation Attributes	Values Used	Simulation Attributes	Values Used
Size of network	500-1000	Sensing time of Channel	0.5 sec
Simulation area	1000 x 1000 m^2	Size of Control Packet	30 bits
Simulation Time	500 seconds	Size of Data	3000 bytes

Path loss exponent	0.5	Highest velocity of node	100 ms^{-1}
MAC Type	802.11	Lowest velocity of node	1 ms^{-1}
Traffic Model	CBR / VBR	Range of Transmission	10 meter
Capacity of Channel	300 mbps	Initialized energy of node	10 Joule

The uncertain traffic condition will always lead to delay in the network. Owing to the distributed nature of the traffic, it is quite difficult to compute the delay factor which also leads to massive energy consumption among the nodes. The description of the algorithms used in design of EEDCR techniques is given below:

i) Algorithm for Delay Compensation

The initial operation of this algorithm begins by exploring size of the network and its relationship with the domain. We assume that the size of the network *S* is limited between *D* to (*D*-1) domains in each of the iteration, where *D* will represent last domain and (*D*-1) will represents its previous domain. The algorithm than finds the highest number of the nodes in i^{th} domain that can be represented as $h_i = 2h_{i-1}$. Therefore, line-2 of the algorithm will represent the maximum number of domains in progression. The size of the network is controlled by lower limit of 2^{D-1}-1 and highest limit of 2^{D}-1. During the transmission, it is quite natural that data packets will be compressed by the forwarding nodes. The proposed work doesn't focus on compression much, but still it considers a hypothetical rate of compression in order to justify the utilization of generic compression technique. The EEDCR structures and arranges its nodes in the form of a tree in *D* domain. The proposed algorithm focuses on the allocation of the nodes particularly in order to minimize the delay by allocating the initial (*D*-1) domains with certain number of nodes. Hence, the time interval scheduled for a gateway node is represented as

$$T_{s(gn)}=S-2^{D-1}.$$

Whereas the second line of step-3 shows the time interval for a specific schedule. This will also mean that $T_{s(gn)}$ is the time interval spent for gathering the data by the gateway node and $T_{s(on)}$ is the time interval for transmitting the data to the operator node.

In this section we calculate the time required for domain-based communication (T_1). A closer look into the time calculation T_1 is the addition of (2^{D-1}-1) and (S-2^{D-1}+1). Therefore, we use a function *f*(x) in equation(1) to compute the time T_1, which is cumulative period of time for the initial (D-1) domain to gather data and then forward the same to the operator node considering the rate of compression.

$$T_1 = f(x) + |\delta^{T_{s(gn)}} + \int_{x=1}^{T_s(gn)} \delta^x dx| \qquad (1)$$

The second component of the equation (1) is cumulative rate of compression starting from gateway node to operator node considering the maximum time duration to be equivalent to $T_{s(gn)}$ expressed in setp-3 of algorithm. The delay calculation is carried out using equation (2) considering another function *g*(*x*) that considers number of selected operator nodes.

$$T_2 = h_{i(D-1)} - 1 + g(x) \qquad (2)$$

It means that EEDCR can control the overhead in the selection of operator node by the gateway node. We performed majority of the testing using 3-7% of selection criteria in order to compute T_2. Finally, delay to be compensated is computed using equation (3) which uses further two functions i.e. α(x) and β(x).

$$T_3 = \alpha(x) + \beta(x) \qquad (3)$$

The function α(x) is computed using maximum time consumed on respective domain and β(x) is computed by summing up all the rate of compression being considered for each gateway nodes. The computation will lead to generation of T_3, which is the delay associated in the network during the peak traffic condition. The computation of the delay is carried out within the gateway node *gn* considering its associated connectivity with its nodes in one domain and other operator nodes.

Algorithm for Delay Compensation

Input: S (Size of Network), D (Domain), δ (rate of compression).

Output: Delay duration (T_2) and delay to be compensated (T_3).

Start:

1. Init S

2. Allocate nodes in domain

$h_i=2^{i-1}$

3. Perform Domain Allocation

$T_{s(gn)}=S-2^{D-1}$

$T_{s(on)}=S-2^{D-1}+1$

4. Calculate time for Domain Communication

$$T_1 = f(x) + |\delta^{T_{s(gn)}} + \int_{x=1}^{T_s(gn)} \delta^x dx|$$

5. Calculate total delay duration

$$T_2 = h_{i(D-1)} - 1 + g(x)$$

6. Calculate delay to be compensated

$$T_3 = \alpha(x) + \beta(x)$$

End

ii) Algorithm for Energy Efficiency

The computation of energy is strongly associated with the selection procedure of the operator node. As there are various operator nodes and heterogeneous domains, the complexity could be further increased resulting in unwanted energy exhaustion. Hence, we formulate an algorithm where a novel technique is applied to compute the amount of energy drained and a simple threshold-based technique to control energy dissipation of the nodes.

Algorithm for Energy Efficiency

Input: θ_1 (size of data to be sent), θ_2 (rate of transmission) , I_1 (current required to perform transmission), I_2 (current require to perform algorithm processing), I_3 (current required in sleep mode of node), Th (Threshold)

Output: Minimized Transmittance Energy

Start

1. init θ_1 and θ_2.

2. Estimate time to transmit data

   $t_s = \theta_1 / \theta_2$

3. Compute power drawn

   $I = I_1 + I_2 + I_3$

4. Perform Energy computation

   $E_{TX} = \gamma.I + E_{overhead}$

5. do(

   $E_{TX} = \arg_{min}[E_{TX} + \log(E_{TX})]$

6. While $E_{TX} > Th$

7. Perform communication

End

The processing of this algorithm is carried out in the gateway node that has to select the operator node with maximum residual energy based on preliminary evaluation (before performing the routing).The algorithm for energy efficiency considers the time required to transmit selected data from one node to gateway by dividing size of data to be sent with respect to rate of transmission. Hence, it is quite clear that the value of t_s in step-2 will differ in all iterations of routing. The novelty of this technique the emphasis lied on the power used during the communication process. For analysis purpose, we consider usage of three different sources of power i.e. power required for performing transmission, power required to perform algorithm processing, and power required in sleep mode of the node.

The transmitted energy is calculated using equation (4) that also includes energy that is expended due to overhead ($E_{overhead}$).

$$E_{TX} = \gamma.I + E_{overhead} \qquad (4)$$

A simple logic is applied to control the energy dissipation by using threshold energy. As majority of the electronic and communication devices comes with standard specification of energy efficiency [24], hence, it is possible to generalize the cut-off energy of all such devices in order to perform optimal communication operations. The computation explained in step-5-6 is carried out by gateway node during selection of operator node. The probability of selecting an operator node by a gateway is around 3-7%. The gateway chooses the operator node through iteration in-order to find the one which yields lower value of E_{TX}. Hence, the logic

written in sep-5 in do statement is about lowering the value of E_{TX} by the log value of it. For an effective computation we use a scale that represents 10 Joules (initialized energy) using 1 unit of probability. This means that the algorithm considers lowest value of E_{TX} as 0 while highest value equivalent to 0.9 units. This consideration is intentionally done so that log of any value within the limit of 0-0.9 will generate negative values that lower the energy consumption during routing.

VII. RESULTS & DISCUSSION

This section discusses the outcomes of the proposed system. As the proposed EEDCR system emphasizes on delay minimization and energy conservation techniques the performance parameters viz. delay estimates, throughput, energy consumption, traffic load are chosen so that they can scale effectively on peak traffic condition with respect to increasing iteration rounds.

Rationale of Comparative Analysis

The proposed system is reviewed by considering the study that has similar concept of delay minimization and very closely associated with it. In order to measure the effectiveness of the proposed EEDCR, we compare the outcomes with the work done by Asadi [25] and Hong [26]. Asadi have developed a technique called as DRONEE (Dual-radio opportunistic networking for energy efficiency), which is essentially meant for achieving energy efficiency for futuristic networking applications e.g. LTE (Long-Term Evolution) with opportunistic scheduling. DRONEE is also characterized by distributed network architecture and is designed for enhancing the quality of signal focusing on energy saving features. The design of this technique also allows a node residing in one cluster to communicate with node located in another cluster. However, the study [25] chooses homogeneity in designing nodes, clustering, and routing process with a target to conserve maximum possible energy in uplink transmission. The study carried out by Hong [26] has focused on using distributed computing for retaining maximum throughput, delay, and energy efficiency. The author have used Slepian-Wolf bound [27], Markov chain, and adaptive MAC for controlling delay, energy in the communication process. The next section discusses about the outcomes accomplished from the simulation study.

A. *Analysis of Delay Estimates*

The delay is estimated by considering the time duration for a data to travel from sender node in one domain to destination node in another domain via gateway and operator node. The outcome highlighted in Fig.6 showcase the comparative analysis of the delay estimates which shows that EEDCR has a superior delay performance as compared to Asadi [25] and Hong [26]. The method presented by [26] is based on the Slepian-Wolf technique that performs source coding in a distributed nature and ALOHA protocol for enhancing the throughput. The problem with this method is the adopted sequential encoding mechanism which consumes more time for processing the algorithm especially on heterogeneous distributed networking system. Another problem of this approach is the usage of linear recursive operation which consumes more processing time in order to identify the skew in the delay vector and then perform minimization of it. Hence, it has maximum delay.

The DRONEE algorithm has mainly focused on performing clustering mechanism on LTE-based network considering the effect of stationary fading. DRONEE doesn’t have any form of recursive-based operation during clustering and hence it can assists in reducing the time duration of the data transfer as compared to Hong [26] approach. Alternatively, his method uses modulation and coding techniques in order to maintain

better fairness in its throughput. The modeling also includes computation of various power-related factors resulting in increased delay when compared to proposed EEDCR. Therefore, EEDCR reduces its computational process through multiple threading processes that results in simultaneous balance between data delivery and energy efficiency. Moreover, channel errors are less in EEDCR as the delay incurred in routing is estimated before the original routes are established. As a result it is evident from the graph depicted below that after 300ith iteration the delay is almost linear in case of EEDCR whereas the other approaches experiences maximum delay. Therefore, proposed EEDCR provides an efficient routing mechanism with a better delay performance as compared to existing techniques.

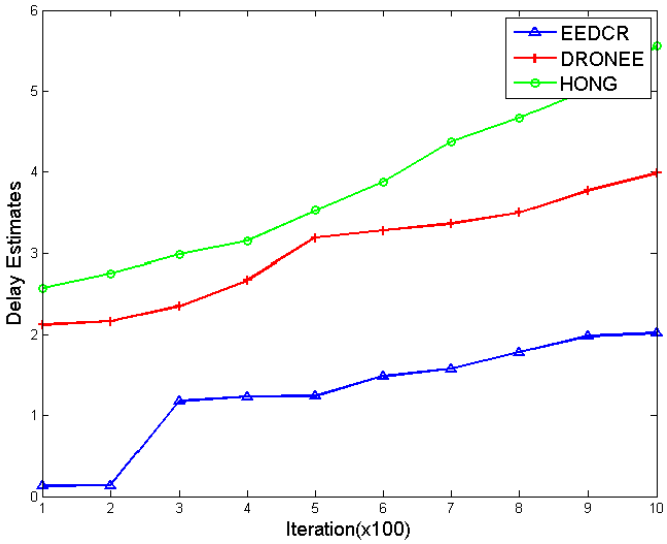


Figure 6 Analysis of Delay Estimates

B. *Analysis of Average Throughput*

The QoS parameter throughput is considered for measuring the effectiveness in communication in the scenario of an additional operator node in EEDCR. It is computed as the cumulative amount of the data relayed from the gateway and the same being processed by the operator node per unit time (second). The result depicted in Fig.7 reveals that the average throughput of EEDCR has comparatively better performance in contrast to existing approaches.

Although the Hong [26] approach was developed for improving the throughput, but its design principle was found to take the decision of enhancing the throughput only based on probability analysis. In order to perform probability analysis, the approach is highly dependent on continuous arbitrary data and lacks capturing of events that occur at preliminary routing stage. Another significant impediment is the dependency on static sink for data processing, which reduces the throughput in a heterogeneous environment. On the other hand, the DRONEE algorithm makes use of the cooperative communication that assists in faster delivery of data packets in shortest interval of time for LTE networks. However, it also performs this task considering the probability of scheduling using enhanced round robin mechanism by selecting the channel that can assure QoS. Hence, both these approaches do not have significant change in their throughput. EEDCR addresses this problem by introducing operator node that is designed for performing data conversion from one domain to another domain during the routing process. Here, because of the pre computation technique associated with the EEDCR system in identifying the best route at the gateway and operator node which does not have any resource constraint, there will be significant enhancement in throughput quality of the network. From the

average throughput analysis graph it is noticed that at after 500th iteration the EEDCR ensures improvement it the throughput where as the other two approaches do not show specific changes. Hence, EEDCR has maximum throughput performance.

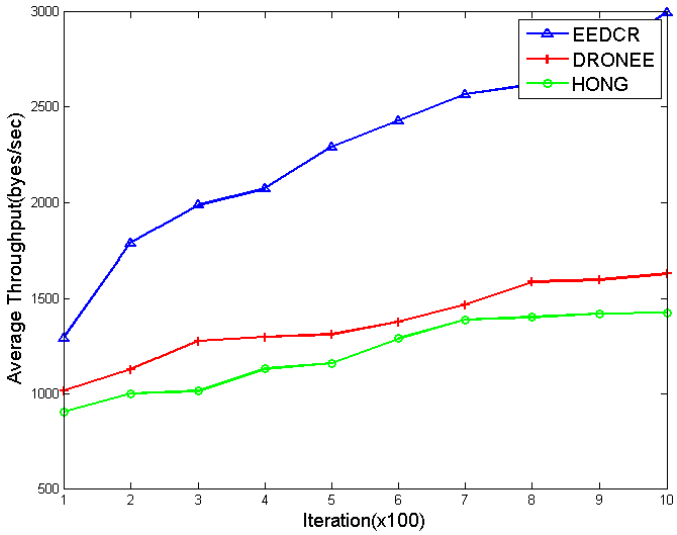


Figure 7 Analysis of Average Throughput

C. *Analysis of Energy Consumption*

The Future Internet Architecture may operate with low power devices. The nodes are operated at maximum energy level for routing process. It is to extend the life time of the network and enhance the quality of service. The comparative analysis of energy optimization in EEDCR and existing approaches is presented in the fig 8.

The approach presented by HONG [26] has emphasized on clustering process in wireless sensor networks to reduce the energy consumption, it does not implement any optimization technique. It is due to the demand of the algorithm (recursive operation and probability theory) for recursive operation. When this approach was implemented on our scenario considering a scale 0 (lower battery state] and 1(higher batter state), we found that it is incompatible with the heterogeneous devices in perspective to energy conservation. The Slepian-Wolf method has proved efficient for small sensor operations, but fails to retain energy. This is because of the utilization of energy towards data processing. The results depicted in the graph shows that the performance of the prior approach with respect to energy consumption was better till 500th level of iterations owing to usage of Markov chain process. As the iteration process was increased, the energy dissipation also increased, because, in this system maximum number of operations are used for counting the state of the messages. On the other side, DRONEE [25] used clustering mechanism for energy optimization but when it was applied in our environment it spends most of its energy in performing opportunistic scheduling. Hence, we conclude that there is no significant difference in the outcomes of both these approaches. The proposed EEDCR adopts a simple threshold-based technique to compute the delay on the routes. It involves less overhead and results in accurate delay calculation that leads to significant conservation of energy. The behavior of EEDCR (Fig.8) with respect to energy is quite predictive in nature after 600th round of simulation in heterogeneous distributed networking system where as the other two approaches exhibits irregularities in their performances.

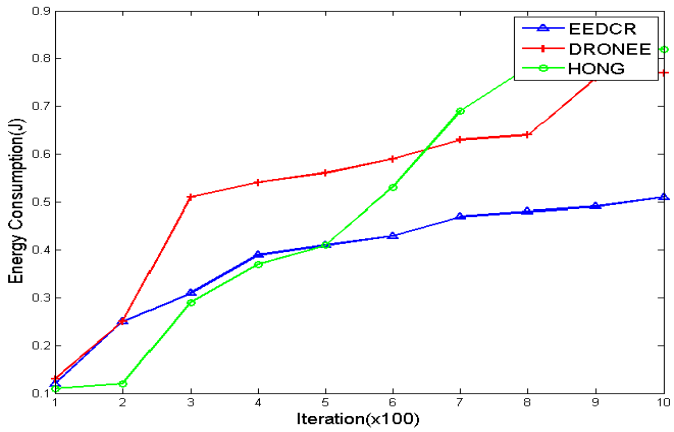


Figure 8 Analysis of Energy Consumption

D. *Analysis of Average Traffic Load*

The traffic load is estimated by considering the number of request being generated over a period of time and the mean time required to process the same. It measures the sustainability of the system with increasing traffic load. It is assumed that if a system has the property of normalizing the traffic, then it exhibits a curve with gradient decent to prove its effectiveness in managing the traffic load under peak traffic condition.

Fig.9 shows the analysis of the average traffic load. The Hong [26] uses arbitrary access mechanism and distributed source coding in a homogeneous network. However, when the simulation environment is changed to a system with heterogeneous nodes, this approach is found to consume more time in encoding the groups. From the average traffic load analysis graph it is proved that the existing mechanism is good for sensor network but not ideal for dynamic networks and hence the system is having lesser capability to sustain increasing traffic load with increasing rounds of iterations.

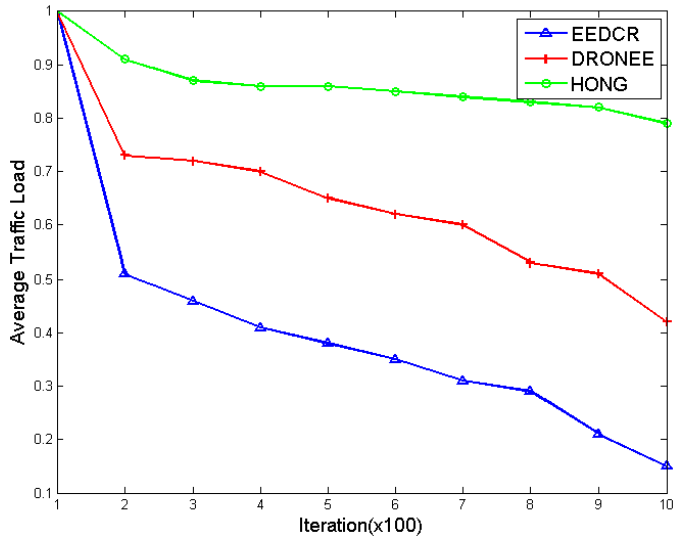


Figure 9 Analysis of Average Traffic Load

The DRONEE algorithm used cooperative communications scheme among wireless relay nodes to achieve high rate of data transmission. This feature was not enabled to compute delay during or before the routing process. The pre computation technique associated with the EEDCR system in identifying the best route for communication at the gateway and operator node assists in leveraging the capability of the node to sustain

increasing traffic load. The simulation results depicted in Fig 9 reveals that the performance of EEDCR decreases linearly with increasing rounds of iteration and traffic load. The following section presents the quality analysis of the proposed EEDCR to conclude that it excels better in its performance when compared to existing techniques with respect to multiple performance parameters.

Quality Analysis

The performance parameters considered for quality analysis are delay, average traffic load, average throughput and energy consumption. The following are the observations arrived from results discussed in section VII.

The delay estimates depicted in the fig 6 shows that at the initial iteration rounds EEDCR have the minimum delay of 0.2 milliseconds. This is because the delay compensation algorithm helps in finding routes with less delay before the routing process is initiated. Absence of this technique in the existing approaches increases the delay to 2 milliseconds in DRONEE algorithm and 2.6 milliseconds in Hong. As the number of iteration rounds increases the delay in EEDCR rises to 1 millisecond and later it almost increases linearly by 10%. In case of the existing approaches they exhibits random variations and therefore the behavior of these systems are unpredictable in future. The average throughput of EDDCR presented in the fig 7 revels that as the number of iteration rounds increase the throughput also increases by 15% for every round of iteration. This is because of the additional operator node used in EEDCR to identify the best routes. The approaches adopted by DRONNE and Hong do not show remarkable changes with respect to average throughput. The comparative analysis of energy consumption by EEDCR showed in fig 8 scales linearly by 0.05 joules for every one step increase in iteration. This controlled dissipation on energy improves the lifetime of the network. In case of DRONNE algorithm the energy consumption increases to 0.5 joules at the 300th iteration and later increases linearly by 0.02 joules for every increase in the iteration round. At 800th iteration it rises to 0.88 joules resulting fast drainage of the battery power and leading to low network lifetime. The Hong approach exhibits high energy consumption in every simulation round resulting in degradation in QoS. Hence both these approaches do not show improvements in the life time of the communicating nodes in the network. Finally the average traffic load depicted in the fig 9 infers that the EEDCR system has about 80% capability to withstand the increasing traffic load due to the presence of operator nodes for sharing the load of gateways and domain nodes ,whereas the DRONNE algorithm and Hong approach has 40% and 20% respectively.

VIII CONCLUSION

This paper has laid an emphasis on the fact that the delay minimization is only possible if the persistent problems associated with traffic and communication protocols are addressed efficiently. In existing internet architecture, the gateway node is overburdened with various networking and authentication responsibilities that degrade the data transmission to certain extent. Hence, we share some of the responsibilities of gateway to a third party node called as operator node. In the proposed model, a requesting node may reside within a particular domain (out of many domains) or move to another domain. Each domain is also assumed to follow its own routing techniques. The request messages of the nodes will be passed on to their respective gateways which are directly communicating with the operator nodes. There are many operator nodes, which keeps addresses of other respective gateways in order to access or forward the data to another node in different domains. The gateway node computes the routes based on minimal delay and minimal energy and perform data transmission, which gives guaranteed results of data transmission. Our result shows better outcomes with respect to data transmission and energy conservation in comparison to the existing techniques.

REFERENCES

[1] A. Galis, A. Gavras, The Future Internet: Future Internet Assembly, Springer,, 2013

[2] M. Villari, A. Celesti, M. Fazio, A. Puliafito, "A Secure Self-Identification Mechanism for Enabling IoT Devices to Join Cloud Computing", *Springer Journal,Lecture Notes of the Institute for Computer Sciences, Social Informatics and Telecommunications Engineering*, vol.151, pp 306-311, 2015

[3] J. Wan, C. Zou, K. Zhou, R. Lu, D. Li, "IoT sensing framework with inter-cloud computing capability in vehicular networking ", *Springer- Electronic Commerce Research*, vol.14, Iss.3, pp.389-416, 2014

[4] S. Yue, Y. Xiao, G. G. Xie, "Experiments on an Election Algorithm for Decision Element Failures in 4D Future Internet Architecture", *Springer-Communication and Networking, Communications in Computer and Information Science*, vol.56, pp 250-258, 2009

[5] D. Papadimitriou, T. Zahariadis, P. M. Julia, I. Papafili, V. Morreale, F. Torelli, B. Sales, P. Demeester, "Design Principles for the Future Internet Architecture", *Springer-The Future Internet, Lecture Notes in Computer Science*, vol.7281, pp.55-67, 2012

[6] T. Zahariadis, D. Papadimitriou, H. Tschofenig, S. Haller, P. Daras, G. D. Stamoulis, M. Hauswirth, "Towards a Future Internet Architecture", *Springer-The Future Internet, Lecture Notes in Computer Science*, vol.6656, pp.7-18, 2011

[7] H. Fujinoki, A. Hauck, "Analysis on the current and the future Internet structure regarding multi-homed and multi-path routing", *Springer- Journal of Internet Services and Applications*, Vol.2, Iss.3, pp 257-270, 2011

[8] J. Strassner, S-S. Kim, J. W-K Hong, "Semantic Routing for Improved Network Management in the Future Internet", *Springer- Recent Trends in Wireless and Mobile Networks*, Series Communications in Computer and Information Science, Vol.84, pp 163-176, 2010

[9] M. Varela, L. S. Kapov, T. Ebrahimi, "Quality of Service Versus Quality of Experience", *Springer-Quality of Experience Part of the series T-Labs Series in Telecommunication Services*, pp 85-96, 2014

[10] Y. Biea, C. Maob, M. Yang, "Development of Vehicle Delay and Queue Length Models for Adaptive Traffic Control at Signalized Roundabout", *Elsevier- Procedia Engineering*, vol.137, pp.141 – 150, 2016

[11] R. Tiwaria, N. Kumar, "Minimizing Query Delay using Co-operation in IVANET", *Elsevier- Procedia Computer Science*, vol.57, pp.84 – 90, 2015

[12] J. Gancet, D. Urbina, P. Letier, "DexROV: Dexterous Undersea Inspection and Maintenance in Presence of Communication Latencies", *Elsevier-ScienceDirect*, vol.2, iss.2, pp.218-223, 2015

[13] W. Zhang, G. Zhang, L. Gou, B. Kong, and D. Bian, "Delay Minimization Topology Control in Planetary Surface Network: An Autonomous Systems Approach", *Hindawi Publishing Corporation International Journal of Distributed Sensor Networks*, 2015

[14] I. Byun, B. H. Ko, K. J. Jeon, K. S. Kim, "Delay and Energy Constrained Random Access Transport Capacity", *IEEE Transactions On Wireless Communications*, vol. 13, no. 8, August 2014

[15] R. Rom, M. Sidi, "An Enhanced Dynamic Framed Slotted ALOHA Anti-Collision Algorithm"*Springer-Multiple Access Protocols Part of the series Telecommunication Networks and Computer Systems*, pp 47-77, pp.47-77, 1990

[16] S. Park, J. Park, D. Shin, Y. Wang, "Accurate Modeling of the Delay and Energy Overhead of Dynamic Voltage and Frequency Scaling in Modern Microprocessors", *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, vol.32, pp.695-708, 2013

[17] M. H. Raza, A. Nafarieh, B. Robertson, "Indirect Estimation of Link Delays by Directly Observing a Triplet of Network Metrics", *Elevier-ScienceDirect Procedia Computer Science*, vol.32, pp.1030 – 1036, 2014

[18] G. Saleh, A. E. Keyi and M. Nafie, "Joint power and rate scheduling for cognitive multi-access networks with imperfect", *Springer- EURASIP Journal onWireless Communications and Networking*, vol.198, 2015

[19] W. Li, M. Bandai, T. Watanabe, "Discussion on Tradeoffs among Delay, Energy and Accuracy of Data Aggregation for Multi-Robot Sensor Networks", *Proceedings of the 2010 IEEE International Conference on Mechatronics and Automation*, 2010

[20] G. Motoyoshi, K. Leibnitz and M. Murata, "Proposal and evaluation of a future mobile network management mechanism with attractor selection", *Springer-EURASIP Journal on Wireless Communications and Networking*, vol.259, 2012

[21] K.Vinodha, R. Selvarani, "Congestion Control in Distributed Networking System-A Review", *International Journal of Computer Applications*, Vol.83, No 6, December 2013

[22] K.Vinodha, R. Selvarani, "Congestion Control in Distributed Networks - A Comparative Study", *Springer- Advances in Computer Science and Information Technology. Computer Science and Engineering*, Vol.85 of the series Lecture Notes of the Institute for Computer Sciences, Social Informatics and Telecommunications Engineering pp 115-123, 2012

[23] R. Selvarani ,K.Vinodha, "An efficient traffic regulation mechanism for distributed networks", *Springer-EURASIP Journal on Wireless Communications and Networking*, vol.154, 2015

[24] https://www.ihs.com/products/electronic-standards-collection.html

[25] A. Asadi, V. Mancuso, "DRONEE: Dual-radio opportunistic networking for energy efficiency", *Elsevier Journal of Computer Communication*, vol.50, pp.41-52, 2014

[26] Y.W. P Hong, Y-R Tsai, Y-Y Liao, "On the Throughput, Delay, and Energy Efficiency of Distributed Source Coding in Random Access Sensor Networks", *IEEE Transactions On Wireless Communications*, vol. 9, no. 6, June 2010

[27] N. Gehrig, P. L. Dragotti, "Symmetric and a-symmetric Slepian-Wolf codes with systematic and non-systematic linear codes", *IEEE communications letters*, 2005

Deep Learning in Text Summarization - A Survey

Athira S[*],SruthyManmadhan[**]
[*] PG Scholar, Department of CSE, NSS College of Engineering, Palakkad
[**]Assistant Professor, Department of CSE, NSS College of Engineering, Palakkad

Abstract- In this modern world dealing with enormous amount of data, text summarization plays a crucial role in extracting meaningful content and presenting precise, understandable information from large text. Many approaches to summarize text have been introduced over the years. Conventional methods create summary from text directly by extracting words that leads to redundancy and neglect document summary relationship. Deep learning techniques are proved to be effective in generating summaries. The paper focuses on deep learning based techniques for text summarization introduced over the years.

Index Terms- Text Summarization, Deep Learning, Auto Encoder, RNN, LSTM, GRU.

I. INTRODUCTION

The textual information available has been flooding with the increasing number of articles and links over the past few years making it difficult to search over the data to collect valuable information and present the information in a concise and clear way. Increase in data increases importance for semantic density thus there arise the need to recognize the most important things in the shortest amount of time. The generated summary helps to decide whether the textual content condensed in the article is relevant or not.

The idea of text summarization is finding a subset of the information to represent the entire document. Text summarization is taken as a task for condensing some textual information to a shorter version of itself which may contain all relevant and important information related to that document. It can be considered as a form of compression and hence suffer from information. Loss. Text summarization is effectively used in generating medical record summaries, weather data summary, news summaries etc.

Text summarization is broadly classified into the following categories:

1. Extractive text summarization: In this text summarization task objects are extracted from the documents without modification.
2. Abstractive Summarization: Different from extractive summarization as a word are modified, perform rephrasing or uses word that are not in the original document to create the summary hence making it more complex.

1.1. Deep Learning

Deep learning[1] is considered to be a type of representation learning method that uses cascade of multiple nonlinear processing units for performing transformations and feature extractions in such a way that output of one layer if feed as an input to next layer. Deep learning algorithms are capable of learning from the inputs in a supervised or unsupervised manner through multiple levels called as feature layers. The features layers are not described and designed by humans but are automatically learned from generalized learning process.

Conventional methods for text summarization includes directly extracting words from the textual content to represent the summary, Text summarization include removing stop words, identifying noun groups, lemmatization etc. The major disadvantage of conventional methods is that the summary generated may contain redundant words. As there are no record of the words that are already been selected it is possible that words may repeat itself in the summary as it does in the main text. Also, in conventional methods the relation between the summary that are generated and the document is very low. Thus, making it difficult for the users to have clear understanding of the document from the summarized content. Thus to overcome the disadvantages Deep learning techniques are employed to text summarization.

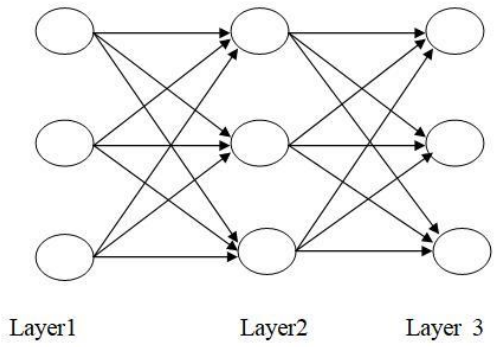


Figure 3: Three layered neural network with one input layer,one output layer and one hidden layer

The paper is organized as follows: Section II explains the various models of deep learning that are employed in text summarization and Section III compares various text summarization techniques using Deep learning models. Section IV discusses evaluation techniques. Section V pro-vides the conclusion and VI the references.

II. MODELS IN TEXT SUMMARIZATION

Various deep learning models are used in text summarization some of them are explained below:

2.1Auto Encoder

Auto encoder is an unsupervised learning algorithm that is used to learn data coding efficiently. Auto encoder aims at learning the representation of a set of data in order to reduce its dimensionality as well as complexity. The auto encoder consists of three layers that are the input layer, the encoding layer and the decoding layer (output layer). The input and the output layer are basically the same in such a way that the algorithm learns to compress the input data into an encoded format in the encoding layer which is later decoded to its original self by decoding layer. Auto encoders perform dimensionality reduction by compressing the input data by removal of noises.

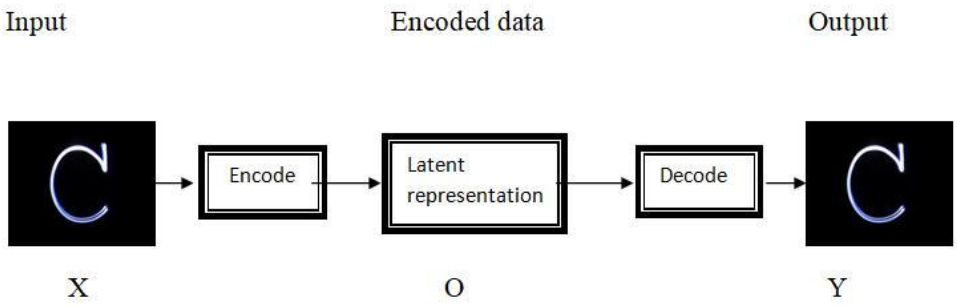


Figure 4: Basic Architecture of Auto Encoder [2]

For each input X given at the input layer, the input is encoded into an latent representation O such that O = f(X) and at the output layer the encoded representation is reconstructed as Y=g(O).

2.2. **Recurrent Neural Network**

Recurrent neural network [11] belongs to the class of artificial neural network that are represented using graphical models. Nodes belong to the part of directed graph along a sequence that allows the exhibition of temporal dynamic behavior. In traditional neural network input and output are considered to be independent of each other thus it does not take into account the previous information. In feed forward neural network [12] the connections between nodes never make up a cycle, information moves in one direction from input states through hidden states to output state. Also, the outputs are independent of each other such that output at time step t is not dependent on the out of time step t-1. In a scenario such as predicting the next word in a sentence feed forward network are observed to be inefficient. In RNN, the internal states can be used to process a sequence of input through back propagation. RNN are found to be effective where the input and outputs are dependent on each other.

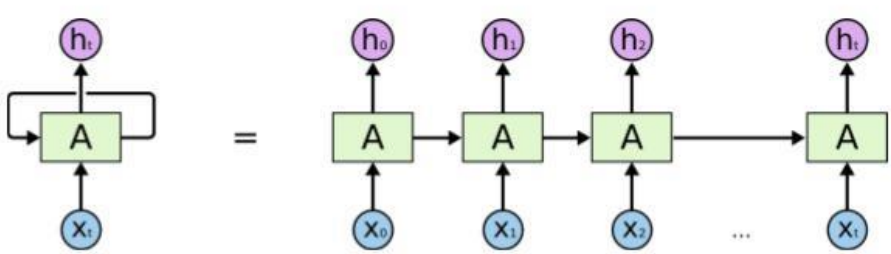


Figure 5: Recurrent Neural Network Unit [13]

Here in each time step t, the RNN unit takes the input vector x_t and the hidden state vector h_{t-1} to form the output of the hidden state h_t the process is thus repeated until all the inputs are processed.RNN can be formulated as a function given as:

$$h_t = f(x_t, h_{t-1}) \quad (1)$$

2.2.1. Limitation in RNN

The major drawback of RNN are termed as vanishing gradient problem[14] and exploding gradient problem[15].While performing back propagation in RNN it tends to calculate the error that is square of the difference between the actual output and the output observed from a model. With the calculated value for error the weight function for the next time step is calculated as:

$$W = n\frac{\partial e}{\partial w} \quad (2)$$

where w is the change in weight, $\frac{\partial e}{\partial w}$ is the rate of change of error with respect to weight that is called the gradient and n is the learning rate. Thus from this the new weight is calculated by adding w to the old weight. Now, if the value of the gradient becomes very small than 1 then Δw becomes negligible resulting in no greater difference in the weight calculation for the next time step. This is referred to as the Vanishing gradient problem .Similarly, If the gradient value becomes too large and there are long term dependencies then in each time step the weight value increases drastically. This is referred to as Exploding gradient problem and can be solved by using truncated BTT[16] or by clipping gradient at threshold. As for the former, the solution is to use LSTM and GRU.

2.3. Long Short Term Memory (LSTM)

LSTM's [24] are a special form of RNN that are capable of learning long term dependencies. In some scenarios, only recent information are required to perform a given task such as language models trying to predict the last word in a sentence .In situations where the gap between relevant information and the place where it is needed is small RNN learns to use the past information without the occurrence of problems discussed earlier But there are cases where more context are needed where the gap between relevant information and where it is needed are large. In such cases LSTM network are proven to be efficient.

LSTM has a chain like structure where the repeating module has four interacting modules that are the cell state, output gate, update gate and the forget gate. The cell state is represented using the horizontal running layer in the LSTM repeating module and act as a convener belt with minimal interactions. The basic structure of an LSTM module is given as follows:

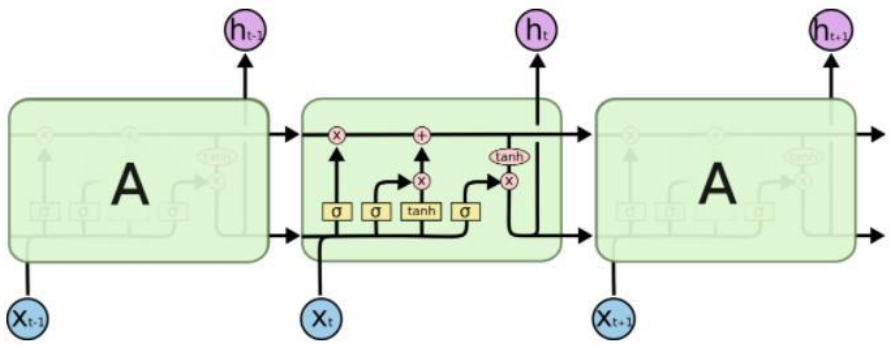


Figure 6: Basic Architecture of LSTM [25]

2.3.1 Forget Gate

The forget gate identifies the information that need to be eliminated from the cell state so that only relevant information are taken forward. The gate take into account the output from the previous time stamp h_{t-1} the new input $x_t$and produces an output between 0 and 1 for each cell state describing whether to keep the state or not. The following equation defines the sigmoid function carried out by the forget gate.

$$f_t = \sigma \ (w_f \ [h_{t-1}, x_t]; \ b_f \ ) \qquad (3)$$

2.3.2. Update Gate

Update gate decides on what information to be stored in the memory cell. The gate works in two layers, one of which is the input sigmoid layer that decides upon the value to be updated and an tanh layer that creates the vector of the new candidate value $\widetilde{C_{t-1}}$ that can be added to the states later.

At first the sigmoid function take into account the input coming from the precious time stamp $h_{t\ 1}$ and the new input x_t to calculate the value of i_t as:

$$i_t = \sigma \ (w_i \ [h_{t-1}, x_t]; \ b_i \ ) \qquad (4)$$

and also the inputs are passed through the tanh layer to create $\widetilde{c_t}$ as follows:

$$\widetilde{C_{t-1}} = tanh \ (w_c \ [h_{t-1}, x_t]; \ b_c \ ) \qquad (5)$$

now the values of i_t , f_t and are used to update the value of new cell state c_t as :

$$c_t = f_t * c_{t-1} + i_t * \widetilde{C_t} \qquad (6)$$

2.3.3. Output Gate

Output gate contains sigmoid layer that decides what parts of the cell state is given to the output. The cell state is made to pass through a tanh activation function to push its values between -1 and 1. After activation, the cell state is multiplied with the output of sigmoid layer to decide the output. The mathematical representation is given by the two equations:

$$o_t = \sigma\ (w_o\ [h_{t-1}, x_t], b_o\ ) \qquad (7)$$

$$h_t = o_t * \tanh(c_t) \qquad (8)$$

The major disadvantage with an LSTM network is that there are lots of operations performed within a single repetitive unit. Which when considered for a big network, the training process consumes greater amount of time. In order to overcome this limitation GRU were introduced.

2.4. Gated Recurrent Unit

Unlike LSTM, GRU does not account to any cell state it uses two different gates namely reset gate and an update gate. The update gate decides how much of the past information needs to passed on to the future states. Where, Reset gate decides in how much past information to forget. These two gates account to the output produced by the repetitive unit in each time step.

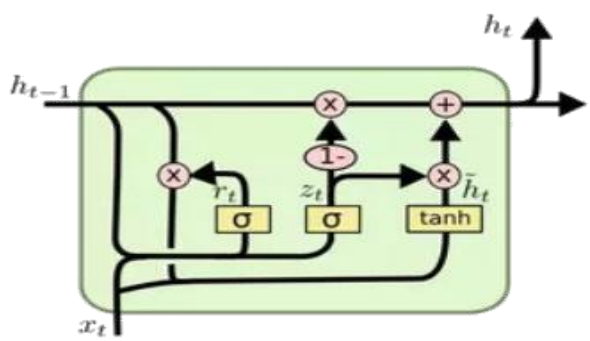


Figure 4: Basic Architecture of GRU [25]

III. COMPARISON

Table 1: Comparison of Text Summarization Techniques Using Auto Encoder

Paper	Pros	Cons
[3]	Local representation reduces sparsity and multiple runs improve efficiency.	Only works well with small vocabulary and mostly support query based search.
[4]	More descriptive feature space and improve recall on average.	Under performed system, need improvement in accuracy.
[5]	Account to summarization and	Computational needs are high. Relation

	reconstruction of text and aims at efficient semantic representation of variable size text.	between summary and document are poor.
[6]	Uses attention model to reduce redundancy.	Needs improvement in performance.
[7]	Improves summary quality and outperforms state-of-art models.	Not proven to be efficient.
[8]	Document–Summary pair training and more interpreted summaries	Sometimes fails to arrange word in correct order
[9]	Achieves state-of-art performance in benchmark datasets and better internal representation	Requires improvement in terms of sentence segmentation representation
[10]	Splits sentences based on position and is a hierarchical model	Fails to outperform baseline attention decoder

Table 2: Comparison of Text Summarization Techniques Using RNN

Paper	Pros	Cons
[17]	Structurally simple and performs end to end training.	Efficient alignment and consistency in generation are challenges.
[18]	Provides state-of-art performance and promising results.	Factual data incorrectly reproduced and replace uncommon words with alternatives.
[19]	Adress the modeling issue of preserving meaning and key content.	Doesnot account to previous information
[20]	Trained on human generated reference summaries.	Less Rouge value.
[21]	Reduces inaccuracy and repetition and out performs state-of-art model	Performance and high level abstraction needs to be achieved
[22]	Captures notations of salience and repetition .Easily interpretable.	Structured summarization problem.
[23]	Semi supervised technique outperforms standard seq2seq.	Less accurate

Table 3: Comparison of Text Summarization Techniques Using LSTM &GRU

Paper	Pros	Cons
[26]	Capture compositionality better without complex Architecture.	Timescale constant can be optimized further.
[27]	Efficient document summary scoring	Requires large scale training corpus.
[28]	Computational efficiency is prominent and has good Accuracy	Redundant information.
[29]	Generate natural sentences	Training is time consuming and also determining semantic similarity between phrases is difficult

IV. EVALUATION METHOD

Recall-Oriented Understudy for Gisting (ROUGE) [30] is a evaluation method for text summarization it automatically determines a summary quality by comparing it with ideal summaries created by humans called the gold standards. The measure determines the count of overlapping unit between the generated and the ideal summaries. There are four different Rouge measures described as follows:

1. ROUGE-N :N-Gram co-occurrence statistics It is a n-gram recall calculated between the generated candidate summary and the set of referenced summaries.
2. ROUGE-L: Longest Common Subsequence Given two sequence P and Q , the longest common subsequence of P an Q are the common subsequence with the maximum length.
3. ROUGE-W: Weighted Longest Common Subsequence Improves ROUGE-L values by remembering the length of the consecutive matches that are encountered so far.
4. ROUGE-S: Skip-Bigram Co-Occurrence Statistics Calculates the overlapping Skip-Bigram between the generated summaries and the set of referenced summaries

V. CONCLUSION

The paper summarizes the various model of deep learning that are employed in text summarization process and also the techniques that are developed over the years. It is observed Auto encoders, RNN and GRU are three main models that are widely employed in text summarization process and are proven to be more efficient than conventional text summarization methods.

ACKNOWLEDGMENT

This research was supported by NSS College of Engineering,Palakkad. We are thankful to our colleagues who provided expertise that greatly assisted the research, although they may not agree with all of the interpretations provided in this paper.

REFERENCES

[1] G. E. Hinton, S. Osindero, Y.-W. Teh, *A fast learning algorithm for deep belief nets*, Neural computation 18 (2006) 1527-1554.

[2] F. Li, H. Qiao, B. Zhang, *Discriminatively boosted image clustering with fully convolutional auto-encoders*, Pattern Recognition 83 (2018) 161-173.

[3] M. Y. Azar, K. Sirts, L. Hamey, D. M. Aliod, *Query-based single document summarization using an ensemble noisy autoencoder*, in: Proceedings of the Australasian Language Technology Association Workshop 2015, pp. 2-10.

[4] M. Youse -Azar, L. Hamey, *Text summarization using unsupervised deep learning*, Expert Systems with Applications 68 (2017) 93-105.

[5] B. Oshri, N. Khandwala, *There and back again: Autoencoders for textual reconstruction*, 2016.

[6] P. Nema, M. Khapra, A. Laha, B. Ravindran, *Diversity driven attention model for query-based abstractive summarization*, arXiv preprint arXiv:1704.08300 (2017).

[7] P. Li, W. Lam, L. Bing, Z. Wang, *Deep recurrent generative decoder for abstractive text summarization*, arXiv preprint arXiv:1708.00625 (2017).

[8] Y.-S. Wang, H.-Y. Lee, *Learning to encode text as humanreadable summaries using generative adversarial networks* (2018).

[9] S. Ma, X. Sun, J. Lin, H. Wang, *Autoencoder as assistant super-visor: Improving text representation for chinese social media text summarization*, arXiv preprint arXiv:1805.04869 (2018).

[10] Y. Zhang, Y. Wang, J. Liao, W. Xiao, *A hierarchical attention seq2seq model with copynet for text summarization*, in: 2018 International Conference on Robots & Intelligent System (ICRIS), IEEE, pp. 316-320.

[11] S. Grossberg, *Recurrent neural networks*, Scholarpedia 8 (2013) 1888.

[12] G. Bebis, M. Georgiopoulos, *Feed-forward neural networks*, IEEE Potentials 13 (1994) 27-31.

[13] M. T. Nayeem, et al., *Methods of sentence extraction, abstraction and ordering for automatic text summarization*, Ph.D. thesis, Lethbridge, Alta.: Universtiy of Lethbridge, Department of Mathematics and Computer Science, 2017.

[14] S. Hochreiter, *The vanishing gradient problem during learning recurrent neural nets and problem solutions*, International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems 6 (1998) 107-116.

[15] R. Pascanu, T. Mikolov, Y. Bengio, Understanding *the exploding gradient problem*, CoRR, abs/1211.5063 (2012).

[16] R. J. Williams, J. Peng, *An efficient gradient-based algorithm for on-line training of recurrent network trajectories*, Neural computation 2 (1990) 490-501.

[17] A. M. Rush, S. Chopra, J. Weston,*A neural attention model for abstractive sentence summarization* ,arXiv preprint arXiv:1509.00685 (2015).

[18] R. Nallapati, B. Zhou, C. Gulcehre, B. Xiang, et al., *Abstractive text summarization using sequence-to-sequence rnns and beyond*, arXiv preprint arXiv:1602.06023 (2016).

[19] G. Rossiello, *Neural abstractive text summarization*., in: DC@ AI* IA, pp. 70-75.

[20] R. Nallapati, F. Zhai, B. Zhou, *Summarunner: A recurrent neural network based sequence model for extractive summarization of documents*., in: AAAI, pp. 3075-3081.

[21] A. See, P. J. Liu, C. D. Manning, *Get to the point: Summarization with pointer-generator networks*, arXiv preprint arXiv:1704.04368 (2017).

[22] R. Nallapati, B. Zhou, M. Ma, *Classify or select: Neural architectures for extractive document summarization*, arXiv preprint arXiv:1611.04244 (2016).

[23] C. Khatri, G. Singh, N. Parikh, *Abstractive and extractive text summarization using document context vector and recurrent neural networks*, arXiv preprint arXiv:1807.08000 (2018).

[24] H. Sak, A. Senior, F. Beaufays, *Long short-term memory recurrent neural network architectures for large scale acoustic modeling*, in: Fifteenth annual conference of the international speech communication association.

[25] C. Olah, *Understanding lstm networks*, GITHUB blog, posted on August 27 (2015) 2015.

[26] M. Kim, M. D. Singh, M. Lee, *Towards abstraction from extraction: Multiple timescale gated recurrent unit for summarization*, arXiv preprint arXiv:1607.00718 (2016).

[27] Y. Hou, Y. Xiang, B. Tang, Q. Chen, X. Wang, F. Zhu, *Identifying high quality document-summary pairs through text matching*, Information 8 (2017) 64.

[28] Y. Zhang, J. Liao, J. Tang, W. Xiao, Y. Wang, *Extractive document summarization based on hierarchical gru*, in: 2018 International Conference on Robots & Intelligent System (ICRIS), IEEE, pp. 341-346.

[29] S. Song, H. Huang, T. Ruan, *Abstractive text summarization using lstm-cnn based deep learning*, Multimedia Tools and Applications (2018) 1-19.

[30] C.-Y. Lin, *Rouge: A package for automatic evaluation of summaries*, Text Summarization Branches Out (2004).

Web Usage Mining Algorithms: A Survey

Sowmya H.K.[*] , Dr. R.J. Anandhi [**]

[*] Department of Computer Science and Engineering, The Oxford College of Engineering, Bangalore, India
hk.sowmyakiran@gmail.com

[**]Department of Information Science and Engineering, New Horizon College of Engineering, Bangalore, India
rjanandhi@hotmail.com

***Abstract*—**The rapid progress of WWW has created gigantic amount of web data. Web usage mining can be applied to determine valuable information from the user communication with the WWW. The main objective of mining of web usage is to interpret the response of web site customers via the approach of data mining. Insight gained from web usage mining perhaps employed to boost designing of website, grant personification service and clear the way for more productive browsing. To attain this, it is required to secure pattern of user access in the form of log files. Web Usage mining is a way of determining communication of user with distinct web application. It is a process, which mainly consists of three inter-dependent phases like preprocessing of data, discovery of pattern and analysis of pattern. This survey paper objective is to study various techniques and algorithms for data preprocessing and data discovery stages. The authors have identified the gaps that are present in various web usage mining algorithms and proposed techniques to handle few of them. They also studied and analyzed frequent itemset mining algorithm like Apriori, FP-growth and SSIFM (Single-scan Frequent Itemset Mining Algorithm) on web usage data.

***Keywords*—** Data mining; web usage mining; preprocessing; pattern discovery; pattern analysis

I. INTRODUCTION

Mining of data is a method of extracting knowledge from massive volume of input. Repository of data perhaps in the form of database, data warehouse or data marts. Process of data mining involves three major stages specifically data pre-processing, pattern extraction and pattern analysis to discover useful patterns.

Web mining is a branch of data mining, helpful in discovering hidden pattern form large volume of web data repository. World Wide Web is an information system which holds huge data in the form of links, images, audio, video and graphics files. Web mining is somewhat distinct from data mining because it majorly pertain to unstructured or semi structured data, while data mining concerned with structured data.

Web mining conceivably on the whole subdivided to form three categories: The headmost subgroup is web content mining. It tries to obtain utility facts or comprehension from the content of web page. Web page may consist of text, audio, video, image, metadata and hyperlinks etc. Investigation in web content mining comprises acquiring facts from the web, classification of document and grouping, and drawing out information from web pages. The second subgroup is web structure mining. It tries to uncover knowledge from the hyperlink structure. It is categorized in to two kinds such as intra page structure and inter page structure. In Intra page structure, same page holds the link. It does not support opening of different page. Inter-page structure assist linking of one page to some other page. The third subgroup of web mining is called as web usage mining. It relates to the exploration of browsing pattern of user from web log. It gives prominence for diverse techniques of data mining to perceive and interpret identified patterns.

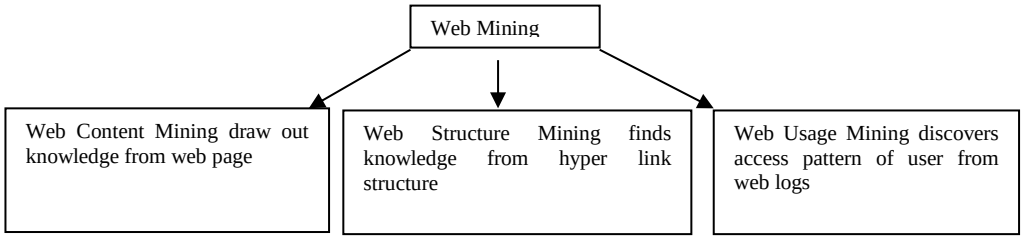


Fig. 1: Categories of Web Mining

Web Usage Mining

Web usage mining pertains to the preprogrammed exploration and investigation of patterns in web log. Discovered usage pattern is helpful to perceive and more effectively support the requirements of Web-based applications. The main purpose is to acquire, design, and analyze the user access patterns and their profiles, when they communicate with website. In mining of web usage, several approaches are employed to extract data produced by proxy server, web server and cookies to arrive at navigation pattern and browsing nature of the user.

Fig. 2 shows the three distinct phases of Web Usage Mining. Preprocessing of data is the first phase under this category. The web log file consists of huge amount of incomplete and unnecessary information, so straight away employing of such primitive log to perform web usage mining is not possible. Quality of web log file can be improved by applying preprocessing technique. Different techniques are applied in preprocessing phase that is cleaning of data, identification of user, sessionization, data integration, data transformation.

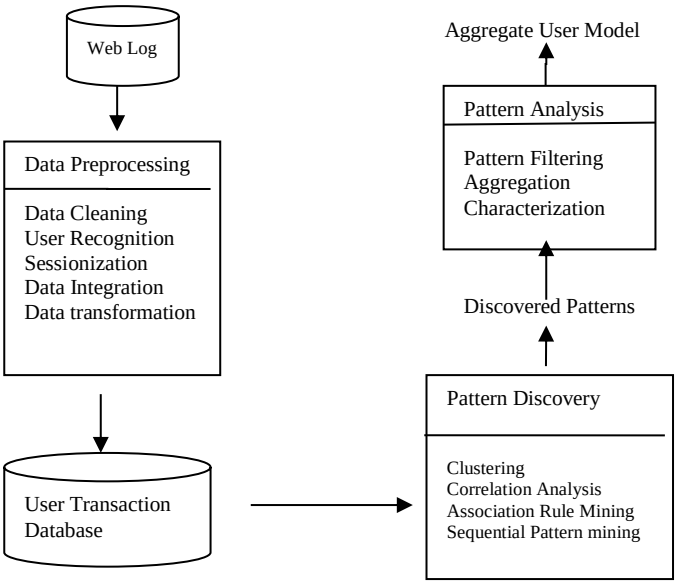


Fig. 2: Web Usage Mining Architecture

A. Data Cleaning

It is a mechanism to recognize and eliminate rows from web log which are not necessary and conformant. Web log consists of several attribute fields, from that only necessary fields are selected and others are

removed. First step in data cleaning is to remove records with file extension JPEG, GIF, CSS and so on, as the web pages are not of user interest, and instead it is implanted objects in the web page. Second step is to discard the entries which are marked with error status code, in case of user request for the web page which is not available on web server. Third step is to remove the web log entries gathered from the crawlers or machines as they do not reflect the way in which actual user move across the website. Several search engines announce them as an agent and thus can be recognized without any difficulty by string matching.

B. User Identification

It is a process to find out the user one who contacted web site and which pages are secured. If users have registered their details, then identification of such users is easy. There are several users who did not sign up their information and access web sites via, agent. Moreover, the same person can use different systems, and different users can use the same system as well. Additionally, proxy server conceals suitable facts about individual users as different computers take part on the internet work with the same IP address with the help of proxy server. All the above complications make the effort of user identification greater difficulty and complexity. Cookies can be used to correctly track users' behaviors. But taking privacy in to account, many users do not use cookies. Hence it is required to determine other approaches to provide solution to this problem

C. Session Identification

Sessionization, is the assignment to recognize the sessions from the raw data. But the challenge here is that, the log file of server do not hold entire essential facts. Modify data in to a suitable form, so that they can be treated as data for the algorithms. Data mining techniques will be applied to the appropriate data to achieve predetermined goal.

The second stage of web usage mining is Pattern Discovery. After identifying user sessions, apply various mechanisms and process originated from various fields namely data mining, machine learning, statistics and pattern recognition. Based on the requirement of the investigator, different kinds of pattern access, such as finding association rules, path analysis, discovery of sequential patterns, and clustering and classification can be performed. Identifying of desired patterns and to extract understandable knowledge from them is a challenging task.

Analysis of pattern is the last phase in the web usage mining. Main objective of this phase intend to remove the guidelines or arrangements which are not relevant. And also retrieve the relative rules or patterns from the outcome of pattern discovery phase. Structured Query Language (SQL), Online Analytical Processing (OLAP) are the tools used to perform pattern analysis.

D. Data Sources

There are three important data sources for web usage mining which includes data repositories such as 1. Server log 2. Client log and 3. Proxy log.

Server log provides details corresponding to page asked by the user. Details of the current requests are entered at the end of the file. This log file holds information related to the request made by the user such as, IP address, date and time, URL of requested page, HTTP code, bytes served, user agent and referrer. All this input conceivably aggregated in to one file or segregated in to various log files, such as error, access and referrer log. Server log file is projected as not complete file since it do not document the

cashed pages. These are mainly user admin purposes such as, constructive web site administration, established resources management.

The browser window of the client holds client log file within it. Recording of details to the log file are made by the Web server. These files are used to represent user behavior as they are most genuine and precise. But it is complex exercise to alter the browser for each user and demands users' concentration and cooperation.

Proxy log file hold the HTTP requisition from various users to various Web servers. This log file is a source of data to uncover the pattern of use of unidentified users who share a same proxy server. As Proxy server log files are more susceptible and composite, it is more difficult to deliver true picture user behavior.

II. LITERATURE SURVEY

The best approach for preprocessing of web usage mining is proposed by K. Sudheer Reddy [7]. Authors have proposed algorithms for user and session identification. This method not only decreases the volume of the log file further improves the quality of data. Mitali Srivastava et.al [3] made an observation of extraction of data and cleaning of data in web usage mining. They suggested data extraction and data cleaning algorithm, which works on web log file. Result of the experiment in this paper showed that raw log data reduced to 80% after cleaning process.

Neha Goel et.al [4], designed a tool for preprocessing web logs. It is a complete tool for reducing data which are not relevant and appropriate. And also modify it into a suitable form so that it can be applicable for analysis. The proposed methodology also figures out some enhanced attributes in statistics form namely the amount of time that has passed in getting the outcome, hit count by corresponding IP. K. Sudheer Reddy et.al [10] developed algorithms for retrieving log file from web server and combining log file. In the same paper, authors also focused on the preprocessing approaches realized on Web Sift tool.

B. Uma Maheswari et.al. [5], suggests Greedy clustering algorithm using belief function for user profile creation. This algorithm uses belief function similarity measure to assist clustering task that has capability to apprehend the unreliability between web user's navigation attainments. This study presented a method based on referrer information for making error free transactions in data preprocessing. Outcome of review is provided to the group of routines which makes use of Dempster's rule.

P. Nithya et.al [9] presented preprocessing approach for mining web log to eliminate noise and robots in the web. The output of preprocessing is an error free input for the later stages of data mining. For measuring the proposed preprocessing technique anonymous web data set is used, and it discloses the number of records. P. Dhanalakshmi et.al. [1], developed algorithms for static log cleaning static user and session identification.

Priyanka S. Panchal et. al. [8], proposed composite method for predicting users acquire pattern rest on Markov model. This technique performs grouping of user exploration build on their affinity measure along with concept of Markov model. Here the fundamentals of apriori algorithm can be set for guessing web link. Web link prevision is a way to predict the user visit to web page based on their past visiting record.

Yakhchi S et.al. [12], proposed improved ARMICA which deals with numerous attributes such as the count of accomplished rules, database scans, and the aspect of produced rules. Authors compared the novel approach with the existing algorithms and outcome of the experiment showcased, ARMICA-Improved is quicker, and produces reduced count of rules with more aspects. It has taken reduced count of database search to produce more exact results.

Enhanced hybrid system proposed by Janisa Colaco et.al. [14], for predicting user navigation pattern. These novel algorithms are used for locating chronic pattern and grouping the identified patterns. This hybrid algorithm shows enhanced results compared to the existing system based on Apriori. Smaller amount of time is taken by this technique, which resulted in improvement on computational and clustering efficiency. Mining Least Association Rule (MILAR) algorithm [20] is applied on student exam dataset to obtain the secondary association rule. Inference shows that the count of implied association rules are downsized as there is increase in CRS (Critical Relative Support) value. Further it reduced the number of rules that are apathetic.

Abdelghani Guerbas et.al [22] applied DBSCAN and OPTICS clustering algorithm for productive mining of web log and prevision of navigational pattern accordingly. Authors also recommended an outline for the best online navigational behavior prediction. It assists in decreasing the reaction time of server by accumulating pages that are most probably appealed by a user.

III. GAPS IDENTIFIED IN EXISTING DATA PREPROCESSING ALGORITHMS

An algorithm for static log cleaning, static user and session identification by P. Dhanalakshmi et.al. [1], has greater attainment in terms of error rate and F-measure only on restricted data volume and fields. Session identification algorithm proposed by P. Sukumar et.al [2], has not explored accuracy metric for the algorithm to identify user and session.

Precision benchmark for user and session identification algorithm is not tackled in preprocessing method proposed by K. Sudheer Reddy et.al. [7], for web usage mining. Effectiveness of the method proposed in this paper is not measured. Algorithms recommended for cleaning data, data fusion and data extraction by Mitali Srivastava et.al [3], are not efficient as they are inappropriate for huge data set. Time taken by algorithms is increasing for large data set. So the proposed algorithm does not have the ability to scale up with growing data set.

The tool that was designed for preprocessing web logs presented by Neha Goel et.al. [4], is restricted to a few records and to a specific website. So it is required to test the algorithms proposed by authors in this paper with different website for measuring its performance. K. Sudheer Reddy et.al [10], suggested algorithms for extracting and joining log file from web server. These algorithms require web log files of longer duration and more rules to improve server performance.

Greedy clustering algorithm using belief function, suggested by B. Uma Maheswari et.al [5], is experimented on web log of smaller size. Only few number of user identified and session formed by this method. As this algorithm cannot scale up to work on huge web log, its usage is limited to small data set. Algorithm proposed by Nisarg Pathak et.al.[6], examines only unchanging log files. It needs to enhance in identifying navigation pattern of user for unsteady website.

New clustering technique proposed by Priyanka S. Panchal et. al [8], computes fixed number of clusters which affects the accuracy of web page prediction. Accuracy and efficiency of the preprocessing algorithms for discarding provincial and widespread noise and web vehicles, is not measured in the work stated by P. Nithya et. al. [9]. In this paper, author has not compared time taken for user interested pattern prediction with other widely used prediction techniques.

IV. GAPS IDENTIFIED IN EXISTING DATA DISCOVERY ALGORITHMS

Efficiency of the algorithm proposed by D.Kerana et.al. [11], is similar to Apriori algorithm. Imperialist Competitive Algorithm presented by Shahpar Yakhchi et. al.[12], requires a predefined parameter and a proper mechanism to set this value. Frequent item set mining using pattern growth approach developed by Rashmi et.al [13], is not tested for large volume of data.

Janisa Colaco et. al [14], developed hybrid algorithm for predicting pattern of user behaviour is tested only for EPA data set. The experimental outcome of the newly introduced system is not distinguished with various alternative current systems. Suggested algorithms by the author perchance evaluated with the help of other set of data. Additionally, the suggested system can be exercised to various operation fields to evaluate its performance and applicability.

Comparison of Dclat algorithm with other algorithms suggested by Vijayarani et. al [15], is not measured the quality of association rule. Novel fitness functions defined using multi-objective GA-PSO [16], need to be scrutinized for improving the kindness of the rules. Manju et.al. presented Ant Colony Optimization approach is employed to produce association rule in only one step [18]. In this paper association rule's quality is measured only with respect to high confidence. New improved FP-tree algorithm developed by Ashika et.al. [19], tested only for fixed number of transactions of size 1000.

Algorithm proposed by Zailani Abdullah et. al.[20], not tested with other benchmark data set. Scalability of the association rule mining algorithm is not addressed in the Association rule mining using Apriori algorithm [21]. Abdelghani Guerbas et.al [22], applied DBSCAN and OPTICS clustering algorithm for persuasive mining of web log and anticipation of online navigational pattern. It uses two-step process for mining navigational patterns, OPTICS and then DBSCAN instead of combined density-based clustering algorithm.

V. PROPOSED METHOD

Collect user log data from web log of the servers, perform data preprocessing for data reduction, user identification, session identification. Design an efficient pattern discovery technique for grouping the web users based on their behavior (click streams). Develop an optimized clustering technique to group the users with similar browsing preference and web pages that appear to be theoretically associated with respect to the users' perception. Develop an organized and qualified mining method based on association rule to find the co-occurrence relationship among the identified patterns to extract information. Efficiency factors to be considered are time factor and number of database scans.

VI. RESULTS AND DISCUSSSION

The present work carried out in Java Eclipse platform, for the study and analysis of Apriori, FP-Growth and SSFIM frequent itemset mining algorithm to structure highest-n frequent item set. Experiment is

conducted on the web server log file of msnbc.com crunched from the website, "http://kdd.ics.uci.edu/databases/msnbc/msnbc.html". This website contains pre-processed web log data that reports the page visited by users on September 28, 1999.

Numerous experiments have been performed applying two types of data sets to examine the algorithm's performance for frequent item set mining such as Apriori, FP-Growth and SSFIM. The first dataset is a collection items, with average size of 4-16 having number of transactions 693. The second instance is a collection of medium-sized database instances, with 3196 transactions. All three algorithms in the experiments have been realized in java Eclipse Platform and experiments run on a desktop machine equipped with Intel I7 processor and 4GB memory. The following table present execution time of various frequent item set mining algorithm for MSNBC data set by varying the values of minimum support.

Table 1. Execution Time of Frequent Itemset Mining on MSNBC Dataset

MSNBC Dataset			
	Execution Time in ms		
Min Sup	FP Growth	Apriori	SSFIM
5	23	41	39
10	21	37	33
15	19	32	32
20	19	31	31

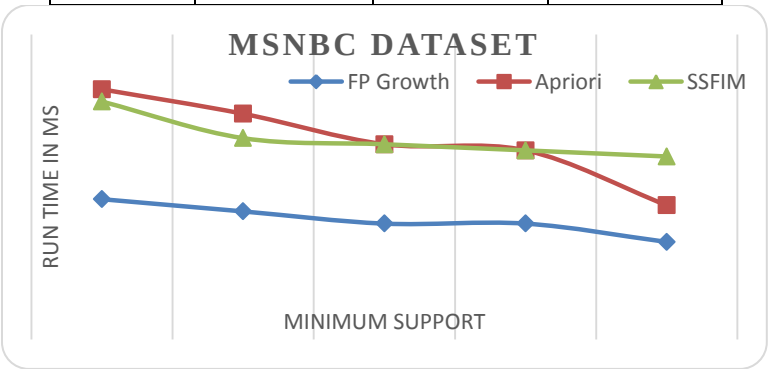


Fig. 3: Execution Time of Frequent Itemset Mining on MSNBC Dataset

Table 1 shows that, for dataset with small instances, FP-growth, Apriori algorithm surpasses SS-FIM. However, for dataset with large number of instances, SS-FIM clearly surpasses FP-Growth and Apriori. Fig. 3 shows the runtime performance of the FP-growth, Apriori and SSFIM using the msnbc datasets described above.

Table 2 approves that SS-FIM algorithm is finer than FP-Growth and Apriori when we handle sparse and large dataset. It presents execution time of various frequent item set mining algorithm for chess data set by varying the values of minimum support.

Table 2. Execution Time of Frequent Itemset Mining on Chess Dataset

Chess Dataset			
Min Sup	Execution Time in ms		
	FP Growth	Apriori	SSFIM
60	926	123419	99
70	310	19829	91
80	178	3872	83
90	138	485	77

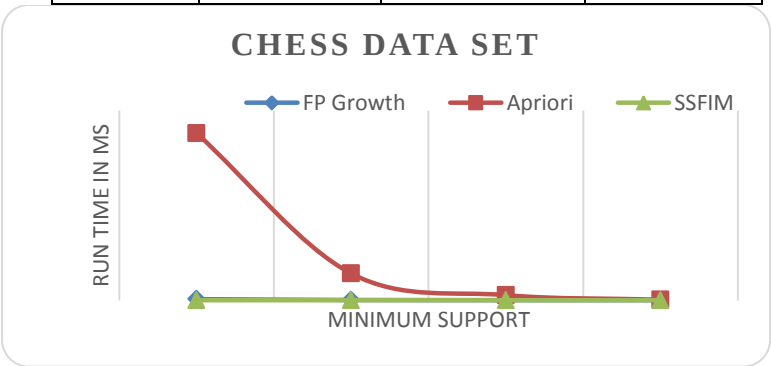


Fig. 4: Execution Time of Frequent Itemset Mining on Chess Dataset

Fig. 4 shows the runtime performance of the FP-growth, Apriori and SSFIM using the chess datasets described above. It shows that SS-FIM is not sensible to any variation of minimum support threshold value. The number of item set generated by this algorithm is fixed. Alternatively, Apriori algorithm performs multiple scanning of the dataset. It produces enormous candidate itemsets when minimum support threshold value is fixed with low value which declines the runtime of the algorithm. FP-Growth algorithm decreases the number of dataset scanning contrasted to Apriori, but expend more memory when it operates on large dataset instances.

VII. CONCLUSION AND FUTURE WORK

Due to huge advancement of web-based applications, most of the research is happening in web usage mining. It considers interpreting the look through behavior of website users and employing the perceived knowledge to elevate perfection of browsing experience. This paper presented an outline of all the aspects of web usage mining method. This paper described the recent approaches for preprocessing of web server log. This work also highlighted techniques and algorithms used in pattern discovery and analysis. Present work identified gaps that are present in several existing preprocessing and pattern discovery algorithms. Study and analysis of frequent item set mining algorithm reveals that for small instances of dataset, FP-growth, Apriori algorithm is better than SS-FIM. However, SS-FIM performs well for dataset with large number of instances.

Scalability of preprocessing algorithm for web usage mining is the focus of the future work. Design and implementation of efficient hybrid method in pattern discovery and analysis phase to extract more knowledge from web log file.

ACKNOWLEDGMENT

We are greatly thankful to our respective departments for providing excellent lab facilities that made this work possible.

REFERENCES

[1] P. Dhanalakshmi, K. Ramani, B. EswaraReddy, "*The Research of Preprocessing and Pattern Discovery Techniques on Web Log Files*", IEEE 6th International Conference on Advanced Computing (IACC), pp.139 – 145, 2016.

[2] P. Sukumar, L. Robert, S. Yuvraj, "*Review on Modern Data Preprocessing Techniques in Web Usage Mining*", International Conference on Computational System and Information Systems for Sustainable Solutions, 2016.

[3] Mitali Srivastava, Rakhi Garg, and P K Mishra, "*Analysis of Data Extraction and Data Cleaning in Web Usage Mining*", ICARCSET - 15, March 06 - 07, 2015, Unnao, India.

[4] Neha Goel, C. K. Jha, "*Preprocessing web logs: A critical phase in web usage mining*", International Conference on Advances in Computer Engineering and Applications, pp. 672 – 676, 2015.

[5] B.UmaMaheswari, P.Sumathi,"A New Clustering and Preprocessing for Web Log Mining", World Congress on Computing and Communication Technologies, 2014, pp.25 – 29.

[6] Nisarg Pathak , Viral shah, Chandramohan Ajmeera, "*A Memory Efficient Algorithm with Enhance Preprocessing Technique for Web Usage Mining*", ICTCS'14 November 14 - 16 2014, Udaipur, Rajasthan, India.

[7] K. Sudheer Reddy, G. Partha Saradhi Varma, and M. Kantha Reddy, "*An Effective Preprocessing Method for Web Usage Mining*", International Journal of Computer Theory and Engineering, Vol. 6, No. 5, October 2014.

[8] Priyanka S. Panchal ,Prof. Urmi D. Agravat, "*Hybrid Technique for User's Web Page Access Prediction based on Markov Model*", IEEE - 31661, 4th ICCCNT 2013, July 4-6, Tiruchengode, India.

[9] P. Nithya, P. Sumathi, "*An Enhanced Preprocessing Technique for Web Log Mining by Removing Web Robots*", IEEE International Conference on Computational Intelligence and Computing Research, pp.1-4, 2012.

[10] K.Sudheer Reddy, Dr. G. Partha Saradhi Varma, Dr. I. Ramesh Babu, "*Preprocessing the Web Server Logs – An illustrative approach for effective usage mining*", ACM SIGSOFT Software Engineering Notes, Volume 37, May 2012.

[11] D. Kerana, Kaliyamurthie K. P. "*Mining Frequent Item sets in Association Rule Mining Using Improved SETM Algorithm*", Artificial Intelligence and Evolutionary Computations in Engineering Systems, Advances in Intelligent Systems and Computing, vol 394. Springer, New Delhi, India 2016.

[12] Yakhchi S., Ghafari S.M., Tjortjis C., Fazeli M., "*ARMICA-Improved: A New Approach for Association Rule Mining*", International Conference on Knowledge Science, Engineering and Management, Springer, pp 296-306, July 2017.

[13] Rashmi V. Mane, V.R. Ghorpade, "*Predicting Student Admission Decisions by Association Rule Mining with Pattern Growth Approach*", International Conference on Electrical, Electronics and Optimization Techniques (ICEEOT), Dec 2016 , Mysuru, India.

[14] Janisa Colaco, Jayashri Mittal, "*Improvised Hybrid model for User Navigation Pattern Prediction*", International Conference on Communication and Electronics Systems (ICCES), 2016.

[15] S. Vijayarani, S. Sharmila, "*Comparative Analysis of Association Rule Mining Algorithms*", International Conference on Inventive Computation Technologies (ICICT), 26 - 27 August 2016, Coimbatore, India.

[16] Aashna Agarwal, Nirali Nanavati, "*Association rule mining using hybrid GA-PSO for multi-objective optimization*", International Conference on Computational Intelligence and Computing Research (ICCIC), 2016, Chennai, India.

[17] Deepti Sahu, Rishi Soni, "*A New Method for detecting User Behavior from Web Server logs*", International Conference on Computational Intelligence and Communication Networks, December 2015, Jabalpur, India.

[18] Manju, Chander Kant, "*Mining Association Rules Directly Using ACO without Generating Frequent Item sets*", International Conference on Energy Systems and Applications, November 2015, Pune, India.

[19] Ashika Gupta, Rakhi Arora, "*Web Usage Mining Using Improved Frequent Pattern Tree Algorithms*", International Conference on Issues and Challenges on Intelligent Computing Techniques, April 2014, Ghaziabad, India.

[20] Zailani Abdullah, Tutut Herawan, Noraziah Ahmad, "*Mining Indirect Least Association Rule from Students 's Examination Datasets*", ICCSA 2014, pp. 783–797.

[21] Avadh Kishor Singh, Ajeet Kumar, "*Association Rule Mining for Web Usage Data to Improve Websites*", International Conference on Advances in Engineering & Technology Research, August 2014, Unnao, India.

[22] Abdelghani Guerbas, Omar Addam, Omar Zaarour, Mohamad Nagi, Ahmad Elhajj, Mick Ridley,Reda Alhajj, "*Effective web log mining and online navigational pattern prediction*", Knowledge-Based Systems Volume 49, September 2013, pp. 50-62.

[23] http://kdd.ics.uci.edu/databases/msnbc/msnbc.html

An Improved Method for Measuring the Properties of Aluminium Oxide Nanopore FESEM Images

Parashuram Bannigidad*, Jalaja Udoshi*, C. C. Vidyasagar**
*Dept. of Computer Science, Rani Channamma University, Belgaum, Karnataka, India,
** Dept. of Chemistry, Rani Channamma University, Belgaum, Karnataka, India

Abstract- Characterization of fabricated nano-template plays a vital role in nano engineering and technology. The geometrical properties of the nanopores in synthesized nano membrane depend on the anodizing parameters like concentration of acid, anodizing time, temperature of the bath and voltage applied. The present study exhibits the automated tool to measure the nanoporous media by extracting the geometrical features; nanopore radius and nanopore circularity of anodic aluminium oxide (AAO) by varying anodizing parameters. The experimental results depicts that as the anodization time is increased (5mins, 9mins, 20mins and 30mins), the nanopore radius have gradually increased from 7.9nm to 11.6nm and nanopore circularity has decreased from 0.73 to 0.56 by keeping the constant value in concentration (5%), temperature (20^0C) and voltage (50 V). Similarly, as the voltage is increased (35V, 40V, 45V), the nanopore radius increased gradually from 8nm to 13.3nm and the nanopore circularity increased initially from 0.68 to 0.78 and decreased to 0.66 by maintaining the concentration (4.7%), time (8min) and temperature (5^0C) constant. The increase in nanopore radius from 55.6nm to 67.4nm and increase in nanopore circularity from 0.55 to 0.67 is witnessed when the concentration (4% and 5%) and temperature (20^0C and 25^0C) has been altered keeping time (20 min) and voltage (50 V) constant. The extracted feature values of nanopore radius and circularity are verified by a chemical expert, which proves the exhaustiveness of the proposed results.

Keywords- *AAO*, Aluminium Nanopore, Characterization, FESEM, Pore radius, Pore circularity, Nanomaterial, Nanotechnology.

I. INTRODUCTION

Anodic aluminium oxide (AAO)[1] has drawn extensive attention in nanotechnology due to its well-defined pore architecture and suitable corrosion resistance, thermal stability, hardness, abrasion resistance and insulation properties. As a nano-template or host material, AAO plays an essential role for various surface engineering applications, e.g. molecular separation, catalysis, energy storage, electronics, sensors, drug delivery and template synthesis, and it is a component in a diverse range of nanostructured materials in the form of nanodots, nanowires and nanotubes [2-6]. Pore geometry is found useful in describing important flow and transport mechanisms and in predicting flow properties of different porous media relevant to numerous fundamental and industrial applications [10-11]. The benefit of nanotechnology depends on the fact that it is possible to tailor the structures of materials at extremely small scales to achieve specific properties, thus greatly extending the material science toolkit. Using nanotechnology, materials can effectively be made stronger, lighter, more durable, more reactive, more sieve-like, or better electrical conductors, among many other traits. There are many daily used commercial products in the market based on the nanoscale materials and its process.

The importance of nanoscale materials has drawn attention of many nano researchers for the last couple of decades. The application of nanoporous templates was presented by C. Sousa [7]. Solar desalination through aluminium nanoparticles was studied by L. Zhou [8]. X Yang worked on the topological parameters of nanopore [9]. Computational geometry was studied by Joost H. [12]. The geometrical features of Al_2O_3 nanopore images was proposed by P. Bannigidad and Jalaja U. [13-15].

The objective of the present study is to extract the geometric feature values of nanopore radius and nanopore circularity from the Al_2O_3 FESEM images using global thresholding method. These values are useful in describing the important flow and transport mechanism and in predicting flow properties of different porous media relevant to numerous fundamental and industrial applications.

II. Materials And Method

Aluminium nanoporous FESEM images (A, B, C and D) are obtained at regular intervals of time (5 mins, 9mins, 20 mins and 30 mins), keeping constant the concentration (5%), temperature (20^0C) and voltage (50 V). The images (E, F and G) are obtained at varying voltage (35V, 40V, 45V) keeping the concentration (4.7%), time (8min) and temperature (5^0C) constant. Images (H and I) are obtained at 4% and 5% concentration, 20^0C and 25^0C temperature respectively keeping time (20 mins) and voltage (50V) constant.

III. Proposed Method

The objective of the present study is to develop an automated tool to determine the effect of changing anodization parameters on the geometrical features; namely, the nanopore radius and nanopore circularity of Al_2O_3 nanopores in the experimental FESEM images. In contrast with the current manual staining techniques the proposed method is more efficient, accurate, reliable and robust. The top views of anodized Al_2O_3 FESEM images are shown in Fig.1. Images labeled A, B, C and D are obtained at regular intervals of time (5 mins, 9 mins, 20 mins and 30 mins), keeping constant the concentration (5%), temperature (20^0C) and voltage (50 V). The images E, F and G are obtained at varying voltage (35V, 40V, 45V) keeping the concentration (4.7%), time (8min) and temperature (5^0C) constant. Image H and I are obtained at 4% and 5% concentration, 20^oC and 25^oC temperature, 20 mins time and 50V voltage.

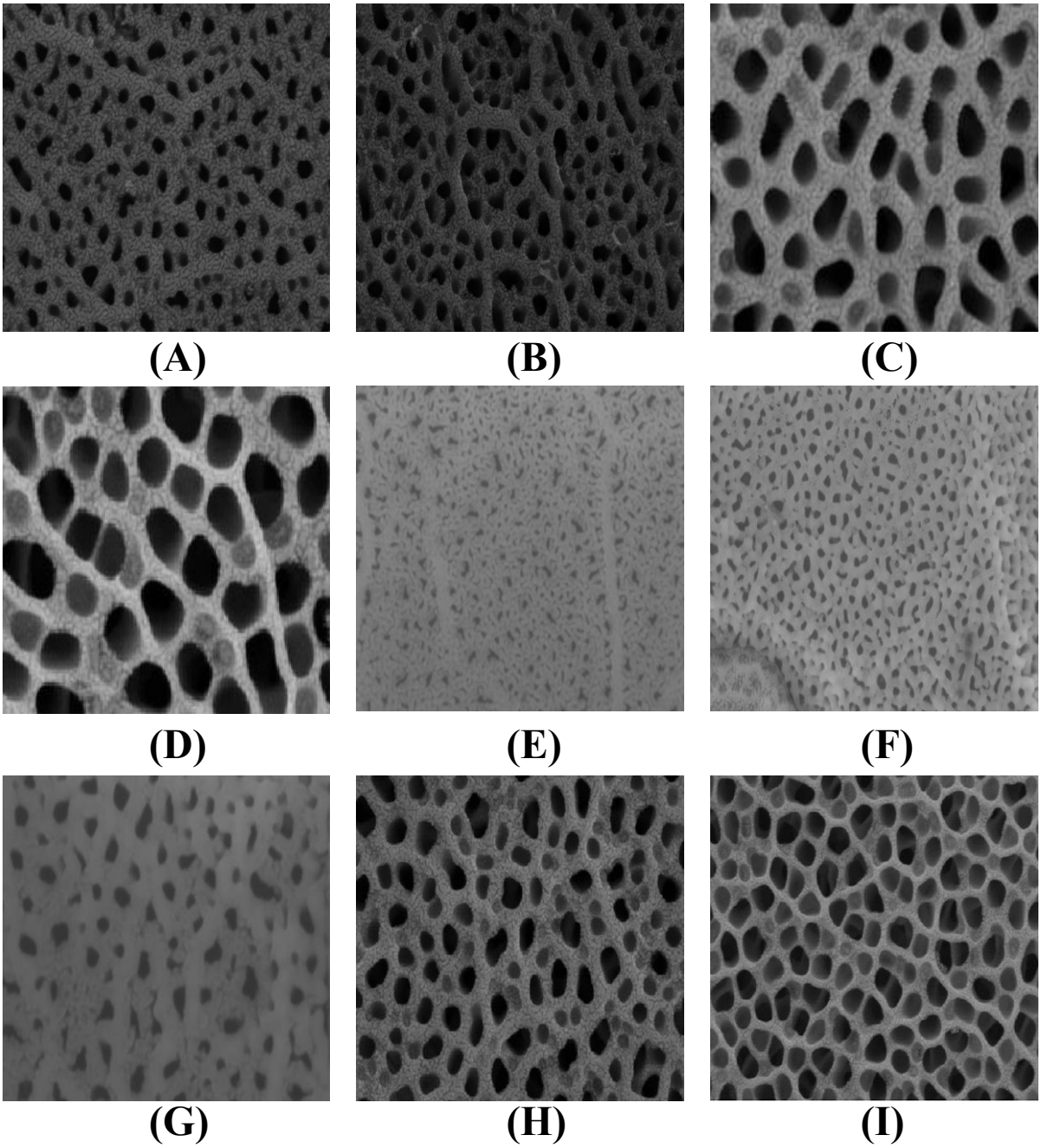


Fig. 1 Top view of aluminium oxide FESEM images

The geometric features extracted from the images; nanopore radius and nanopore circularity are defined as below:

1. Nanopore radius: Radius is defined as half, the average of nanopore major axis and minor axis.
2. Nanopore circularity: The nanopore circularity is defined by using the below equation:

$$C=(4\pi.S)/L^2$$

Where S=Surface area occupied by single nanopore.

L= Nanopore perimeter.

The following facts stand true with nanopore circularity (C):

- C = 1, when the nanopore is an ideal circle.
- C < 0, when the nanopore deformation occurs.
- C value is close to 0, when nanopore is similar to elongated polygon.

The algorithm of the proposed method:

1. Read aluminium nanopore FESEM image.
2. Convert the given input image to RGB image using the function rgb2gray()
3. Perform pre-processing operations on step 2:
   a. Applied image enhancement to enhance the quality of the image
   b. To obtain binarized image, the following steps are applied:

```
for  i=1:size(gray_image,1);
  for j=1:size(gray_image,2);
     if gray_image(i,j)>121
        binary_image(i,j)=1;
     else
        binary_image(i,j)=0;
     end
  end
end
```

   c. Perform morphological operations on the binarized image.

4. Segment the individual nanopores from the binarized image using the below method.

```
filteredpours=ones(1,num);
for segno=1:num
    if length(find(L==segno))<100
        L(find(L==segno))=0;
        filteredpours(segno)=0;
    end
 end
```

5. Extract the geometric features; i.e., nanopore radius and nanopore circularity and store them as knowledge base.
6. Finally, analyse and interpret the results using the following conditions:
   If nanopore circularity = 1 then, the nanopore is an ideal circle.
   Else if nanopore circularity < 0, then, the nanopore deformation occurs.
   Else if nanopore circularity value is close to 0, then nanopore is similar to elongated polygon.

The flow diagram of the proposed method is depicted in the below Fig. 2:

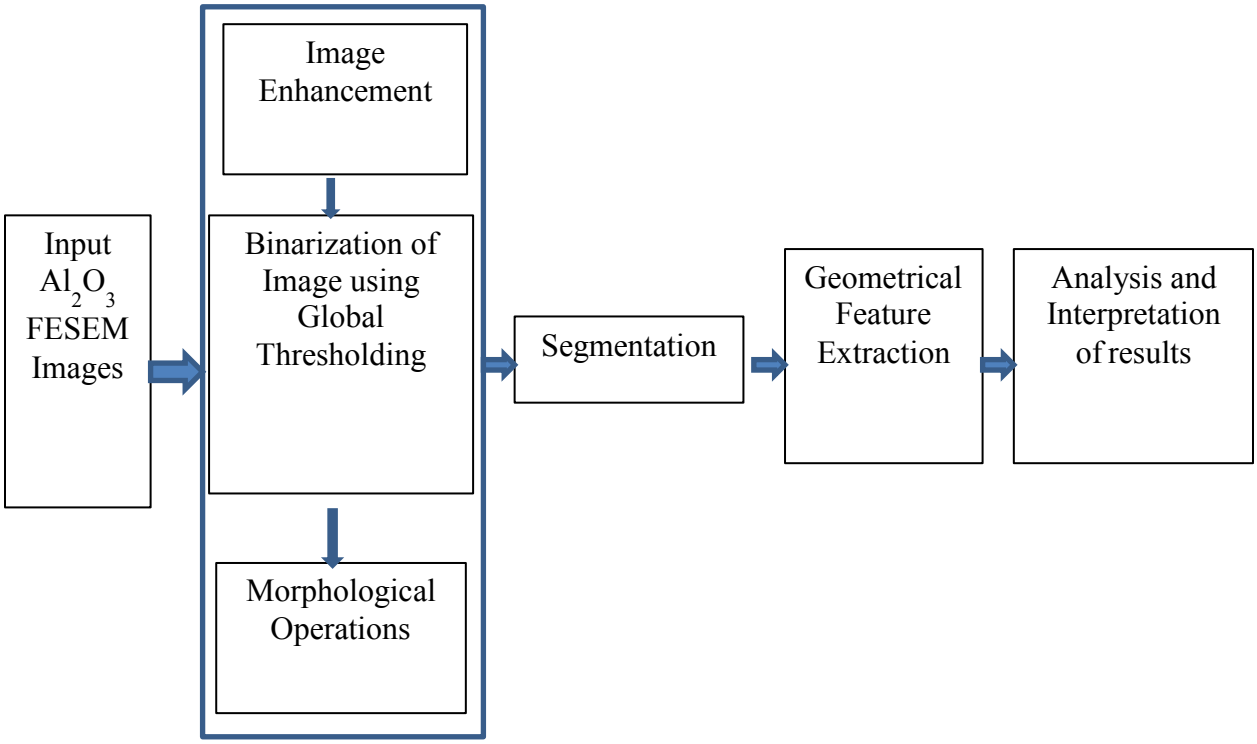


Fig. 2 Flow diagram of proposed method

IV. Experimental Results and Discussion

The experimentation is carried out on Intel(R) Core(TM) Duo T6670 @ 220GHz with 2 GB RAM using MATLAB R2014a software tool. The images used in the experimentation are of 500*500 dimension with 205.24KX magnification and are obtained from the Department of Chemistry, Rani Channamma University, Belagavi at varying anodizing properties. The details of these Al_2O_3 FESEM images are depicted in Fig 1.

.

The FESEM images are initially enhanced and then converted to binary image using global thresholding method (Fig2. (ii), Fig. 3 (ii) and Fig. 4(ii)). The binarized images will undergo the morphological operation and then the nanopores are extracted through segmentation (Fig2. (iii), Fig. 3 (iii) and Fig. 4(iii)). The nanopore radius and nanopore circularity is computed for every nanopore. Finally, the results are interpreted, analysed and categorised based on the values of nanopore circularity criteria (discussed in section 2).

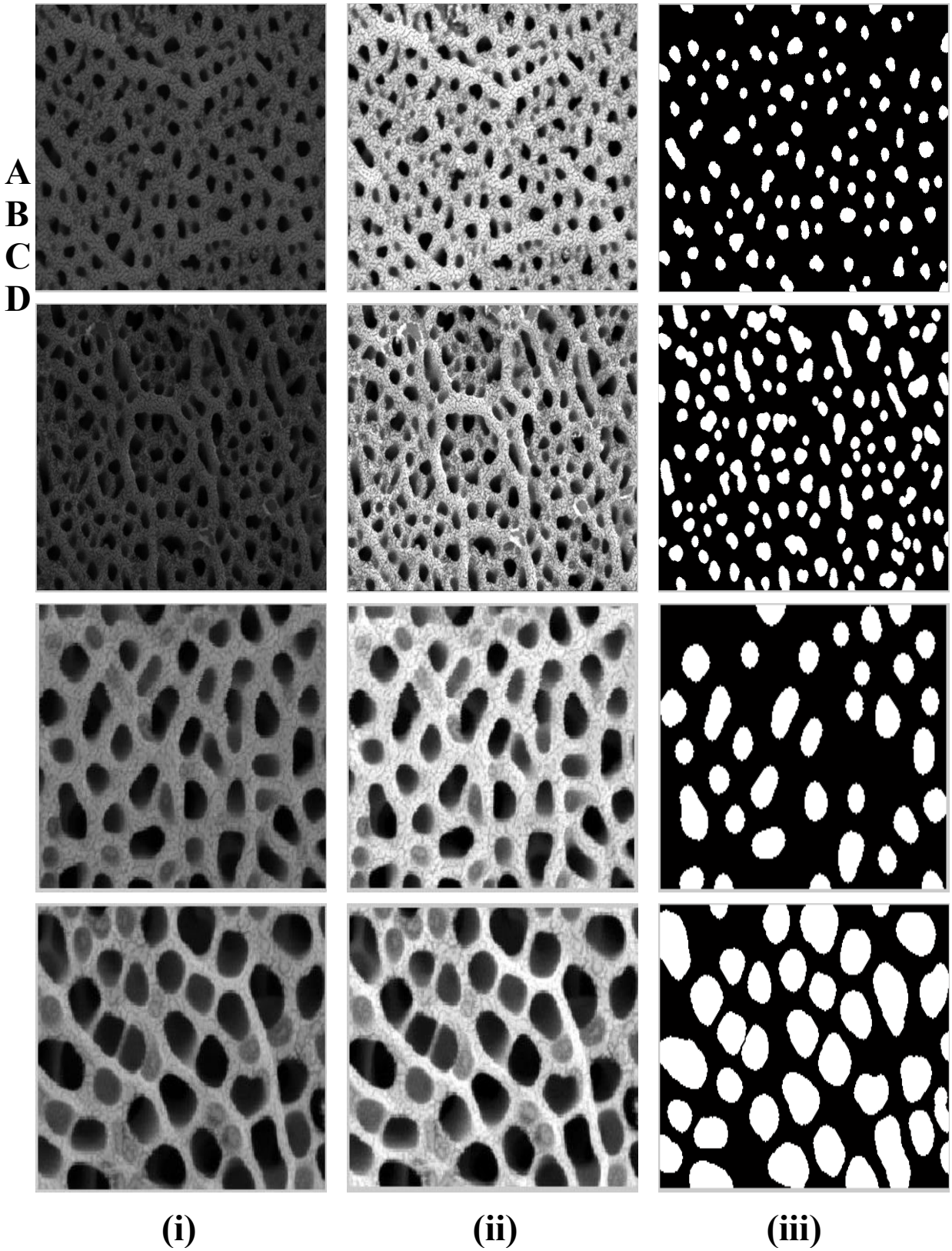


Fig. 3 (i) Original FESEM images captured at different intervals of time, (ii) Grayscale images, (iii) Segmented images

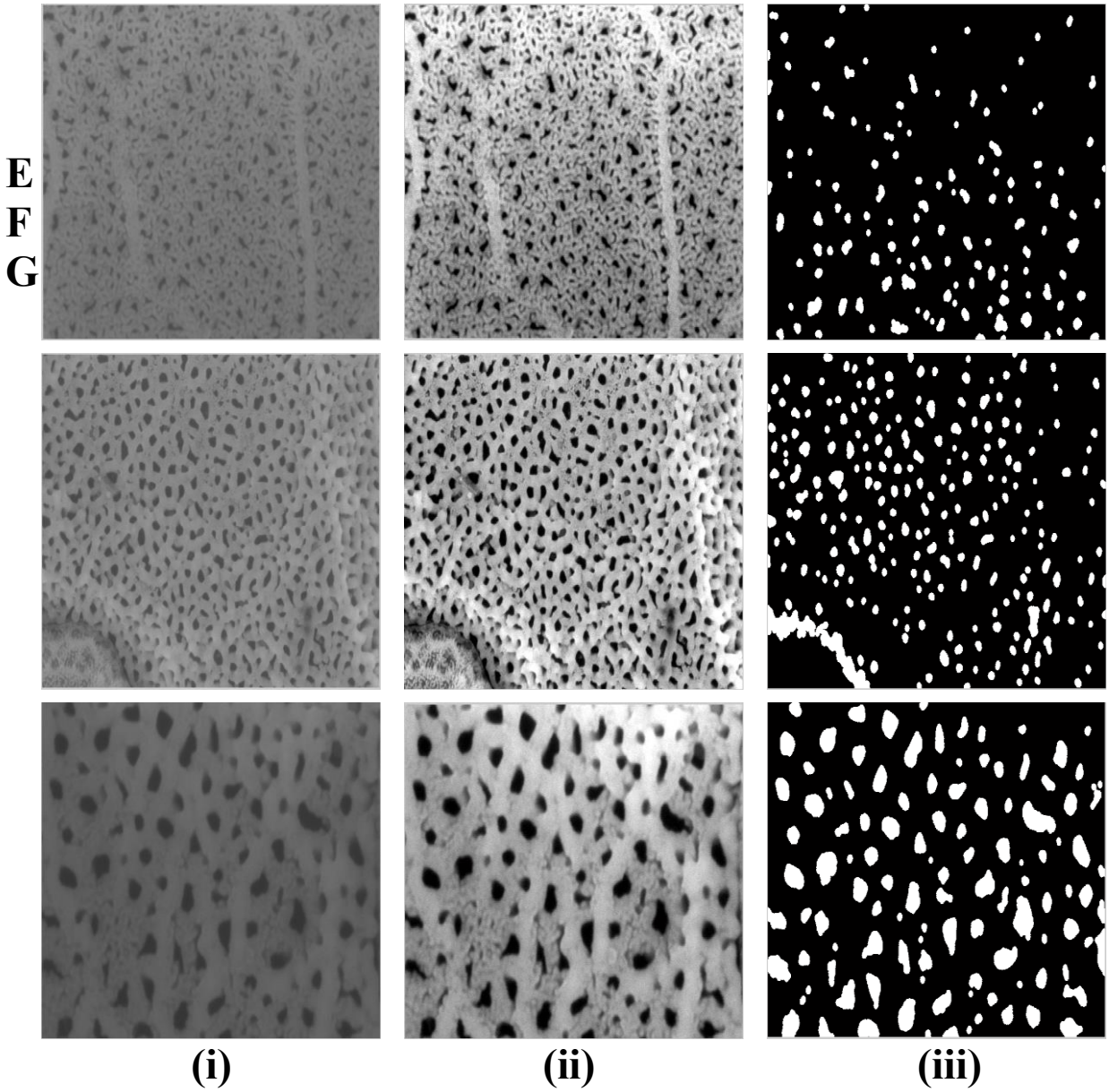


Fig. 4 (i) Original FESEM images captured at different voltage, (ii) Grayscale images, (iii) Segmented images

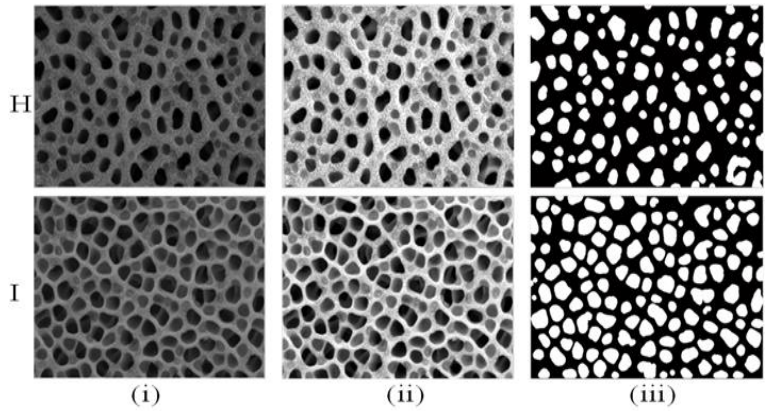


Fig. 5 (i) Original FESEM images captured at different temperature and concentration, (ii) Grayscale images, (iii) Segmented images

The average results of Al_2O_3 nanopore images categorised based on nanopore circularity is shown in the Table 1.

Table 1. The average results of aluminium oxide nanopore images categorised based on circularity

FESEM Image	Concentration (%)	Time (min)	Temperature (^{0}C)	Voltage (V)	Average Nanopore Radius (nm)	Average Nanopore Circularity (C)
A	5	5	20	50	7.9	0.73
B	5	9	20	50	8.9	0.61
C	5	20	20	50	9.1	0.57
D	5	30	20	50	11.6	0.56
E	4.7	8	5	35	8.0	0.68
F	4.7	8	5	40	9.6	0.78
G	4.7	8	5	45	13.2	0.66
H	4	20	20	50	12.7	0.55
I	5	20	25	50	14.5	0.67

FESEM Images	Number of Nanopores	Nanopores categorised based on the nanopore circularity (C)		
		C near to 1 (C>=0.8 and C<=1)	C close to 0 (C>=0.3 and C<0.8)	C<0
A	109	46	63	0
B	139	51	88	0
C	35	19	16	0
D	37	11	26	0
E	129	63	66	0
F	221	125	96	0
G	74	30	44	0
H	99	33	66	0
I	86	34	52	0
Total	929	412	517	0

The aluminium oxide nanopores, categorised based on the geometric feature values of nanopore circularity extracted from the investigated samples are given in the Table 2.

Table 2. The aluminium oxide nanopores are categorised based on the geometric feature values of nanopore circularity extracted from the investigated samples

- As the anodization time is increased (5 mins, 9 mins, 20 mins and 30 mins), keeping constant the concentration (5%), temperature (20^0 C) and voltage (50V);
    - The nanopore radius has gradually increased from 7.9nm to 11.6nm (Fig. 6(1)).
    - The nanopore circularity has decreased from 0.73 to 0.56 (Fig. 7(1)).
- Similarly, as the voltage is increased (35V, 40V, 45V), keeping the concentration (4.7%), time (8 min) and temperature (5^0C) constant;
    - The nanopore radius increased gradually from 8nm to 13.3nm (Fig. 6(2)).
    - The nanopore circularity increased initially from 0.68 to 0.78 when the voltage is increased from 35V to 40V but decreased to 0.66 when the voltage was further increased to 45V (Fig. 7(2)).
- When the concentration (4% and 5%) and temperature (20^0C and 25^0C) was altered keeping time (20 min) and voltage (50 V) constant;
    - The nanopore radius increased from 55.6nm to 67.4nm (Fig. 6(3)).
    - The nanopore circularity increased from 0.55 to 0.67 (Fig. 7(3)).

The extracted feature values of nanopore radius and nanopore circularity are verified by a chemical expert, which proves the exhaustiveness of the proposed results. Further, the total number of nanopores based on the geometric feature values of nanopore circularity which are closer to 1 are 412 nanopores and 517 nanopores are closer to 0, out of total 929 nanopores from all the image data sets (A to I).

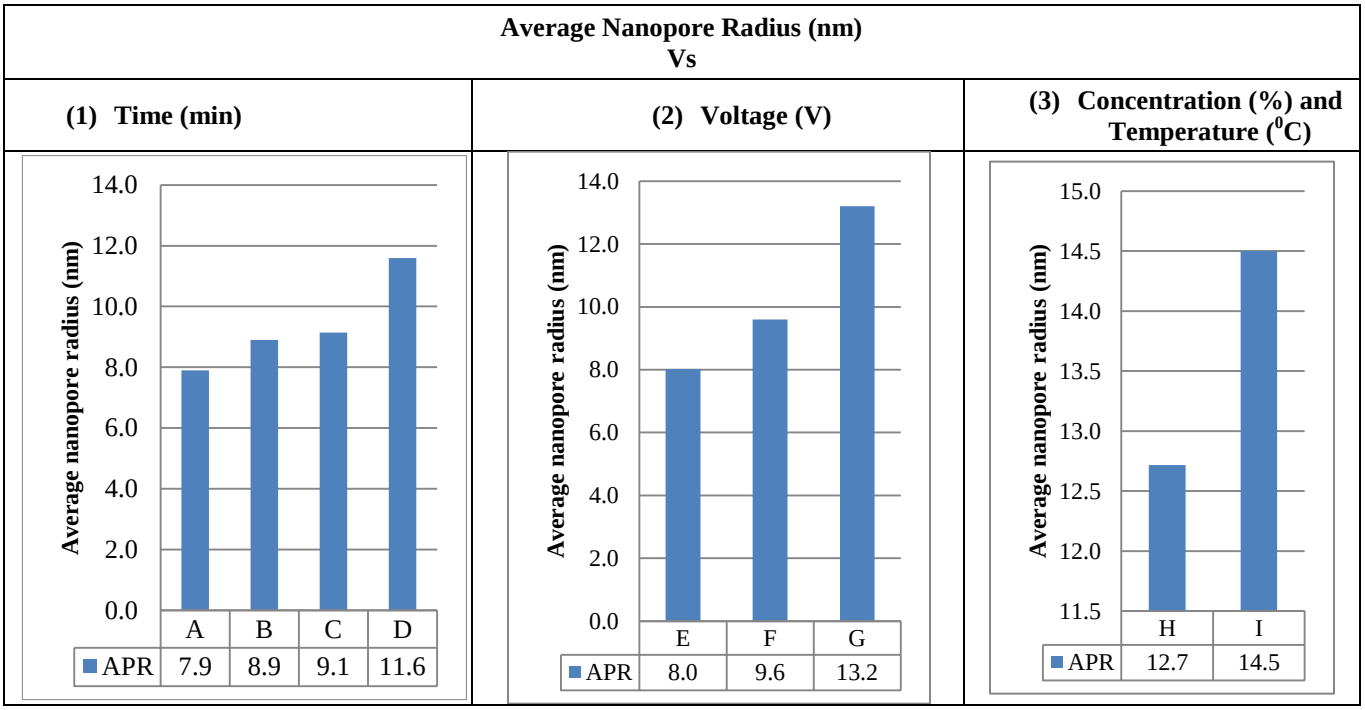


Fig. 6 Nanopore radius in FESEM images with varied anodization time, voltage, temperature and concentration

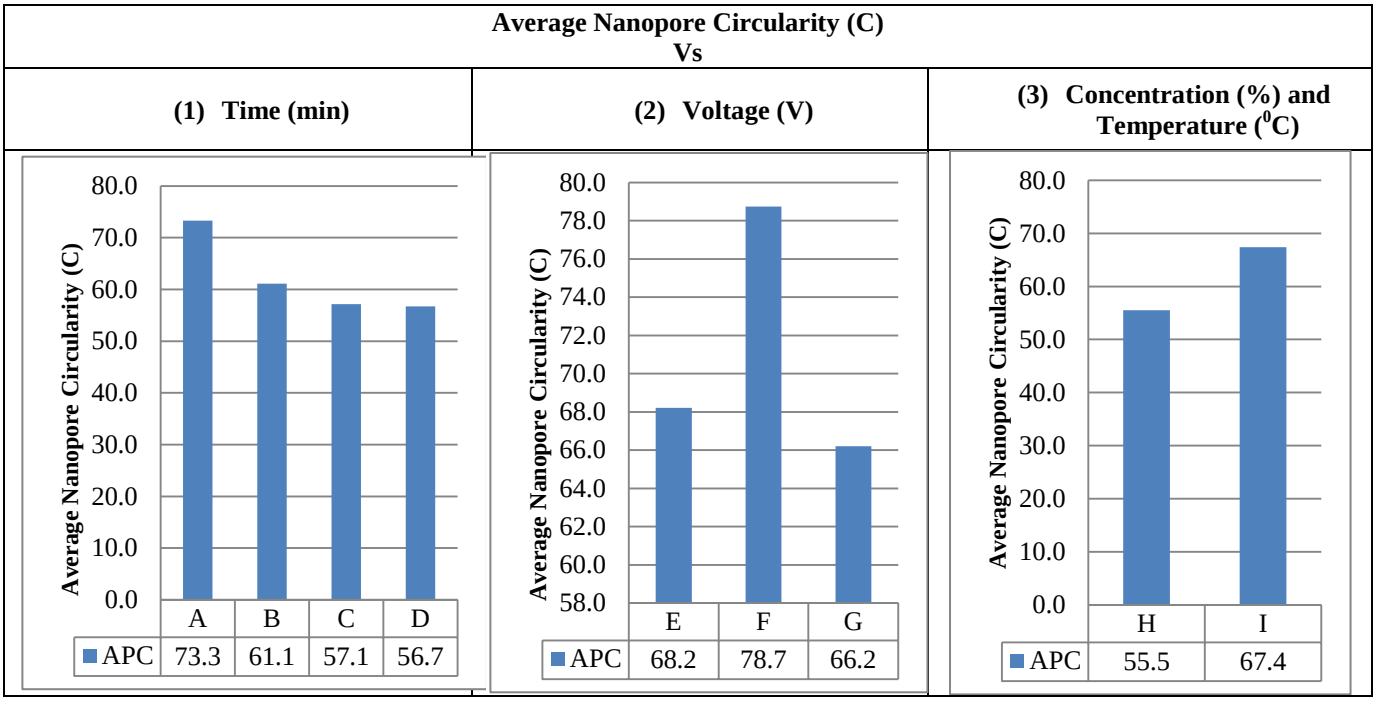


Fig. 7 Nanopore circularity in FESEM images with varied anodization time, voltage, temperature and concentration

V. Conclusion

An automated tool is developed to measure the geometrical features; nanopore radius and nanopore circularity of AAO with the varying anodising parameters. The experimental results depicts that as the anodization time is increased (5mins, 9mins, 20mins and 30mins), and the nanopore radius have gradually increased from 7.9nm to 11.6nm and nanopore circularity has decreased from 0.73 to 0.56 by keeping the constant value in concentration (5%), temperature (20^{0}C) and voltage (50 V). Similarly, as the voltage is increased (35V, 40V, 45V), the nanopore radius increased gradually from 8nm to 13.3nm and the nanopore circularity increased initially from 0.68 to 0.78 and decreased to 0.66 by maintaining the concentration (4.7%), time (8min) and temperature (5^{0}C) constant. The increase in nanopore radius from 55.6nm to 67.4nm and increased in nanopore circularity from 0.55 to 0.67 was witnessed when the concentration (4% and 5%) and temperature (20^{0}C and 25^{0}C) was altered keeping time (20 min) and voltage (50 V) constant. The extracted feature values of nanopore radius and nanopore circularity are verified by a chemical expert, which proves the exhaustiveness of the proposed results. Further, the total number of nanopores based on the geometric feature values of nanopore circularity, closer to 1 are 412 nanopores and 517 nanopores are closer to 0, out of total 929 nanopores from all the image data sets (A to I). The extracted properties of aluminium nanoporous images are may be useful in describing the important flow and transport mechanism and in predicting flow properties of different porous media relevant to numerous fundamental and industrial applications.

Acknowledgments

The authors are grateful to Indian Institute of Science (IISc), Bengaluru and Department of Chemistry, Rani Channamma University, Belagavi for providing the Al_2O_3 FESEM images and visualizing the manual results.

REFERENCES

[1] Marta M., Wojciech J. , Leszek R., "Characterization of nanopores arrangement of anodic alumina layers synthesized on low-(AA1050) and high-purity aluminium by two-step anodizing in sulfuric acid with addition of ethylene glycol at low temperature", *J Porous Mater*, Vol. 24, pp. 79–786, 2017.

[2] Guo Y., Zhang L., Han M., Wang X., Xie J., Deng L., "The effect of ethylene glycol on pore arrangement of anodic aluminium oxide prepared by hard anodization", R. Soc. open sci., Vol. 5, pp. 171412, 2018.

[3] Ghicov A., Schmuki P. "Self-ordering electrochemistry: a review on growth and functionality of TiO_2 nanotubes and other self-aligned MOx structures", Chem. Commun (Camb), Vol. 10, pp. 2791–2808, 2009.

[4] Ren Y., Ma Z., Bruce P. G., "Ordered mesoporous metal oxides: synthesis and applications", Chem. Soc. Rev, Vol. 41, pp. 4909–4927, 2012.

[5] Vega V., Javier G., Montero-Moreno J. M., Hernando B., Bachmann J., Prida V. M., Nielsch , "Unveiling the hard anodization regime of aluminum: insight into nanopores self-organization and growth mechanism", *ACS Appl. Mater. Interfaces*, Vol. 7, pp. 682–28 692, 2015.

[6] Md. Jani A.M., Losic D., Voelcker N. H., "Nanoporous anodic aluminium oxide: advances in surface engineering and emerging applications", *Prog. Mater. Sci*, Vol.58, pp. 636–704, 2013.

[7] C. T. Sousa, D. C. Leitao, M. P. Proenca, J. Ventura, A. M. Pereira, and J. P. Araujo, "Nanoporous alumina as templates for multifunctional applications", *App. Phy. Rev*., Vol. 1, pp.031102, 2014.

[8] Lin Zhou, Yingling Tan, Jingyang Wang, Weichao Xu, , Ye Yuan, Wenshan Cai, Shining Zhu, Jia Zhu, "3D self-assembly of aluminium nanoparticles for plasmon-enhanced solar desalination", Nature Photonics, Vol. 10, pp. 393-398, 2016.

[9] X. H. Yang, T. J. Lu and T. Kim, "A simplistic model for the tortuosity in two-phase close-celled porous media", *J. Phys. D: Appl. Phys*, Vol. 46, No. 12, 2013.

[10] Miao, X., Gerke, K. M., & Sizonenko, T. O., "A new way to parameterize hydraulic conductances of pore elements: A step towards creating pore-networks without pore shape simplifications", "Advances in Water Resources", Vol. 105, pp.162–172, 2013.

[11] Van der Linden, J. H., Sufian, A., Narsilio, G. A., Russell, A. R., Tordesillas, A., "A computational geometry approach to pore network construction for granular packings", Computers and Geosciences, Vol. 112, pp. 133-143, 2018.

[12] Joost H., Van Lindena, Adnan Sufianbc, Guillermo A., Narsilioa Adrian R. Russellb Antoinette Tordesillasd, "A computational geometry approach to pore network construction for granular packings", Vol 112, pp. 133-143, 2018.

[13] Parashuram B., Jalaja U., C. C. Vidyasagar, "Effect of Concentration and Temperature on Aluminium Oxide Nanopore FESEM Images", IJCEA, Vol. 11(1), pp. 222-229, 2018.

[14] Parashuram B., Jalaja U., C. C. Vidyasagar, "Aluminium Oxide (Al_2O_3) Nanopore Image Analysis Using Digital Image Processing Techniques", Proc. of IEEE Xplore. Int. Conf. on Computing, Communication, Control and Automation (ICCUBEA 2017), held at Pimpri Chinchwad College of Engineering, Pune, India, on 17th and 18th Aug. 2017.

[15] Parashuram B., Jalaja U., C. C. Vidyasagar, "Effect of Voltage on Aluminium Nanopore Images using Digital Image Processing Techniques", Journal of Nanoscience Nanoengineering and Applications, Vol. 8(3) , pp. 8-14, 2018.

Analysis of Software Reusablity Concepts Used In Automotive Software Development Using Model Based Design and Testing Tools

Dr. P. Sivakumar[*], Mrs. R.S. Sandhya Devi [**], Dr. B. Vinoth Kumar [***], Mr. R. Balaji [*], Mr. Dominic savari raj X[*]

[*] Department of EEE, PSG College of Technology, Coimbatore, India
[**] Department of EEE, Kumaraguru College of Technology Coimbatore, India
[***] Department of Information Technology, PSG College of Technology, Coimbatore, India

Abstract- Automotive electronics has gained very high significance over the past three decades as electronics in today's cars has already exceeded 20% of the total vehicle value, and is expected to increase further in the forthcoming years. Hence is an increasing demand to reduce the automotive Electronic Control Unit (ECU) development cost, enhance reliabilty, and shorten the development cycle time. Modularity facilitates code reuse, integration of third-party components is likely to reduce the development cost by introducing new features. In this paper, MATLAB is used to generate Application layer and RTE layer which in compliance with AUTOSAR. The generated AUTOSAR compliance RTE layer in MATLAB is completely independent of hardware, so the generated RTE is portable to any different hardware platform. Therefore the reuse of SWCs in automotive system is achieved using AUTOSAR standard. Model Based Design (MBD) is a promising way to meet the above demands. Most of the future innovations in the car industry will happen in the area of electronics, where the role of software is going to be very significant. Further research work in this field facilitates automobile manufacturers to develop vehicles that satisfy new government norms that focus on specific parameters like road safety and pollution

Index Terms- Reuse, Model Based Design, Model Based Testing, AUTOSAR.

I. INTRODUCTION

Ever increasing market demand and newer technology affecting today's competitive business environment globally. To enhance customization process several strategies such as; modular design concept, common platform, parts standardization, components commonality etc. can be adopted [1].

Automobiles in the last three to four decades have seen an unprecedented rise in use and importance of software [2]. Being lightweight is important when automobile weight is reduced then fuel consmpation is also reduced by 0.5 litres [3]. Therefore, another solution is necessary. Because software has no unit cost, the idea to reduce complexity by using standardized software suggests itself. Prominent standardized automotive software architecture is the AUTOSAR.

Model Based development under real-time constraint helps to launch new vehicles by reducing the valuable time required for testing. The Model Based methodology approach proposes the use of an architecture exploration tool to facilitate the rapid exploration of various CPUs, memories, and peripherals [16].

II. Reuseablity of Software in Automotive

Typically, functionalities changes only to a small amount from one vehicle generation to the next. Most of the old functionalities remains remain same can be found in the new car generation. Today the process of software reuse is not systematically planned between OEMs and Suppliers, as required, say, for software product lines [4].

In the end, a lot of the code could be generated from high-level models, which can be reused in product line approaches [5]. The organizational structure of the development process, with its interplay between OEMs and suppliers and the resulting conflicting desires for reuse, must also be taken into account, and could, in a second step, possibly be reshaped [6].

A. Reuse of Software: A Challenge for Automotive OEMs

More and more highly connected functions must be developed to series-production readiness, while at the same time development cycles become shorter and shorter. The importance of software in the automotive industry is shown in Figure 1. According to this study, in the year 2010, 13% of the production costs of a vehicle will go in software.

The main prerequisite for reusing software in the automotive domain is to separate the hardware of an ECU from the embedded software that runs on it. In order to fulfill the increased communication needs of these electronic systems, the ECUs communicate via different bus systems [7]. In requirements for the development of safety critical functions are stated for example, verification process, time-triggered architectures, and software architecture models [8].

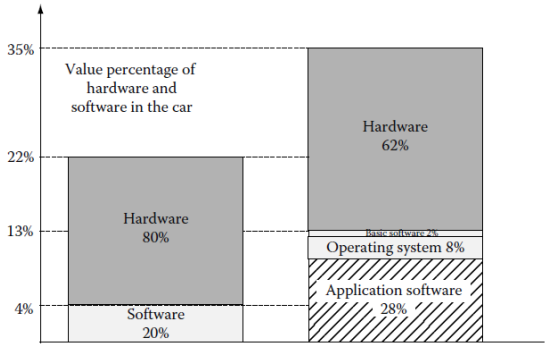


Figure.1 Rise of importance of software in the car

The increasing complexity of automotive electronics, the platform, distributed real-time embedded software, and the need for continuous evolution from one generation to the next has necessitated highly productive design approaches [9].

One of the most valuable aspects of Model-Based Design is the availability of executable models to perform Verification, Validation, and Test (VV&T) throughout the development process, especially its earliest stages [17].

III. Model Based Software Development Process for Automtoive Software Systems

Using Model Based Design the overall of cost- and time-consumption is heavily reduced because errors can be detected and corrected in early design phases [11]. Today, more than 3000 functions are realised by software running on ECUs, intelligent sensors and actuators [12] [13].

The reason for building those intermediate levels is the fact that it is much cheaper and faster to modify a model than to change the final product. The entire process is called MBD [14]. Additionally, methods such as Correct-by-Construction (CbyC) exist [15]. This method enables automatic code generation and ensures that the parameters that are verified on the model level can also be verified on the embedded code level [32].

MBD can be defined as a visual way to create complex control systems. The main target is to move a designer's focus from code programming to system design. An example on how to define it as a process is the V-model form, shown in Figure. 2 [31].

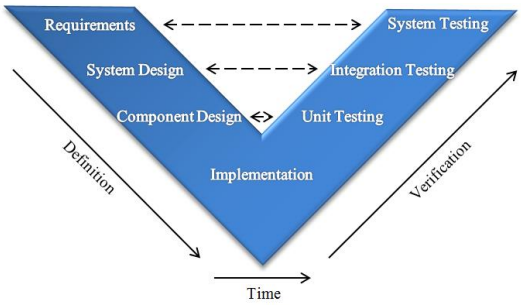


Figure.2 A V-model of the Model Based Design process

The big challenge for today's engineers is to manage complex design requirements. This is why the system design phase is often divided into two parts. The first part is to create an overall architecture consisting all the input, output and inner parameters of the designed system. The second part is to design the individual components. Dividing the system design into smaller subsystems gives an ability to divide the workload amongst a design team [18].

IV. Model Based Testing Approach for Software Development in Automotive

Here comes the most crucial step for yourresearch publication. Ensure the drafted paper is critically reviewed by your peers or any subject matter experts. Always try to get maximum review comments even if you are well confident about your paper.

Model based Testing (MBT) aids to link the process of creating test oracles [19]. MBT relates to a process of test generation from models related to a SUT by applying a number of sophisticated methods. MBT is defined as testing in which test cases are derived in their entirety or in part from a model that describes some aspects of the SUT based on selected criteria [19]. In the automotive industry, MBT describes all testing activities in the context of MBD.

Commercial tools such as Conformiq, Reactis have been developed to support MBT which can be used as an integrated part of the testing process [20] [33]. Currently the challenges present in automotive software development are to meet the brand survival demands, competition, quicker time to market, increased complexity advanced functionality, tight performance constraints and high reliability demands [21].

A. Reuablity & Implementation of Reusablity through AUTOSAR

The applications are decoupled from the hardware as shown in Figure.3. With a standardized interface it is possible to buy software and hardware from different manufacturers, and all work together [24].

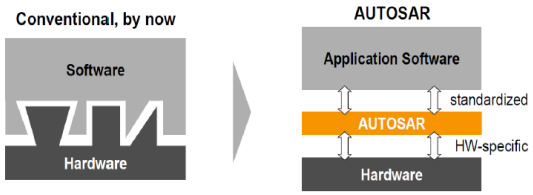


Figure. 3 Application software independent from the hardware using AUTOSAR

The complexity of automotive software systems continues to increase [21]. Modularity facilitates code reuse and integration of third-party components which is likely to reduce the development cost by introducing new features. ISO-26262 requires robust testing of the software units of a system [23] [31] [32]. The industry requires the need for tools to perform the robust testing of AUTOSAR components [10].

V. REUSE OF AUTOMOTIVE SOFTWARE

The MathWorks approach based upon MATLAB®, Simulink, and Real-Time Workshop Embedded Coder for generating AUTOSAR-compliant code follows a transparent and intuitive process [25] [32].

In the top-down workflow use an architecture design tool to design the architecture ECU network then export an XML-description of the corresponding components: the AUTOSAR Software Component Description. AUTOSAR System Target in Real-Time Workshop Embedded Coder needs to be selected [26]. The bottom-up workflow requires the same AUTOSAR configurations as described in the top-down workflow.

A. Software Environments for Automotive System

OSEK was developed to provide standard Real-Time OS and software architecture for various automotive ECUs [27]. The application software can be described in the form of a hardware independent model; advantage of this approach is the easy migration of proven function models from one vehicle to the next.

B. AUTOSAR Architecture

AUTOSAR is a development partnership of Automobile manufacturer's world wide, companies from the electronics, semiconductor, suppliers and software industry formed in 2003. This partnership has developed an industry wide standard for the automobile electronics, which is headed by the core partners – BMW Group, Bosch, and Continental [22].

In AUTOSAR, the ECU software is sub-classified as application layer, software BSW layer, and run time environment RTE. The layer of the architecture is shown in Figure.4.

The ECU Abstraction Layer offers consistent access to all features of an ECU like communication, memory or I/O. The drivers for such outside peripheral components reside in this layer.

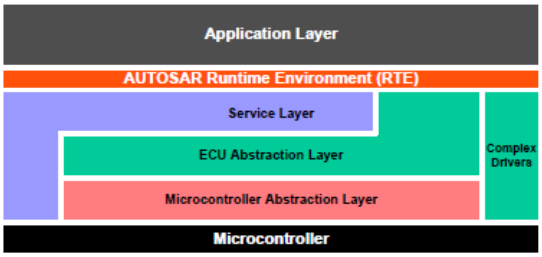


Figure. 4 Different layers of AUTOSAR

The Service Layer provides various types of background services such as vehicle network communication and memory management, diagnostic services, management services, ECU state management [35].

The RTE implements the data exchange and controls the integration between the application software section and BSW [35].

The AUTOSAR Software Component SWCs is a fundamental device concept of AUTOSAR, which is the fundamental structure of an AUTOSAR application Each SWCs is deployed, or mapped, on one ECU. For example, an automatic light adaptive application may consist of three AUTOSAR SWCs which are outside brightness detective component, a light request component and a light control master component [28]. The drivers do not follow the normal layered AUTOSAR architecture, but instead access both the microcontroller and RTE layers directly [29]. Detailed Comparsion metric with common software architecture based on different paramets as shown in Table1.

Table 1. Quality Attribute table for AUTOSAR.

Quality Attributes	**Architecture Metric**	**AUTOSAR / Common Architeture**	**Existing Architecture**
Complexity	Software component Complexity	1.688	2.164
Portablity	Dependency on Basic Software	0.249	0.3579
Testablity	Average Input Interface Complexity	1.123	3.183
Maintanablity	Average output Interface Complexity	0.734	2.367

From the table infered that, there is clear statisical differentiation between Existing Architecture and the Common Software Architecture [34]. Fig. 5 Shows the Statistical comparision of Existing Architecture and the Common Software Architecture.

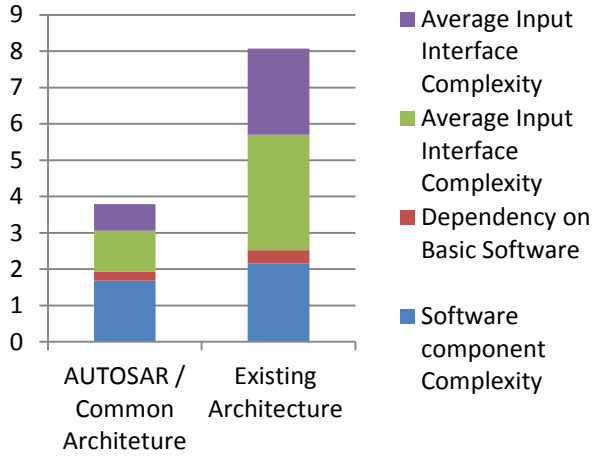


Figure.5 Statistical comparision of Architecture Metric

VI. Kernel Development Process

The AUTOSAR porting procedure on Raspberry Pi is described [36] [31]. Kernel devlopment of AUTOSAR OS consider this Raspberry Pi board as Hardware for the kernel development. This kernel expansion process includes initialization, exception handling, memory modeling and context switch. Initialization is the primary code that is executed when the OS starts up.

A Raspberry Pi's boot process starts with a small ROM based primary boot loader copying a file containing a secondary boot loader from SDHC card into the L2 cache of the microcontroller. The boot

loaders compute SDRAM and load a third boot loader into SDRAM which starts up the OS [30]. The bootloader of Raspberry Pi consist of bootcode.bin, loader.bin and start.elf. It boots at the ARM address space from the 0x00008000, which is the ARM entry point and location of the first ARM instruction.

A. Cross Compilation Results

After the completion of cross compiling a kernel image file along with bootloader files are generated. This kernel image corresponds to AUTOSAR OS. Kernel.img is the heart of the AUTOSAR OS which is be used for data communication through CAN protocol and also for sensor elements. Bootloader.bin and start.elf files are necessary for booting ARM processor from 0x00008000 of ARM address space and that is the ARM opening point and location for the initial ARM instruction.

B. Reusability of Software Components

In order to provide these reusabilty functionalities, a set of 6 SWCs is defined as shown in Figure.6 showing the SteerSensor and WheelSensor sensor SWCs; the SteerManager and WheelManager application SWCs; and the SteerActuator and WheelActuator actuator SWCs. The type of the ports used is sender-receiver. Data is acquired at the sensor SWCs level, transformed and then yielded to the 2 application SWCs. Data is also exchanged between application SWCs. Finally the data is sent to the actuator SWCs from the application SWCs. Each SWCs is composed of one or more Runnables.

The SteerManager SWCs is composed of 3 Runnables: Today car manufacturers produce large number of Customized Cars to meet customer demands. The objectives are to increase the ability to reuse software and to meet customization requirements.

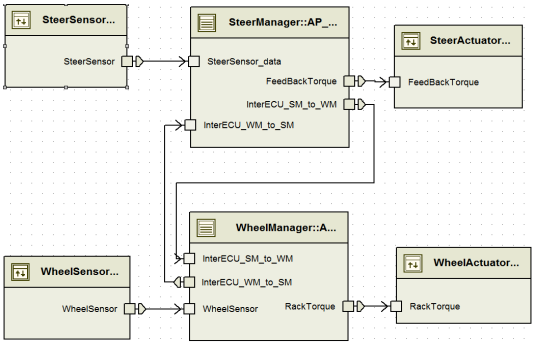


Figure.6 Typical steer-by-wire systems

This system contains two Actuator SWCs for the two direction indicators. The system also contains two Sensor SWCs to read the state of the two switches. The Direction Indicator Switch component sends the state of the switch to a service, while the Warning Light Switch component uses a Sender-Receiver communication to send the state of the switch to another component. The communication between the components is designed only virtual by using the VFB.

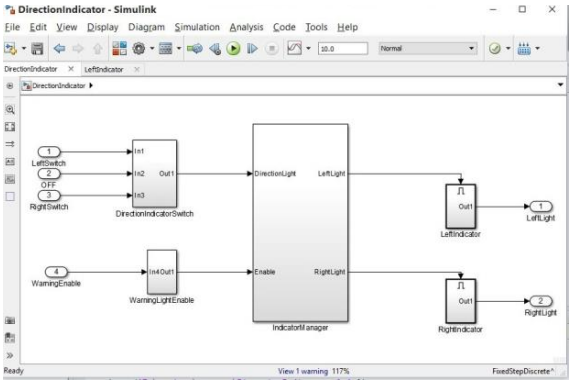


Figure.7 Design of a car direction indicator system

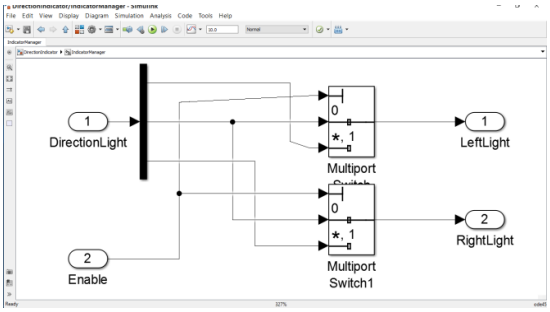


Fig.8 Internal component of indicator manger

The model is developed using Simulink block as shown in Figure.7. DirectionIndicatorSwitch to process the signal from the switch is used to select the direction Light to turn ON by means of input switch. Light blinks only if WarningLightEnable is enabled. IndicatorManager is used to issue the signal to blink the Indicator based on the input. Left/Right Indicator contains pulse generator and based on the logic issued to the manager it blinks the light. The internal component of indicator manager is shown in Figure.8. Baesd on the logic issued by Manager blinks the indicator.

C. Integration

MATLAB generates Application layer and RTE layer with AUTOSAR compliance code. Driver files for Raspberry Pi are readily available. RTE interface should be manually written, which interfaces the MATLAB generated RTE and Driver file. Finally integrate the RTE which is generated by MATLAB, for Raspberry Pi board to form a kernel.img file. Copy that kernel.img into the SD card and insert into the Raspberry Pi board and turn it ON. Creating more than two runnables i.e. tasks and integrating these runnables with OS service i.e. Real – Time OS is also possiable. The MATLAB generated RTE i.e. runnables and OSEK OS are integrated with the Hardware driver file and then converted it into kernel image. The kernel image is then loaded into SD card of the Rraspberry pi.

AUTOSAR uses its own notation for modeling applications. An application consists of one or more SWCs based in the Application layer. In order for SWCs to communicate with each other the Virtual Functional Bus (VFB) is used from SWC's point of view all it sees is the VFB and not the hardware dependent BSW and the hardware itself; in reality this is provided by the RTE.

In this paper, resuse of automotive software component is implemented using hardware and software tools it follows AUTOSAR-Compliant Code, AUTOSAR architecture it support both MBD and MBT development.

VII. CONCLUSION

One of the advantages of modularity is the ability to focus specialized engineering talents that results a higher quality product or services. Modularization can be applied as a framework that splits the activities of end products among OEMs and suppliers. It is a concept of design and develops smaller sub-systems independently which are able to function properly after assembled and tested with an endproduct. In auto industries, expanded use of modularity concept could create ample opportunities for the production of customized vehicles. In shorter term, modularity may not be very much effective but in longer run this has been proven to be beneficial for industries. The paper deals with the reuse of components, or subsystems, which is one way to achieve time to market. Thus AUTOSAR is developed for Raspberry Pi. Therefore AUTOSAR act as a platform for different embedded systems to speed up the development process and also provide an open hardware platform for experimentation on advanced automotive ECUs. In this paper, AUTOSAR compliance OS has been developed and the generated AUTOSAR compliance RTE layer using MATLAB, is completely independent of the hardware, which makes it portable to any different hardware platform. Therefore the reuse of SWCs in automotive system is achieved using AUTOSAR standard. The generated RTE is completely independent of hardware however the driver file depends on the specific hardware used in the design.

REFERENCES

[1] Liggesmeyer, P & Trapp, M, " Trends in embedded software engineering," *IEEE software*, vol. 26, no. 3.2009.

[2] Schlegel, J, "Technical foundations for the development of automotive embedded systems," *Technical report, Hasso-Plattner-Institut für Softwaresystemtechnik GmbH*. 2009

[3] Voget, S & Becker, M, "Establishing a software product line in an immature domain," *In International Conference on Software Product Lines*, Springer Berlin Heidelberg, 2002, pp. 60-67.

[4] Ahluwalia, J, Krüger, IH, Phillips, W & Meisinger, M, "Model-based run-time monitoring of end-to end deadlines," *In Proceedings of the 5th ACM international conference on Embedded software*, 2005, pp. 100-109.

[5] Dannenberg, J, & Kleinhans, C, "The coming age of collaboration in the automotive industry," *Mercer Management Journal*, vol. 17, no. 66, 2004, pp. 88-94.

[6] Krüger, A, Hardung, B & Kölzow, T, "Reuse of software in automotive electronics," *In Automotive Embedded Systems Handbook*, CRC Press , 2008, pp. 214-233.

[7] Simonot-Lion, F, "In car embedded electronic architectures: How to ensure their safety," *In 5th IFAC International Conference on Fieldbus Systems and their Applications*-FeT, 2003, pp. 1-8.

[8] Broy, M, Kirstan, S, Krcmar, H, Schätz, B & Zimmermann, J, "What is the benefit of a model-based design of embedded software systems in the car industry?," *Software Design and Development: Concepts, Methodologies, Tools, and Applications*, IGI Global,USA, 2013, pp. 310-334.

[9] Yu, H, Joshi, P, Talpin, JP, Shukla, S & Shiraishi, S, "The challenge of interoperability: Model-

based integration for automotive control software," *In Proceedings of the 52nd Annual Design Automation Conference*, ACM, Article 58, 2015, pp. 58:1- 58:6.

[10] Mitschang, J, "Evaluation of a Model-Based Development Process for Automotive Embedded Systems," *Doctoral dissertation, Kaiserslautern, Technische Universität Kaiserslautern, Diplomarbeit*, 2009.

[11] Kugele, S, "Model-Based Development of Software-intensive Automotive Systems," *Doctoral dissertation, München, Technische Universität München, Diss*., 2012

[12] Ebert, C, & Jones, C, "Embedded software: Facts, figures, and future," *Computer*, vol. 4, no. 4, 2009, pp. 42-52

[13] Yoon, H, "A Multi-layered Statechart Diagram Including Mitigation Behavior,*" International Journal of Computer Science and Network Security (IJCSNS)*, vol.15, no. 4, 2015, pp. 44-47.

[14] Baleani, M, Ferrari, A, Mangeruca, L, Sangiovanni-Vincentelli, AL, Freund, U, Schlenker, E, & Wolff, HJ, "Correct-by-construction transformations across design environments for model-based embedded software development," *In IEEE Design, Automation and Test in Europe, Proceedings*, 2005, pp. 1044-1049.

[15] Conrad, M, "Verification and validation according to ISO 26262: A workflow to facilitate the development of high-integrity software," *Embedded Real Time Software and Systems. MathWorks*, Inc. Natick, MA, USA, 2012.

[16] Lumpkin, E & Gabrick, M, "Hardware/software design and development process," (No. 2006-01-0170). *SAE Technical Paper*, 2006

[17] Murphy, B, Wakefield, A, & Friedman, J, "Best practices for verification, validation, and test in model-based design," (No. 2008-01-1469). *SAE Technical Paper*.2008.

[18] Skruch, P & Buchala, G, "Model-Based Real-Time Testing of Embedded Automotive Systems,*" SAE International Journal of Passenger Cars-Electronic and Electrical Systems*, vol. 7, no. 2, 2014, pp. 337-344.

[19] Bringmann, E & Kramer, A, "Model based testing of Automotive Systems," *International Conference on Software Testing, Verification and Validation*, 2008, pp. 485-493.

[20] Broy, M, "Challenges in automotive software engineering," *Proceedings of the 28th international conference on Software engineering*, ACM, 2008, pp. 33-42.

[21] Marinescu, R, Saadatmand, M, Bucaioni, A, Seceleanu, C, & Pettersson, P, "A model-based testing framework for automotive embedded systems," *In Software Engineering and Advanced Applications (SEAA)*, 40th EUROMICRO Conference. 2014, pp. 38-47.

[22] Mössinger, J, "Software in Automotive Systems," *IEEE Software*, vol. 27, no. 2, 2010, pp. 92-94.

[23] Melin, J & Bostrom, D, "Applying AUTOSAR in Practice Available Development Tools and Migration Paths," *Master Thesis, Computer Science*, School of Innovation Design and Engineering, Mälardalen University., 2010.

[24] Köhler, A & Volkswagen, AG, "AUTOSAR-Compliant Functional Modeling with MATLAB®, Simulink®, Stateflow® and Real-Time Workshop® Embedded Coder of a Serial Comfort Body

Controller," *In MathWorks Automotive Conference*, 2007.

[25] Sandmann, G & Thompson, R, ":Development of AUTOSAR software components within model-based design," (No. 2008-01-0383). *SAE Technical Paper*. 2008.

[26] Cho, J, "Software Development Environment for Automotive SoC. In Algorithm & SoC Design for Automotive Vision Systems," *Springer Netherlands*, 2014, pp. 231-261.

[27] Park, G, Ku, D, Lee, S, Won, WJ, & Jung, W, "Test methods of the AUTOSAR application software components," *In IEEE ICCAS-SICE*, 2009, pp. 2601-2606.

[28] Wu, R, Li, H, Yao, M, Wang, J & Yang, Y, "A hierarchical modeling method for AUTOSAR software components," *In IEEE Computer Engineering and Technology (ICCET), 2nd International Conference*, vol. 4, 2010, pp. V4-144 –V4-184.

[29] Bo, H, Hui, D, Dafang, W & Guifan, Z, "Basic concepts on AUTOSAR development," *In IEEE Intelligent Computation Technology and Automation (ICICTA), International Conference*, vol. 1, 2010, pp. 871-873.

[30] Voget, S & Becker, M, "Establishing a software product line in an immature domain," *In International Conference on Software Product Lines*, Springer Berlin Heidelberg, 2002, pp. 60-67.

[31] Sivakumar, P. and Vinod, B. and Devi, R.S. Sandhya and Divya, R, "Novelty Testing Measures and Defect Management in Automotive Software Development," *Australian Journal of Basic and Applied Sciences*, 2016, Vol. 10(1), Pages: 607-613.

[32] Sivakumar, P , Vinod, B. , Sandhya Devi, R. and S.Poorani Nithilavallee, "Model Based Design Approach In Automotive Software and Systems," *International Journal of Applied Engineering Research*, 2016, ISSN 0973-4562 Volume 10, Number 11,pp. 29857-29865.

[33] Sivakumar, P, Vinod, B, Sandhys Devi, R. and Divya, R, "Deployment of Effective Testing Methodology in Automotive Software Development," *Circuits and Systems*, 2016, 7, 2568-2577. doi: 10.4236/cs.2016.79222

[34] Sreekanth, A., Srikanth, K., Aditya, C., Satish, T. and Ramchandran, R, " Deploying common software system for hybrid electric vehicles in AUTOSAR way," *In Transportation Electrification Conference (ITEC-India)*, 2017 IEEE, pp. 1-6.

[35] Sandhya Devi R.S., Sivakumar P., Balaji R, "AUTOSAR Architecture Based Kernel Development for Automotive Application," *In: Hemanth J., Fernando X., Lafata P., Baig Z. (eds) International Conference on Intelligent Data Communication Technologies and Internet of Things (ICICI)* 2018. Lecture Notes on Data Engineering and Communications Technologies, vol 26. Springer, Cham.

[36] M. M. Srihari and P. Sivakumar, "Implementation of Multi-Function Printer for Professional Institutions," *International Conference on Inventive Research in Computing Applications* (ICIRCA), Coimbatore, 2018, pp. 1-3.doi: 10.1109/ICIRCA.2018.8597283

Security and Privacy in Smart Systems

Manohar Mohith*, Naresh E, Vijaya Kumar B.P****
Dept. of Information Science and Engineering, M. S. Ramaiah Institute of Technology, MSR Nagar, Bengaluru, INDIA

Abstract- Smart systems are those that have the capability to modernize the way that perform daily activities, the applications of smart are extended but not limited to the way in which do some of the basic activities like start a car. Nowadays, even the government equipment are coming up with fascinating technologies to improve efficiency of its performance. These systems however also present the same amount of vulnerability as that of advantage. This is precedent to the fact that all of these smart systems have data vulnerability and many other privacy breach issues based on the type of its application. However, this should not be a factor to hinder innovation and this brings up the necessity to protect and preserve privacy.

This paper discusses about a smart system, it's technological stacks and security issues which evidently provides extensive basis to the Security and Privacy of these systems. It starts with the definition of a smart systems in an ecosystem considering a car as an example, it examines the various vulnerabilities in details considering the recent breaches by Cambridge Analytica as well. A mathematical approach to the cause of data breach has been considered to analyse its effects and finally the technological approaches to solving this issue. This paper explains not only about the existing approach to security but also the modern day practices to resolve it.

Index Terms- Smart Systems, Internet of things, Artificial Intelligence, Vulnerability, Security.

I. INTRODUCTION TO SMART SYSTEMS

What are smart systems? When the word 'smart' is used as an adjective with any of the common things present in this world, the general idea is that it has modern technologies integrated with it which may lead to, but, is not limited to reducing costs, improving conditions of living, decreasing time of processing, increasing scalability, feasibility, etc. It is the cluster of these smart systems that constitute the rapid growth in technology in today's modern world [1].

A smart system in the field of Information Technology is with respect to the Internet of Things(IoT) that maybe, in simple words, either controlled by humans in order to achieve something smart or autonomously programmed in order to solve some of the emerging problems due to the rapid technological advances. To do a set of inbuilt commands in either of the cases is IoT, while on the other hand, if it has the ability to learn from itself then it has Artificial Intelligence integrated into it. Nowadays, 'smartness' has been integrated into almost every small thing around us that eventually has built up to form one large smart ecosystem [11, 7].

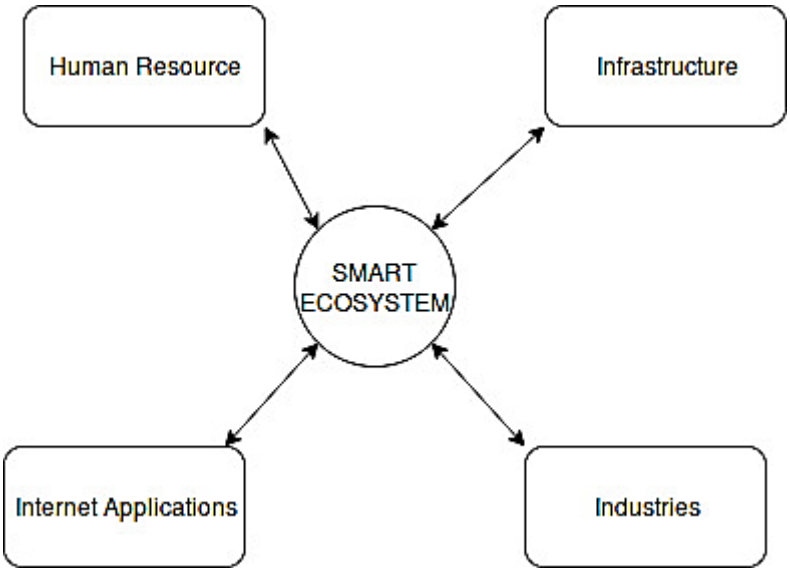


Fig 1. Smart ecosystem

II. WHY SMART SYSTEMS (IN THE FIELD OF IT)?

In the field of Information Technology, in order to ensure only forward growth in the technologies and no repetitions or redundancies, the connectivity or network plays an important role in this case as well as to integrate cognitive abilities in the systems used in daily life thereby to give it 'smartness'. This connectivity, as shown in Fig 1, which denotes the ecosystem as a whole plays an important role in improving the efficiency of the system which is the main intention [11, 7].

The two main reasons in general for using smart systems are:

- User ease of experience
- Machine adaptability and efficiency
  [Here, efficiency refers to multiple parameters like speed, time, computing power, etc.] [7, 1]

Each system has its own perspective of how it can be improved in order to either improve its efficiency or improve the user experience, thus justifying its requirement in today's world.

III. SECURITY IN SMART SYSTEMS

First of all, the main question that has to be addressed is, why security? If the systems are so smart and integrated with intelligence (Artificial Intelligence), then why do they require security? To understand this, let us consider a case study of a smart car and then summarize few of the general points of security breaches in smart systems. There are many safety issues with respect to the driver and passengers as well during the journey, however, we will only be quoting the software vulnerability. [11, 3]. As a side note, to be clear, these vulnerabilities are only with respect to non-autonomous cars, however, autonomous cars have more vulnerability considering higher connectivity with the internet and other connected devices [16, 8].

3.1. Case study- Smart car: [3]

Let us consider all its applications with respect to every aspect of the car that makes it smart and what are the vulnerabilities in each case [7, 3].

- Interface
- Physical components
- Software Applications

Fig 2. Interconnectivity between smart devices

3.1.1. Interface: [8]

This is the software interface used to connect the user to the hardware of the car thus allowing him/her to make use of the smartness of the car to its maximum extent. As shown in Fig 2, which denotes the smart links nowadays, the connectivity makes the vulnerability of this part of the network to be the highest percentage which can also be linked to the movie 'The Fate and the Furious' that might not be too much of science fiction after all. The interface vulnerabilities could be of the following types: [7, 6]

- Infotainment system & connected networks: [6, 5]
  This is the radio/entertainment system present in the car thus allowing the user to get news updates, play shows and even run many other applications that are related to the infotainment field on the display system screen installed in the car or through the audio system installed (the network has great vulnerability if left unprotected). [10] The main hacks according to recent estimates have been mainly through the smartphones that are connected to the car which makes it easier for the hackers to access and control any of the physical parts of the car through the connected network.

- Navigation: [1]
  This is one of the globally used interface, whether it is to go to an unknown location or a known location. [12, 10] However, the vulnerabilities of this application are numerous especially considering the fact that the current location of the car is being publicly vulnerable to hackers. [5, 3] The attacks can be directly through the data obtained through the cloud, through manually accessing the smart car at its location or even remotely gaining access to the car through the location range at which it is currently on route.

- Advanced unlocking systems: [13, 10, 3]

With the arrival of keyless entry into this world, the vulnerability of the cars has gone up much more than ever expected. The vulnerabilities are shown accordingly in Table 1.

Table 1. Vulnerability of smart applications [6, 3]

Type of smart system	Journey	Vulnerability	Medium of attack	Risk level
Infotainment system + connected network	Pre journey	Data leak/loss	Cloud stored data/Account hack	Low
	During	Car access + data loss/leak	Intruder through car offline/cloud online	Medium
	Post	Easy access of data + full control of insecure car	Through the unknown connected network/ parking area hack	High
Navigation	Pre journey	Destination/relevant data loss	Through destination and house location details noted when connected to cloud	Low
	During	Location/full control of smart car	Cloud based location and other connected networks	High
	Post	Destination/relevant data loss	Through destination and physically vulnerable equipment at the destination	Medium
Advanced unlocking systems	Pre journey	Unlocking of car through connected devices at home	Access to the cloud network connected at home	High
	During	Cloud data loss/sensors hack	Sensors in range of the car	High
	Post	Complete control of unattended car	Sensors in range of the car	High
Smart car control(voice/remote)	Pre journey	Voice tracking/connected remote's vicinity	Manually in range of remote, bugging the car's network	High
	During	Cloud data loss/sensors hack	Not much chances unless manually accessible	Low
	Post	Voice tracking/connected remote's vicinity	Manually in range of remote, bugging the car's network	High

- Smart car control: [13, 8, 5]
  When the term smart car control arises, it is not only with respect to the voice controlled key that is at vicinity and in easily accessible range with respect to the hackers but also the variety of devices that are connected with the car that makes it more vulnerable and easier

to access [13, 12, 9]. The devices themselves can also be hacked to gain access to the car through various sensors and devices at its vicinity.

3.1.2. Physical car components:

When it comes to the physical components of the car, every part of the car that is part of the 'smart' network is at huge risk if left unattended without proper security measures. Especially considering the plethora of functions that can voice/app controlled from devices such as the mobile, laptops, etc [13, 12, 10].

- Radio
- Digital display
- Air conditioning
- Windshield wipers
- Wiper fluid
- Transmission
- Brakes

All these car components are left to become extremely vulnerable when connected to the cloud however, what really is disturbing to know, are all the controlling components of the car are at the attacker's vicinity [12, 10]. This means that even how our car goes physically on the road is also at stake which raises a lot of safety issues as well [8].

3.1.3. Software Applications: [3]

These can be a wide range of software starting from the messaging app used in the car to the keyless control that is used in the car [5]. Any app basically connected to the car in turn to the cloud is vulnerable to attack either through the cloud or through the app used in the car. The access to this is partly our fault as well, considering that the data given to these apps are in abundance basically, pointing out where and when to attack. This clearly indicates the importance of using trusted apps over networks [13, 12].

Despite trust of the apps over a network, [6] the user is still liable for managing their own private data which is clear from the data that was being misused through the Cambridge Analytica incident which clearly proves, anything is possible if data is not protected.

The country-wise data, as per Fig 3, which denotes the Vulnerability data provided by Facebook during the recent data breach by Cambridge Analytica, clearly distinguishes the fact that smart connected networks are highly vulnerable to attacks since the United States is clearly abundant with the smart cars network. An example to prove the previous statement lies with Tesla [4]. which has integrated the smart network in such a way that if one car is successfully penetrated, the entire network of cars is open to vulnerability [9, 6, 5].

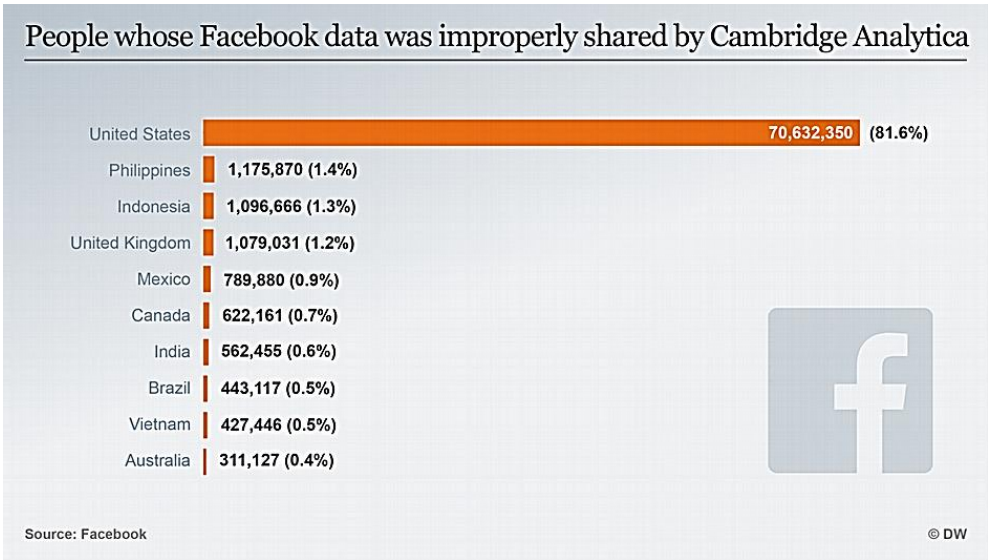


Fig 3. Vulnerability of data due to Cambridge Analytica (Country wise), by Facebook

The further part of the paper clearly discusses the general points of the security of smart systems, i.e., Why, what happens and how to stop it?

IV.CAUSES – VULNERABILITY AND THE NEED FOR SECURITY IN SMART SYSTEMS

With the exponential growth in the technology in today's world, it is very important to keep the smart devices safe. Also, considering the fact that all smart devices are interconnected. [12, 3]

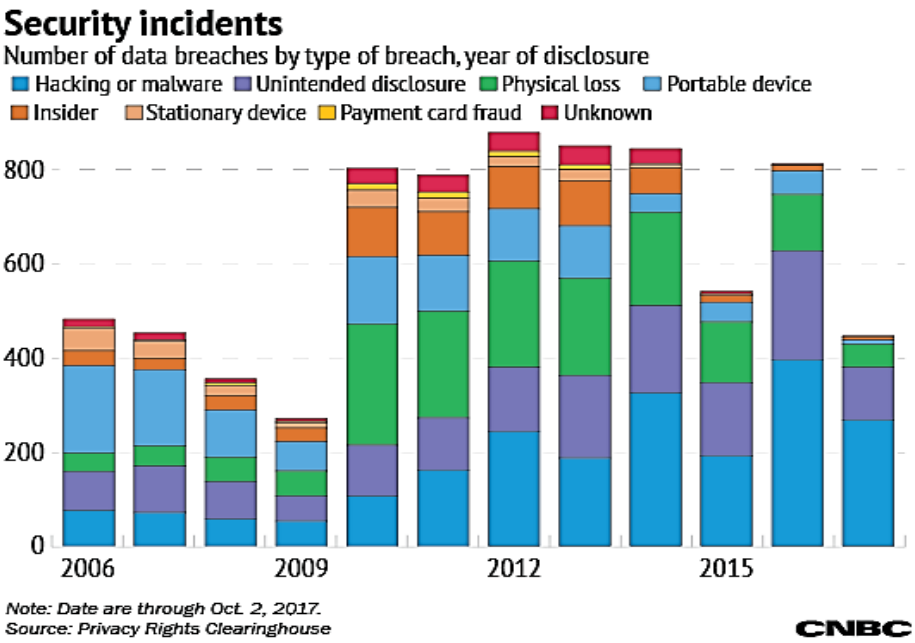


Fig 4. Security vulnerability of smart systems by type, by CNBC

Let us now consider the 'v' as amount of vulnerability of smart devices and 'y' the year under consideration. It is clear from Figure 4 that one conclusion can be drawn: [9]

$$\mathbf{v \ \alpha \ y} \qquad \ldots(1)$$

Converting it into an equation, we need to assume a constant 'k' which represents the amount of vulnerability based on the smart device.[14]

i.e.,

$$\mathbf{v = ky} \qquad \ldots(2)$$

k is lesser for non-connected network smart devices than the connected ones. Also , if 'n' represents the number of smart devices, then, from Fig 4, which shows the statistics of data breaches, based on type provided by CNBC in the year 2017, we get

$$\mathbf{n \ \alpha \ 2v} \qquad \ldots(3)$$

This clearly proves that, as the years go by, due to the increasing availability of smart devices to every individual, the vulnerability is going nowhere and on the contrary it is essentially doubling its rate [16, 10].

V. EFFECTS – WHAT DOES BREAKING INTO SMART SYSTEMS DO TO YOUR DATA

The main thing to be addressed are the concerns to what are the consequences of break-ins to the smart devices. With the existing technology of cloud connected devices, the vulnerability issue is now a major concern considering that all the devices are synced over the cloud, in other words, if one system is hacked, all the systems are doomed. However, what does it mean to be 'doomed'? Some of the major concerns to smart hacks are listed below: [16, 2, 1]

- Data breach: [5]
  The main loss that occurs due to an attack is the loos of sensitive data that leads to a lot of other interconnected losses. [17, 15, 14] For example, loss of the car password through an attack might lead to the loss of the control towards the car and makes the user prone to safety.

- Control of smart devices: [15, 14, 12, 8]
  With the rising of use of Tech Giant's latest cloud sync technology, it makes all the smart devices connected over a cloud vulnerable, meaning that if one of the device is attacked, it could mean an entire takeover of all the smart devices owned.

- Financial loss: [5, 1]
  This is the most unfortunate consequence of all the effects of an attack on any smart device. Devices may be targeted and attacked for financial benefits by the hackers, which is a nightmare if any of the smart devices contain sensitive financial information stored in it [15].

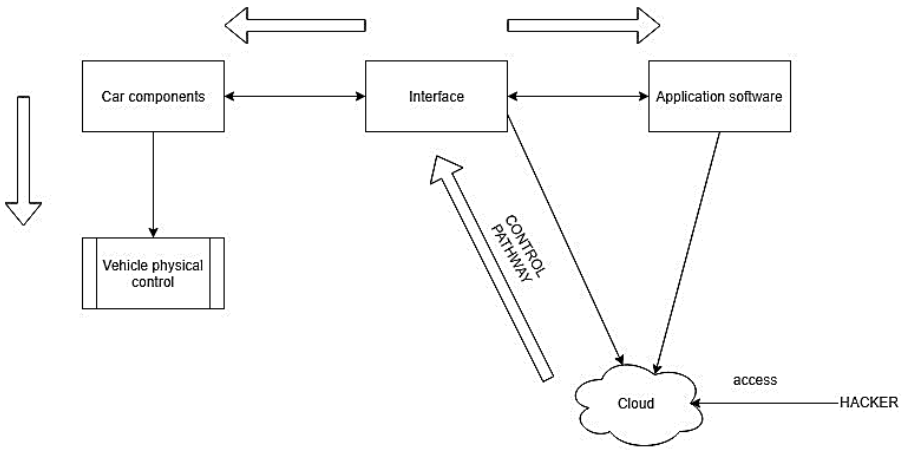


Fig 5. The vulnerability pathway access in smart cars [2]

Fig 5, which denotes the pathway of travel of control through the entire Smart network based on self-research, is an example and insight into what might happen if a hacker is able to get hold of one part of a network and thus increase its exposure, especially that of sensitive data. [17, 6]

VI. MEASURES – SECURITY THAT IS CUTTING EDGE AND SHOULD BE USED TODAY

Existing security technology for smart systems: [3, 1]

The existing although might not be that advanced in protection, but protects maximum amount of vulnerability of users against attackers. Some of the basic self-protection techniques include: [18, 10, 3, 1]

- Ensure connection with the manufacturer
- Regular updating of interface/application software's
- Keep smart devices in close vicinity
- Do not disclose sensitive data [2, 1]
- Ensure the genuineness of devices interconnected

Some of the existing technologies to protect these smart devices in pre and post attack include: [1]

- Cryptography/Encryption techniques: [16, 5]
  To encrypt the data within certain parameters that are known only to the user thereby ensuring privacy of data.

- Intrusion detection systems: [5, 3, 1]
  This is to mainly ensure that the smart systems are not taken advantage by detecting intrusion in the first place as a preventive measure to avoid breach in networks. This technique is the most effective measure in terms of costs. It also includes the domain of hash lock mechanism.

- Cyber-Physical Smart Systems: [19]
  This is to provide the smart system an AI based approach to prevent cyber attacks. This follows a concept of integrating both human and system with other systems as well which
  is wholly the entire concept of Tesla that ensures security of its systems through its connected network by learning from each other.

- IoT:
  This is a huge field of applications which if used in the right way, can be pulled over to the security field. From common knowledge, this fits in the field of communication thus involving a lot of the protocols to transfer (E.g.: TCP/IP). This being said, this technique involves providing security while it is being transferred.

All these technologies can be used in order to ensure safety of the smart devices at bay, however they do not ensure the safety of the devices themselves, which ultimately relies on the main fact that the owner has to be responsible [18, 14].

If the device does get hacked despite all these measures, the cybercrime control is the best approach in dealing with the situation. However, the details of which are out of scope in this paper.

VII. CONCLUSION

With the appearance of wireless communication in the car, and the connection of the under-the-hood elements with the outer world, new security threats arise on the area of Connected Cars. ICT security needs penetrate into the vehicles. Because of this, privacy and security of the owner must be handled and the well-known security principles must be taken into account during the design of such system: confidentiality, integrity, authenticity, availability and non-repudiation [23]. With the appearance of wireless communication in the car, and the connection of the under-the-hood elements with the outer world, new security threats arise on the area of Connected Cars. ICT security needs penetrate into the vehicles. Because of this, privacy and security of the owner must be handled and the well-known security principles must be taken into account during the design of such system: confidentiality, integrity, authenticity, availability and non-repudiation [23] With the appearance of wireless communication in the car, and the connection of the under-the-hood elements with the outer world, new security threats arise on the area of Connected Cars. ICT security needs penetrate into the vehicles. Because of this, privacy and security of the owner must be handled and the well-known security principles must be taken into account during the design of such system: confidentiality, integrity, authenticity, availability and non-repudiation [23] With the appearance of wireless communication in the car, and the connection of the under-the-hood elements with the outer world, new security threats arise on the area of Connected Cars. ICT security needs penetrate into the vehicles. Because of this, privacy and security of the owner must be handled and the well-known security principles must be taken into account during the design of such system: confidentiality, integrity, authenticity, availability and non-repudiation [23]

With the rise in the ‘smart’ era, however exciting it is to the people to keep up with the trending technology, [1] it is also important to keep any data or vulnerable equipment related to those smart devices at close quarters and under utmost discretion to avoid it from being misused for various unlawful purposes. This is clearly stated proven from the recent Facebook & Cambridge Analytica data breach incident [9].

In the technical perspective, when we start from the very basics of a system consisting of many smart devices that are connected together in an ecosystem as long as we keep trying to improve the ease of use and machine adaptability, vulnerability always increases exponentially. Whether it is a smart car or any other smart device, it is always open to a data or control breach for either personal/financial reasons of the attacker towards the victim. However, modern day technologies do provide solutions to cope with these attacks that range from cryptography techniques, intrusion detection to cyber-physical systems and IoT.

REFERENCES

[1] Carsten Maple, “Security and privacy in the internet of things”, 2017.

[2] Huichen Lin & Neil W. Bergmann , “IoT Privacy and Security Challenges for Smart Home Environments”, 2016.

[3] Mauro Conti, Ali Dehghantanha, Katrin Franke & Steve Watson, “Internet of Things Security and Forensics: Challenges and Opportunities”, 2018.

[4] Zhen Ling , Junzhou Luo , Yiling Xu , Chao Gao , Kui Wu & Xinwen Fu, “Security Vulnerabilities of Internet of Things: A Case Study of the Smart Plug System”, 2017.

[5] Sarah J. Darby, “Smart technology in the home: time for more clarity”, 2018.

[6] Tamás Bécsi, Szilárd Aradi & Péter Gáspár, “Security issues and vulnerabilities in connected car systems”, 2015.

[7] P. Varaiya, “Smart cars on smart roads: problems of control”, 1993.

[8] Fei-Yue Wang, Daniel Zeng & Liuqing Yang, “Smart Cars on Smart Roads: An IEEE Intelligent Transportation Systems Society Update”, 2016.

[9] JC Reindl, “With more tech comes more hacking risk for smart cars connecting vehicles to one another and the world around them poses a substantial cybersecurity hurdle that automakers must overcome.”, 2018.

[10] Clayton Weeks, “How to prevent your car from being hacked”, 2018.

[11] NiluferTuptuk & StephenHailes, “Security of smart manufacturing systems”, 2018.

[12] Amara Dinesh Kumar, Koti Naga Renu Chebrolu, Vinayakumar R & Soman KP, “A Brief Survey on Autonomous Vehicle Possible Attacks, Exploits and Vulnerabilities”, 2018.

[13] Scott l. Wenzel, “Not even remotely liable: smart car hacking liability”, 2017.

[14] Ishtiaq Roufa , Rob Millerb , Hossen Mustafaa , Travis Taylora , Sangho Ohb Wenyuan Xua , Marco Gruteserb , Wade Trappeb & Ivan Seskarb, “Security and Privacy Vulnerabilities of In-Car Wireless Networks: A Tire Pressure Monitoring System Case Study”, 2010.

[15] A K M Bahalul Haque, “Need for critical cyber defence, security strategy and privacy policy in bangladesh - hype or reality?”, 2019.

[16] Hamidreza Damghani, Heliasadat Hosseinian & Leila Damghani, “Privacy Risks of Hybrid Broadcast Broadband TV (HbbTV)”, 2019.

[17] Randall Young,Lixuan Zhang & Victor R. Prybutok, “Hacking into the Minds of Hackers”, 2007.

[18] Meenaakshi N. Munjal, “Ethical hacking: an impact on society”, 2014.

[19] Prof. Dr.-Ing. Reiner Anderl, “Industrie 4.0 - Advanced Engineering of Smart Products and Smart Production”, 2014

A Survey on Web Page Recommender Systems

Bhoomika AP [*], Dr. Selvarani R [**]

*Department of Computer Science, ACED, Alliance University
**Department of Computer Science, ACED, Alliance University

Abstract- The huge information in the web has created a big challenge for users to find relevant and useful information. Web page recommender systems solve this problem by suggesting the products or web pages of user interest. There are different types recommender systems generate useful recommendations based on navigational behavior of user. FAST algorithm is one of the sequential frequent pattern mining algorithms that is capable of mining complete set of patterns by greatly reducing the effort for support counting and candidate sequences generation. This paper discusses different types of recommender systems and proposes a hybrid context aware recommender system that is based on integrating semantic knowledge at different stages of web usage mining and FAST algorithm, coupled with clustering and sequential association rule mining.

Index Terms- web page recommendation, web usage mining, sequential patterns, association rule mining.

I. INTRODUCTION

World Wide Web (WWW) is the biggest source of information. The size and complexity of web is getting increased as it tends to grow at an exponential rate [1]. Although web contains huge amount of information that is useful to the users, there is a problem in finding desired information easily. Often users find it difficult to extract most relevant information in the right time from large information space. This is very crucial in e-commerce web sites where they can lose the customers very easily. So, it is necessary for users to make use of automated tools such as recommender systems in order to obtain desired information.

There are different approaches in recommending web pages.

1. Content based filtering
2. Collaborative filtering
3. Web usage based recommender system

In content based filtering the key words in the web pages are used to describe the items. In this approach, a user profile is built based on his preferences to indicate the type of items he likes [2]. Most similar items are recommended by comparing various items in the web pages with the items that are previously rated by the user. Collaborative filtering methods generate recommendations to a given user based on their similarity to other users. This method works by collecting and analyzing information on users' behaviors, activities or preferences in order to generate recommendations.

Web usage mining is a technique that analyses user information to find access pattern of web pages. The user related information can be usually obtained from web server log. The web server log is further mined to obtain usage patterns. Recommendations are generated to a given user by comparing the active user session with previously mined usage patterns.

There is hybrid recommender systems, which are the combination of the different approaches discussed above. Sparsity and scalability problems are the disadvantages of collaborative filtering. Content based filtering fail to predict the future interests of user since solely the content of the webpage is considered for recommendation. Recommendations generated from web usage patterns suffer from new item problem, where the recommender system fails to recommend newly added items to user. In recent days researches have been carried out to combine semantic information at various stages of web page recommender system based on web usage mining to generate meaningful and accurate recommendations.

Prefix span algorithm is used to generate frequent sequential navigational patterns. These patterns are in turn used to mine association rules and to generate web page recommendations. The proposed approach uses FAST algorithm to generate sequential patterns. Researches have showed that FAST algorithm outperforms prefix span algorithm and other existing sequential pattern mining algorithms in generating sequential patterns in terms of speed and memory. Lastly sequential rules can be mined from generated patterns. Meaningful and accurate recommendations are generated for given user by considering active user session, context information and previously mined sequential rules.

II. RELATED WORK

Robin van Meteren and Maarten van Someren presented [3] a recommender system that makes use of content-based filtering techniques to suggest items to users. A content-based filtering system selects items based on the correlation between the content of the items and the user's preferences. Explicit feedback and implicit feedback from user and tf-idf scheme are considered to find the document they are interested in. The recommender system provides dynamic hyperlinks to web pages that contain the items of user interest. The presented system failed to predict the future interests of user.

Badrul Sarwar et al presented a recommender system [4] based on item based collaborative filtering. The conventional collaborative algorithms suffered with scalability and sparsity problem. To alleviate these problems, item-based collaborative filtering considers the relationships between items over relationships between users. Cosine based similarity and adjusted cosine based similarity are used to find the similarity of items. The presented system produced much faster recommendations when compared to conventional user based collaborative filtering systems.

Sule Gunduz et al presented an approach [5] that considers order of pages in a session, the distance between identical pages, and the time spent on these pages for providing recommendations. The authors introduced a similarity metric to find pairwise similarities between user sessions. Clustering of user sessions is carried out based on pair wise similarity and the clusters are represented by using a click – stream tree. The click-stream tree is used to generate recommendations to the users.

Jia Li et al [6] present a framework for a combined web recommender system, in which users' navigational patterns are automatically learned from web usage data and content data. These navigational patterns are then used to generate recommendations based on a user's current status. The items in a recommendation list are ranked according to their importance, which is in turn computed based on web structure information. The presented system overcomes limited information problem, incorrect information problem and persistence problem observed in existing recommendation systems. There was an increase in recommendation accuracy and up to date recommendations were obtained.

Kim et al presented an approach that combines content-based filtering and collaborative filtering [7] to utilize the strengths of both approaches for achieving good performance. The recommender system used group rating matrix obtained from user profiles for generating web page recommendations. The presented system could overcome the disadvantages of individual content and collaborative filtering approaches and makes better use of strengths offered by each approach.

Taowei Wang et al presented an approach [8] for generating recommendations based on collaborative filtering and web usage mining. To enhance recommending quality, the recommender

system made use of Uniform Resource Locator (URL) related analysis and K-means algorithm. The presented system could overcome the sparsity problem of collaborative systems.

Reza Samizadeh et al presented an approach [9] to overcome the drawbacks of user-based collaborative filtering such as sparsity, scalability, new item problem etc., by combining of semantic knowledge with web usage mining. Scalability problem was alleviated by using web usage mining in collaborative component. New item and sparsity problems were overcome by extracting and utilizing semantic information in the web pages. The semantic patterns were created by integrating domain knowledge in the form of ontology with navigation patterns. The presented system provided good recommendation accuracy.

C. Ramesh, et al presented a recommender system [10] that integrates semantic information at different stages of web usage mining process. The main advantage of presented approach is incorporating semantic information into web usage mining process which could provide more interesting patterns that consequently makes the recommendation system more functional, smarter and comprehensive. While the limitation is, web pages with semantic information cannot be validated as no standard parsers are available for validation of such web pages.

Soheila Abrishami et al presented a web page recommender system [11] based on semantic web usage mining. The presented system integrated sematic information at different stages of web usage mining. Frequent sequential patterns were generated using prefix span algorithm. Sequential association rules were generated by using Rule gen algorithm. The presented system generated accurate, more meaningful recommendations and showed high precision and coverage compared to other systems.

Mehdi Hosseinzadeh Aghdam [12] presented a context aware recommender system based on hierarchical hidden Markov model. Context aware recommender systems consider contextual information that affects user information and states. Many model-based recommender systems, such as matrix factorization or neighborhood-based methods, do not consider changes in user's interests to recommend items. Context-aware recommender systems take into account changes in user preferences by modeling them in time. Hierarchical hidden markov model identifies the changes in user’s preferences over time by modeling the latent context of the users. Using the user-selected items, the proposed method models the user as a hidden Markov process and considers the current context of the user as a hidden variable. The latent contexts are automatically learned for each user utilizing hidden Markov model on the data collected from the user’s feedback sequences. The incentive to use the latent context is privacy, data access, and usability considerations. The proposed model has a better performance ad showed diversity in recommendations compared to existing methods.

III. MOTIVATION

As discussed above, there are different types of recommender systems which generate relevant webpage recommendations to the user. In content based recommender systems, recommendations are provided solely based on content of the web pages, but these systems predict future interests of users. In

collaborative filtering, products are recommended based on user- user or item- item similarity, but these systems do suffer from scalability and sparsity problems.

Table 1- Authors and their approaches for web page recommendation

Sr. No	Author Name and Title	Approach	Algorithms and Techniques	Advantages	Disadvantages
1	Robin van Meteren and Maarten van Someren Using content Based Filtering for Recommendation.	Content Based Filtering	tf-idf	User independence, Transparency, No cold start	Unable to predict future interests of user
2	Badrul Sarwar, George Karypis, Joseph Konstan, and John Riedl Item – Based Collaborative Filtering Recommendation.	Collaborative Filtering	K nearest neighbor approach,	No cold start and sparsity problem	Expensive model building
3	Sule Gunduz and M. Tamer Ozsu Web page prediction model based on click- stream tree Representation of User behavior.	Web Usage Mining	Click stream tree, clustering	No cold start and sparsity problem.	Unable to generate meaningful recommendations to user
4	Jia Li and Osmar R. Zaian Combining Usage, Content, and Structure Data to Improve Web Page Recommendation.	Content Based Filtering, Web usage Mining	Missions, HITS algorithm	Overcomes persistence problem, incorrect information problem	Unable to generate meaningful recommendations to user
5	Kim, Byeong Man, Qing Li, Chang Seok Park, Si Gwan Kim, and Ju Yeon Kim Combining content-based and collaborative filters in web page Recommendation.	Content based filtering, Collaborative filtering	Group rating matrix, Adjacent cosine algorithm	Overcomes the disadvantages of content based filtering and collaborative filtering	Scalability problem
6	Taowei Wang and Yibo Ren Web Page Recommendation Based on Web Usage Mining Using Collaborative Filtering Technique.	Collaborative Filtering, Web Usage Mining	K means clustering algorithm	Good recommendation accuracy	Scalability and sparsity problems
7	Reza Samizadeh and Babak Ghelichkhani Improving Web Page Recommendations using web usage Mining and Semantic information.	Semantic Web usage Mining	BOW	Good recommendation accuracy	No efficient standard parameters are defined for evaluating semantic similarities.
8	C.Ramesh, Dr. K. V. Chalapati Rao, and Dr. A. Goverdhan A Semantically Enriched Web Usage Based Recommendation Model	Semantic Web usage Mining	Clustering and association rule mining	Good recommendation accuracy	No standard parsers available for validation of web pages
9	Soheila Abrishami, Mahmoud Naghibzadeh, and Mehrdad Jalali Web Page Recommender System Based on Semantic Web Usage Mining	Semantic Web usage Mining	Prefix span, Rule gen	Good recommendation accuracy	No standard parsers available for validation of web pages
10	Mehdi Hosseinzadeh Aghdam Context-aware recommender systems using hierarchical hidden Markov model	Collaborative filtering, Context aware recommendations	Hierarchical Hidden Markov Model	Accurate and Diverse Recommendations	Scalability problem

In recent days, web usage mining has gained notable consideration for finding user behavioral patterns. Web usage mining is a technique for mining and analyzing the web server logs for finding

interesting usage patterns. Despite of this, one of the most important disadvantages of this approach is result is produced in the form of web page addresses so that common navigation profile does not have any semantic meaning. The patterns often fail to reason about user's underlying interests and preferences. Thus recommendations generated may not be accurate and the system would not be able to interpret and explain the recommendations. New item problem is the another disadvantage of web usage mining; it is the failure to recommend newly added pages or products to the visitors since these products or pages are not in the current common navigation profiles. This arises the need of a strategy that alleviates problems of existing recommender systems.

The combination of web usage mining and semantic web has created a new and fast emerging research area called semantic web usage mining, where semantic web is incorporated at different stages of web usage mining process. Recommender systems that are based on semantic web usage mining have capability to overcome all the drawbacks of existing recommender systems thereby providing accurate and meaningful recommendations.

Context is any information that can be used to characterize the situation of an entity. An entity is a person, place, or object that is considered relevant to the interaction between user and system. The context can be defined by a vector of features such as time and location.

In this paper a new system is proposed to generate comprehensive, diverse and accurate recommendations to user. The semantic information is incorporated at different stages of web usage mining to generate more meaningful recommendations to user. FAST algorithm is used to generate frequent sequential patterns. Context information along with sequential association rules mined from sequential patterns is used in generating recommendations.

IV. PROPOSED METHOD

The proposed system has four steps in recommending web pages to users.

A. Data preprocessing
B. Sequential Association Rule mining
C. Clustering webpages based on their semantic similarity.
D. Web page Recommendation

A. Data preprocessing

Data preprocessing is the first step in web page recommendation. Input for Pre-processing is raw web server logs [13]. Web log files are files that contain information about website visitor activity. Log files are created by web servers automatically. Each time a visitor requests any file (page, image, etc.) from the site, information on his request is appended to a current log file.

As a first step in data pre-processing, data cleaning is carried out. In data cleaning invalid log entries are removed.[14] After data cleaning, the navigation history of each session from log files extracted. For extraction of navigation history user and session identification is carried out. In the next stage ontology mapping is done, where each page access in every transaction/session is mapped to the ontology instances defined in the ontology for the website. The Web server does not register semantic information about the request in the log file; instead it registers only the address of the request. Therefore, before starting the

frequent sequence finding, mapping between ontology and the Web site address should be carried out. At the end of data pre-processing transactions consisting of ontology individuals are obtained.

B. Sequential pattern generation using FAST algorithm

After the preprocessing phase, the next step is the extraction of frequent navigation patterns. In this step sequential navigation patterns are created by using FAST algorithm. FAST algorithm quickly mines the complete set of patterns in a sequence database, reducing the effort for support counting and candidate sequence generation phases[15]. It employs new data representation of the dataset based on sparse id-lists and indexed vertical id-lists, which allows to quickly access an element and count its support without database scans. Researches have showed that FAST algorithm overcomes the limits of existing pattern mining methods like 1)the need of multiple scans of database 2)the generation of a potentially huge set of candidate sequences 3)the inefficiency of handling very long sequential patterns. The algorithm is mainly divided into two steps:

a. Item set extension
b. Sequence extension

FAST algorithm is applied on transaction containing ontology individuals to obtain semantically rich sequential patterns. In the next step association rules are mined from generated patterns. The main advantage of this algorithm is recommendation time and accuracy can be improved.

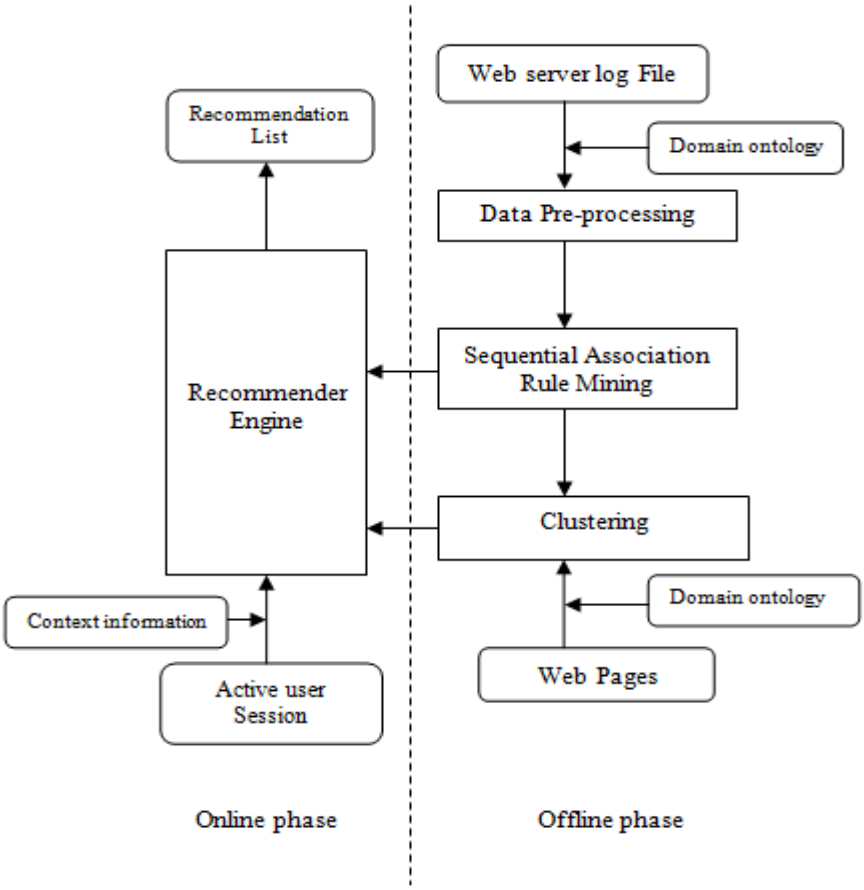


Fig. 1 General Architecture of proposed Hybrid Recommender System

C. Clustering webpages based on their semantic similarity.

In this step webpages are clustered based on semantic similarity. The distance between the two ontology concepts in calculated by concept match, which is based on upward cotopy – a semantic similarity measure [16]. By using these distance metric web pages are clustered using k- means clustering algorithm based on their semantic similarity.

D. Web page Recommendation

This is the final step in generating web page recommendations. This is an online phase in which active user session is matched with sequential association rules to generate web page recommendations to user by considering user context information as well. Each of the recommended webpage is checked to determine to which cluster they belong to. Finally the cluster with maximum number of webpages is added to the final recommendation set.

V. Conclusion

Web page recommender systems are the tools which recommend the web pages based on user interests. There are several recommender systems based on content based filtering, collaborative filtering and web usage mining. Researches have showed that incorporating semantic information at different stages of recommender system can generate meaningful and accurate recommendations. In this paper a new approach is proposed to recommend webpages, which used FAST algorithm and k-means clustering to generate recommendations. This hybrid recommender system improves the recommendation time and accuracy.

References

[1] Dietmar Jannach and Gerhard Friedrich, "Tutorial: Recommender Systems", In *Proceedings of International Joint Conference on Artificial Intelligence Beijing*, pp. 1-144, 2013.

[2] http://en.wikipedia.org/wiki/Recommender_system.

[3] Robin van Meteren and Maarten van Someren, "Using Content-Based Filtering for Recommendation", *Proceedings of the Machine Learning in the New Information Age: MLnet/ECML2000 Workshop*, pp. 26 – 28, 2000.

[4] Badrul Sarwar, George Karypis, Joseph Konstan, and John Riedl, "Item-Based Collaborative Filtering Recommendation Algorithms", *Proceedings of the 10th international conference on World Wide Web*. *ACM*, pp. 285– 295, 2001.

[5] Sule Gunduz and M. Tamer Ozsu, "A Web Page Prediction Model Based on Click Stream Tree Representation of User Behavior", *Proceedings of the ninth ACM SIGKDD international conference on Knowledge discovery and data mining*, pp. 535 – 540, 2003.

[6] Jia Li and Osmar R. Zaiane, "Combining Usage, Content, and Structure Data to improve Web Site Recommendation", *5th International Conference, EC-Web, Zaragoza, Spain*, pp. 305 – 315, 2004.

[7] Kim, Byeong Man, Qing Li, Chang Seok Park, Si Gwan Kim, and Ju Yeon Kim., "A new approach for combining content-based and collaborative filters", *Journal of Intelligent Information Systems*, Vol. 27, no. 1, pp. 79 – 91, 2006.

[8] Taowei Wang and Yibo Ren, "Research On Personalized Recommendation Based On Web Usage Mining Using Collaborative Filtering Technique", Wseas Transactions on Information Science and Applications, Vol. 6, no. 1, pp. 62 - 72, 2009.

[9] Reza Samizadeh and Babak Ghelichkhani, " Use of Semantic Similarity and Web Usage Mining to Alleviate the Drawbacks of User-Based Collaborative Filtering Recommender Systems", *International Journal of Industrial Engineering & Production Research*, Vol.21, no.3, pp. 137 – 136, 2010.

[10] C. Ramesh and Dr. K. V. Chalapati Rao and Dr. A. Goverdhan, "A Semantically Enriched Web Usage Based Recommendation Model", *International Journal of Computer Science & Information Technology*, Vol. 3, no. 5, pp. 193 – 202, 2011.

[11] Soheila Abrishami, Mahmoud Naghibzadeh, and Mehrdad Jalali, "Web Page Recommendation Based on Semantic Web Usage Mining", *Social Informatics; Springer Berlin Heidelberg*, Vol. 7710, no. 1, pp. 393 – 405, 2012.

[12] Mehdi Hosseinzadeh Aghdam, "Context-aware recommender systems using hierarchical hidden Markov model", Physica A (2018), https://doi.org/10.1016/j.physa.2018.11.037

[13] Vijayashri Losarwar and Dr. Madhuri Joshi, "Data Preprocessing in Web Usage Mining", *In Proceedings of International Conference on Artificial Intelligence and Embedded Systems*, pp. 1 - 5, 2012.

[14] Kim Shaily Langhnoja, Mehul Barot and Darshak Mehta, "Pre-Processing: Procedure on Web Log File for Web Usage Mining", *International Journal of Emerging Technology and Advanced Engineering*, vol. 2, no. 12, pp. 419 – 423, 2012.

[15] Alexander Maedche and Valentin Zacharias,"Clustering Ontology-based Metadata in the Semantic Web", *Springer, Heidelberg*, vol. 2431, no. 2, pp. 383–408, 2002.

[16] Eliana Salvemini, Fabio Fumarola, Donato Malerba, and Jiawei Han, "FAST Sequence Mining Based on Sparse Id-Lists".

MARVIN - The Intelligence Evaluator

Kumar Nityan Suman, **Aisha Begam, Abhinish Kumar, Aparna Singh**

Information Science and Engineering, PES University

Abstract- Conducting examination is a hectic process. It is an assessment intended to measure a test-taker's knowledge, skill, aptitude, and so on. Standard approach of conducting an examination is expensive, resource consuming and time taking. Major tasks involved in conducting a successful examination includes questions paper generation, answer key generation, fair conduction of test and standardized evaluation. Question paper setting takes a lot of time and requires a skilled human work. Same goes with the answer key. Manual grading of answers takes up a significant amount of valuable time, money and other resources. Air conduction of examination is another challenge that organization faces with current infrastructure. Standardized evaluation of answer sheets are another concern which will always have human bias playing a part in the current scenario. Our work aims to solve this problem by building an automated examination platform using cutting-edge machine learning, natural language processing and web technologies. We aim to provide an inexpensive alternative to the current examination system.

Index Terms- Automated Examination Platform, Answer Evaluation, Exam Assessment Analysis, Keyphrase extraction, Question Identification, Text Similarity Ranking

I. INTRODUCTION

An examination (informally, test, exam or evaluation) is an assessment intended to measure a test-taker's (candidates) knowledge, skill, aptitude, or classification in many topics. In the current modern mass-education system, the style of examination is mostly fixed, with the stress on standardized papers to be sat by large numbers of students. Both World War I and World War II demonstrated the necessity of standardized testing and the benefits associated with these tests. Tests were used to determine the mental aptitude of recruits to the military. The US Army used the Stanford–Binet Intelligence Scale to test the IQ of the soldiers. Thus it is very important that examination needs to be conducted in a fair and standardized manner for any unfair advantage or bias against the other.

The purpose of this work (MARVIN - **M**achine **A**ssessment using **R**eactive **V**iew **IN**telligence) is to build an inexpensive, accurate and efficient software platform which automatically forms questions along with their respective answers to conduct an online examination (or evaluation). Automated examination, if proven to match or exceed the reliability of human evaluators then it will reduce costs, resources and time required for conducting manual examination. Also it will be easier to understand the underlying patterns behind how students respond to a test.

II. RELATED WORK

The literature survey was done to gain insights on prior work done on question formation using classical natural language processing, answer evaluation and platform dependent challenges.

Manvi Mahana, Mishel Johns, Ashwin Apte's paper, ***"****Automated Essay Grading Using Machine Learning",* took a shot on grading essays which were categorized into 8 classes based on the context. The approach was able to achieve a *kappa score* of 0.73 across all 8 essay sets. Total number of essays taken into consideration for this experiment was ~13K from *kaggle.com* . They used 5-fold cross validation to train and test the model rigorously. Jason Zhao's attempt to score essays in the paper, *"Essay Scoring using Machine Learning",* was very different than others. Their focus for this essay grading was the style of the essay, which is an extension on the studies conducted determining the quality of scientific articles by adding maturity to the feature set (Louis and Nenkova, 2013). The dataset used is from *kaggle.com*, containing ~13K, categorized into 8 topics based on the context.
As far as our knowledge is concerned, there are various online examination platform available in present day. All of them fail to create a question on the fly and also to identify answers. They mainly maintain a collection of question - answer pairs, which are manually created and identified. No work has been done in this manner to automate the entire examination system using machine learning and natural language processing to remove human intervention and its resulting bias in the system.

III. OUR MODEL

Our platform consists of four parts: keyphrase extraction module, question formation module, answer identification module and response evaluation module. Suppose the input document (standard .txt format) is S which consists of multiple sentences = $\{s_1, s_2, s_3, \ldots, s_T\}$, T being the total number of sentences in the document. The goal here is to form questions, identify answers to the respective questions, which is then used to evaluate candidate responses. A detailed report is generated based on the number of correct responses. Two types of test can be generated: objective and subjective. Objective test will have questions with four probable answers given as option, only one being correct. On contrary, subjective tests have long type answers. No probable solution is provided in this case. User/candidate is expected to form the answer all by him/herself. The answers are evaluated on the basis of contextual similarity and not naive text similarity metric.

Keyphrase Extraction

The goal of keyphrase extraction is to get most important keyphrases, where length of keyphrase i.e., number of words allowed in a keyphrase, *1 <= number_of_words <= 3*. Keyphrase extraction module uses TF-IDF with ELMo based sentence embeddings to prune out the unnecessary keyphrases. TF-IDF, short for term frequency–inverse document frequency, is a numerical statistic that is intended to reflect how important a word is to a document in a collection or corpus. ELMo is a deep contextualized word representation that models both (1) complex characteristics of word use (e.g., syntax and semantics), and (2) how these uses vary across linguistic contexts (i.e., to model polysemy). These word vectors are learned functions of the internal states of a deep bidirectional language model (BiLM), which is pre-trained on a large text corpus. ELMo is also used in answer evaluation. Contextual vector representation

of original answer and the response is generated for similarity measure using a siamese BiLSTM model architecture.

Bidirectional Language Model

Given a sequence of N tokens, (t_1, t_2, ..., t_N), a forward language model computes the probability of the sequence by modeling the probability of token t_k given the history (t_1, ..., t_{k-1}):

$$p(t_1, t_2, \ldots, t_N) = \prod_{k=1}^{N} p(t_k \mid t_1, t_2, \ldots, t_{k-1}).$$

Figure 3.1: Forward Language Model

A backward LM (language model) is similar to a forward LM, except it runs over the sequence in reverse, predicting the previous token given the future context:

$$p(t_1, t_2, \ldots, t_N) = \prod_{k=1}^{N} p(t_k \mid t_{k+1}, t_{k+2}, \ldots, t_N).$$

Figure 3.2: Backward Language Model

A biLM combines both a forward and backward LM.

ELMo

Unlike most widely used word embeddings (Pennington et al., 2014), ELMo word representations are functions of the entire input sentence, as described in this section. They are computed on top of two-layer biLMs with character convolutions, as a linear function of the internal network states. This setup allows us to do semi-supervised learning, where the biLM is pre-trained at a large scale and easily incorporated into a wide range of existing neural NLP architectures. ELMo is a task specific combination of the intermediate layer representations in the biLM. For each token tk, a L-layer biLM computes a set of 2L + 1 representations:

$$\begin{aligned} R_k &= \{\mathbf{x}_k^{LM}, \overrightarrow{\mathbf{h}}_{k,j}^{LM}, \overleftarrow{\mathbf{h}}_{k,j}^{LM} \mid j = 1, \ldots, L\} \\ &= \{\mathbf{h}_{k,j}^{LM} \mid j = 0, \ldots, L\}, \end{aligned}$$

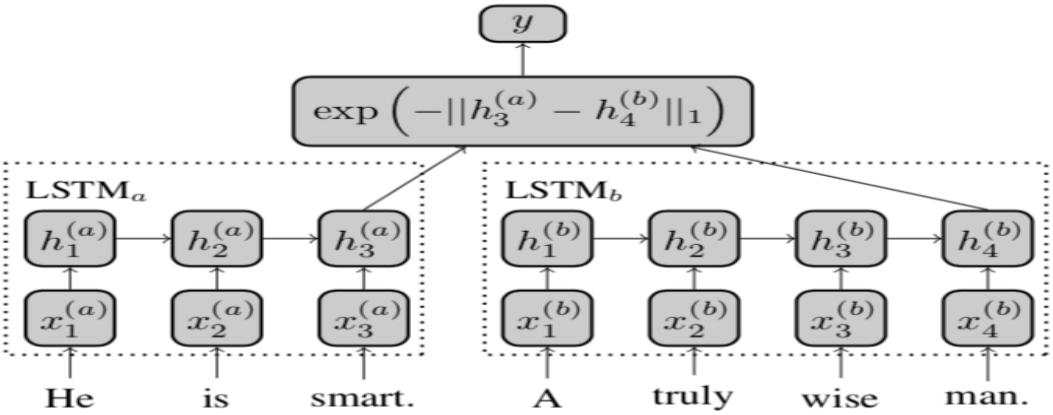
Figure 3.3:
Where, $h_{k,0}$ is layer and $h_{k,j}$ = [For. $h_{k,j}$;Bac. $h_{k,j}$], for each biLSTM layer.

The identified phrases are used by question formation module to create question based on the predefined question templates. Once questions are created, answer identification module takes both questions and phrases to identify appropriate answers. Answer module uses pre-trained BiDAF (*Bidirectional Attention Flow*) from *Microsoft Inc.*, to identify answer for the given question from a set of probables. Ideally BIDAF architecture is used in machine

comprehension but here we have used it as a question - answering model. It selects an answer from the given probables. The questions formed are used to test candidates and the received responses are evaluating w.r.t the identified answer using cosine distance or Euclidean between features vectors from Siamese LSTM neural network architecture.

Siamese LSTM

The model is outlined in the Figure below. There are two networks LSTMa and LSTMb which each process one of the sentences in a given pair, but we solely focus on Siamese architectures with tied



weights such that LSTMa = LSTMb in this work.

Figure 3.4: Siamese LSTM Model Architecture

The LSTM learns a mapping from the space of variable length sequences of d_{in}-dimensional vectors into R_{drep} (d_{in} = 300, drep = 50 in this work). More concretely, each sentence (represented as a sequence of word vectors) x1 , . . . , xT , is passed to the LSTM, which updates its hidden state at each sequence-index via equations (1)-(7). Below are the updates performed at each t ∈ {1,...,T} in an LSTM parameterized by weight matrices Wi, Wf , Wc, Wo, Ui, Uf , Uc, Uo and bias-vectors bi,bf,bc,bo:

$$h_t = \text{sigmoid}\left(Wx_t + Uh_{t-1}\right) \qquad (1)$$

$$i_t = \text{sigmoid}\left(W_i x_t + U_i h_{t-1} + b_i\right) \qquad (2)$$

$$f_t = \text{sigmoid}\left(W_f x_t + U_f h_{t-1} + b_f\right) \qquad (3)$$

$$\widetilde{c}_t = \tanh\left(W_c x_t + U_c h_{t-1} + b_c\right) \qquad (4)$$

$$c_t = i_t \odot \widetilde{c}_t + f_t \odot c_{t-1} \qquad (5)$$

$$o_t = \text{sigmoid}\left(W_o x_t + U_o h_{t-1} + b_o\right) \qquad (6)$$

$$h_t = o_t \odot \tanh(c_t) \qquad (7)$$

The two vector representation are measured on their similarity using the Euclidean distance between the two. Here we have an exponent of a negative (in this case) the output will be between 0 and 1. Based on the similarity score, answers are evaluated. Higher the score, higher the performance of the candidate. A detailed analysis report is generated based on the candidates' performance. It includes number of wrong attempts, number of correct attempts, type of mistakes made and similarity of mistakes.

IV. IMPROVEMENT AND FUTURE WORK

Future works include question generation using deep recurrent neural networks. Answer evaluation is done using a Siamese LSTM model. A more accurate result can be achieved by using bidirectional LSTM to counter the forward and backward context. One of the limitations of the proposed framework is its inability to generate computational problems and evaluate the same. Computer vision using deep convolutional neural networks comes to the rescue here and can play a vital in detecting mathematically formulas and converting them to a format more understandable to the model. Due to limited computational resources (GPUs), the models are not tuned to the best of its capacity. A good hyper-parameter tuning can help enhance the performance further.

V. CONCLUSION

In this paper, we proposed a platform framework to automate the current examination system. The platform forms questions based on a text corpus, then identifies answer to the formed questions and gives the candidate the opportunity to test his/her skills using an objective or subject type of examination. We achieved an accuracy of 88% in terms of subjective answer evaluation and a whooping 97% for objective answer evaluation, ~80% in terms of objective question formation and a good accuracy of 71% in terms of subjective question formation. The question formation is based on classical keyphrase extraction. The system developed is user-friendly and can be easily used by a naive user with little or no overhead knowledge. The system performs unbiased and consistent evaluation, which is a common problem encountered in offline examination systems. The platform can be accessed through any device as long as it supports a client browser and has internet connectivity.

REFERENCES

[1] *Automated Essay Grading Using Machine Learning* - Manvi Mahana, Mishel Johns, Ashwin Apte CS229 Machine Learning - Autumn 2012 Stanford University.

[2] Shihui Song & Jason Zhao (2012) , "*Automated Essay Scoring Using Machine Learning"*.

[3] Peters, Matthew E. and Neumann, Mark and Iyer, Mohit and Gardner, Matt and Clark, Christopher and Lee, Kenton and Zettlemoyer, Luke, "*Deep contextualized word representation"*.

[4] *Siamese Recurrent Architectures for Learning Sentence Similarity* - Jonas Mueller, Aditya Thyagarajan, AAAI Conference on Artificial Intelligence (AAAI -16).

[5] Livescu. 2016. *Charagram: Embedding words and sentences via character n-grams*. In *EMNLP*.

[6] Bennani-Smires, Musat, Hossmann, Baeriswyl, Jaggi, "*Simple Unsupervised Keyphrase Extraction using Sentence Embedding",* Swisscom AG.

[7] Quoc Le, Tomas Mikolov, and Tmikolov Google Com. 2014. *"Distributed Representations of Sentences and Documents"*, ICML.

[8] Shang, Liu, Jiang, Ren, R Voss, Han -*"Automated Phrase Mining from Massive Text Corpora"*, 2017.

[9] Sepp Hochreiter; Jürgen Schmidhuber (1997). *"Long short-term memory"*. Neural Computation 9(8): 1735-1780, 1997.

STUDY ON REVERSIBLE LOGIC CIRCUITS AND ANALYSIS

A P Sooriamala[1], Aby K Thomas[2], Reeba Korah[3]

[1]Department of ECE, Hindustan Institute of Technology and Science, Chennai,
[2] Department of ECE, Hindustan Institute of Technology and Science, Chennai,
[3]Department of ECE, Alliance College of Engineering and Design, Alliance University, Bangalore

Abstract- Low power is the main agenda in analog and digital VLSI circuits design, in today's digital world. This study focuses on low power circuits using reversible logic gates. Reducing the number of constant inputs and minimizing the garbage outputs, in existing design is the objective. New Universal Reversible Logic gates Twin SJ and SMJ have been designed. Another gate by name DINV gate is also designed. Their inputs are suitably configured so that they perform various logic functions. Using these gates and other reversible logic gates, combinational circuit and sequential circuits are developed. Reversible D Flip Flop, Reversible T Flip Flop, and Reversible 5:32 decoder and Reversible 8:1 multiplexer circuit are developed. All circuits are coded on Xilinx tool using VHDL coding, simulated and verified.

Keywords: constant inputs, garbage outputs, Reversible logic combinational circuits, Reversible logic sequential circuits

I. INTRODUCTION

Reducing power is a magic word that has gained importance with development of deep sub micron and nanotechnologies. Low power consumption, which is requirement of the modern world as more PDAs - Personal Digital Assistant and other gadgets are being used by everyone irrespective of the age or area of business or study. There are many techniques for Low Power design. The "reversible logic design" attracts more interest. In this logic reduction in energy dissipation is achieved without destroying the information bits. In 1961 Rolf Landauer [7] proved that traditional logic gates or binary logic gates dissipate a certain amount of energy during computation. This is caused by loss of each bit of information. [6] kTln2 joules is the loss for each bit of information. In this K is the Boltzmann's constant and T is the temperature of operation. Charles Henry Bennett [7] showed in 1973, that [6] in order to avoid KTln2 joules of energy dissipation in an irreversible circuit, reversible logic gates will be of use, since there is no information loss in reversible circuits. However, fan-out is not possible in reversible Logic, as one to many cannot be reversed. If we assume proper technology, a reversible logic circuit can realize the inverse operation simply by applying the gates in the reverse order. Synthesis is possible either from the outputs toward the inputs or from the inputs toward the outputs.

In this paper [2] decoder 2:4 is built using Feynman and Fredkin gates. Using this 3:8 decoder is built. [1] in this paper decoders up to 4:16 are built. This work on 4:16 DECODER using, BVF F2G and FRG Gates was an improvement on Decoder using only FRG Gates. [4] HL gate and Rgate are built to construct the decoder circuits. [8] decoders are built on xilinx.

[1] The authors gave possible research directions to optimize other combinational and sequential circuits and to improve on low power and quantum cost. [1] The authors also suggested, other modifications in this research work to reduce garbage output value and number of constant inputs/outputs and quantum cost using other reversible gates such as Modified Fredkin gate, DPG or DKG gates.

[5] In this paper power efficiency is discussed. Main disadvantage is high number of constant inputs is used. Hence it has scope on improvement. Improvement on following the basic rule for Reversible Gates, i.e number of inputs should equal the number of outputs is possible by optimizing the constant inputs and garbage outputs.

[9] [10] Sequential circuits – the authors Himanshu Thapliyal and M.B Srinivas, developed D Flip Flop using modified Fredkin Gate. RS Flip Flop, JK Flip Flop, T Flip Flop, Master slave JK Flip Flop are built. The authors had built D Flip Flop with 7 gates and number of garbage outputs are 8 which is very high.

II. REVERSIBLE GATES

All digital functions in conventional method are irreversible. At any point of time, previous computations cannot be recovered. Operations required in computation could be performed in a reversible manner, thus dissipating no heat. One of the conditions for any circuit to be reversible is that its input and output be uniquely retrievable from each other or mapped one to one. The other condition is if a device can actually run backwards then it is called physically reversible. Basic Reversible Gates are Feynman Gate 2 x 2, Double Feynman Gate 3 x 3, Toffoli Gate 3 x 3, Fredkin Gate. 3 x 3, Peres Gate 3 x 3, TSG gate 4 x 4, Sayem Gate 4 x 4, to name a few. There are 3 x 3 and 5 x 5 gates built for unique applications in this study. The same gates can be used for other applications also.

Reversible Logic has the following Rules,

- Number of outputs equal that of inputs. n x n
- Unique input to output pattern
- Garbage outputs to be minimized
- Gate count should be minimum
- Constant value for inputs [‘0’ or ‘1’] should be reduced
- Fan-out should be NIL
- Feedback / Loop back circuit is not allowed

III. PROPOSED STUDY

This study focuses on building (i) Reversible sequential circuits such as REV D Flip Flop and REV T Flip Flop and (ii)Combinational circuits such as 5:32 REV decoder using 4:16 REV decoder, 8:1 REV mux and REV 16 Bit Latch using D Flip Flop. This study involves VHDL coding for the reversible logic circuit and to analyze the power consumption and reduce garbage outputs. In this study, designing of new reversible gates, Universal gate (Twin SJ gate),D INV gate, and building different reversible logic circuits - Flip Flops, decoders, multiplexer and Latch circuits using new design gates and other basic reversible gates is carried out. These circuits are built, simulated and the results are analyzed using Xilinx 14.7

NEW REVERSIBLE GATES

(I) DINV Gate This DINV GATE is reversible gate with 3 x 3 configurations. The outputs are mostly of AND configuration. The Outputs are A'.B, A.B and A.C. This gate is used in Decoder circuits.

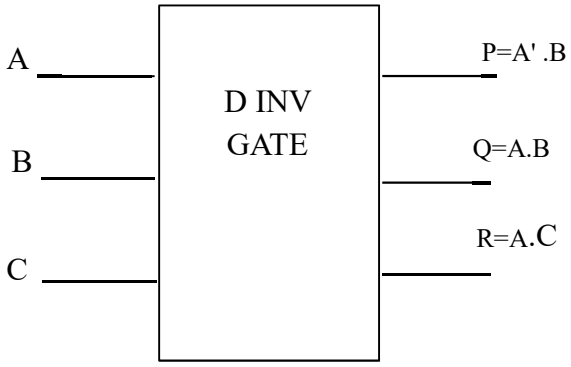


Figure 1 : DINV GATE

(II) TWIN SJ Gate This TWIN SJ GATE is reversible gate with 5 x 5 configuration with Outputs P=A'C+AD, Q=[(A+B).(C+D)]⊕E, R=B⊕E, S=A+E, T=B+E

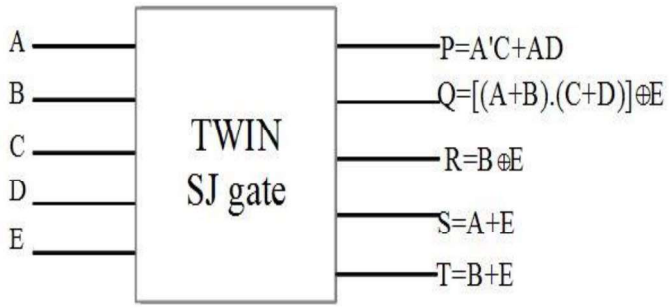


Figure 2 :TWIN SJ GATE

DINV gate and TWIN SJ gate are used in 4:16 and 5:32 DECODERS. Toffoli gate, Feynman Gate, Fredkin Gate, Peres gate and Dual AJ gate are common reversible gates used in Flip Flop, mux and Latch design.

REVERSIBLE D FLIPFLOP Flip flop is used for storage of data as it latches data. It stores the bit that is available at 'D' input. Change in D input is updated on rising edge of the next clock. I.e. a delay of one clock. Hence it is also Delay (D) flip Flop. Reversible D Flip Flop is built in VHDL coding. Also reversible gate FEYNMAN is used as a buffer or for duplicating the signals as fan out is not possible in Reversible Logic. D Flip Flop is simulated and verified in enabled and disabled conditions. Results are given in Figure 12

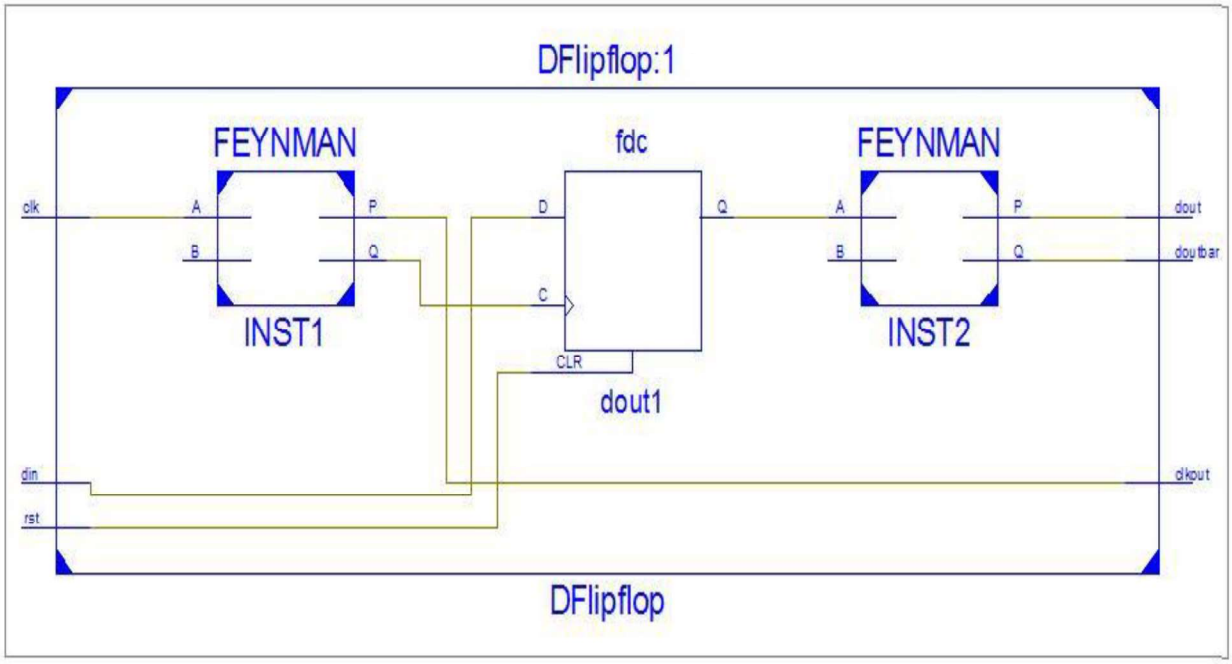


Figure 3: D FLIP FLOP RTL SCHEMATIC

REVERSIBLE T FLIPFLOP T Flip Flop changes its output on every clock i.e. it toggles. Hence the frequency of output is half of the input frequency. Output toggles on every rising edge of the clock. Reversible T Flip Flop is built in VHDL coding. Reversible gate FEYNMAN is used as a buffer or for duplicating the signals as fan out is not possible in Reversible Logic. T Flip Flop is simulated and verified in enabled and disabled condition. Results are given in Figure. 13

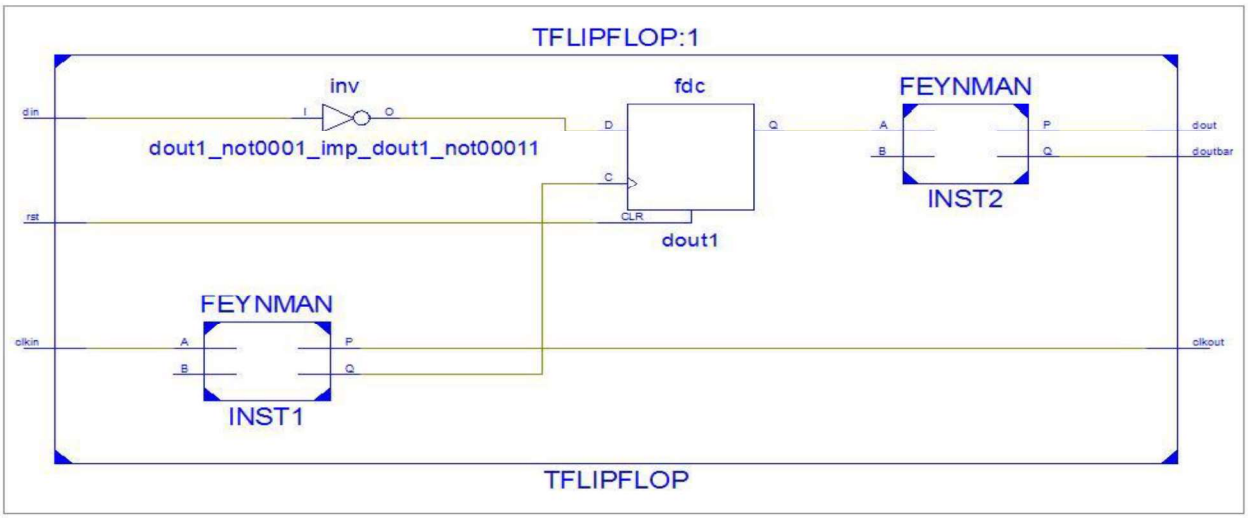


Figure 4: T FLIP FLOP RTL Schematic

REVERSIBLE 5: 32 DECODER

N:2^N Decoder is to select one out of 2^N output lines using suitable inputs. Decoders are normally used to select a particular device or memory location using suitable signal lines generally, Address lines. In this study, 5:32 Decoder is built using two 4:16 decoders and suitably enabling them with 5th Input signal (I_4).

Block diagram Fig.5 and Fig.6, show the over all view of the Decoder. RTL schematic of 5:32 decoder with exploded view of 4:16 decoder is given Figure 7. Four numbers of Feynman gates are used for buffering the Input signals I_0 to I_3. I_0 to I_3 are connected to both 4:16 decoders. Enable signal 'EN' and I_4 are inputs to DINV gate. DINV is a new gate, used to select / enable either of the 4:16 decoders based on I_4. Outputs of DINV gate are enable signals for the 4:16 decoders. If I_4 is '0', then Decoder 1, for D0-D15 is enabled. If I_4 is '1', then Decoder 2, for D16 - D31 is enabled. Circuit developed for 5:32 Decoder is simulated and verified in enabled and disabled condition. Results are given in Figure. 14 and Figure 15. When 'EN' signal is '0', then all the bits are 'zero', for different states if I_0 to I_4. This is clearly visible in the results. 4:16 decoder is built with 3:8 decoder and Feynman Gates. 3:8 decoder is built with 2:4 decoder, DINV gate and Feynman Gates.

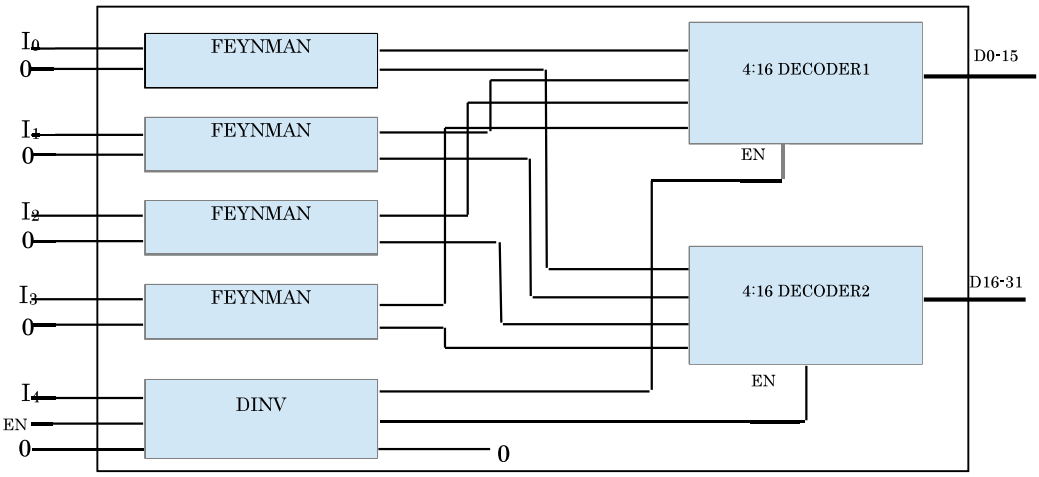


Figure 5 : 5:32 DECODER BLOCK DIAGRAM

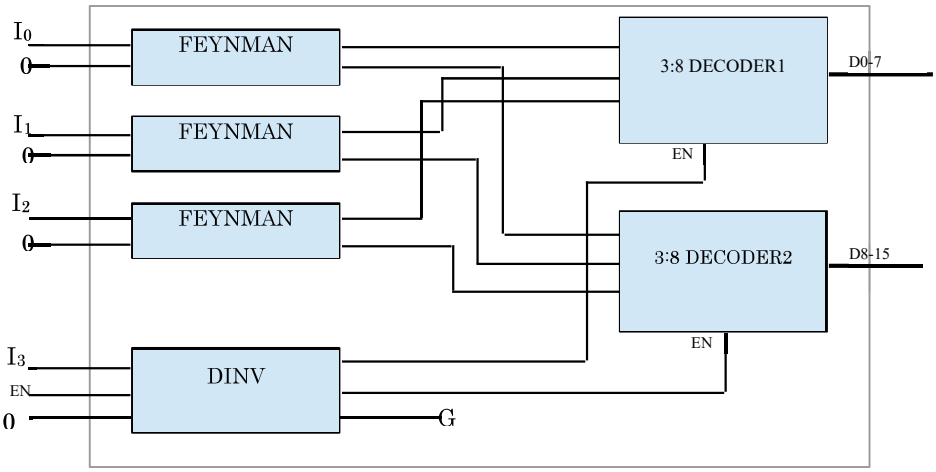


Figure 6: 4:16 DECODER

RTL Schematic of Decoder 5:32 with the exploded view of 4:16 decoders is given in Fig.7

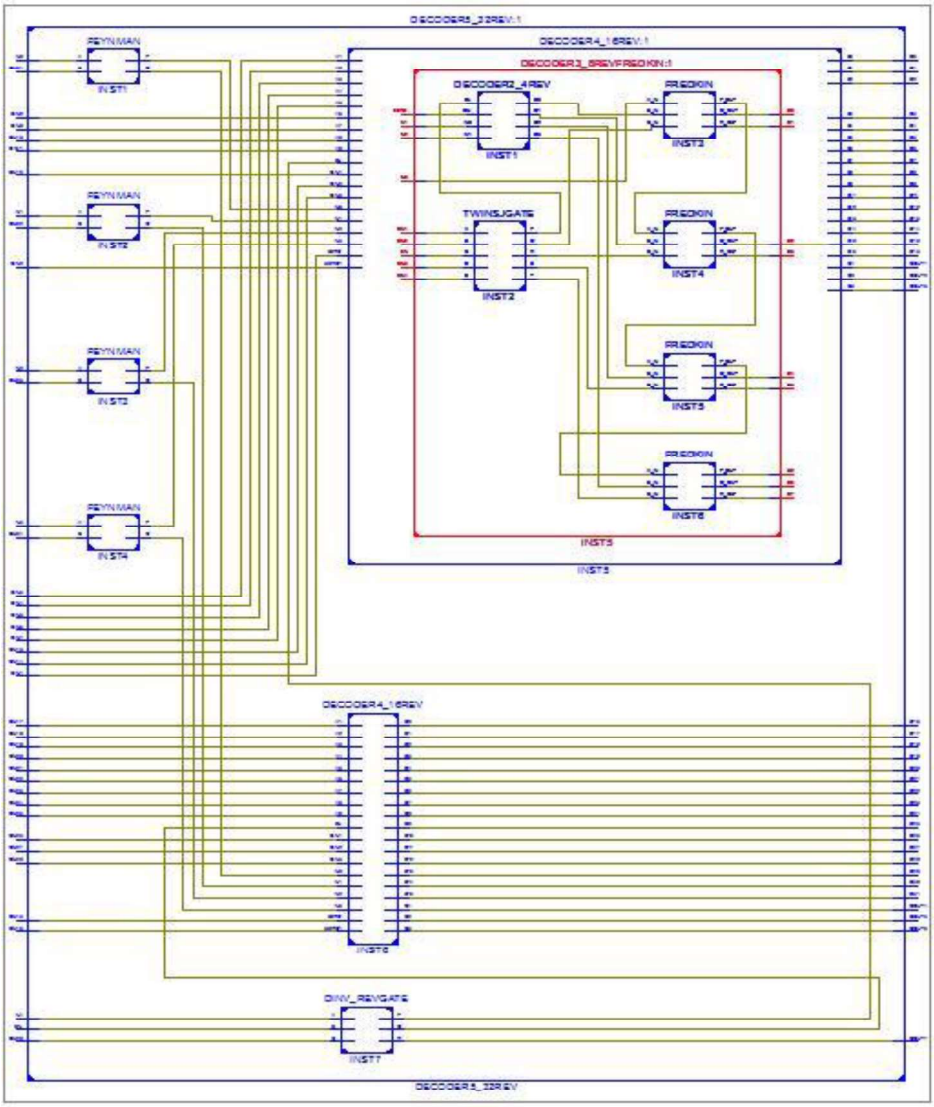


Figure 7: 5:32 DECODER RTL SCHEMATIC

REVERSIBLE MULTIPLEXER 8:1

Multiplexer is a combinational circuit used to connect one of the 2^n input lines witn 'n' select lines. In this study, one of the 8 inputs get connected to the output by three select lines. 8:1 Multiplexer has 3 select lines (SEL0 to SEL2) to choose one out of 8 inputs.

In reversible circuits, garbage outputs need to be optimized. Number of useful output is minimum in case of multiplexers. Two 4:1 multiplexers are used to build 8:1 multiplexer. 4: 1 multiplexers are built using 2:1 multiplexers and Feynman Gates. This circuit uses basic gates. No special gate is used. Reversible Feynman gate is a basic 2 x 2 gate. Outputs A and A XOR B (inputs A and B), enable to repeat or invert the signale based on input B. Giving '0' to B repeats A at both outputs.

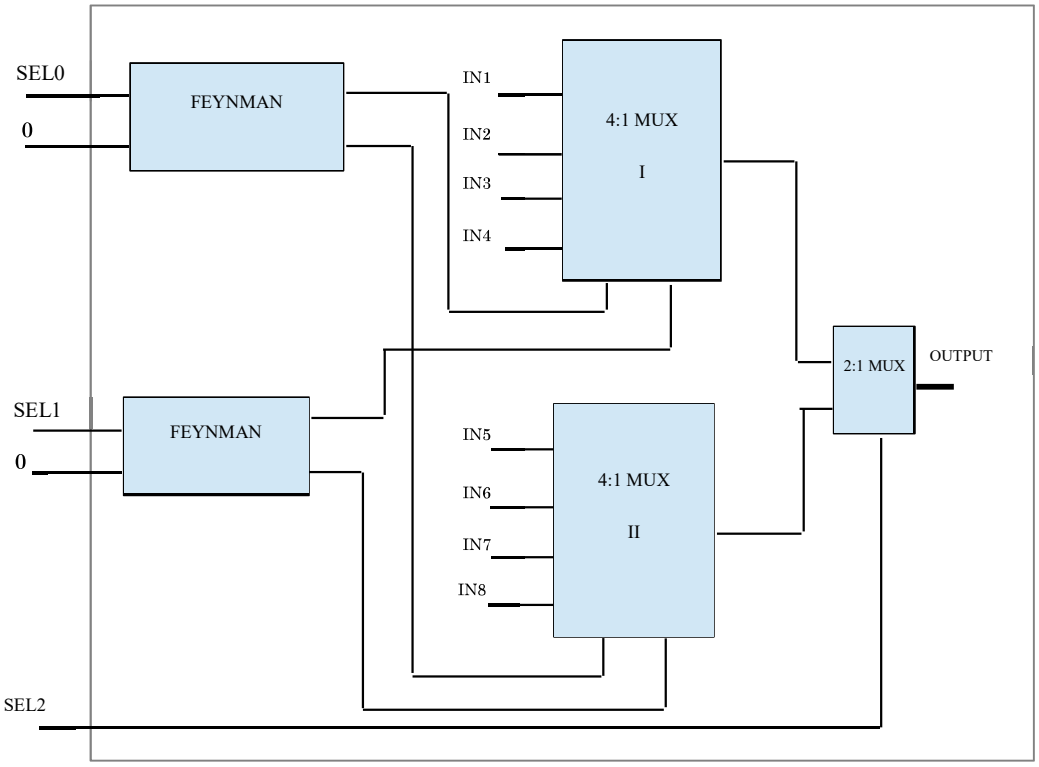


Figure 8: MULTIPLEXER 8:1 BLOCK DIAGRAM

Feynman gate is used to copy the select bits SEL0 and SEL1. Inputs v_in1 to v_in4 are connected to 4:1 mux I and Inputs v_in5 to v_in8 are connected to 4:1 mux II. Using selection bits SEL0 and SEL1, corresponding input is selected. Mux output from these 4:1 mux are connected to a 2:1 mux whose selection bit is SEL2.

For example, if the selection bits are 000, IN1 should be available at OUTPUT. IN1 is at the output of Mux I and IN5 is at the output of Mux II. These are inputs to 2:1 Mux. From these inputs, IN1 is selected by SEL2 (0) of the 2:1 Mux. As another example, if the selection bits are 101, IN6 should be available at OUTPUT. IN2 is at the output of Mux I and IN6 is at the output of Mux II. These are inputs to 2:1 Mux. From these inputs, IN6 is selected by SEL2 (1) of the 2:1 Mux. During simulation, To visualize the results, Inputs v_in1, v_in3, v_in5 and v_in7 are given '0' and v_in2, v_in4, v_in6 and v_in8 are given '1'. By giving combination of Selection bits with different timing, the output corresponding to the selection is shown at output. Ref. Fig 18 for the results of 8:1 Multiplexer

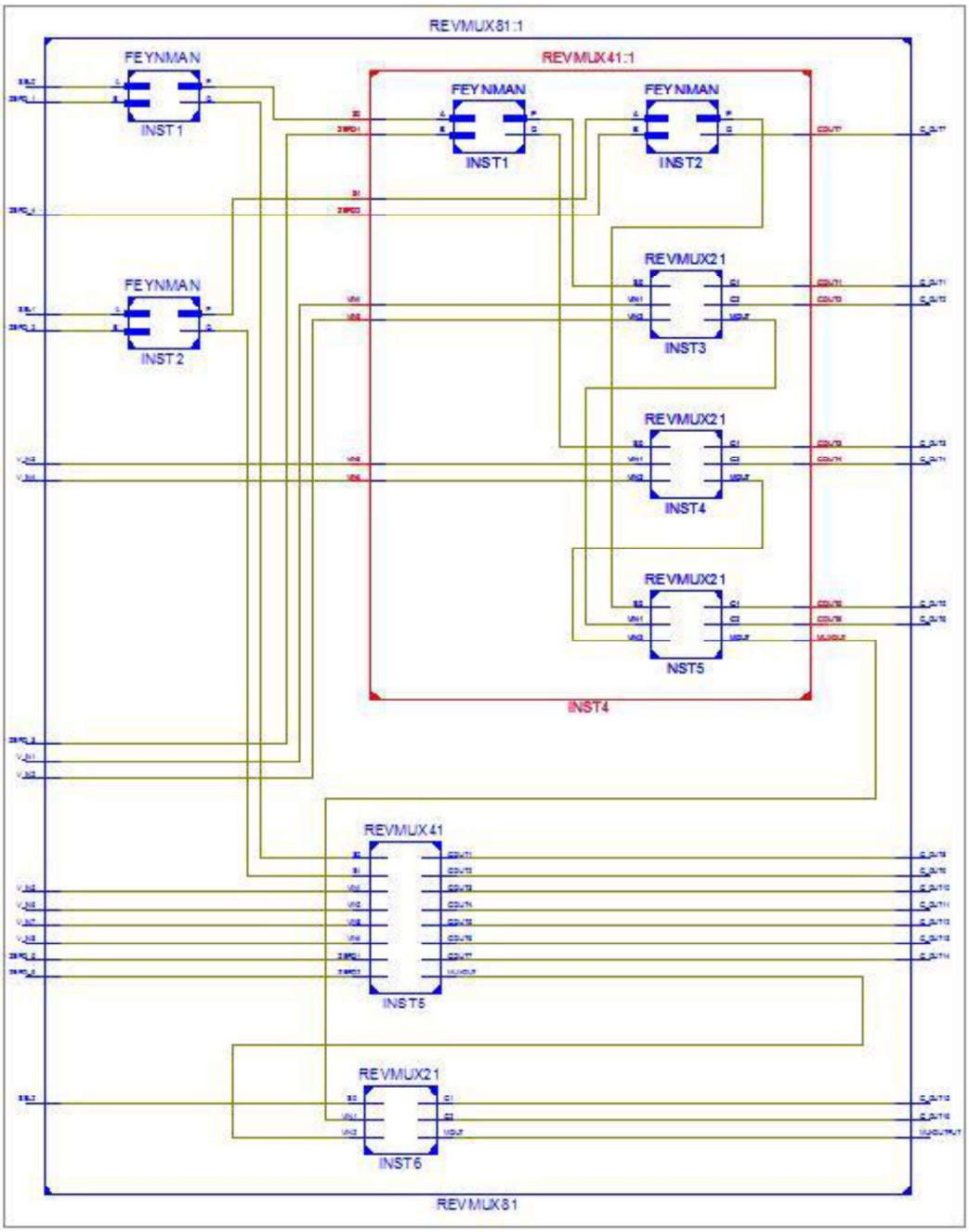


Figure 9: MULTIPLEXER 8:1 RTL SCHEMATIC

LATCH is a sequential logic circuit. 16 bit Latch circuit is used to hold data. D Flip Flop (latch) is used for storage of one bit. It gets updated on each clock. 16Bit latch has 16 D Flip Flops. Selection of Flip Flop is done using Decoder. By decoding the Address lines, particular memory bit can be accessed (Read / write). A block called 'Interface' is built using Feynman and Toffoli gates. RST of all D flip flops is connected from Interface Block. Input to Interface block, Toffoli gate is suitably given to set RST of particular D FlipFlop to '0'. Clock output (clkout) of one D flip flop is connected to the clock input of next D Flip Flop. This reduces the number of Garbage outputs. Din (D0 to D15) of D flip flops is the actual data to be stored. Data is updated at the rising edge of the input clock. Exploded view of Interface Block is given in Figure 10.

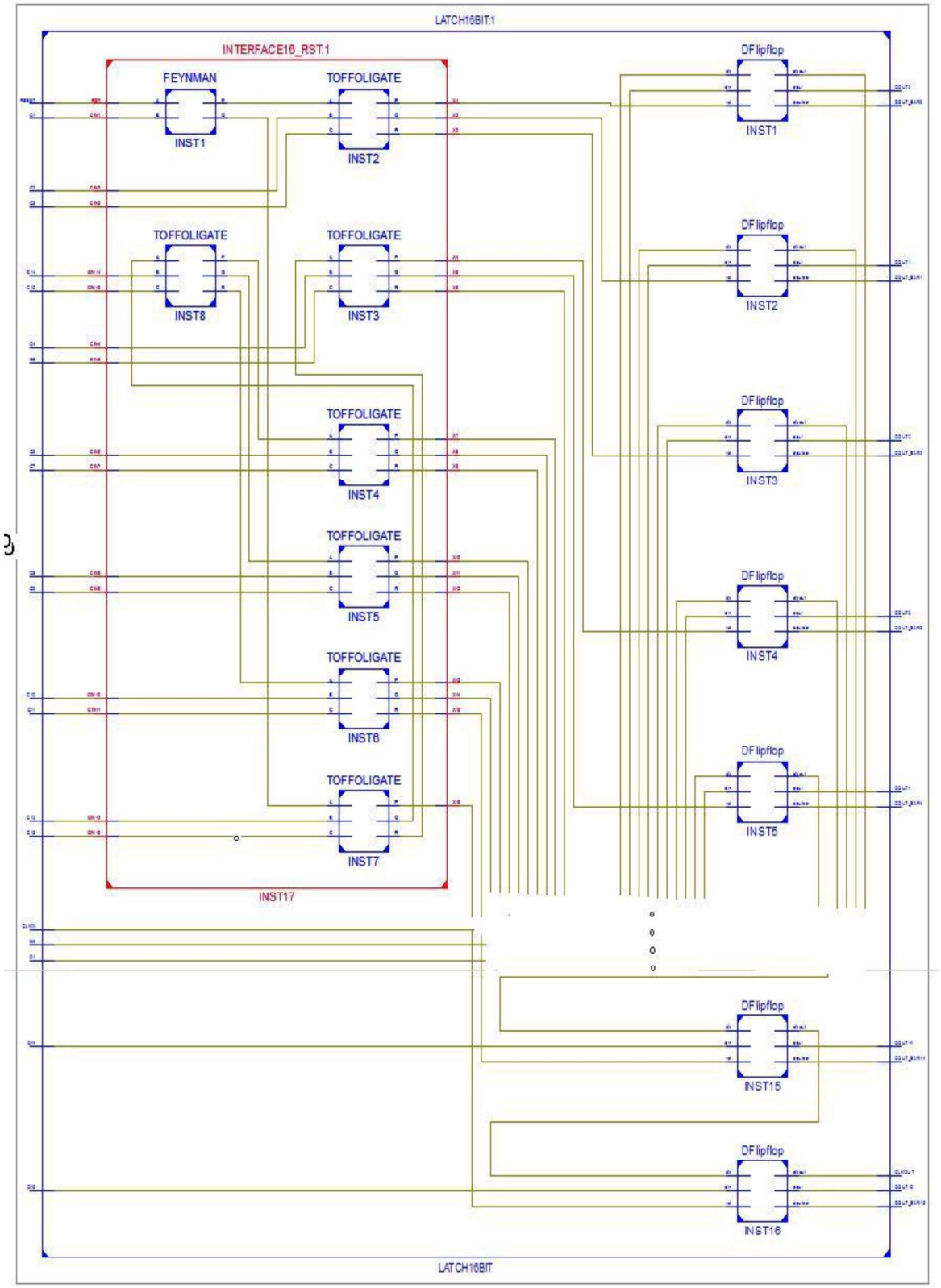


Figure 10: 16 BIT LATCH RTL SCHEMATIC (exploded view)

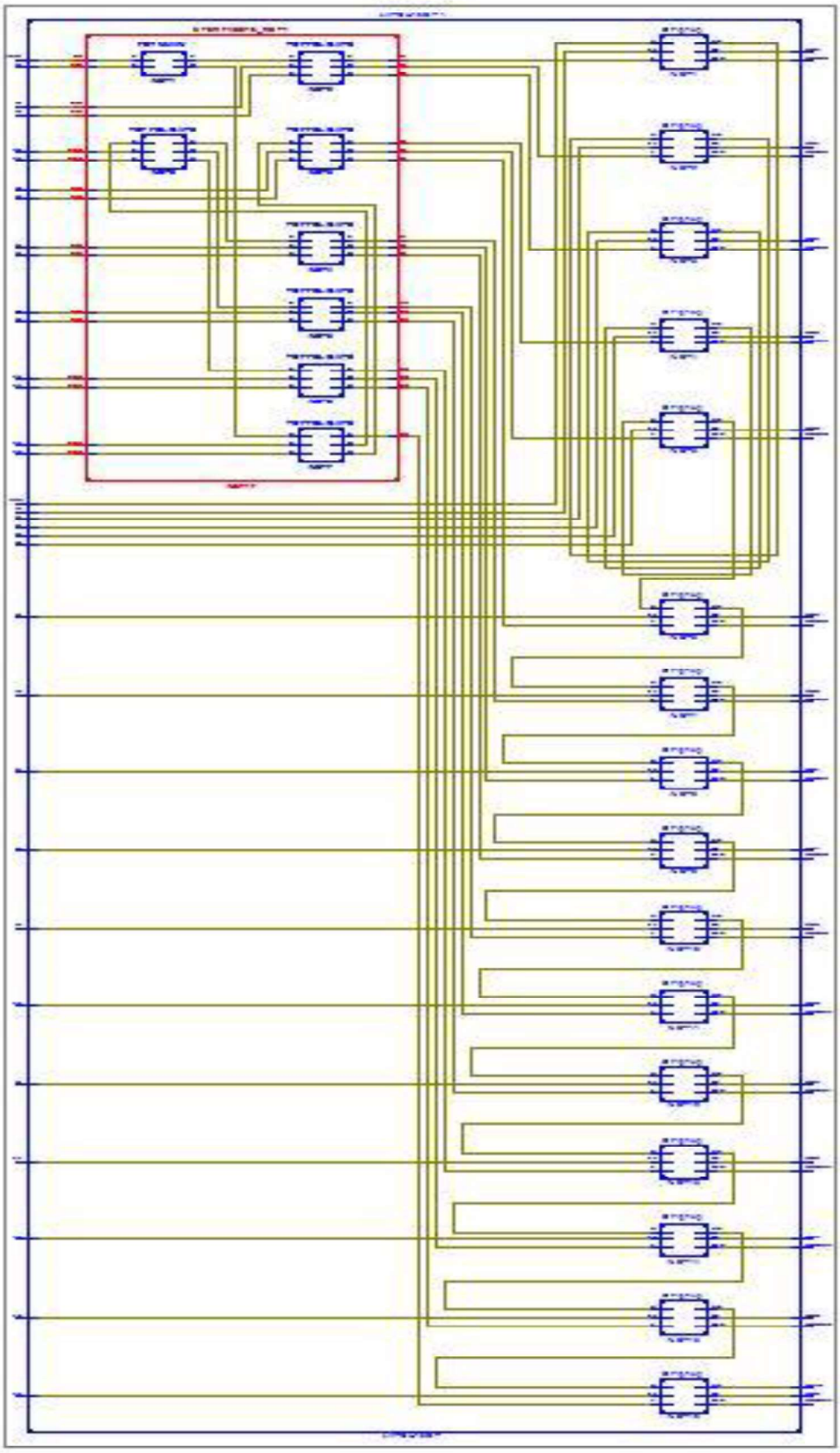


Figure 11: 16 BIT LATCH RTL SCHEMATIC

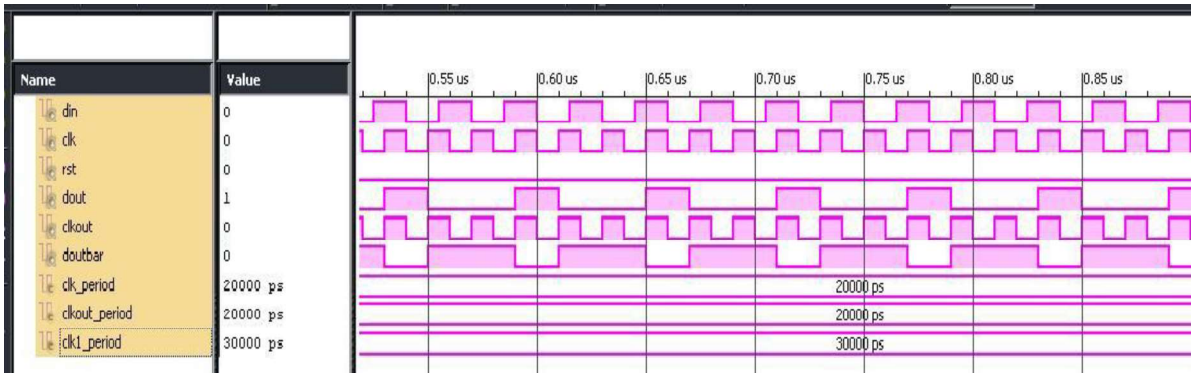
IV RESULT & CONCLUSION

Software Tool: In this study all the gates and circuits are built on Xilinx 14.7 tool. Xilinx ISE is a design environment software tool for synthesis and analysis of HDL designs. This enable us to synthesize (compile) the design and to

- perform timing analysis
- examine RTL diagrams
- analyse the responses for different stimuli

This tool has a limitation of not to be used with FPGA products of other manufacturers. Xilins ISE (circuit synthesis and design) + ISIM or modelsim logic simulator (used for system level testing).

D FLIPFLOP Initially reset (rst) is made to '1' and after a delay it is made to '0'. A rising pulse is given to reset, to clear. 20ns pulse is given as clock and 30ns pulse is given as data. The state change of data (din) is updated at the output (dout) during next rising edge of the clock (clk). In the given figure 12., at 0.60us, 'din' goes low and the 'clk' is low. Hence there is no change in 'dout'. At the next rising edge of the



clock(0.62us), 'dout' goes low.

Figure 12: D FLIP FLOP RESULT

T FLIPFLOP Initially reset (rst) is made to '1' and after a delay it is made to '0'. A rising pulse is given to reset, to clear. 10ns pulse is given as clock. In the Figure 13., the output 'dout' changes state (toggles) when there is state change of data (din) during next rising edge of the clock (clk). In the given at 0.45us, 'din' goes low and the 'clk' is low. The 'dout' continues to be low. There is no change in 'dout'. At the next rising edge of the clock, at 0.455us, 'dout' goes High (toggled). At 0.6us 'din' goes high and the 'clk' is low. The 'dout' continues to be high. There is no change in 'dout'. At the next rising edge of the clock, at 0.655us 'dout' goes low (toggled).

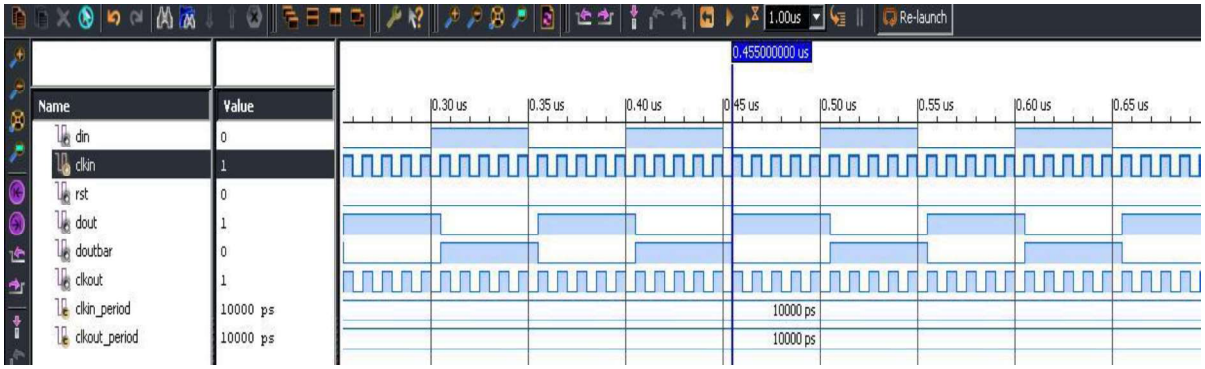


Figure 13: T FLIP FLOP RESULT

DECODER5:32 Results are given in Figure. 14 to Figure 17. When 'EN' signal is '0', then all the outbit bits d0 to d31 remain 'zero', for different states of I_0 to I_4. In Figure.14, at 0.81us in0 and in1 are '1' and in2, in3 and in4 are '0', '00011' corresponding bit 'd3' is selected. Similarly in Figure.15, at 2.39us, in3, in1 and in0 are '1', in2 and in4 are '0',01011 is the input and corresponding bit 'd11' is selected. In Figure.16, at 3.99us, in0, in1 and in4 are '1' and in2 and in3 are '0', '10011' corresponding bit 'd19' is selected. Similarly in Figure.17, at 5.6us, in4, in3, in1 and in0 are '1', in2 is '0', 11011 is the input and corresponding bit 'd27' is selected. It can be noted that when 'EN' is low, though the input signals were active, no output is available.

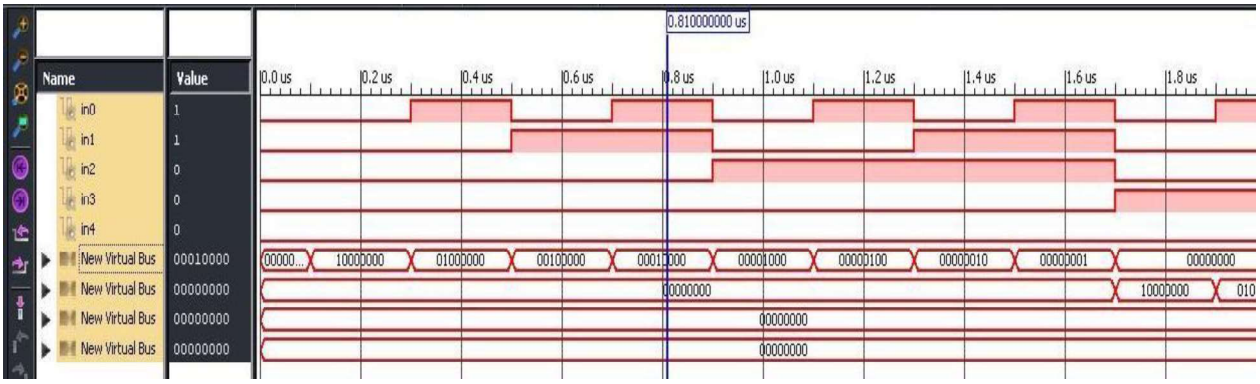


Figure 14: 5:32 DECODER RESULT (A)

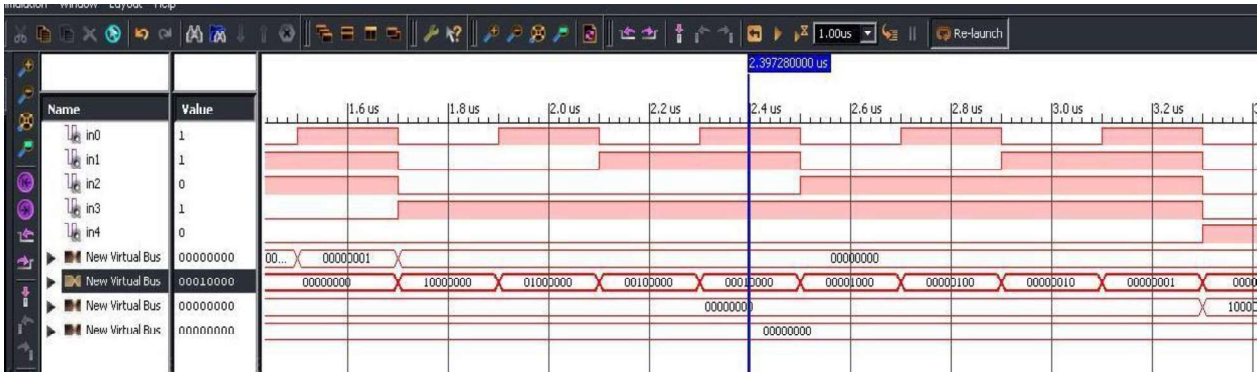


Figure 15: 5:32 DECODER RESULT (B)

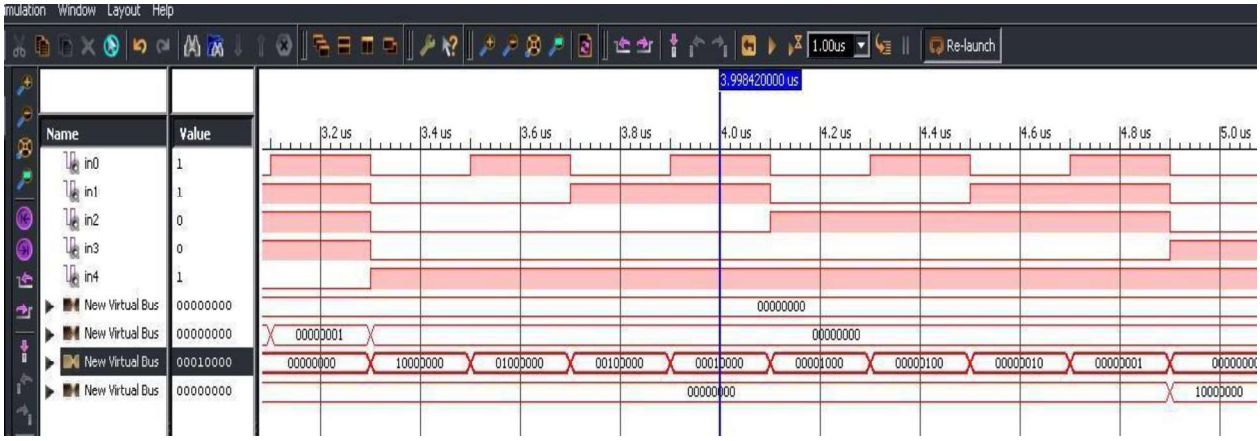


Figure 16: 5:32 DECODER RESULT (C)

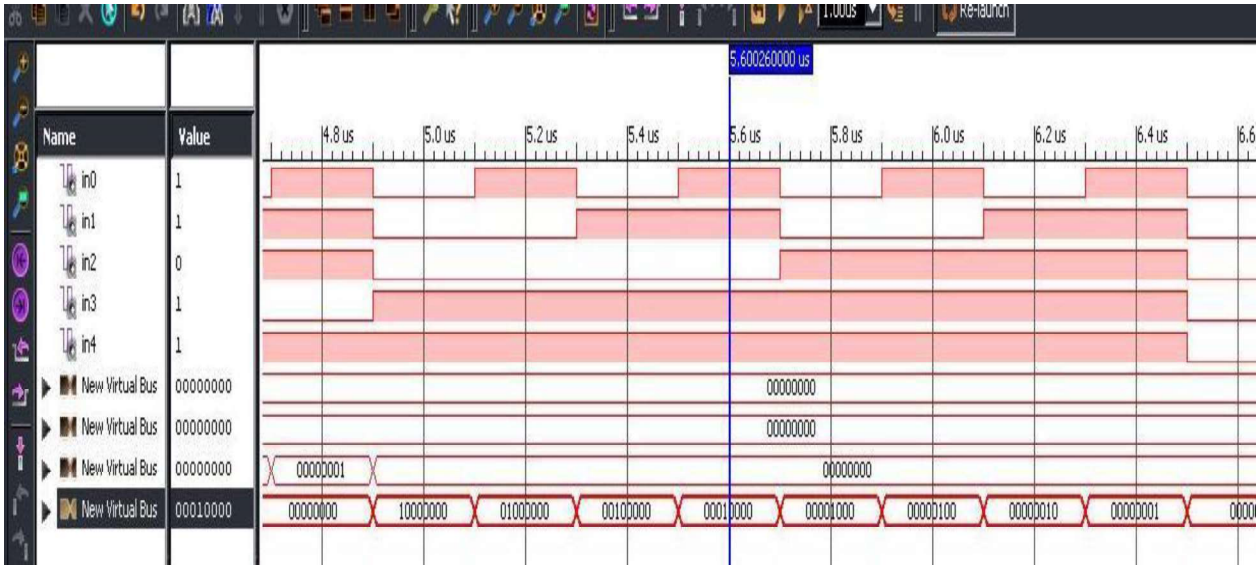


Figure 17: 5:32 DECODER RESULT (D)

8:1 MUX Feynman gate is used to copy the select bits SEL0 and SEL1, by giving B input '0'. Inputs v_in1 to v_in4 are connected to 4:1 mux1 and Inputs v_in5 to v_in8 are connected to 4:1 mux2. Using selection bits SEL0 and SEL1, corresponding input is selected. Mux output from these 4:1 mux are connected to a 2:1 mux whose selection bit is SEL2.. To visualise the results, Inputs v_in1, v_in3, v_in5 and v_in7 are given '0' and v_in2, v_in4, v_in6 and v_in8 are given '1'. By giving combination of Selection bits with different timing, the output of mux is proved. In the figure16., at 500ns, sel0 is '0', sel1 is '1' and sel2 '0' corresponding to v_in3 (low), connected to muxoutput. Muxoutput has be linked to the timing of select bit timings, to verify the results.

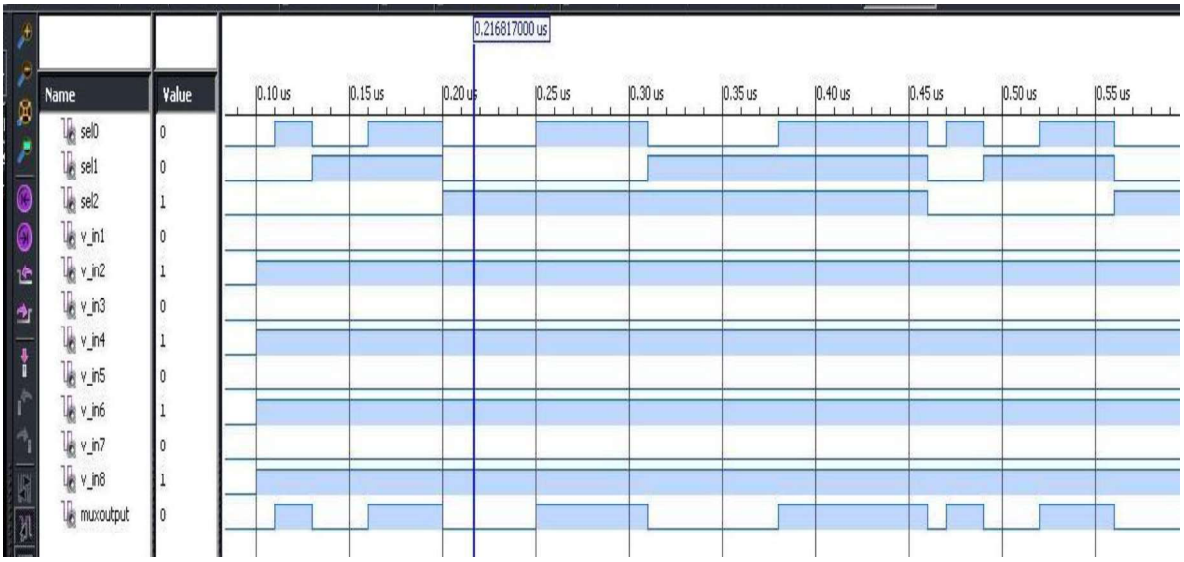


Figure 18: 8:1 MULTIPLEXER RESULT

16 BIT LATCH A short pulse is given as reset. 20ns clock is there at the input (clkin). Inputs D0 to D15 are given with '1' (11111111 11111111) at 0.280us. It is the falling edge of the input clock (clkin) and the DOUT becomes 11111111 11111111 on the risining edge of the input clock, at 0.290us. Similarly different16bit inputs 01010101 01010101, 10101010 10101010, 1111111111 00000000 and 00000000 1111111111 are given at datain and the DOUT is checked. It is observed that the DOUT matches the datain and results are verified.

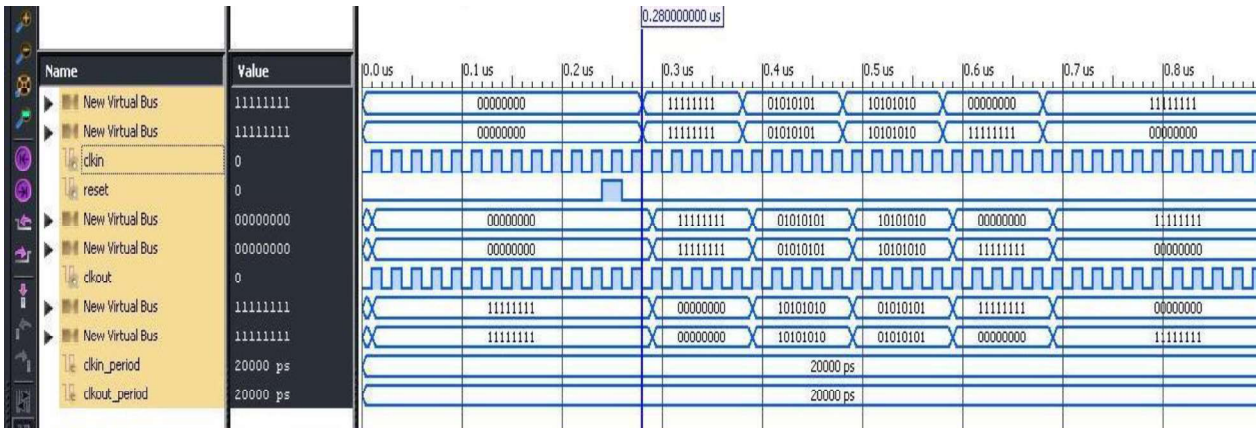


Figure 19: 16 BIT LATCH RESULT

The circuit has been simulated and functionality is verified using Xilinx. The circuit under study is designed [8] in Xilinx using VHDL code, simulated and verified. It is built on Tanner Tool. Garbage output of the designed circuits with existing design are compared. These circuits are simulated on Field Programmable Gate Arrays (FPGAs) using VHDL code and simulated on Isim. Reduced Garbage outputs are achieved. In future more complex circuits can be designed using these simple sequential and combinational circuits. Circuits like Dual port RAM, memory arrays shall be built using decoders and Latch circuits. Reversible counter, comparator circuits can be improved in garbage output and constant input. Power reduction is more visible when the circuit is built on CMOS.

CIRCUIT	GARBAGE OUTPUTS		
	PREVIOUS METHOD	EXISTING METHOD	PROPOSED METHOD
D FLIPFLOP	NOT AVAILABLE	8	4
2 to 4 decoder[3]	3	2	0
3 to 8 decoder[4]	4	3	1
4 to 16 decoder	5	4	3
5 to 32 decoder	NOT AVAILABLE	NOT AVAILABLE	7
4×1 multiplexer	17	16	7
8×1 multiplexer	NOT AVAILABLE	NOT AVAILABLE	16
16bit latch	NOT AVAILABLE	NOT AVAILABLE	16

Table 1. COMPARISON OF GARBAGE OUTPUT

V REFERENCES

[1] International Journal of Research (IJR) e-ISSN: 2348-6848, p- ISSN: 2348-795X Volume 2, Issue 11, November 2015 Performance Analysis of Low Power Decoders Using Reversible Computing Manjinder Pal Singh, Er. Birinderjit Singh, Er.Amandeep Singh Bhandari.

[2] Pandey, N., Dadhich, N., & Talha, M. Z. (2014, February). Realization of 2: 4 reversible decoder and its Applications. Signal Processing and Integrated Networks (SPIN), 2014 International Conference on (pp. 349-353). IEEE.

[3] M. Nachtigal, and N. Ranganathan, "Design and Analysis of a Novel Reversible Encoder/Decoder", 11th IEEE International Conference on Nanotechnology, Portland Marriott, pp. 1543-1546, 2011.

[4] 290| Page DESIGN OF 3:8 REVERSIBLE DECODER USING R-GATE Sweta Mann1, Rita Jain (ESHM-17) ISBN:978-93-86171-25-2 pp. 149-157, 2017.

[5] International Journal of Electronics, Electrical and Computational System IJEECS ISSN 2348-117X Volume 7, Issue 4 April 2018, Design and Synthesis of Combinational Circuits using Reversible Logic G.Sreekanth1, C. Venkata Sudhakar2 VLSI,Sree Vidyanikethan Engineering College, Chittoor, India

[6] C.H.Bennet, "Logical Reversibility of computation", IBM Journal of Research and Development, pp 525-532, 1973.

[7] R. Landauer, "Irreversibility and heat generation in the computing process". IBM Journal of Research and Development, vol. 5, pp. 183-191, 1961.

[8] "Design and synthesis of combinational circuits using reversible decoder in Xilinx" Gopi Chand Naguboina ; K. Anusudha , 2017 IEEE International Conference on Computer, Communication and Signal Processing (ICCCSP) 2017 DOI: 10.1109/ICCCSP.2017.7944062

[9] ”A Beginning in the Reversible Logic Synthesis of Sequential Circuits” Himanshu Thapliyal and M.B Srinivas, Mark Zwolinski Electronic System Design Group Electronics & Computer Science University of Southampton, U.K (mz@ecs.soton.ac.uk) Thapliyal MAPLD 2005/10121

[10] Thapliyal, H., & Ranganathan, N. (2010). Design of reversible sequential circuits optimizing quantum cost, delay, and garbage outputs. ACM Journal on Emerging Technologies in Computing Systems, 6(4), 1–31. doi:10.1145/1877745.1877748

Objected Oriented Design Defect Analysis and Refactoring-Overview.

Nagaraj MS[1] Dr. R.Selvarani[2]
[1]Honeywell Technology Solutions, Bangalore, India
[2]Director –Doctoral Program – ACED, Alliance University, India

Abstract- Software design defects often lead to bugs, runtime errors and software maintenance difficulties. They should be systematically prevented, found, removed or fixed all along the software lifecycle (Design, development and maintenance stages). However, detecting and fixing these defects is still to the greater extent a difficult, time-consuming and manual process. Identifying and fixing the defects at earlier part of software life cycle will reduce the significant maintenance cost. In this paper, we propose detecting the design defects at design phase and software defects at implementation phase of the software life cycle. Detecting defects in early stage of design cycle is useful from the perspective of cost quality and schedule reduction.

Index Terms- Software design defects, software bugs, run time errors, object oriented defects, anti-patterns, code smells

I. INTRODUCTION

Object-oriented programming (OOP) is industry adapted a programming paradigm which consists of “objects" that have data fields (attributes that describe the object) and associated procedures known as methods. Objects, are instances of classes, are used to interact with one another to design applications and computer programs. Object oriented design, today, is becoming more popular in software development environment and Object Oriented design metrics is an essential part of software environment. The main objective of analyzing these metrics is to improve the quality of the software. Design defects, anti-patterns, code smells defects at the architectural level and software coding errors must be detected and corrected to improve software quality, automatic detection and correction of these software architectural defects, which suffer of a lack of tools. The contribution of this paper is to present issues related to the detection and correction of design defects at design level and software coding errors at implementation phase.

II. DIFFERENCES BETWEEN DESIGN PATTERNS AND DESIGN FLAWS.

Design patterns is a reusable solution to commonly occurring design problems in software. Each class in design pattern plays specific role and interactions between the classes are well defined. Anti-pattern represents undesirable design structure that is difficult to maintain and understand the software. The complexity of the classes that constitute design patterns are comparatively less than compared to classes in anti-patterns. Roles and complexity and interaction with other classes can be used to identify the design pattern because their important features can be easily defined and recognized, which is not possible in case of some anti-patterns [1] as the classes that constitute the anti-pattern higher in complexity and consists of multiple roles. Anti-patterns have very large variety of characteristics (e.g., number of methods, naming of methods/classes, method parameters, class functionality etc.), therefore it is harder to apply general detection rules to all of them.

III. DESIGN-DEFECTS OR ANTI-PATTERNS

Design defects are design structures that are complex, difficult to understand and maintain. They are bad practices in software design. Design solution that is initially appears to be a good solution for the problem to solve results in the creation of conflicts because of its implementation. Having knowledge of design defects, the developer is equipping with the knowledge needed to avoid or fix errors before writing any code or designing the software. A design defect or anti-pattern is a literary form that describes a commonly occurring solution to a design problem, solution which generates negative consequences in maintaining the software. Design defects are bad solutions to recurring design problems. The idea of design defect is to show what not to do. The Blob, the Spaghetti, the Poltergeist, the Lava Flow are among well-known design defects. For example, the Blob represents single complex controller class that monopolizes the processing and is surrounded by simple data classes. The Spaghetti code, which is one of the most famous design defect, describes a program or system with a software structure that lacks clarity and hard to maintain.

IV. RELATION BETWEEN OBJECT ORIENTED MATRICES AND DESIGN DEFECT

Design defects can be identified by measuring the objected oriented metrics coupling, cohesion, complexity and inheritance. Below table lists the relationship between the different object oriented metric categories and most commonly occurring design defects [6].

Anti-patterns	Metrics Category			
	Coupling	Cohesion	Complexity	Inheritance
Blob	High	low	High	Low
Lava Flow	Low		High	
Functional Decomposition		high	Low	Very Low
Poltergeists	High		Low	
Swiss Army Knife	High		High	

V. BLOB DESIGN DEFECT

The Blob class also known as God class of the design. The Blob class violates the “Single-responsibility principle” [14], Single-responsibility principle states there should one only one reason to change the class, The Blob class is responsible for all (or most of the) behavior of an application while the rest of the classes (the data classes) are only responsible for Encapsulating data, hence it monopolizes the processing and acts as controller class that performs majority of system responsibilities. The basic form of a god class is defined in Figure 1.

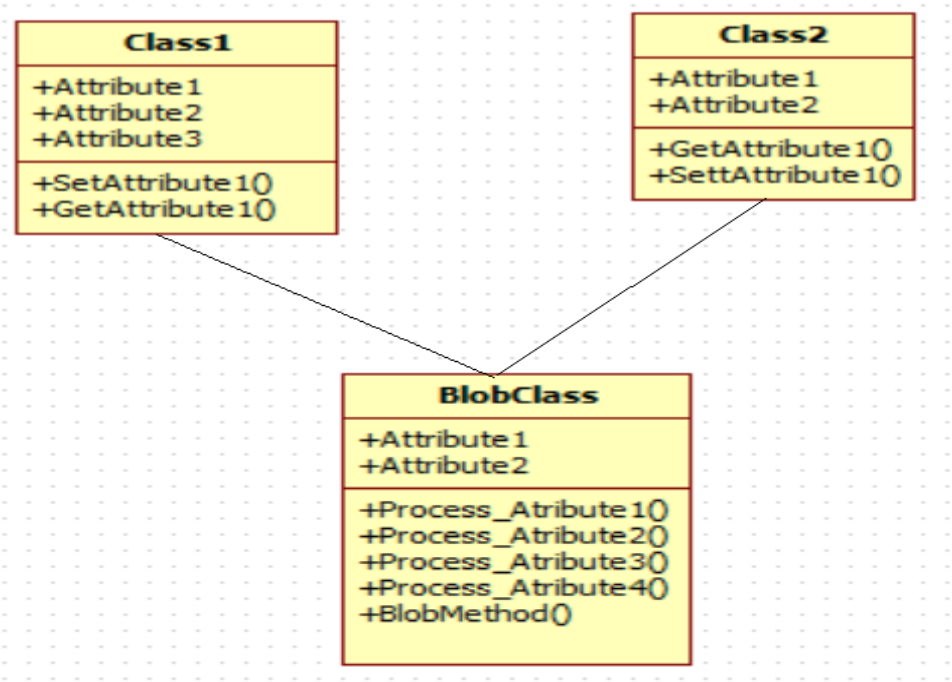


Figure 1 Blob or God Class Architecture

Blob class will consist of one or more Blob method that obtains their data from classes different from the class they belong to [2] and perform the complex computations or operations. The blob method can also be visualized as complex method that performs more than one functionality i.e. there will be more than one reason to change the method. The basic form of a god method is defined in Figure 2 by the diagram modelled with continuous lines. This shows that the god method accesses attributes from the other classes through its method which expose the attributes of the class.

Note: Arrow marks ⟶ in all the diagrams indicate the function call. Where the pointing to Called from caller function.

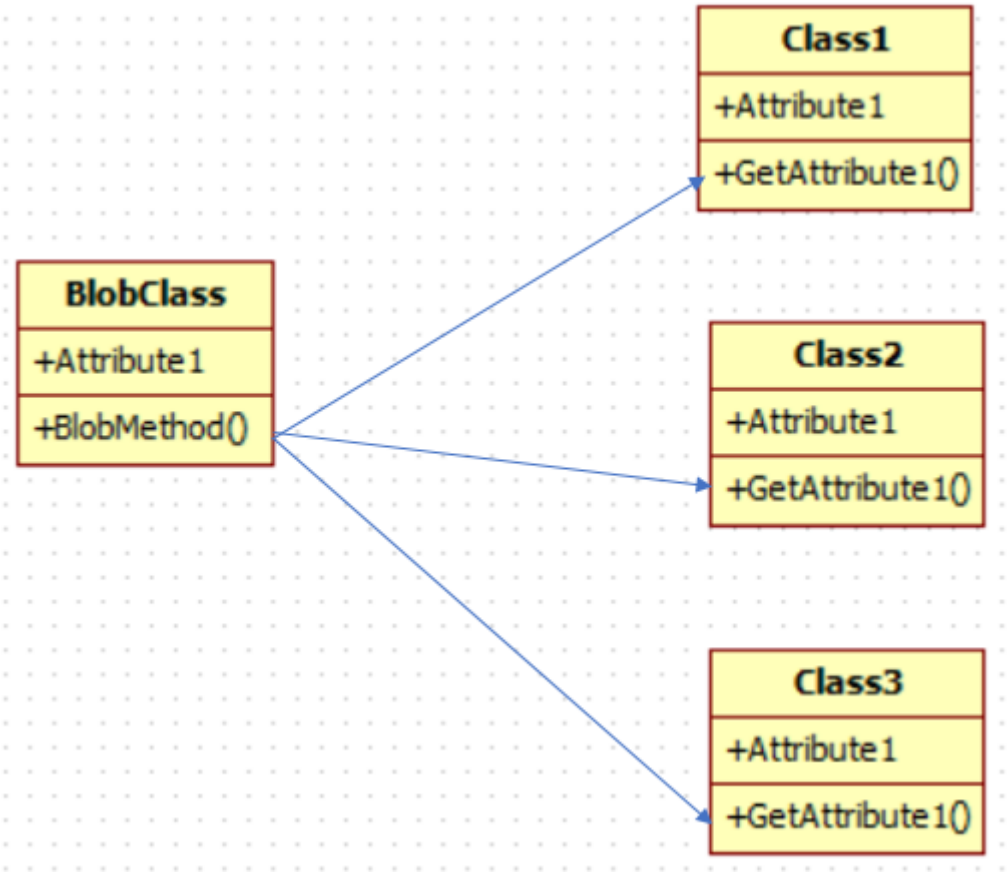


Figure 2 Blob Method Architecture

Detecting Blob pattern

As Blob method performs the complex computations, the complexity of the function is high and, they access the data from other classes resulting high coupling and low Cohesion. The Blob defect can be identified by measuring the object-oriented metrics complexity, coupling and Cohesion. Object oriented metrics are captured through software metrics and properties are expressed in terms of valid values for these metrics [11]. The most widely used metrics are the ones defined by Chidamber and Kemerer [3]. These include: Weighted methods per class, WMC, Coupling between objects, CBO.

Refactoring a Blob Class

Solution: Refactoring of Blog class exposes the operation rather than the attributes as shown in Figure 3.

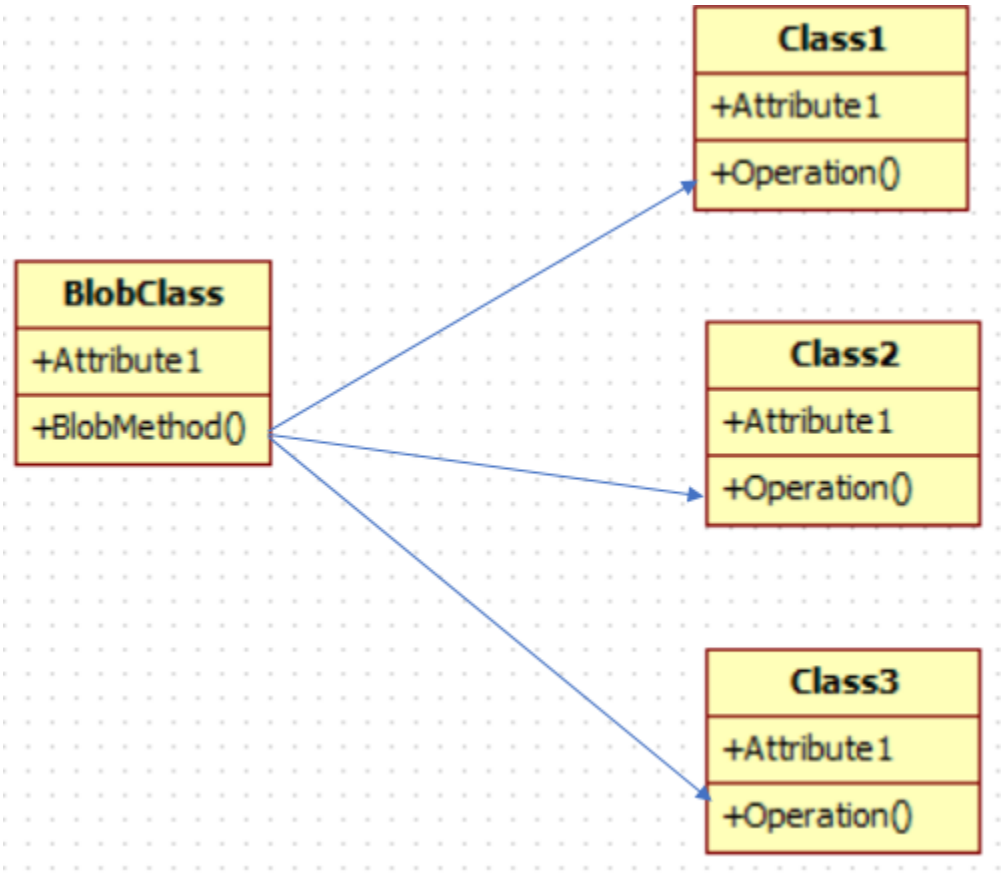


Figure 3 decomposing the responsibilities

Refactoring a Blob Method.

The Blog method consists of multiple responsibilities; this method can be decomposing until complexity reduces and becomes a function with single responsibility with different classes as shown in the Figure 3.

VI. POLTERGEISTS DESIGN DEFECT

Poltergeist is a class with minimal or limited responsibilities and roles to play in the software system; therefore, their effective life cycle is quite brief; they clutter software designs, creating unnecessary abstractions [6];

Poltergeists can occur in four different forms as follows

- Irrelevant classes:
- Agent classes
- Operation classes
- Object classes

Irrelevant classes:

An irrelevant class does not have any meaningful behavior in the software design. They composed of only of get (class data member accessor), set (sets the class data member) methods. Figure 5 shows the UML notation of design defect Irrelevant classes. The concreate class will access the irrelevant class attribute through the get method and sets attributes using set method.

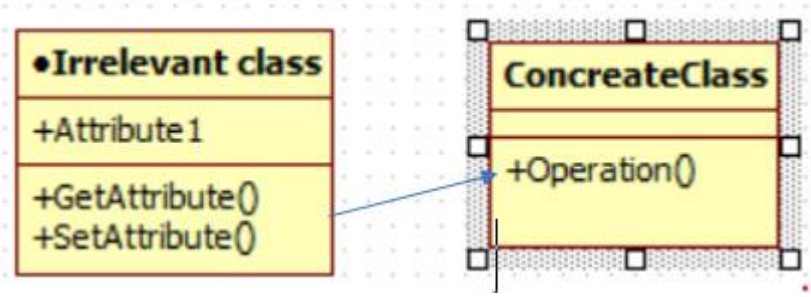


Figure 4 Irrelevant Classes

Refactoring Irrelevant Class.

Although the behavior of irrelevant classes is meaningless, the data that it may contain is not. The correction of irrelevant classes consists in both eliminating them from the design and placing the data they contain with the respective accessor class.

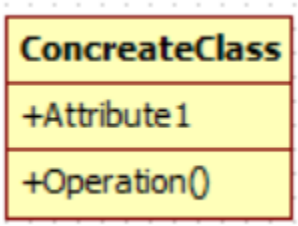


Figure 5 Refactoring irrelevant class

Agent classes:

Agent design defect are classes that are responsible for only passing the messages from one class to another, i.e., methods that offer redundant paths to access operations of other classes in the design.

The UML specification of this design defect is shown in Figure 6.

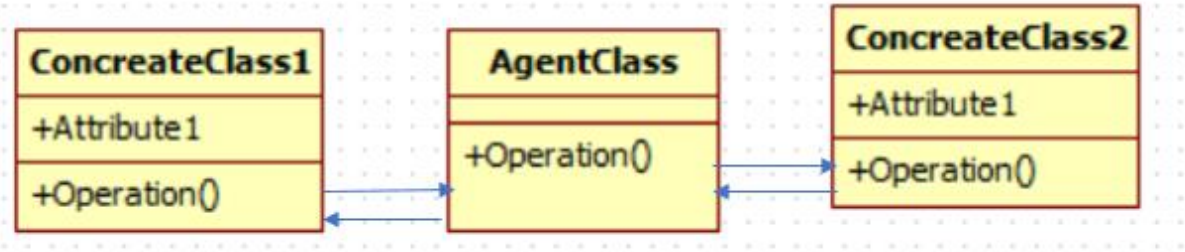


Figure 6 Agent classes

Refactoring Agent Class

Refactoring agent class involves removing the agent method from the design and replacing the communication it performs to be done directly between the other two classes involved in the anti-pattern [4].

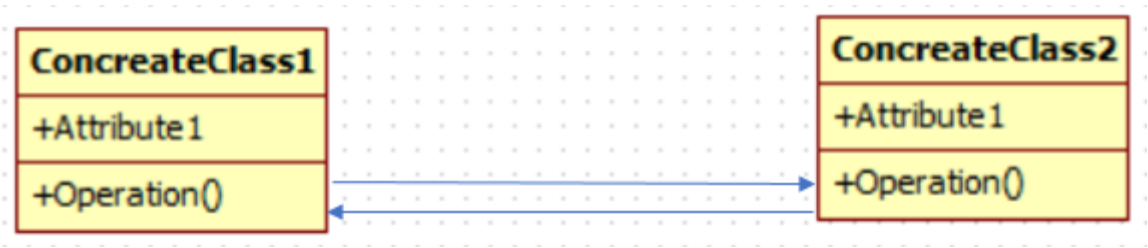


Figure 7 Refactoring Agent class

Operation classes:

Operation design defect are classes with only one meaningful behavior and for having a short life cycle. The main idea of an operation class is that an operation that should have been a method within a class has been turned into a class itself. UML notation of Operation classes is shown in Figure 8.

Figure 8 Operation classes

Refactoring operation class.

Refactoring of operation class design defect involves moving the attributes and functionality to suitable class.

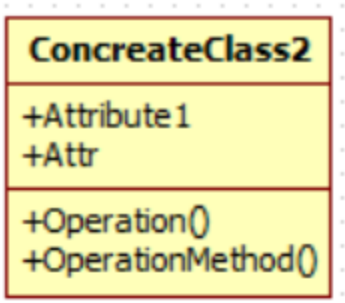


Figure 9 Refactoring operation class

Object classes:

Object classes are subclasses representing exactly the parent classes with no additional functionalities or attributes. In figure 11 classes SbClass1 and SubClass2 are Object classes as they are exactly same as their parent class ConcreateClass2.

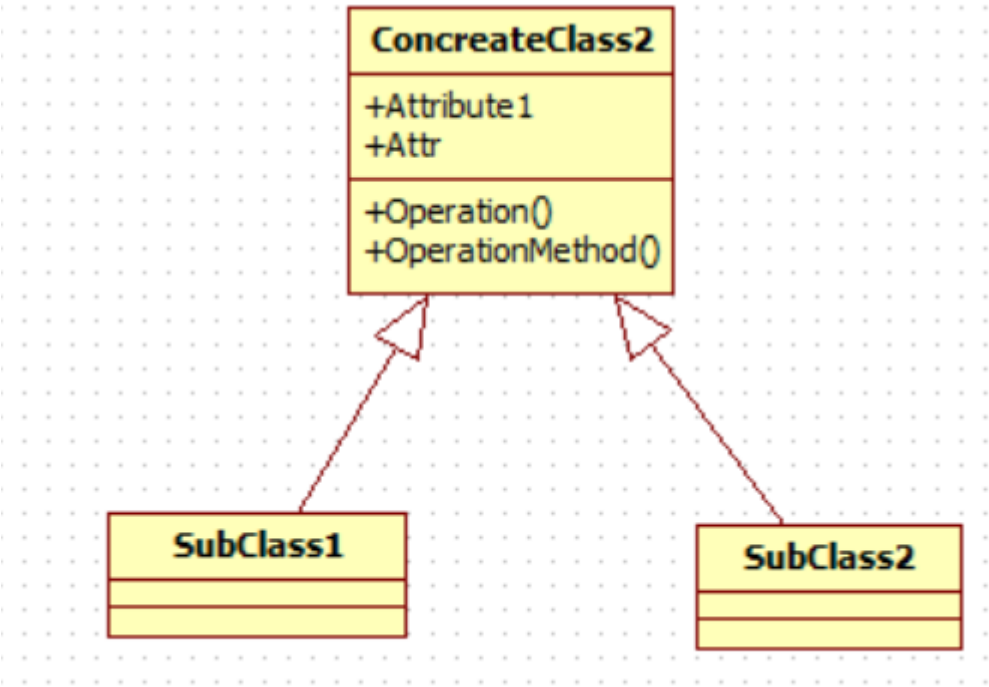


Figure 10 Object classes

Refactoring Object class.

Object class does not override any behavior or functionality of their parent class and they do not have additional behavior, they are unnecessary and therefore, must be remove them from the class hierarchy altogether.

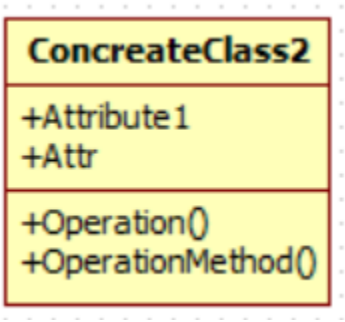


Figure 11 Refactoring Object class

VII. LAVA FLOW DESIGN DEFECT OR FUNCTIONAL DECOMPOSITION

Lava flow design defect is a class with single action such as function which makes it simple [4].

Lava flow design defect is created by designer when he creates each class for function Figure 13, resulting multiple classes in the design where the functionality not logically grouped.

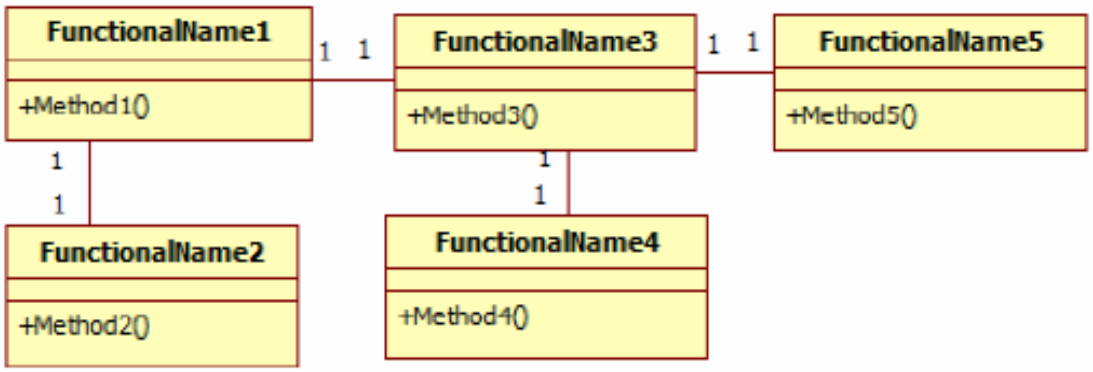


Figure 12 Lava Flow

Refactoring Function Decomposition design defect

If the class has a single method are helper classes with single functionality, remove this class by moving the method to part of an existing class base class. The goal is to consolidate the functionality of several types into a single class that captures a broader domain concept than the previous finer-grained classes. For example, rather than have classes to manage device access, to filter information to and from the devices, and to control the device, combine them into a single device controller object with methods that perform the activities previously spread out among several classes. If the class does not contain state information of any kind, consider rewriting it as a function. Potentially, some parts of the system may be best modeled as functions that can be accessed throughout various parts of the system without restriction.

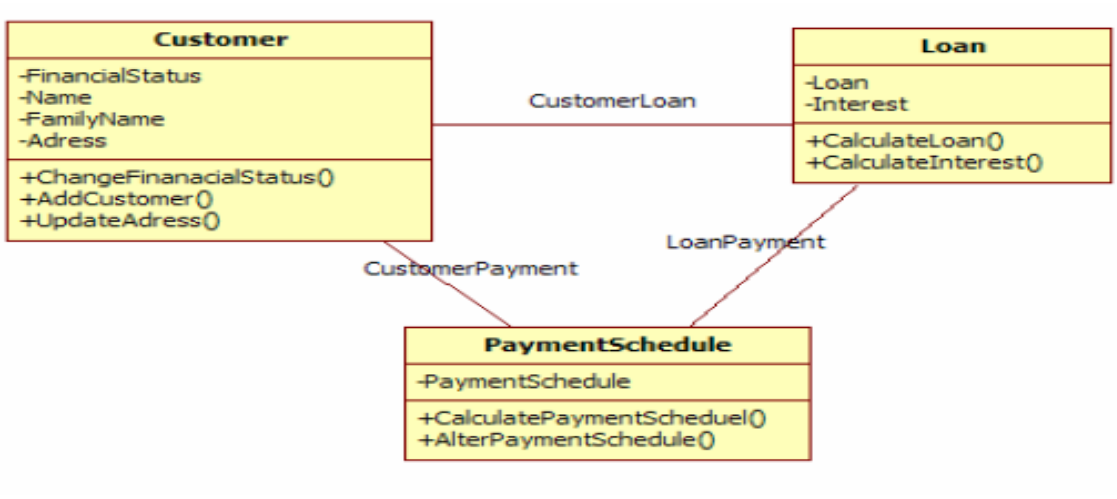


Figure 13 Refactoring Function Decomposition Design Defect

VIII. SWISS ARMY KNIFE DESIGN DEFECT

Swiss army knife class implements many interfaces to expose the maximum possible functionalities. As it implements may interface it becomes complex class exposing many functionalities. The difference between Swiss army knife and the Blob is that the Swiss army knife exposes a high complexity to address all foreseeable needs of the class, whereas the Blob is a single large multifunctional object that monopolizes all the treatment and the system data. The symptoms of the presence of Swiss Army Knife anti-pattern is : Complex interfaces with no clear abstraction or purpose for the class, which is represented by the lack of focus in the interface.

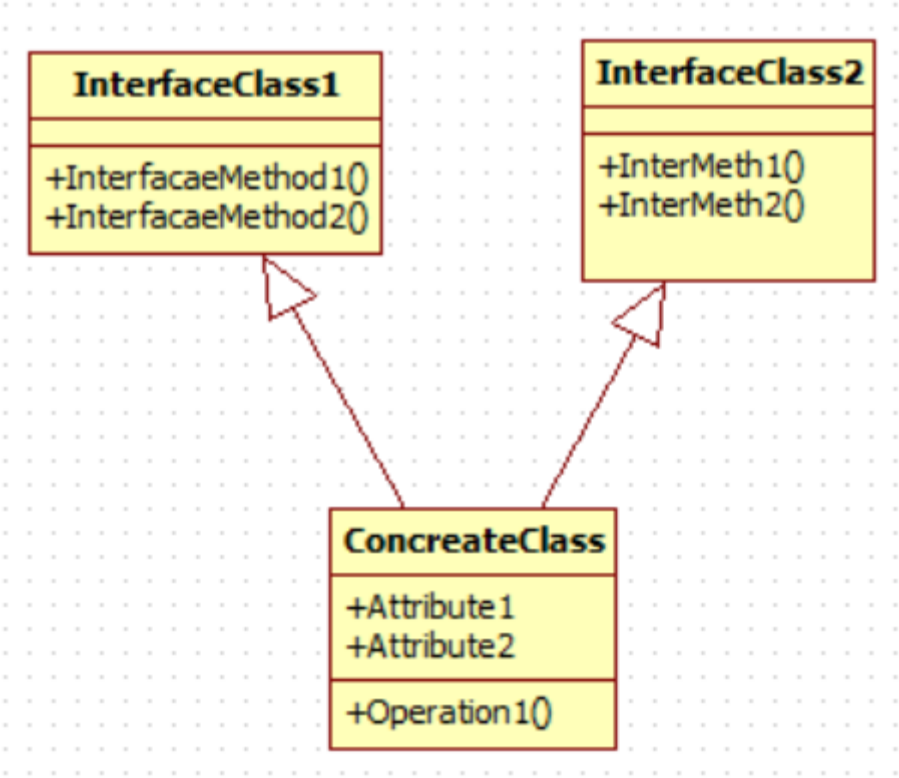


Figure 14 Swiss Army Knife

Refactoring Swiss Army Knife

Refactoring the swiss army knife involves reducing the complexity of the interfaces.

Apply Extension Interface Patterns:

- Introduce a common protocol for all provided interfaces (incl. Interface navigation)
- Integrate additional functionality so that clients can discover existing component interfaces and navigate between them.

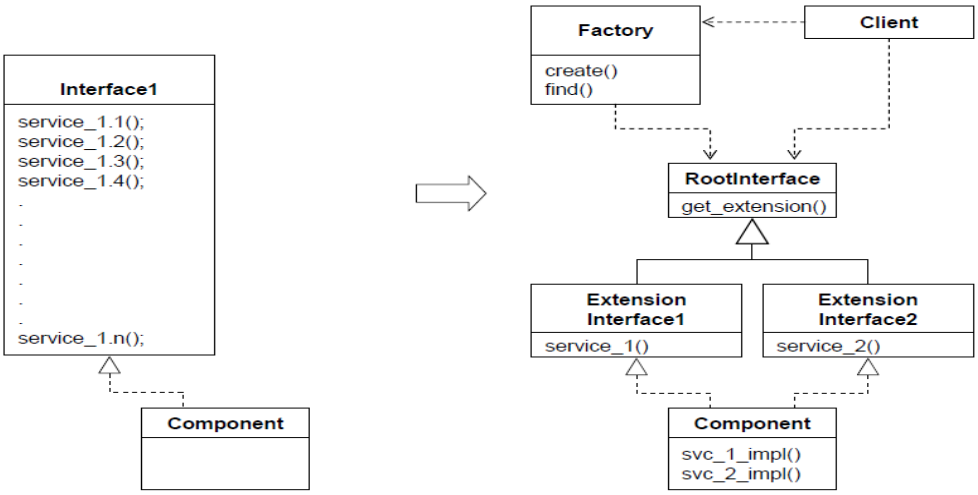


Figure 15 Refactoring of Swiss Army Knife anti-patter

IX. PROGRAMMING ERRORS OR DEFECTS.

Defects for programming coding errors, assignment versus equality operators, type mismatch, wrap around, string arrays. They occur due to the bad programming below table lists defects and their symptom and correction. Defects are under Numerical Defects, Programming Defects and Programming Defects.

Numerical Defects [8]		
Defect Type - Defect Description	**Defective Code Sample**	**Corrective Code Sample**

Float overflow - Overflow from operation between floating points	float square(void) {
	float val = FLT_MAX;
	return val * val;
	}
		double square(void) {
		float val = FLT_MAX;
		return val * val;
		}

Invalid use of standard library floating point routine - Wrong arguments to standard library function	double arccosine(void) {
	double degree = 5.0;
	return acos(degree);
	}
		double arccosine(void) {
		double degree = 5.0;
		double radian = degree*180/(3.14159);
		return acos(radian);
		}

Float division by zero - Dividing floating point number by zero	float fraction(float num){
	float denom = 0.0;
	float result = 0.0;
	result = num/denom;
	return result;}
		float fraction(float num){
		float denom = 0.0;
		float result = 0.0;
		if(((int)denom) != 0)
		result = num/denom;
		return result;}

Integer conversion overflow - Overflow when converting between integer types	char convert(void) {
	int num = 1000000;
	return (char)num;
	}
		long convert(void) {
		int num = 1000000;
		return (long)num;
		}

Integer overflow - Overflow from operation between integers	int plusplus(void) {
	int var = INT_MAX;
	var++;
	return var;
	}
		long plusplus(void) {
		long lvar = INT_MAX;
		lvar++;
		return lvar;
		}

Invalid use of standard library integer routine - Wrong arguments to standard library function	int absoluteValue(void) {
	int neg = INT_MIN;
	return abs(neg);
		int absoluteValue(void) {
		int neg = INT_MIN+1;
		return abs(neg);

```
}
```

```
}
```

Integer division by zero - Dividing integer number by zero

```
int fraction(int num){
    int denom = 0;
    int result = 0;
    result = num/denom;
    return result;
}
```

```
int fraction(int num){
    int denom = 0;
    int result = 0;
    if (denom != 0)
        result = num/denom;
    return result;
}
```

Shift of a negative value - Shift operator on negative value

```
int shifting(int val){
    int res = -1;
    return res << val;
}
```

```
int shifting(int val){
    unsigned int res = -1;
    return res << val
}
```

Shift operation overflow- Overflow from shifting operation

```
int left_shift(void) {
    int foo = 33;
    return 1 << foo;
}
```

```
long left_shift(void) {
    int foo = 33;
    return 1 << foo;
}
```

Sign change integer conversion overflow - Overflow when converting between signed and unsigned integers

```
char sign_change(void) {
    unsigned char count = 255;
    return (char)count;
}
```

```
int sign_change(void) {
    unsigned char count = 255;
    return (int)count;
}
```

Unsigned integer conversion overflow - Overflow when converting between unsigned integer types

```
unsigned char convert(void) {
    unsigned int unum = 1000000U;
    return (unsigned char)unum;
}
```

```
unsigned long convert(void) {
    unsigned int unum = 1000000U;
    return (unsigned long)unum;
}
```

Unsigned integer overflow - Overflow from operation between unsigned integers

```
unsigned int plusplus(void) {
    unsigned uvar = UINT_MAX;
    uvar++;
    return uvar;
}
```

```
unsigned long plusplus(void) {
    unsigned uvar = UINT_MAX;
    unsigned long ulvar = uvar++;
    return ulvar;
}
```

Programming Defects [8]

Invalid use of == (equality) operator - Equality operation in assignment statement

```
    for (j == 5; j < 9; j++) {
        array[i] = j;
        i++;
    }
```

```
  for (j = 5; j < 9; j++) {
        array[i] = j;
        i++;
    }
```

Invalid use of = (assignment) operator - Assignment in control statement

```
    if(alpha = beta){
        printf("Equal\n");
    }
```

```
    if(alpha == beta){
        printf("Equal\n");
    }
```

Invalid use of floating point operation - Imprecise comparison of floating point variables

```
    float flt = 1.0;
    if (flt == 1.1)
        return flt;
    return 0;
```

```
    float flt = 1.0;
    if (fabs(flt-1.1) < Epilson)
        return flt;
    return 0;
```

Dead code - Code cannot be reached along any execution path

```
 int table[5];/* Create a table */
 for(int i=0;i<=4;i++)
   table[i]=i^2+i+1;
 if(table[ch]>100) return 0;
 /*Defect: Condition always false */
 return table[ch];}
```

```
   int table[5];
  /* Create a table */
  for(int i=0;i<=4;i++)
    table[i]=i^2+i+1;
  /* Fix: Remove dead code */
  return table[ch];
}
```

Non-initialized variable - Variable not initialized before use

```
    int command;
    int val;
    command = getsensor();
    if (command == 2)
      {
        val = getsensor();
      }
```

```
    int command;
    /* Fix: Initialize val */
    int val=0;
    command = getsensor();
    if (command == 2)
      {
        val = getsensor();
      }
```

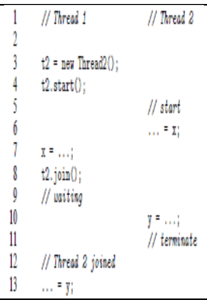
	return val;	return val;

Uncalled function - Function with static scope never called in file	static int Initialize(void) {…
	}
	void main()
	{
	int num;
	num=0;
	printf("The value of num is %d",num);
	}
		void main()
		{
		int num;
		/* Fix: Call static function Initialize */
		num=Initialize();
		printf("The value of num is %d",num);
		}

Variable shadowing - Variable hides another variable of same name with nested scope	int fact[5]={1,2,6,24,120};
	int factorial(int n){
	int fact=1;
	/*Defect: Local variable hides global array with same name */
	return(fact);
	}
		int fact[5]={1,2,6,24,120};
		int factorial(int n){
		/* Fix: Change name of local variable */
		int f=1;
		for(int i=1;i<=n;i++)
		f*=i;
		return(f);
		}

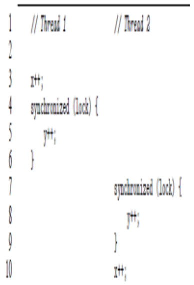
Defects in Multi-Threaded [8]

Data race - A data race is a situation where events from different threads execute without ordering and read and write the same data. Data races can lead to data inconsistency and unintended nondeterminism.



```
// Thread 1            // Thread 2

t2 = new Thread2();
t2.start();
                       // start
                       ... = x;
x = ...;
t2.join();
// waiting
                       y = ...;
                       // terminate
// Thread 2 joined
... = y;
```



```
// Thread 1            // Thread 2

x++;
synchronized (lock) {
    y++;
}
                       synchronized (lock) {
                           y++;
                       }
                       x++;
```

violation of atomicity - A violation of atomicity occurs if a sequence of shared data access of one thread is interleaved with access to the same data from other threads.	class Stack {
	int top;
	Object[] arr;
	int size() {
	return top;
	}
	void push(Object o) {
	// assert (top < arr.length);
	arr[top++] = o;
	}
	Object pop() {
	// assert(top > 0);
	return arr[--top];
	}
	}
		class Stack {
		int top;
		Object[] arr;
		int size() {
		return top;
		}
		synchronized void push(Object o) {
		// assert (top < arr.length);
		arr[top++] = o;
		}
		synchronized Object pop() {
		// assert(top > 0);
		return arr[--top];
		}
		}

Deadlock - A deadlock situation occurs at runtime if threads use synchronization so that a cyclicwait condition arises.	deadlock situation occurs at runtime if threads use synchronization so that a mutual wait condition arises	Detected should be detected and corrected

ACKNOWLEDGMENT

In preparation of this paper, I had to take help and guidance of respected persons, who deserve my deepest gratitude. I would like show my gratitude to Mr. Rabindranath Raul from Honeywell who provided the technical support also helped in IP review. In addition, I thank Dr. R. Selvarani, who guided me throughout paper preparation.

X. CONCLUSION

Currently system engineers use different design tools like UML, SYSML, MATLAB, SCADE etc to convert system requirements to architectural design diagrams. Design verification to detect the Design Defects and rectify them at design level and software errors at implementation phase is very impartment to control the flow of defects to the subsequent process steps of SLDC, making design and software more robust. Fixing design defects will make design more maintainable and reduces significant maintenance cost and software errors will reduce the unexpected behavior of the software and those reduce the defects identified at the testing phase, those reducing the cycle time, which intern reduces the manual hours.

REFERENCES

[1]. IVAN POLÁŠEK : Anti-pattern Detection as a Knowledge Utilization

[2] W. Brown, R. Malveau, H. McCormick, and T. Mowbray, AntiPatterns:Refactoring software, architectures, and projects in crisis. John Wiley & Sons, 1998.

[3] Naouel Moha1, Amine Mohamed Rouane Hacene2, Petko Valtchev3, and Yann-Ga¨el Gu´eh´eneuc1 Refactorings of Design Defects using Relational Concept Analysis

[4] Roger Lee Studies in Computational Intelligence,Volume 364

[5] R. Selvarani. S. Aishwarya. “A Formal Model indicating Inter-relationship of CK Metrics in view of Defect Proneness". Accepted for publication in the International Journal of computers and Technologies (IJCT). August 2013

[6] Maria Teresa Llano† and Rob Pooley UML Specification and Correction of Object-Oriented Anti-patterns.

[7]. Brown, W.J., Malveau, R.C., McCormick, H.W., Mowbray, T.J.: Antipatterns: Refactorin Software, Architectures, and Projects in Crisis, 1st edn. John Wily and Sons,West Sussex (1998)

[8]. Matlab defects web page http://www.mathworks.in/help/bugfinder/ref/floatconversionoverflow.html

[9] Christoph von Praun, Dipl.-Inf. TU-M¨unchen
Detecting Synchronization Defects inMulti-Threaded Object-Oriented Programs

[10] R. Selvarani.Areej, R.Bharathi “Software Reliability Estimation at Design Stage based on multifunctional estimation Technique” submitted to ACM TOSEM, August 2013

[11] R. Selvarani. S. Aishwarya , “A frame work for codifying the association of CBO and NOC in the observation of Defect proneness” submitted to Empirical software engineering journal, August 2013

[12] Smith, C.U. and L.G. Williams, Software performance antipatterns, in Proceedings of the 2nd international workshop on Software and performance. 2000, ACM: Ottawa, Ontario, Canada. p. 127-136.

[13] Lanza, M. and R. Marinescu, Object-Oriented Metrics in Practice. 2005: Springer-Verlag New York, Inc.

[14] Gaffney, J. E.: Metrics in software quality assurance, in *Proc. of the ACM '81 Conference*, ACM, 126-130, 1981.

[15]. A. J. Riel. Object-Oriented Design Heuristics. Addison-Wesley, 1996.

[16] Roger S. Pressman. *Software Engineering { A Prac- titioner's Approach*. McGraw-Hill Higher Education, 5*th* edition, November 2001. isbn: 0-07-249668-1.

[17] Robert C. Martin, first five object-oriented design(OOD) principles, https://scotch.io/bar-talk/s-o-l-i-d-the-first-five-principles-of-object-oriented-design

[18] R. Selvarani. T.R.G. Nair. "Defect Proneness Estimation and Feedback Approach for Software Design Quality Improvement" ELSEVIER, Information and Software Technology, volume 54 issue 3, 274 – 285 (2012)

[19] OMG, *OMG Unified Modeling Language (OMG UML), Infrastructure. V.2.2*. Object Management Group, 2009.

[20] N. Moha and Y. Gu´eh´eneuc, "On the automatic detection and correction of software architectural defects in object oriented designs," *ECOOP Workshop on Object-Oriented OUTReengineering*, 2005.

A Stepwise Approach to Automate the Search for Optimal Parameters in Seasonal ARIMA Models

Manisha Mukherjee, Diptarka Saha

WalmartLabs, Bengaluru

Abstract- Reliable forecasts of univariate time series data are often necessary in several contexts. ARIMA models are quite popular among practitioners in this regard. Hence, choosing correct parameter values for ARIMA is an imperative yet task. Thus, a stepwise algorithm is introduced to provide robust estimates for parameters $(p, d, q)(P, D, Q)$ used in seasonal ARIMA models. This process is focused on improvising the overall quality of the estimates and it alleviates the problems induced due to the unidimensional nature of the methods that are currently used such as 'auto.arima'. The fast and automated search of parameter space also ensures reliable estimates of the parameters that possess several desirable qualities and consequently resulting in higher test accuracy especially in the cases of noisy data. After vigorous testing on real as well as simulated data the authors have found sufficient evidence to conclude that the algorithm performs better than current state-of-the-art methods, all the while completely obviating the need for human intervention due to its automated nature.

Index Terms- Time Series, auto.arima, ARIMA parameters, Forecast, R

I. INTRODUCTION

Time Series analysis and forecasting plays an essential role in various fields such as business, finance, economics, science, and engineering. Due to its importance in solving practical problems, several methods have been proposed in the literature to model a time series so that past observations are carefully handled, and future can be predicted accurately with confidence. Time series forecasting is thus nothing but an endeavour to predict the future by an astute scrutiny of the past.

One of the most popular and frequently used time series models is ARIMA (Autoregressive Integrated Moving Average) suggested by George Box and Gwilym Jenkins in their seminal text-book *Time Series Analysis: Forecasting and Control* [1]. The modelling approach is well celebrated in the academic community due to its robust theoretical underpinnings. In fact, under certain assumptions, it has been shown that ARIMA models may yield the *optimal forecasts*, outperforming competing methods such as *Exponential Smoothing* [2]. Variants of ARIMA, such as seasonal ARIMA, has been in use as well, with additional sets of parameters to capture the seasonality present in the series.

When applied on real-world data however, ARIMA originally didn't receive such vogue from industry practitioners. This is partly because business data may not always conform to the necessary assumptions, but mostly due to the difficult, iterative time-consuming, and highly subjective procedure described by Box and Jenkins to identify the proper form of the model for a given data set. There have since been many attempts at automating the search for the optimal ARIMA model. However, as will be discussed in the later sections – all of them show several limitations.

This paper presents a stepwise algorithm which automates the iterative nature of Box and Jenkin's approach to find best seasonal ARIMA model for a particular time series and performs better than current state of the art algorithms in terms of various criterion on which a time series model is judged and thus may fail to eliminate the need of human intervention.

Section II summarizes the existing works which has been done to find solution to automatic ARIMA modelling. Section III describes the ARIMA model and the proposed algorithm. Section IV discusses the advantages of proposed algorithm over existing processes. Section V sums up the performance of the proposed algorithm on various time series data sets and section VI provides a salient conclusion.

II. LITERATURE REVIEW

Several attempts have been made in direction to automate the procedure to find ARIMA model in the past years, especially, in the eighties and nineties. The most recent method is given by Hyndman et. al [3] in which he proposed finding an optimum ARIMA model by minimizing AIC [4] by considering different combinations of model parameters. Presently, this procedure is most used commercially as a back-end algorithm in statistical software R's forecast package's *auto.arima* and in Scikit-Learn's *pmdarima* in Python.

Hannan [5] proposed a method to identify the order of an ARMA model for a stationary series by fitting the innovations as an autoregressive model to the data followed by computation of likelihood of potential models using a series of standard regressions. The asymptotic properties of the procedure under very general conditions were then derived.

Gömez [6] extended the Hannan-Rissanen identification method to include multiplicative seasonal ARIMA model identification. They implemented the automatic identification procedure in the softwares TRAMO and SEATS in which the algorithm fetched the model with minimum BIC.

Mélard and Pasteels [7] proposed an algorithm for univariate ARIMA models which also allows intervention analysis and has been implemented in the software package "*Time Series Expert*" (TSE-AX). Liu [8] also suggested a method for identification of seasonal ARIMA models using a filtering method and certain heuristic rules which is used in SCA-Expert software. Forecast Pro [9], which is partially based on [10] and is famous for its automatic ARIMA algorithm which was used in the M3-forecasting competition [11]. Another proprietary algorithm is implemented in Autobox [12].

Hwang [13] developed an automated time series cost forecasting system (ATMF) including both auto-selected procedures for determining a best-fitting model and an auto-extracting module for forecasting values using the Box-Jenkins approach. Amin et. al [14] proposed an automated forecasting approach based on ARIMA to capture linear and non-linear patterns to predict future values of Quality of Service (QoS) attributes that can assist in controlling in software intensive systems.

However, it needs to be highlighted that not much attempts have been proposed post Hyndman et. al [3] to automate the procedure of modelling in ARIMA. The lack of literature in recent times along with some serious limitations of the existing processes make it imperative to develop an automated forecast method based on the original Box and Jenkins approach.

III. MODEL

A. Framework

A seasonal ARIMA $(p,d,q)(P,D,Q)_f$ process is given by

$$\Phi\left(B^f\right)\phi(B)(1-B^m)^D(1-B)^d y_t \ = c + \Theta\left(B^f\right)\theta(B)\epsilon_t$$

Where, ϵ_t is a white noise process with mean 0 and variance $\sigma^2, \Phi(z)$ and $\Theta(z)$ are polynomials of order P and Q respectively, each containing no roots inside a unit circle, B is a backshift operator and f is the seasonal frequency of the series.

For a seasonal time series, $D = 0 \ or \ D = 1$ is decided on the basis of Canova-Hansen test [15]. This test tests the Null Hypothesis that no seasonal unit root is present. Unit-root tests to test for presence of stochastic trend such as ADF test [16] and KPSS test [17] are performed to choose the optimum value of the parameter d. ADF test tests the Null Hypothesis that unit root is present whereas KPSS test tests the Null Hypothesis that no unit root is present.

After optimum d and D is chosen, optimum $p,\ q,\ P,$ and Q are selected on the basis of the ACF and PACF of the series so that best model with minimum AIC and best performance on test data is obtained.

A time series Y_t has a mean,

$$\mu = E[Y_t]$$

and autocovariance function,

$$\gamma_Y(t+h,t) = Cov(Y_{\{t+h\}}, Y\_t) = E\left[\left(Y_{t+h} - \mu_{(t+h)}\right)(Y_t - \mu_t)\right]$$

It is stationary when both mean and autocovariance function are independent of t.

The Autocorrelation Function (ACF) is

$$\rho_y(h) = \frac{\gamma_Y(h)}{\gamma_Y(0)} \ = \ Corr(Y_{t+h}\ , Y_t)$$

When $d,\ D,\ p,\ q, P,$ and Q are known, a model can be evaluated via an information criterion such as AIC:

$$AIC \ = \ -2log(L) \ + \ 2(p+q+P+Q+k),$$

where, $k = 1$ if $c \neq 0$ and $k \ = \ 0$ otherwise and L is the maximized likelihood of the model fitted to the differenced data $\left(1-B^f\right)^D(1-B)^d y_t$

The performance of a time series model is evaluated on the basis of *Mean Absolute Percentage Error* (MAPE). In a time series, if y_t is the actual value and $\widehat{y_t}$ is the predicted value (forecast) and the series conists of n time points then MAPE is defined as follows:

$$MAPE \ = \frac{100}{n}\sum_{t=1}^{n}\frac{|y_t - \widehat{y_t}|}{y_t}$$

An optimised seasonal ARIMA model should satisfy the following criteria:

1. Optimum difference order, d, should be on the basis of both KPSS and ADF test.
2. Optimum seasonal difference order, D, should be on the basis of Canova-Hensen test.
3. Optimum p, q, P and Q should be on the basis of ACF and PACF.
4. It should have minimum AIC.
5. It should have minimum test MAPE.
6. Model residuals should not have serial autocorrelation

B. Algorithm

Finding the best ARIMA Model has always been subjective and difficult. There is no hard and fast rule suggested in literature to find a best ARIMA model which makes developing an automated solution to it even more cumbersome. Robert Hyndman [3] has suggested finding optimum ARIMA models on the basis of AIC, which is currently being used in popular statistical softwares such as auto.arima in R and pmdarima in Scikit-learn in python. However, we note, this approach is unidimensional in nature as it only looks to find the model with minimum AIC.

We suggest an algorithm which has a multidimensional view and it has been broken down into three separate algorithms in this paper for the understanding of the reader. The Algorithm 1 finds optimum D and d of the ARIMA model. It consumes input data which is a time series data and tests for presence of seasonality in the data. It carries out Canova Hansen test to check for seasonal unit root and passes out the seasonally differenced data to next step. The algorithm automatically skips this test if the data is annual in nature. In next step, Augmented Dickey Fuller test and KPSS test are performed to find optimum d. In case, results of ADF test and KPSS test do not match, it throws a warning to user.

Algorithm 2 gives optimum $p, q, P,$ and Q of the model. The differenced data from algorithm 1 is taken as an input and ACF and PACF of the series are computed. Based on the values of ACF and PACF, optimum values of p, q, P and Q are found. There has been an upper bound put on the values of these parameters so that the resulting model is parsimonious.

Algorithm 3 fits the ARIMA model and performs diagnostic tests. It stores AIC of the resulting model and compute residuals. It then performs Ljung-Box test on residuals to check for presence of serial autocorrelation in the residuals. It throws the optimum parameters and AIC of the model along with the results of Ljung-Box test.

Let us, then, formally introduce the algorithms.

Algorithm 1: Finding Optimum D and d of ARIMA Model

Input: Time Series y_t, explanatory variable x_t, frequency f, cut off value V

Output: Optimum D, d

Initialize:

$$z_t = 0, d = 0, D = 0, w_t = 0, d_1 = 0, d_2 = 0$$

Seasonal Unit Root:

if $f > 1$:

Perform Canova Hansen Test on y_t

if H_0 *is rejected* then $z_t = y_t - y_{t-f}$; $D = 1$

else $z_t = y_t$; $D = 0$

Augmented Dickey Fuller Test:

while $i \in \{0,1,2\}$ do

$u_t = \Delta^i(z_t)$ Perform ADF test on u_t

if H_0 *is rejected* then $w_t = u_t$; $d_1 = i$

else $i = i + 1$

Kwiatkowski Phillips Schmidt Shin Test:

while $i \in \{0,1,2\}$ do

$v_t = \Delta^i(z_t)$ Perform KPSS test on v_t

if H_0 *is rejected* then $w_t = v_t$; $d_1 = i$

else $i = i + 1$

if $d_1 = d_2$ then $d = d_1 = d_2$

else $d = d_1$; output warning message

Algorithm 2: Finding Optimum p, q, P, and Q of ARIMA Model

Input: Time Series y_t, explanatory variable x_t, frequency f, cut off value V

Output: Optimum p, q, P, Q

Initialize:

$$a_t = 0, b_t = 0, p = 0, q = 0, P = 0, Q = 0$$

for $1 \leq h \leq f + 5$

$$a_t[h] = \rho(h)$$

for $1 \leq h \leq f + 5$

$$b_t[h] = \pi(h)$$

Optimum q:

```
for i ∈ {1,2,3,4,5} do
    if a_t[i] < V then
        q = i − 1; j = i + 1

        for j ∈ {1,2,3,4,5} do
            if a_t[j] < V then
                q = j − 1
            else
                q = 0
```

Optimum p:

```
for i ∈ {1,2,3,4,5} do
    if b_t[i] < V then
        p = i − 1; j = i + 1

        for j ∈ {1,2,3,4,5} do
            if b_t[j] < V then
                p = j − 1
            else
                p = 0
```

Optimum P:

```
for i ∈ {f + 1, f + 2, ..., f + 5} do
    if b_t[i] < V then
        P = i − 1; j = i + 1

        for j ∈ {1,2,3,4,5} do
            if b_t[j] < V then
                P = j − 1
            else
                P = 0
```

Optimum P:

```
for i ∈ {f + 1, f + 2, ..., f + 5} do
    if a_t[i] < V then
        Q = i − 1; j = i + 1

        for j ∈ {1,2,3,4,5} do
            if a_t[j] < V then
                Q = j − 1
            else
                Q = 0
```

Algorithm 3: Fitting optimal ARIMA model and performing diagnostic tests

Input: Time Series y_t, explanatory variable x_t, frequency f, cut off value V

Output: Optimum p, d, q, P, D, Q, R, AIC

Initialize:

$$AIC = 0, R = 0$$

Fit ARIMA $(p, d, q)(P, D, Q)[f]$ on w_t and explanatory variable x_t

Compute AIC of the model, store as AIC

Compute residuals $e_t = y_t - \widehat{y_t}$

Ljung Box test on Model Residuals:

if H_0 *is Rejected*

$R = 1$; Serial Autocorrelation Absent

else

$R = 0$; Serial Autocorrelation Present

IV. DISCUSSION

Literature suggests that a best time series model should satisfy all the criterion discussed in section III.A. Existing automated solutions for finding suitable ARIMA model used in popular software packages only take care of one aspect – namely, minimizing the AIC of the model and only conducts KPSS test to find optimum d. They consider different combinations of p, q, P, and Q and choose the one which provides minimum AIC.

The authors have sufficient evidence to believe that due to such neglect in consideration of ACF and PACF current methods fail to produce optimum model, especially in those cases where the data is noisy. It has also been observed that in some cases they give positive values of P and Q even though it is clearly evident from the data (and from ACF and PACF) that there is no seasonality present. In addition to this, they do not conduct Ljung-Box test on model residuals on their own. In literature, it is highly recommended in order to notify the user if the serial correlation is present in the residuals or not.

The algorithm discussed in previous section overcomes all these limitations and consequently shows better performance than existing solutions. First, it takes into account the results of both KPSS test and ADF test to find out optimum difference order and in case of contradictory results, it throws a warning to the user. Second, it iterates over ACF and PACF of the series to find out optimum $p, q, P,$ and Q which makes this algorithm highly effective in case of noisy data to capture all the patterns and nuances present in the series. It tests the model on test data on its own to provide the user test MAPE. Third, this algorithm performs Ljung-Box test on model residuals to notify user if there is a serial correlation present

in residuals generated by the model. Fourth, it gives the user liberty to choose the cut off value i.e. the lag(V) till which ACF and PACF are considered significant on his own, unlike the existing black box solutions. Fifth, in cases where ACF and PACF behaves abnormally, for e.g., ACF at lag 1 and lag 2 is insignificant but ACF at lag 3 is significant, the existing solutions would provide MA order as 1 or 2 but this algorithm goes to lag 3 and beyond to find optimum MA order while making sure that model remains parsimonious.

As a result, the model generated by this algorithm tend to perform better than the models generated by existing solution in terms of both AIC and test MAPE.

V. EXPERIMENTS

We show here three implementations of the proposed algorithm and demonstrate the superiority of our proposed solutions over currently popular solution such as auto.arima. Three cases have been chosen to highlight different aspect of our algorithm as will be discussed later.

A. Dataset 1: Nile Data

It is a standard time series data present in various statistical softwares summarizing annual flow of Nile River from year 1871 to 1970. It is illustrated in figure 1(a).

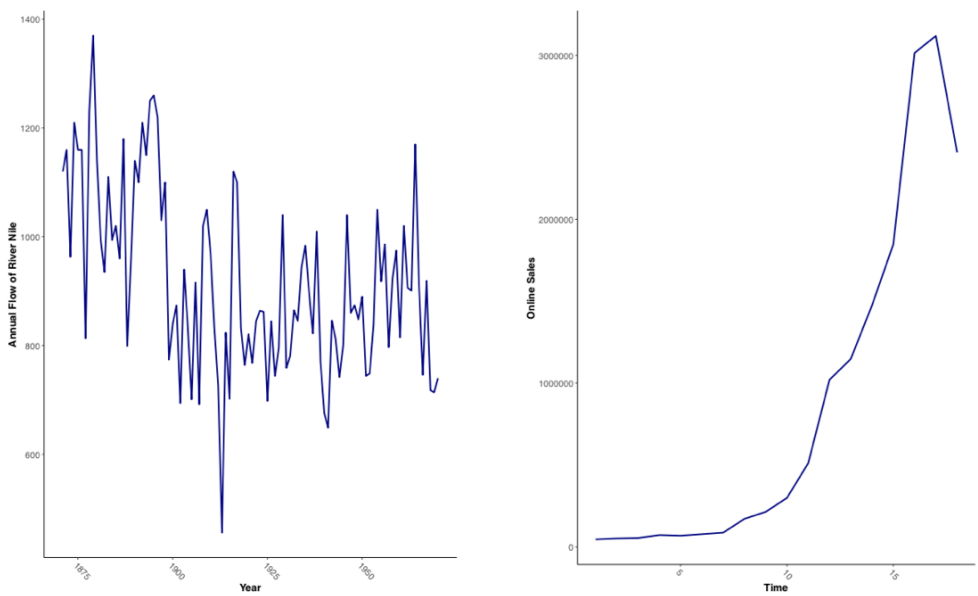


Figure 1(a): Annual Flow of Nile River

Figure 1(b): Quarterly Sales in Walmart

The data is split into train and test data, in which, train data is from year 1871 to 1941 and test data from year 1940 to 1970. To align with notations used in the algorithm in section II, we have, frequency $f = 0$, Train data as y_t and cut off value $V = 0.33$ as input. Canova-Hansen Test is not performed as time series is yearly in nature. Later, ADF test is performed followed by KPSS test, both of which, suggested optimum $d = 1$. ACF and PACF were then computed on the differenced series, w_t, to find optimum values of $p, q, P,$ and Q. The correlograms are depicted in figure 2(a).

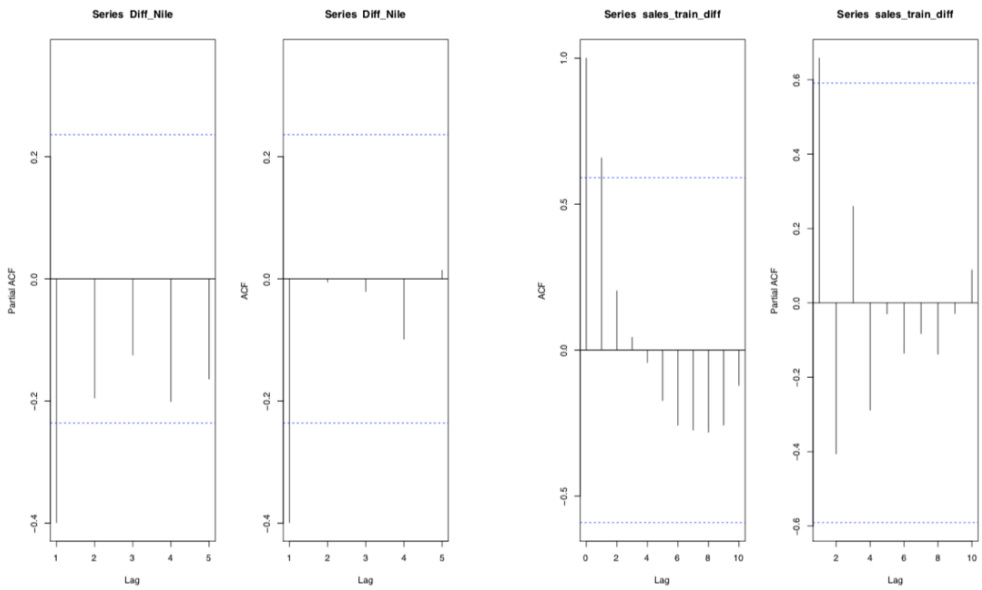


Figure 2(a): ACF, PACF in Dataset 1

Figure 2(b): ACF, PACF in Dataset 2

The algorithm derives ACF and PACF of the differenced series and gives optimum values of $p, q, P,$ and Q as (1,1,0,0) respectively. This finding is consistent with the figure 2(a). There are significant spikes in correlograms of both ACF and PACF at lag 1 which diminish from lag 2 onwards. This step is followed by fitting an ARIMA model and performing Ljung-Box test on the model residuals which suggested no serial autocorrelation in model residuals with $p - value = 0.71$. The final output is $p = 1, d = 1, q = 1, P = 0, Q = 0, D = 0, R = 1, AIC = 894.13$. The test MAPE is 10.56 %.

To compare with existing automated solutions, auto.arima is implemented on the same train and test data. It suggests $p = 1, d = 0, q = 0, P = 0, D = 0, Q = 0, AIC = 910$. The optimum values of q is not consistent with values of ACF of the series. It does not perform Ljung-Box test on its own. When performed manually on model residuals generated by this model, the test suggests no significant serial autocorrelation at p-value 0.87. The test MAPE is 12.68 %. By looking at figure 2(a) and at the results of both KPSS test and ADF test, it is evident that the performance of the search algorithms used in auto.arima is suboptimal. Due to mis-specification of the model, the AIC obtained is not minimum and it performs relatively poor on the test data.

B. Dataset 2: Real-life Retail Data

We proceed by showcasing how the existing automated solutions tend to fail with real life seasonal data. Figure 1(b) presents Quarterly Sales data of a particular department of food section of Walmart U.S. The data has a non-linearity present which is evident from the slightly U-shaped form.

We execute the proposed algorithm on this data by specifying y_t as the Online Sales, $f = 4$, $V = 0.33$ We have following two explanatory variables as input all the hypothesis testing were carried out at 5% significance level.

1. x_{1t}: Real GDP of USA

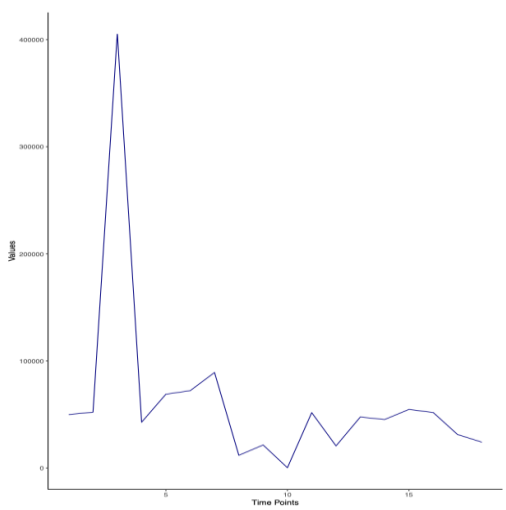
2. x_{2t}: Unemployment rate of USA

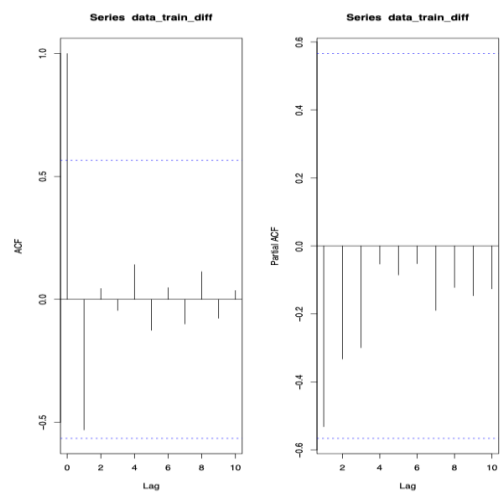
Canova-Hansen test is applied which suggested optimum $D = 0$ with $p - value = 0.17$. Afterwards, both KPSS test and ADF are test applied and both suggested optimum $d = 2$. ACF and PACF were then computed on the differenced series, w_t.The correlograms are presented in figure 2(b).

Optimum $p, q, P,$ and Q are generated followed by fitting the ARIMA model and performing the Ljung-Box test on the model residuals. The output is $p = 2, q = 1, d = 2, P = 0, D = 0, Q = 0, R = 1, AIC = 302.67$. The optimum values of $p, q, P,$ and Q match with the occurrence of spikes in the ACF and PACF of the series presented in fig. 2(b). The test MAPE is 18.67%.

For comparison with the existing process, we implement auto.arima on same train and test data which gives $p = 0, q = 0, d = 0, P = 0, D = 0, Q = 0, AIC = 381.43$. We notice once again, auto.arima fails to produce model with correct parameter values. Model residuals have insignificant autocorrelation according to Ljung-Box test. The test MAPE is 35.39%.

We note here that since auto.arima doesn't perform extensive tests for p, q and hence misses the mark while detecting the non-linearity mentioned previously. This is alleviated our algorithm and consequently,





better results are obtained.

C. Dataset 3: Simulated Data with Noise

Figure 3(a): Plot of Dataset 3

Figure 3(b): ACF, PACF in Dataset 3

A time series is simulated to showcase how the proposed algorithm is performing better when the series is noisy. The simulated data is quarterly in nature.

The data is presented in figure 3(a). An unusually high peak at time point 2 and small trough at time point 10 indicate that the series is noisy and has significant outliers present. We executed the proposed

algorithm on this series with no exogenous variables at $V = 0.33\ and\ f = 4$. All the tests are done at 5% significance level.

First, Canova-Hansen test is executed which gave optimum $D = 0$ with $p - value = 0.33$. Following this, KPSS test and ADF test are executed. Here, the ADF test suggested optimum $d = 1$ and KPSS test suggested optimum $d = 0$, it threw warning to the user that '*ADF and KPSS results are different*' and went on to consider optimum $d = 1$. ACF and PACF are then computed on the differenced series, w_t. The correlograms are shown in figure 3(b).

Optimum $p, q, P,$ and Q are found out based on the values of ACF and PACF of the differenced series and ARIMA model is then fitted. This is followed by performing Ljung-Box test on the residuals generated. The output is $p = 2, d = 1, q = 1, P = 0, D = 0, Q = 0, R = 1, AIC = 320.67$. The optimum values of the parameters of the model match with the plots of ACF and PACF of the differenced series presented in fig. 3(b). The test MAPE is 29.52%.

For comparison purpose, we also implemented auto.arima on same train and test data which produced the output $p = 0, q = 0, d = 0, O = 0, Q = 0, D = 0, AIC = 340.03$. The test MAPE is 91.52% which is magnanimous and indicates that auto.arima fails miserably to read the irregularities present in the data.

To diagnose the reason behind this bump in performance, we must note that the proposed algorithm takes into account the behaviour of ACF and PACF of the series which makes it aware of the outliers present in the data. The resultant optimum value of $p = 2$ dictates that the predicted value of series at any time point depends on the values of past two points and hence the effect of the outlier will be subdued. Adding yo this, the coefficient attached to two AR components are 0.57 and 0.45, both of them less than unity in absolute terms. This implies that, although the occurrence of outlier at time point 10 goes on to affect the predicted values from time point 11 onwards, the effect of the outlier declines significantly as we move far from time point 10. Effectively, this means that when we consider test data from time point 14 onwards, the effect of outlier had already started getting dampened to the point that the predicted values do not get affected much. This results in high accuracy of the forecasts produced.

On the other hand, auto.arima, in order to reduce AIC without considering ACF and PACF of the series produces a model with AR and MA components equal to zero and a non-zero mean. It fails to match the crests and troughs present in a data. The estimated value of mean is obviously affected by the presence of outlier and is equal to 71766.15. It ends up predicting same value of the time series in both train and test data at all time points which are equal to 71766.15. As a result, it fares poorly in terms of MAPE on test data as well in terms of AIC.

The examples described above proves how the existing automated solutions get model specifications wrong at times due to not taking in to account the results of both ADF and KPSS test and the behaviour of ACF and PACF of the time series. The proposed algorithm outperforms these solutions in terms of both model AIC and test MAPE while taking same amount of computational time. It handles the noise as well as non-linearity in the data much better and produce best ARIMA models.

VI. CONCLUSION

Authors would like to reiterate here that the proposed methods will only be successful where ARIMA is able to capture the variability in the data. ARIMA, despite being statistically coherent, suffers from certain limitations. Petrica et al [18] had discussed in detail how ARIMA fares poorly in financial data. This is due to its inability to capture heteroscedasticity. In case of small data, ARIMA fails to capture a seasonal pattern which makes it data hungry and its dependency on assumption of normality of data makes it impractical.

That being said, time series is an indispensable part of research in several domains and seasonal ARIMA is one of the most common models used in all such works. However, as a result of the generality that SARIMA models offer parameter tuning becomes tiresome. As we discussed here, all previous efforts have been focused on optimising these parameters based on one or two criteria. Hence, in practice almost always the practitioner resorts to iterative optimisation through parameter space. We attempt to remedy that by introducing a series of tests in a need statement. As can be understood heuristically at each step we urge to fulfil one of the criteria of a good ARIMA fit as described in section III.A. The superiority of the process over the existing processes is also evident in the examples we have provided. Based on this and several other experiments the authors have conducted, we can conclude not only our model provides optimal parameter tuning resulting in better performance when executed on test data, but also the output parameters are more realistic and meaningful.

REFERENCES

[1] G. E. P. Box and Gwilym M. Jenkins, Time series analysis: forecasting and control, Holden-Day, 1970.

[2] C. C. Holt, "Forecasting seasonals and trends by exponentially weighted moving averages," *International Journal of Forecasting,* vol. 20, no. 1, pp. 5-10, 2004.

[3] R. Hyndman and Y. Khandakar, "Automatic Time Series Forecasting: The forecast Package for R," *Journal of Statistical Software,* vol. 26, 2007.

[4] H. Akaike, "A new look at the statistical model identification," *IEEE Transactions on Automatic Control* , vol. 19, no. 6, pp. 716 - 723, 1974.

[5] E. Hannan and J. Rissanen, "Recursive estimation of mixed autoregressive-moving average order," *Biometrika,* vol. 69, no. 1, pp. 81-94, 1982.

[6] V. Gomez and Maraval, "Statistical and Econometrics Software: TRAMO and SEATS," *Statistical and Econometrics Software,* 1996.

[7] G. Melard and J. Pasteels, "Automatic ARIMA modeling including interventions, using time series expert software," *International Journal of Forecasting,* vol. 16, no. 4, pp. 497-508, 2000.

[8] L.-M. Liu, "Identification of seasonal ARIMA models using a filtering method," *Communications in Statistics - Theory and Methods,* vol. 18, no. 6, pp. 2279--2288, 1989.

[9] R. L. Goodrich, "The Forecast Pro methodology," *International Journal of Forecasting,* vol. 16, pp. 533-535, 2000.

[10] R. L. Goodrich, "FOREX: A Time Series Forecast-ing Expert System," *Proceedings of Fourth International Symposium on Forecasting,* 1984.

[11] S. Makridakis and M. Hibon, "The M3-Competition: results, conclusions and implications," *International Journal of Forecasting,* vol. 16, no. 4, pp. 451-476, 2000.

[12] D. Reilly, "The AUTOBOX system," *International Journal of Forecasting,* vol. 16, no. 4, pp. 531-533, 2000.

[13] S. Hwang, M. Park, H.-S. Lee and H. Kim, "Automated Time-Series Cost Forecasting System for Construction Materials," *Journal of Construction Engineering and Management,* vol. 138, no. 11, pp. 1259-1269, 2012.

[14] A. Amin, L. Grunske and A. Colman, "An automated approach to forecasting QoS attributes based on linear and non-linear time series modeling," in *27th IEEE/ACM International Conference on Automated Software Engineering*, 2012.

[15] F. Canova and B. Hansen, "Are Seasonal Patterns Constant over Time? A Test for Seasonal Stability," *Journal of Business & Economic Statistics,* vol. 13, no. 3, pp. 237-52, 1995.

[16] D. Dickey and W. A. Fuller, "Likelihood Ratio Statistics for Autoregressive Time Series with a Unit Root," *Econometrica,* vol. 49, no. 4, pp. 1057-72, 1981.

[17] D. Kwiatkowski, P. C. Phillips and P. Schmidt, "Testing the Null Hypothesis of Stationarity Against the Alternative of a Unit Root: How Sure Are We That Economic Time Series Have a Unit Root?," Cowles Foundation for Research in Economics, Yale University, 1991.

[18] A.-C. Petrică, S. Stancu and A. Tindeche, "Limitation of ARIMA models in financial and monetary economics," *Theoretical and Applied Economics,* vol. 23, no. 4, pp. 19-42, 2016.

Low Power Dissipation in Adder using Isolated Sleepy Keeper Approach

Sandeep Dhariwal*, Harinath Aireddy*, Prajwal Kumar K V*, Ravi Shankar Mishra**

* Department of Electronics and Communications, College of Engineering & Design Alliance University, Bengaluru (INDIA)

** Department of Electronics and Communications, Symbiosis University of Applied Sciences, Indore (INDIA)

Abstract- Arithmetic adder is basic of all computation or calculation so if it is improved with decreasing power so it will lead to better result. Approaches like sleepy keeper technique, sleepy keeper using isolated sleepy keeper technique are discussed to reduce power and their simulation is done on cadence low power kit to get result. There are disturbances from the input which lead to disturbance in the output, which is assumed to be stable in case of sleepy keeper approach. This disturbance creates swings and introduce noise in such a way that signal cannot be recognized any more. Isolated Keeper approach can retain the output in sleep state even in the case of disturbance. With this approach, a considerable reduced dynamic power is obtained at the cost more power in static state. Low dynamic power leads to good efficiency of device.

Index Terms- Static power, Dynamic power, Sleepy keeper approach, Isolated Sleepy keeper approach.

I. INTRODUCTION

Addition, subtract, multiply, division are some regular operations called arithmetic that are generally utilized as a part of microelectronic frameworks. Addition is a fundamental operation of any arithmetic process and is the fundamental or pillar of numerous other customarily utilized math operations. Accordingly, 1-bit Full Adder cell is the most critical and fundamental square of a number mathematical unit of a framework. Clearly, enhancing its power straightforwardly prompts enhancing the execution of the entire framework The wide utilization of this operation in math capacities, have made numerous analysts anxious to propose a few sorts of various rationale styles for actualizing 1-bit Full Adder cell. The increase in business of convenient gadgets, for example, mobile phones, gaming reassures, and so forth require microelectronic devices plan with low power dissipation. As the extent of chip, combination and multifaceted nature of the chips keep on increasing, and then heat or power is expanding then cool it down, increment the cost and usefulness of the processing frameworks. As the innovation increases to 65nm there is not as much increment in unique power dissipation but rather the static or leakage power is same or surpass when the dynamic power increment past 65nm technology. In this way the procedures used to control power is not restricted to dynamic power [1-3]. At different theoretical levels control streamlining in a processor can be accomplished. Framework or Algorithm or Architecture have a more potential for power sparing even these methods tend to expand greater usefulness on an Integrated Circuit. So improvement at Circuit and Technology level is likewise vital for scaling down of ICs. Add up to Power dissipation in a CMOS circuit is total of dynamic power, static or spillage control. Plan for low-control suggests the capacity to diminish every one of these segments of energy utilization in CMOS circuits amid the improvement of a new low power electronic item.

II. TECHNIQUES FOR LOWERING THE POWER LEAKAGE

2.1 Dynamic Power Suppression

Dynamic/Switching power happens in view of charging and discharging of load capacitors of a circuit [1-2]. Scaling of supply voltage has been the most embraced approach for power optimization, since it yields broad power speculation stores in view of the quadratic dependence of trading/dynamic power on supply voltage V_{DD}. The real disadvantage of this approach is that cutting down the supply voltage impacts circuit speed. So both design or outline and innovative or technology arrangements must be connected to repay the lessening in circuit execution exhibited by reduced voltage. A portion of the systems regularly used to lessen dynamic power are portrayed underneath.

2.2 Adiabatic Circuits

In adiabatic circuits as opposed to dissipating the power, a similar power is reused. By remotely control the shape and length of signal transition power consumed by flip a bit can be diminished or reduce to little values [2-3]. Diodes are not utilized as a part of outline of adiabatic circuit since diodes are thermodynamically irreversible. MOSFETs ought not to be turned ON and OFF when there is noteworthy potential contrast amongst source and drain when there is a huge current passing through the device separately.

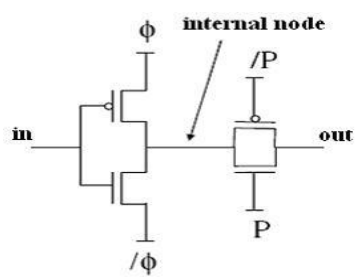


Figure 1 Charge Recovery Logic [3]

In the above adiabatic circuit at first, ø and /ø at $V_{dd/2,}$ P at Gnd, and/P at Vdd. On substantial or valid input and by step by step swinging P and/P, the pass gate is turned on. Rails ø and /ø "split", step by step or gradually swinging to V_{dd} and Gnd. When the output is examined, pass gates are off. Whenever ø and /ø bit by bit swing to $V_{dd/2}$ then the interior node is re-established or restore. Once the device is ON, the vitality move happens in a controlled way so that there is no potential drop over the gadget or a circuit.

2.3 Logic Design for Low Power

During the outline of circuit, different decisions are made, for example, Choices between static versus dynamic topologies, standard CMOS versus pass-transistor rationale styles and synchronous versus Asynchronous planning styles [3-4] . In static CMOS circuits, about the 10% of the aggregate power utilization is because of short out current. In unique circuits there is no such issue, in light of the fact that there is no any prompt dc path from supply voltage. i.e. vdd to ground. There is such a way, exit just in domino-logic circuits so as to decrease sharing; henceforth there is a less measure of short out power dissipation. We can also use the pass transistor logic to exploit reduced swing to lower power.

$$P = CL*Vdd*(Vdd-Vt) \qquad (1)$$

2.4 Reducing Glitches

At the point when two parallel driving common gate at various times then Glitches happen in a logic chain. The output changes to incorrect for a short minute before settling to right outcome. Consider circuit appeared as underneath [4-5]. Give us a chance to expect at first that without buffer, path A is fast and Path B is moderate. So if A=1 and B=1 then Z=0.Next if B is to change to 0 and A to 0 since B is moderate the information 0 touching base at B will be direct and along these lines Z switches towards 1 promptly before changing back to 0 realizing power dissipation.

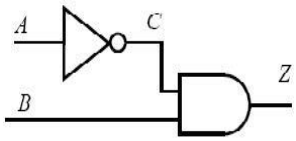


Figure 2 CMOS AND circuit [3]

2.5 Logic level power optimization

During logic optimization for low power, technology parameters, for example, supply voltage are settled or fixed and the degrees of flexibility are in choosing the usefulness and estimating the gates. Path equalization with buffer insertion is a strategy which guarantees that flag signal from inputs to output of a logic network follows ways of comparative length to overcome glitches [3-5]. So when ways are leveled or equalized a large portion of the gates have adjusted moves at their information sources, which limiting the exchanging action or glitches. Another logic level power minimization systems which incorporate neighborhood changes is appeared in figure beneath .In this re-mapping change a high-movement node which is set apart with set apart with x is evacuated and supplanted by new mapping.

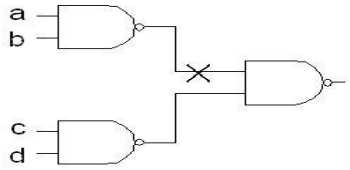


Figure 3 Logic Remapping for Low Power [3]

2.6 Standby Mode Leakage Suppression

Substrate current and sub threshold leakage produces the static and leakage control. For technology of 1 µm or more from 1um, switching force was dominating. Yet, for deep submicron forms underneath 180nm, Leakage control ends up predominant element [6]. Since Leakage control impacts battery lifetime subsequently it is a noteworthy worry in later technology. At the point when transistors are either not exchanging or in remain by mode then CMOS technology has been to a great degree control effective , and is normal by the framework designer ,the low leakage from CMOS chips. To meet our necessities identified with leakage control constraints, numerous edge and variable limit circuit procedures are utilized [5-6]. In various edge CMOS, the procedure gives two different sort of limit transistors.

(1) Low-edge transistors are utilized on speed fundamental sub-circuits and they are fast.
(2) High-edge transistors are slower yet make low sub-edge spillage, and they are used as a piece of noncritical or direct methods for the chip.

As more transistors progress toward becoming planning basic then various limit strategies prompts lose adequacy or effectiveness.

2.7 Variable Body Biasing

Variable body biasing circuits control the limit voltage of transistors utilizing substrate biasing and in this manner deficiency related is overcome with multi threshold design or outline [8-10]. At the point when a variable-limit of circuit is in standby mode, then NMOS transistors substrate is negative biased, and the threshold increments because of the body bias impact. In this manner similarly, PMOS transistors substrate is positive body bias to expand their V tie the threshold in standby mode in a principal, Variable threshold circuits can, take care of the static leakage issue, yet they require controlling circuits that regulate substrate voltage in standby mode. Quick and accurate body bias control with control circuit is a major challenge and it requires carefully designed closed-loop control. At the state when the circuit is idle mode, the bulk or body of both PMOS and NMOS are biased by supply voltage to build the limit voltage, V_t of the MOSFET. However, in typical operation they are changed back to lessen the V_t.

2.8 Dynamic Threshold MOS

In dynamic threshold CMOS i.e. DTMOS, the limit voltage is changed to suit the working condition of the circuit. A high threshold voltage in the standby mode gives the low leakage current, while a low limit voltage in the dynamic mode takes into account higher current drives [7-8]. Dynamic threshold CMOS can be obtained by tying the gate alongside the body. The diode build in potential in mass silicon technology is utilized to restrict the supply voltage of DTMOS. The PN diode amongst source and body ought to be turn reversed biased. Hence, this method is reasonable for ultralow voltage circuits in mass CMOS going from 0.6V and underneath.

2.9 Short Circuit Power Suppression

At the point when sets of PMOS/NMOS transistors are directing at the same time it cause the short circuit current which prompts to short circuit power [8-10]. The short out way exists for direct current spill out of v_{dd} to gnd. One approach to diminish short circuit power is to keep the fall time and rise time same. On the off chance that $V_{dd} < V_{tn} + |V_{tp}|$ then short out power can be dispensed with.

III. TECHNIQUES FOR THE SIMULATION

3.1 Sleepy keeper approach

In sleepy keeper approach, a NMOS transistor having high threshold value limit is associated in parallel with the PMOS transistor to which sleep signal is given and a PMOS transistors having high threshold value limit is associated in parallel with the NMOS to which sleep bar signal is given are associated in a pull down network [6-8].The current situation with the circuit can be held in rest mode and furthermore decreases sub edge leakage control essentially .Result of this paper is that the tired attendant is the best way to deal with limit the static power in computerized VLSI circuit.

3.2 Isolated Sleepy Keeper Approach

This approach includes the techniques as we discussed previously except one that is isolated sleepy keeper technique [6-8]. It states that the sleepy keeper approach is good in getting the output in sleep state but if there is any disturbance in sleep state from input then there will be voltage swing at output therefore to isolate the output from the network in such a way that no disturbance should be transferred from base network or from input to the output during the sleep mode, isolated sleepy keeper approach is introduced. The pass transistor can pass only 1or 0. Therefore instead of pass transistor, transmission gate is used which can pass both 1 and 0 in an efficient manner in active modes is 'set' to 0 and 'not s' is set to 1 then whatever will be result it pass to the output .in sleep mode's' is set to 1 and 'not s' is set to 0 then both transistors of transmission gate are off and output is isolated from the base network.

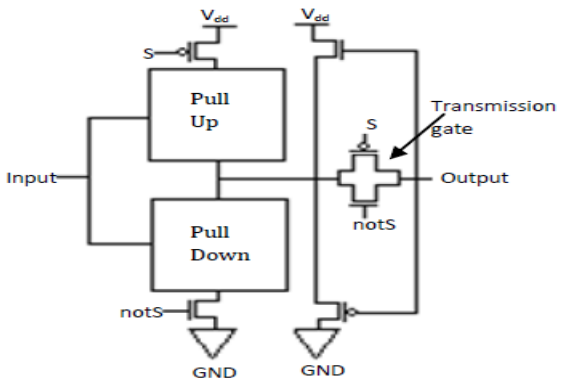


Figure 4 Isolated sleepy keeper network [6]

IV. RESULTS

All the designs are simulated in the spectre circuit simulator in cadence. It creates the netlist for DC, AC or transient analysis. Low power unit empowers groups with restricted involvement in actualizing propelled low-control procedures which help the group to embrace them in their outline with essentially hazard is decreased. It enhances the efficiency noteworthy through provided framework. Through the use of demonstrated procedures the timetable consistency re-establishes or re construct. It helps in keeps away from normal issues in low-control plan through practices and master learning. Low control kit advances flows and trade off to guarantee that the utilized technology is giving best outcome or not which may prompt enhance nature of silicon. It also reduces overall packaging and system cost.

Sleep keeper approach and isolated sleep keeper approach have been applied to inverter and full adder to analyze the reduced power. Tabular comparative analysis is also done for the considered low power techniques with base design and sleepy keeper approach .So inverter and full adder circuits with their results are considered in this section.

4.1 Inverter

Inverter is used to invert the input. It is basically a 'NOT' gate

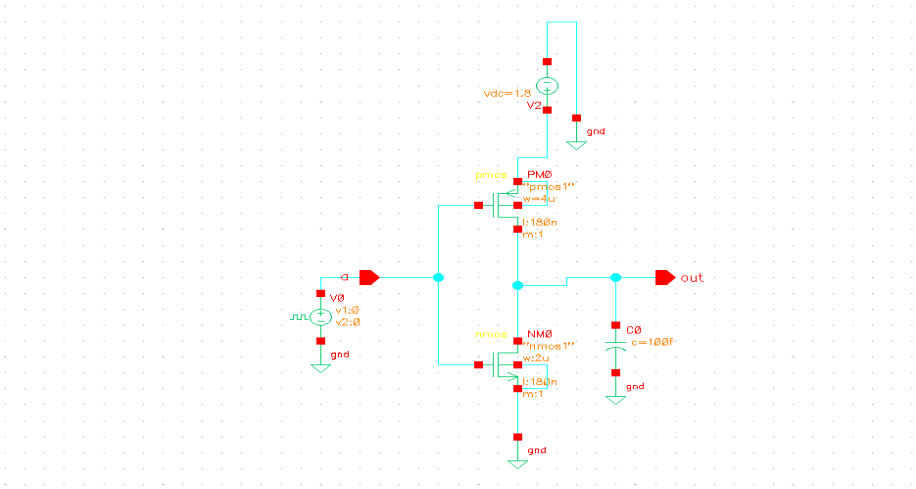


Figure 5 Circuit diagram of inverter

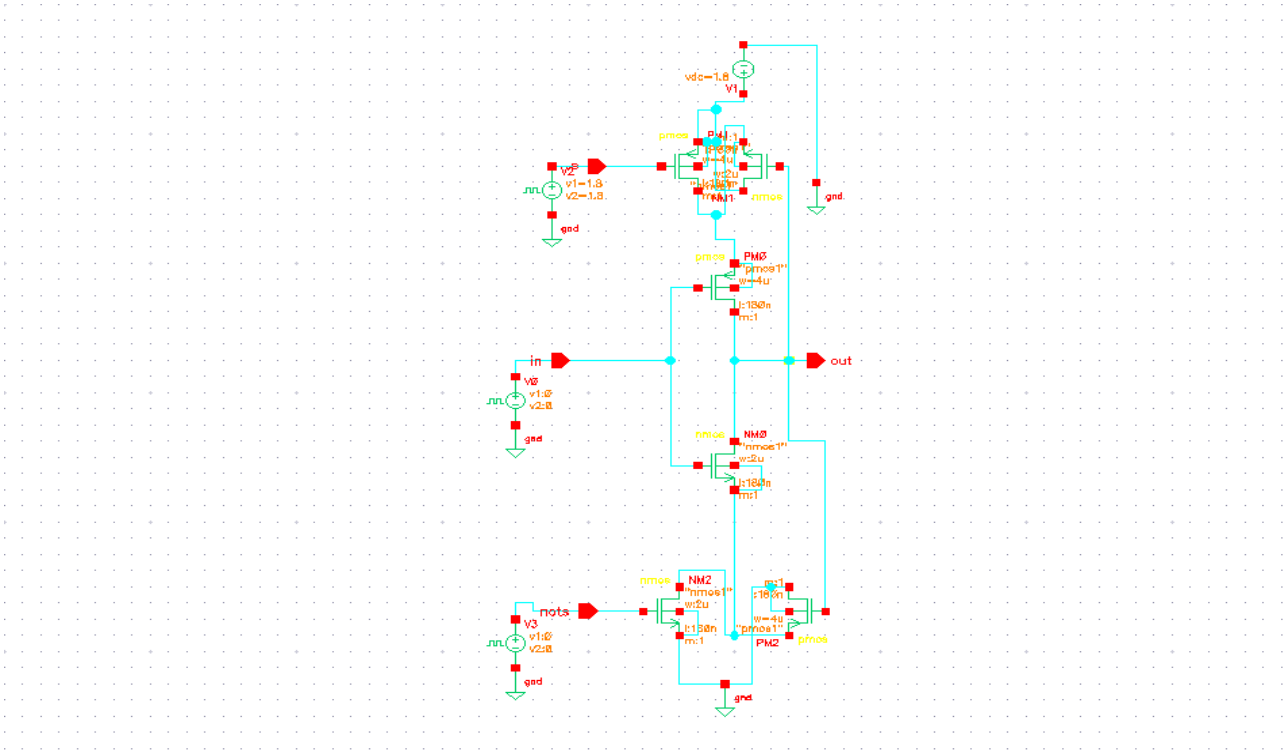


Figure 6 Circuit diagram of inverter using sleepy keeper approach

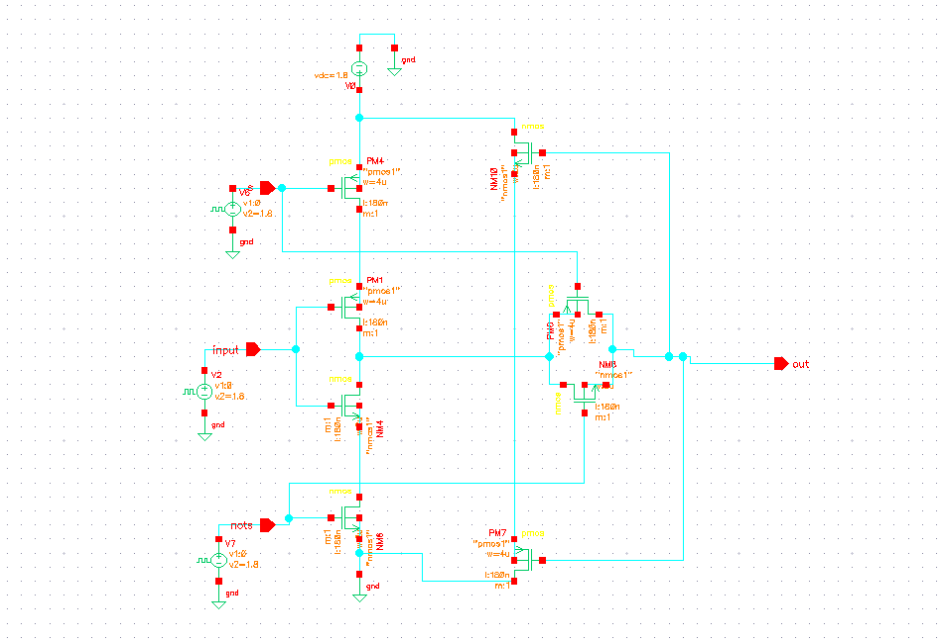


Figure 7 Circuit diagram of an Inverter using isolated sleepy keeper approach

Table 1 Comparison of power of an inverter with different approach

NETWORKS	TOTAL POWER	STATIC POWER	DYNAMIC POWER
BASE INVERTER	7.53E-06	2.73E-11	7.52E-06
INVERTER WITH SLEEPY KEEPER APPROACH	1.0E-6	1.41E-11	1.00E-06
INVERTER WITH ISOLATED SLEEPY KEEPER APPROACH	7.46E-07	2.04E-11	7.47E-07

From table 1, it is clear that isolated sleepy keeper approach is much effective for the total power dissipation in the simulated inverter. In the static power, the change is not significant, but for the dynamic power a significant change has been observed.

4.2 Full Adder

Full adder adds the binary numbers .For example the for one bit full adder it has three input say a, b, cin and it provide output sum and carry .

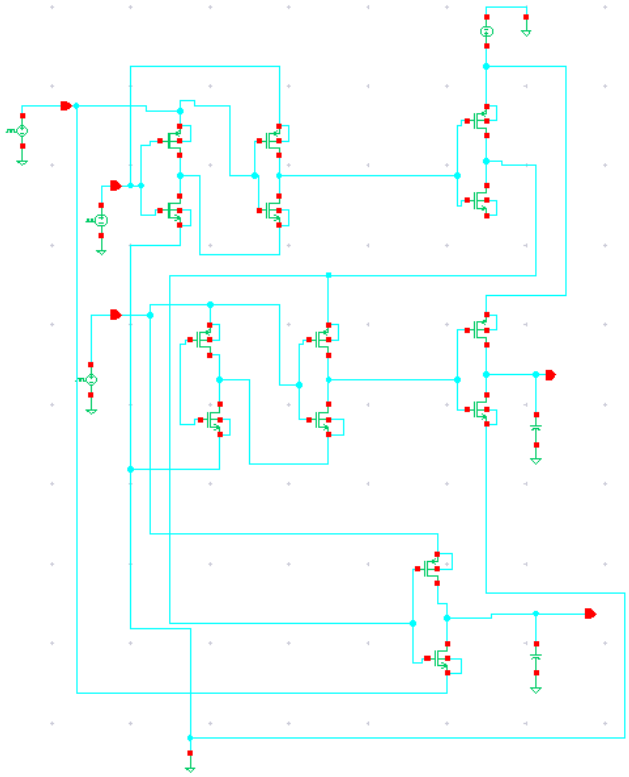


Figure 8 Circuit diagram of full adder

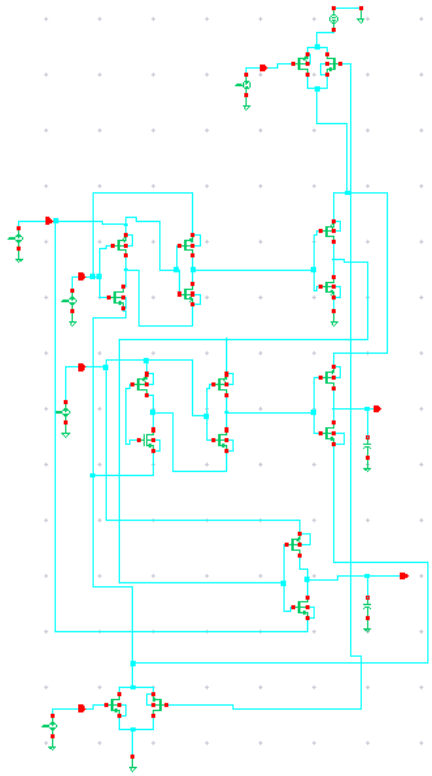


Figure 9 Circuit diagram of full adder using sleepy keeper approach

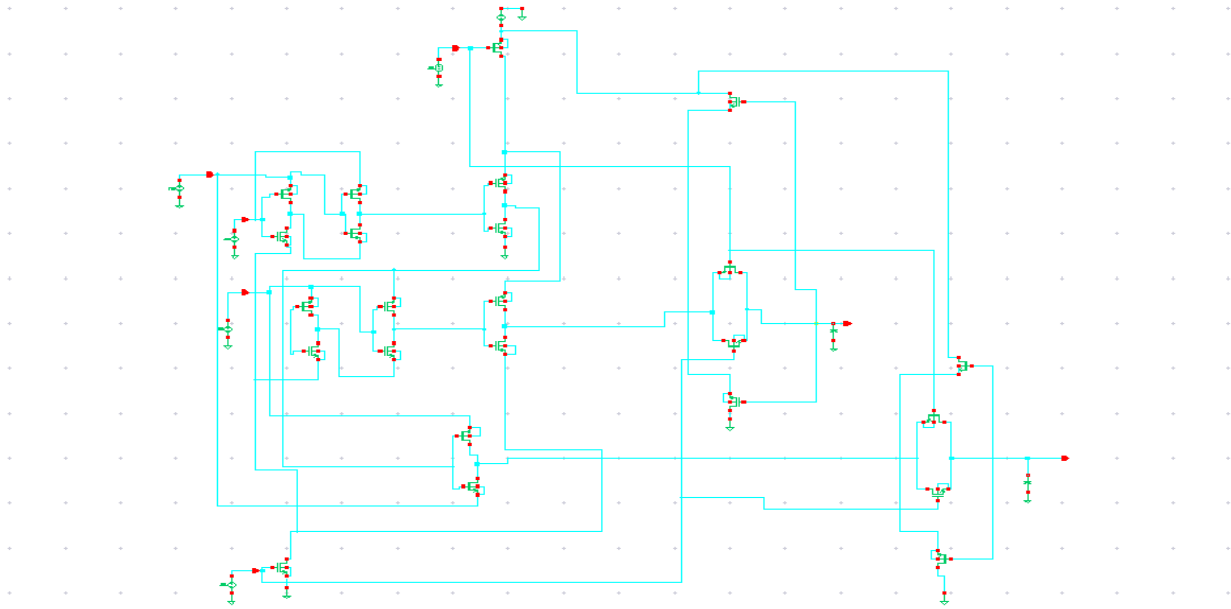


Figure 10 Circuit diagram of full adder using isolated sleepy keeper approach

Table 2 Comparison of power of full adder with different approaches

NETWORK	TOTAL POWER	STATIC POWER	DYNAMIC POWER
FULL ADDER	2.07E-07	7.12E-11	2.07E-07
FULL ADDER WITH SLEEPY KEEPER APPROACH	5.92E-07	9.32E-12	5.92E-07
FULL ADDER WITH ISOLATED SLEEPY KEEPER APPROACH	5.71E-08	1.86E-11	5.71E-08

From table 2, it is clear that isolated sleepy keeper approach is much effective for the total power dissipation in the simulated full adder. In the static power, the change is not significant, but for the dynamic power a significant change has been observed.

V. CONCLUSION

Approaches like sleepy keeper technique, isolated sleepy keeper technique are discussed to reduce power and their simulation is done on cadence low power kit to get results. These approaches are implemented on circuits of Inverter and Full Adder. With the approach of isolated sleepy keeper technique, a significant improvement has been observed in the power dissipation. Adders are the integral part for high speed multipliers and for other arithmetic operations and a significant power reduction in one cell can contributes to high power reduction in the complete digital integrated circuits. This research work further can be expanded to apply this approach on memory designs and advanced high speed adders.

ACKNOWLEDGMENT

I would like to thank the university for providing me the opportunity to work on the cadence tool and carrying this research work for paper publication.

REFERENCES

[1] K. Zhang, U. Bhattacharya, Z. Chen, F. Hamzaoglu, D. Murry, N. Vallepalli, Y.Wang, B. Zheng, and M. Bohr, "SRAM design on 65nm CMOS technology with dynamic sleep transistor for leakage reduction," IEEE J. Solid-State Circuits, vol. 40, no. 4, pp. 895–901, Apr. 2005.

[2] S. Kim and V. Mooney, "The Sleepy Keeper Approach: Methodolgy, Layout and Power Results for a 4 bit Adder," Technical Report GITCERCS- 06-03, Georgia Institute of Technology, March 2006,

[3] Harmander Singh, Kanak Agarwal, Dennis Sylvester, and Kevin J. Nowka, "Enhanced Leakage Reduction Techniques Using Intermediate Strength Power Gating," IEEE Trans. on Very Large Scale Integration (VLSI) Systems, vol. 15, no. 11, pp.12-15, November 2007.

[4] Pardeep Kumar, Susmita Mishra, Amrita Singh, "Analysis of Various Full-Adder Circuits in Cadence", International Journal of Engineering Research and Applications (IJERA) ISSN 2248-9622, Vol. 3, Issue 3, May-Jun 2013, pp.509-513.

[5] Deshpande Akshay, Sanidhya Mohan Sharma, Lochan Anil Vyas and K. Sivasankaran, "Design of Low Power and Area Efficient 4-bit Arithmetic and Logic Unit using Nanoscale FinFET", Indian Journal of Science and Technology, Vol 8(S2), 250-256, January 2015.

[6] Sainiranjan Muchakayala, Owais Shah, "Isolated Sleepy Keeper Approach: An Effective Sleep State Approach in Low leakage Power, VLSI Design", International Research Journal of Engineering and Technology (IRJET), Volume 3, Issue 5, May-2016.

[7] Rohith Nama, Shoban Mude, "Sleep Methods-An Efficient Way to Reduce the Leakage Power", Int. Journal of Engineering Research and Applications, ISSN 2248-9622, Vol. 4, Issue 1 (Version 2), January 2014, pp.321-325.

[8] K. Gnana Deepika, K. Mariya Priyadarshini and K. David Solomon Raj, "Sleepy Keeper Approach for Power Performance Tuning in VLSI Design", International Journal of Electronics and Communication Engineering, ISSN 0974-2166 Volume 6, Number 1 (2013), pp. 17-28.

[9] Md. Asif Jahangir Chowdhury, Md. Shahriar Rizwan, and M. S. Islam, "An Efficient VLSI Design Approach to Reduce Static Power using Variable Body Biasing", World Academy of Science, Engineering and Technology, International Journal of Electrical, Computer, Energetic, Electronic and Communication Engineering Volume 6, No.4, 2012.

[10] Yingtao Jiang, Abdulkarim Al-Sheraidah, Yuke Wang, Edwin Sha, and Jin-Gyun Chung, "A Novel Multiplexer-Based Low-Power Full Adder", IEEE Transactions on circuits and systems—II: express briefs, Volume 51, No. 7, July 2004, 345.

Life Span Studies on Functionally Graded Composite Coatings

Abhinav [*], Rahul Ribeiro [**], Abhijeet Hugar [**]

[*] Department of Mechanical Engineering, Alliance College of Engineering and Design, Alliance University
[**] Department of Mechanical Engineering, Alliance College of Engineering and Design, Alliance University

Abstract- The present paper examines the failure mechanism of the coating system of Al_2O_3-ZrO_2 ·5CaO applied on a Cast iron (CI) substrate. Atmospheric plasma spray coating technique was used for coating applications. Muffle furnace was used to heat the sample at 600±2ºC followed by ambient cooling. Both the heating and cooling cycle was maintained for 30 minutes. Results were compared with the as-sprayed coated specimen with the post thermal cyclic test specimen. It has been noticed that normal stresses developed due to the formation of thermally grown oxide (TGO) at the interface of the top/bond coat and institute the weakest link in the coating system. Also due to differential porosity at the interface led the moisture to penetrate from the top coat to bottom coat thus thermochemical effect commenced underneath the bond coat, leading to the formation of oxides and made the top coat gradually more brittle during the thermal cyclic process.

Index Terms- Muffle furnace, Al_2O_3-ZrO_2 ·5CaO, thermochemical, thermally grown oxide.

I. INTRODUCTION

Thermal barrier coatings (TBCs) often used to protect the components subjected to high-temperature applications viz. gas turbine engines, internal combustion parts, etc. [1]. In the previous studies it has been found that yttria stabilized zirconia failed due to the phase transformation from tetra, t-ZrO_2 to monoclinic, m-ZrO_2 [2]. Attention is required to improve the thermal cyclic behavior by appropriate material composition. Fracture response and thermal resistance of yttria-stabilized zirconia (YSZ)–NiCoCrAlY bond coat (BC) under thermomechanical loads were investigated by Rangaraj and Kokini, and it has been reported that thermal shock resistance is low in brittle materials like ceramic [3]. It has been found that failure mainly governed by the thermos-mechanic failures, chemical failures, erosion failures, oxidation of bond coat, hot erosion effect, CMAS (CaO-MgO-Al_2O_3-SiO_2) attack, inconsistency in the thermal expansion, and due to changes in the thermal conductivity, etc. [4-9]. Thermal shock resistance changes with many properties such as fracture toughness, elastic modulus, Poisson's ratio, thermal expansion coefficient, and thermal conductivity. It has been investigated that thermal stresses that occur due to the temperature difference at the interface of the bond coat and surface of the specimen when cooled with water [10]. Few basic properties such as toughness, low thermal conductivity, stability at high temperature, high thermal expansion coefficient and low elastic modulus values are required to resist the thermal shock failure [11, 12]. Ceramic materials have the high-temperature resistant capacity so that it can be used in TBC systems to reduce the thermal shock failures. It satisfies basic properties such as toughness, low thermal conductivity, the coefficient of thermal expansion and low elastic module. Studies found that elastic modulus value changes with the closing or growth of cracks and plays a crucial role in the determination of TBC life. If crack growth increase, the value of the elastic module will change and that will hamper the TBC life period under service condition [13-15].The failure of the coatings depends on the cracks, show when it propagates and results in failure. [16-18].

In the present investigation, an effort has been made to investigate the causes and mechanism behind the spallation of composite coating Al_2O_3-$ZrO_2 \cdot 5CaO$ applied to cast iron (CI) substrate of the particular thickness. The muffle furnace was used to heat and cool the coated sample uniformly during the thermal cyclic test. It is believed that the above studies are valuable for material scientists for the design of thermal-shock-resistant materials.

II. EXPERIMENTAL METHODOLOGY

2.1.Selection of material and powder trade names

Cast iron (CI) was selected as a substrate material. The selection of powder was based on the thermal coefficient of expansion provided by Sulzer Metco. The trade names of different powders are shown in Table 1.

Table 1: Trade name and composition of the powder

Trade Name	Composition by wt.%
Metco 105 SFP	99.9% Al_2O_3
Metco 201 NS	ZrO_2.5CaO
Metco 452	Fe38Ni10Al

2.2. Coating Methodology

The Atmospheric plasma spray technique was used to coat on the cast iron substrate. Before the coating process, the mixture of Al_2O_3 and ZrO_2.5CaO in 50:50 was prepared using ball mill technique. The substrates were chemically cleaned using tetra chloride-ethylene followed by preheating it to the temperature of 250±50 ºC. This process was done to minimize the thermal mismatch between the substrate and the bond coat. The schematic of a coating system is shown in Fig.1.The top coat thicknesses were about 100µm. The plasma spray machine specification and spray parameters for bond coat and top coat are given in Table 2 and 3 respectively.

Top Coat (100 µm) Pure Ceramic (ZrO_2.5CaO+Al_2O_3)
Bond Coat (50 µm) (Metallic Powder) Fe38Ni10Al
Substrate (Gray Cast Iron)

Fig.1 The schematic of coating system applied on CI substrate

Table 2: Air Plasma machine Specification

Specifications	**Parameters**
Plasma gun	3 Nylon Brush
Nozzle temperature	10,000 ˚C
Current	500 amps
Voltage	65-70 volts
Powder feed	45-50 gms/mint
Spray distance	50 -78 mm

Table 3 Plasma spray parameters for different coating materials.

Materials	Primary gas (Argon) Pressure(Bar)	Secondary gas (Hydrogen) Pressure (Bar)	Carrier gas Argon flow (lpm)	Current (amps)	Voltage (volts)	Spray distance (mm)
Al_2O_3+ ZrO_2.5CaO	3.7	3.45	35	500	65	65-76
Fe38Ni10Al	6.9	3.30	35	500	65	50-76

2.3. Thermal cyclic test procedure

The thermal cyclic test was conducted on a Muffle furnace. The schematic of the test setup shown in Fig.2. The temperature of the furnace was maintained at a temperature of 600±2ºC. The coated substrate was kept on silica ceramic cups to ensure uniform heating all-around refer Fig.2. The heating and cooling cycle time of 30 minutes was maintained throughout the experiment. The cooling was done under ambient conditions.

Fig. 2 Muffle furnace test setup

2.4 Coating characterization

Evaluation of the coating thickness, surface morphology, and determination of elemental composition was carried out using Zeiss Evo 18 special edition machine, and the machine specifications are given in Table 4.XRD analysis was carried out on Bruker, and its specifications are given in Table 5.

Table 4: Machine specifications of Zeiss Evo 18

Filament	Tungsten
Secondary e-image resolution	50 NM
Tilt	0 - 60 Degree
Rotation	360 Degree
EHT	200V - 30KV
Magnification	Up to 50K ~ 100K (Depends on sample)

Table 5: X-ray Diffractometer specifications

Parameters	
Make	Bruker
Model	D8 Advance
Measuring circle diameter	435,500,600
Smallest addressable increment	0.0001°
Reproducibility	0.0001°
Anode	Cu
Detector	Scintillation & Lynxeye

III. RESULTS AND DISCUSSIONS

Morphology of the as-sprayed and after thermal cyclic test shown in Fig. 3(a) and (b) respectively. The composite mixture applied as the top coat was found to be homogeneous, localized agglomeration, uneven microvoids, and localized pin holes. At higher magnification of 1500X, noticeable hair cracks were observed in the case of post-thermal cyclic test sample [Fig.3 b]. The cracks are formed in the alloyed mixture, mainly at the interfaces of the adjoining splats of the two different phases, alumina, and calcia stabilized zirconia, and hair crack impressions acknowledged mostly on the grains of Zirconia phase. The coefficient of thermal expansion value for alumina and zirconia dioxide are,9.6×10-6 K-1& 15.3×10-6 K-1respectively[19]. The nucleation of these crack often attributed to the inconsistency in thermal coefficient of expansion between the two phases [20]. Also, due to the repeated thermal cyclic loading, i.e., heating and cooling residual compressive stresses institute in the composite mixture [21]. It is also evident from the micrograph that the network of cracks is more or less same in both the phases, Zirconia phase/grains (white color) and Alumina phase (dark gray color) identified as shown in Fig.3 b. The bulk stresses present in the bond coat decide the life of the coating, and the residual stresses present at the interface may interact with the micro defects and promote crack growth [22].

It is also evident that the top coat delaminated from the bond coat. The bond coat didn't spall from the substrate, and excellent metallurgical bonding understood. The reason of spallation of the top coat, Al_2O_3-ZrO_2.5CaO from the bond coat, Fe38Ni10Al attributed to the thermal inconsistency at the interface and effect amplified during prolonged heating and cooling cycle. It is also understood that the weakest link formed at the top/bond coat interface. Other reason for failure at the interface attributed to the differential porosity difference between the interfaces and roughness. The roughness at the interfaces act as stress concentrators and to a certain extent decides the lifespan of the coatings [23]. The average porosity for as-sprayed coating in case of the top coat and bottom coat found to be 1.90 & 2.75%, [Fig 4.c & d] respectively. The life of the coatings also found to depend on the porosity, horizontal and vertical cracks at the interface of the top/bond coat.LU et al. [24] had found that the presence of porosity is generally beneficial if a pre-existing crack dominates the failure. However, in the present work due to the porosity difference at the interface and the moisture present in the atmosphere leads to oxidation and can be visualized as dark black spots at the top coat [Fig.5 e]. EDX analysis has also confirmed the oxidation phenomenon. The elemental abundance of the individual elements, traces of iron content 1.51 wt. % along with oxygenated element 50.85 wt. % has been confirmed after thermal cyclic test refers to Fig. 6 h. No traces of any iron element found in case of as-sprayed coatings refer Fig.6 g. The oxidation at the top acknowledges due to the thermochemical reaction between the iron (Fe) elements present in the bond coat with the moisture present when exposed to ambient temperature during the thermal cyclic test [25]. The byproduct, rust formed at the surface can also be understood with the help of thermochemical, redox reaction given in Equation 1.

$$Fe^{2+} + O_2 + \Delta \rightarrow 4\,Fe^{3+} + 2\,O^{2-} \text{----------Eq.1}$$

During the test, it was noticed that the thermally grown oxide layer (TGO) developed at the junction of the bond coat and top coat resulting weakening of the top coat/bond coat interface [Fig.8] The growth of the TGO layer found an increase with an increase in time. At the end of 312 cycles, the top coat starts delaminated from the edges shown in Fig.6 b. The failure of the coating in the present work is in agreement with results obtained by Julian D. Osorio [26].

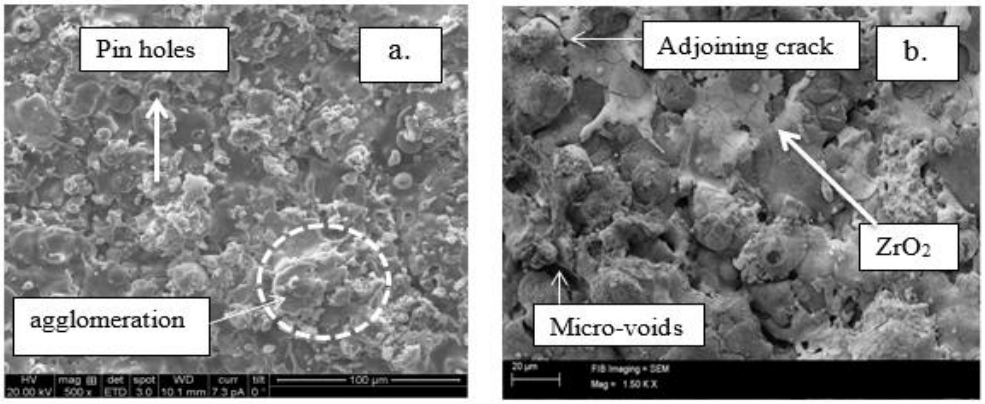


Fig.3 Morphology of the top coat (a) As-sprayed coating (b) After thermal cyclic test

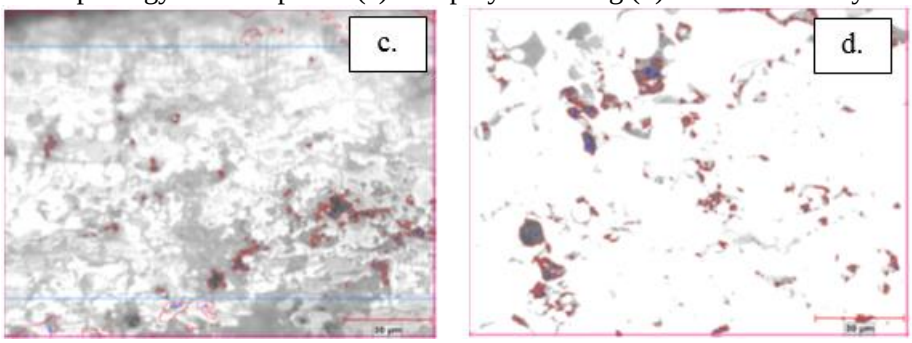


Fig.4 Micrograph of as-sprayed average porosity (c) 1.90 % topcoat and (d) 2.75% bond coat

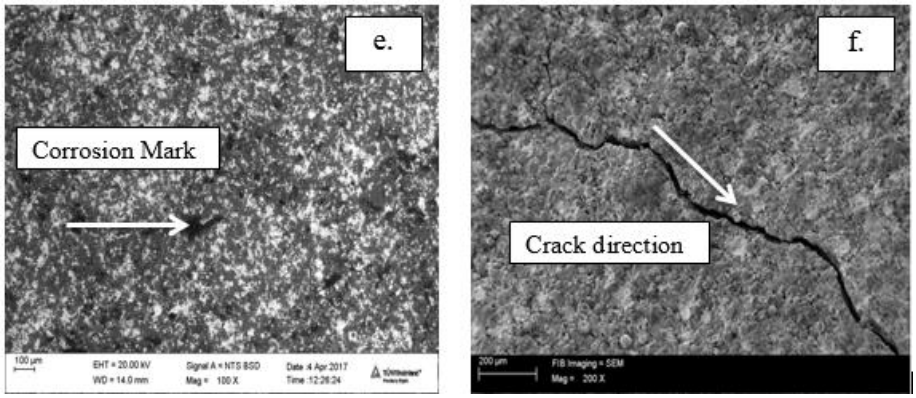
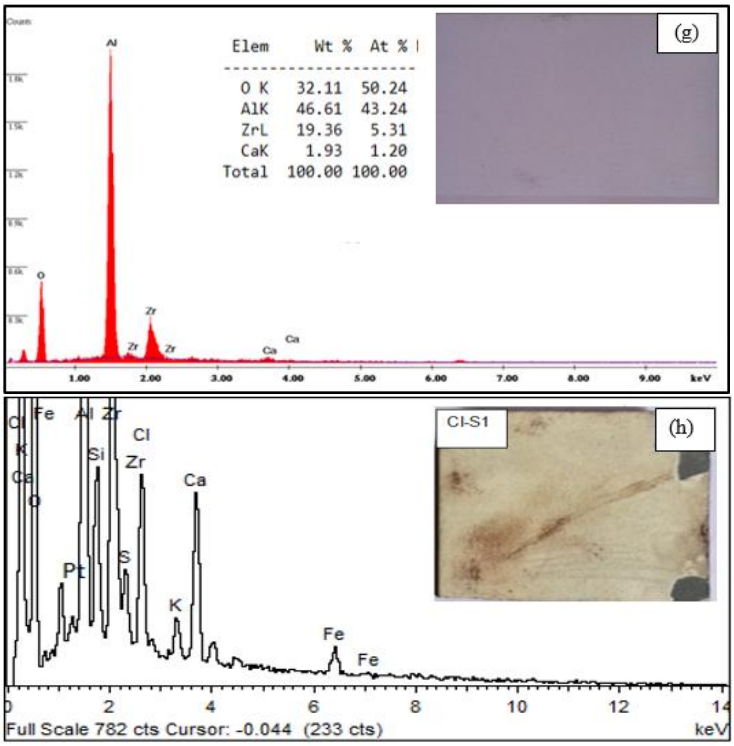


Fig.5 Micrograph of the topcoat post thermal cyclic test e. oxidation status after 295 cycle f. crack propagation after 312 thermal cycles test.



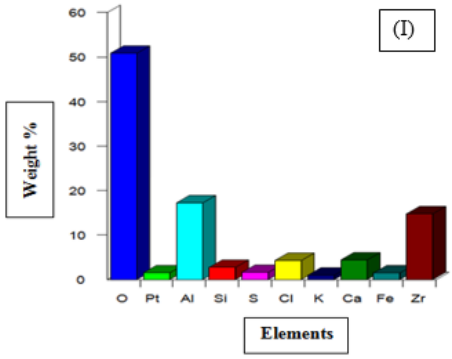


Fig.6 EDX and elemental composition of the topcoat (g): As-sprayed coatings (h)& (i): after thermal cyclic test.

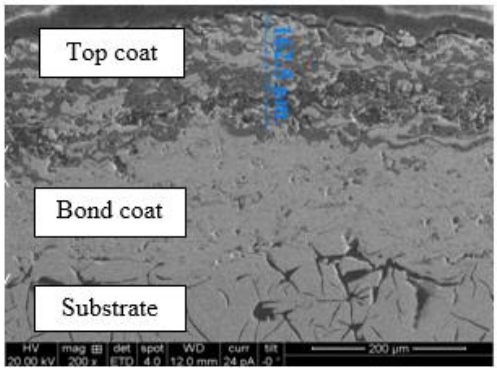


Fig.7 SEM micrograph (200X) at the cross section with no TGO layer between top coat/bond coat.

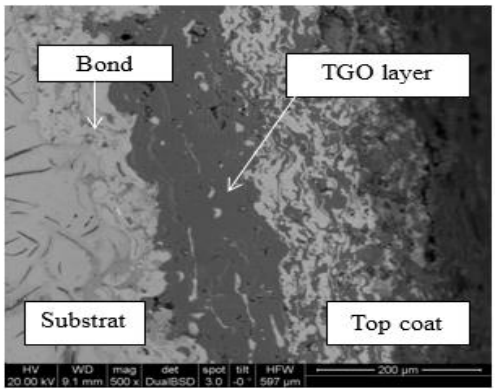


Fig. 8 SEM micrograph (500X) at the cross section with TGO layer between top coat/bond coat.

IV. CONCLUSION

The thermochemical effect found to play a crucial role and is having a detrimental impact on the life of the coating systems. No failure of the coating observed during constant heating as long as the specimen remained in the Muffle furnace. Bond coat shows excellent metallurgical bonding with the substrate. During prolonged heating and cooling cycle it is speculated that TGO, a brittle byproduct formed at the bond (Fe38Ni10Al) /top coat (Al_2O_3-ZrO_2.5CaO) interfaces. Thermochemical and differential porosity at the interface found to develop the weakest link in the coating system also found governing mechanism behind spallation of the top coat. After 312 cycles the visible crack observed and considered as the failure of the top coat.

REFERENCES

[1] Julian D. Osorio et al., " *Thermal barrier coatings for gas turbine applications failure mechanisms and key microstructural features*," Medellin, December,:ResearchGate, 2012, pp. 149-158.

[2] Saket Jaiswal, *"Synthesis of 3 Mol% Yttria Stabilized Zirconia (3YSZ) and Study on Corrosion Behavior of CMAS on Sintered 3YSZ Pellets,"* Bachelor of Technology in Ceramic Engineering, Department of Ceramic Engineering National Institute of Technology, Rourkela, 2015.

[3] Sudarshan Rangaraj et al.,*"Fracture in single layer (YSZ) – Bond Coat Alloy (NiCoCr10Al) composite coating under thermal shock,"*ScienceDirect,: Vol 52, Issue 2,2004pp. 455-465.

[4] Liu J., *"Effects of bond coat surface preparation on thermal cycling lifetime and failure characteristic of thermal barrier coatings,"* Master of Science Thesis, Department of Mechanical, Materials, and Aerospace Engineering, University of Central Florida,Summer Term, 2004.

[5] Verbeek A.T.J., *"Plasma sprayed thermal barrier coatings: Production, characterization and testing,"* PhD Thesis, Technisch Universiteit Eindhoven,1992.

[6] Aygun A., *"Novel Thermal Barrier Coatings (TBCs) that are resistant to high Temperature attack by glassy deposits,"* The Ohio State University, ResearchGate, 2008.

[7] Evans A.G., et al., *"Scaling laws governing the erosion and impact resistance of thermal barrier*

[8] *coatings,"* Wear, 2006, pp.886–894.

[9] Wellman R.G., et al., *"Erosion, corrosion and erosion–corrosion of EB PVD thermal barrier coatings,"* Tribology International, 2008, vol.41 (7), 657-662.

[10] Mao W.G., et al., *"Modeling of residual stresses variation With thermal cycling in thermal barrier coatings,"* ResearchGate, 2006, vol.38, pp.1118-1127.

[11] Kara,I, E., *"TBK Kaplamaların Termal Şok Ozelliklerinin Incelenmesi,"* Master Thesis,Sakarya University," Metallurgical and Materials Engineering Department,Turkish, 2008.

[12] Han J.C., et al., *"Thermal shock resistance of ceramics with temperature-Dependent material properties at elevated temperature,"* Acta Materiala, 2011, vol.59, pp.1373-1382.

[13] Girolama G.D., et al., *"Microstructure, mechanical properties and thermal shock resistance of plasma sprayed nanostructured zirconia coatings,"* Ceramics International, 2011, 37, pp.2711-2717.

[14] Siebert B., et al., *"Changes in porosity and Young's Modulus due to sintering of plasma sprayed thermal barrier coatings,"* Journal of Materials Processing Technology, 1999, 92-93, pp.217-223.

[15] Yoshiba M., et al., *"High-Temperature Oxidation and Hot Corrosion Behavior of Two Kinds of Thermal Barrier Coating Systems for Advanced Gas Turbines,"* Journal of Thermal Spray Technology, 1996, vol. 5 (No. 3), 259-268.

[16] Balint D.S., et al., *"An analytical model of rumpling in thermal barrier Coatings,"* Journal of the Mechanics and Physics of solids, 2005, vol. 53, pp.949–973.

[17] Vaben R, Kerkhoff., et al., *"Development of a micromechanical life prediction model for plasma sprayed thermal barrier coatings,"* Materials Science and Engineering A, 2001, vol.303, pp. 100–109.

[18] Yanar N.M., et al., *"Comparison of the Failures During Cyclic Oxidation of Yttria-Stabilized (7 to 8 Weight Percent) Zirconia Thermal Barrier Coatings Fabricated via Electron Beam Physical Vapor Deposition and Air Plasma Spray,"* Metallurgical and Materials Transactions A, 2011, vol. 42, pp . 905-921.

[19] Guo H., et al., *"Effects of Heat Treatment on Microstructures and Physical Properties of Segmented Thermal Barrier Coatings,"* Materials Transactions, 2005,vol. 46, pp.1775-1778.

[20] X.Q. Cao et al. *"Ceramic materials for thermal barrier coatings,"* Journal of the European Ceramic Society, vol. 24, Year 2004, pp. 1–10

[21] Koshi Takenaka., *"Negative thermal expansion materials technological key for control of thermal expansion,"* Science and Technologyof Advance Materials.,2012, pp .11

[22] D.M.Nisseley., *"Thermal Barrier Coatings Life Modeling in Aircraft Gas Turbine*

[23] *Engines,"* Jurnal of Thermal Spray Technlogy, vol.6, March 1997, pp.91-98.

[24] Xiancheng Zhang et al., *"Effect of spraying condition and material properties on the residual stress in plasma spraying,"* Science and Technologyof Advance Materials., vol.20, 2004.

[25] A.G Evans et al., *"Mechanical Behavior of Brittle Coatings on a Layer*," Oxidation Metals.,vol.20, 1983, pp. 193-216.

[26] T. J. L., et al., *"The thermal shock resistance of solids,"* Acta mater. vol. 46, No. 13, 1998 , pp. 4755-4768,

[27] W.J.Brindley., et al., *"Surf.Coating Technol,"* vol.43/44,1990,pp.446-57.

[28] Julian D. Osorio., et al., *"Thermal barrier coatings for gas turbine applications failure mechanisms and key microstructural features,"* Dyna, year 79, Nro.176, pp. 149-158. Medellin, December, 2012. ISSN 0012-7353.

Acute Lymphoblastic Leukemia diagnosis in microscopic blood smear images using Texture features and SVM classifier

S. Alagu [*], K. Bhoopathy Bagan [**]
[*] Teaching Fellow, Department of Electronics Engineering, Anna University, India
[**] Professor, Department of Electronics Engineering, Anna University, India

Abstract- Acute lymphoblastic leukemia (ALL) is the most common cancer in children and adults. Leukemia produces a large number of immature blood cells in the bone marrow. The detection of leukemia in the earlier stage is important before it spreads into the blood streams and other vital organs. For decades, the diagnosis of leukemia has been done by experienced operators and it is a time consuming task for pathologists. The computer aided detection of acute lymphoblastic leukemia using supervised learning is discussed in this paper. The proposed method reduces the diagnostic time and gives better accuracy. The microscopic blood smear images from the database are preprocessed and segmented as three clusters based on shape, color and texture using k-means clustering algorithm .The texture features are extracted by grey level co-occurrence matrix (GLCM) and local binary Pattern (LBP). Support vector machine (SVM) with Gaussian radial basis function (RBF) as kernel is used for classification. The proposed methodology is tested for 367 images from ALL-IDB. The accuracy of 90.5% for ALL-IDB1 and 95.3% for ALL-IDB2 are obtained using SVM classifier and the results are compared with other standard classifiers such as Linear Discriminant (LD), Ensemble (Bagged trees) and KNN.

Index Terms- Acute lymphoblastic leukemia (ALL), Blood smear images, Acute lymphoblastic leukemia image database (ALL-IDB), Support vector machine (SVM) and K-Nearest Neighbors (KNN).

I. INTRODUCTION

Bone marrow is the soft and flexible tissue available in bone cavities which can generate millions of blood cells every day. Three types of blood cells substantially produced by the bone marrow are platelets, erythrocytes and leukocytes (white blood cells-WBC). WBC is responsible for the human immune system. Acute lymphocytic leukemia known as acute lymphoblastic leukemia is the type of cancer that occurs in white blood cells [4, 6]. Mostly children are affected by acute lymphocytic leukemia rather than adults. The most common types of leukemia are acute lymphoblastic leukemia, acute myeloid leukemia (AML), chronic lymphoblastic leukemia (CLL) and chronic myeloid leukemia (CML) [2]. Several methods have been proposed for the diagnosis of leukemia over the past ten years.

In this paper, the blood smear images of healthy persons and patients with acute lymphoblastic leukemia are obtained from the database available in the internet. The images were preprocessed and segmented using k- means clustering as three clusters based on the shape, color and texture .GLCM and local binary pattern techniques are used as texture operators. SVM classifier is adopted for classification purpose.

Acute lymphoblastic leukemia is diagnosed using texture features and SVM classifier with Gaussian radial basis function. The proposed method gives better accuracy than other standard classification

algorithms such as Linear Discriminant (LD), Ensemble (Bagged trees) and KNN and results are tabulated in the section IV.

The remaining of the paper is organized as follows. Related works are presented in section II and the proposed method is introduced in section III. In section IV, Image acquisition, the segmentation and classification results are shown and the metrics used for performance evaluation are also discussed. Finally, the conclusion and perspectives on future works are given in section V.

II. RELATED WORKS

Several methods have been used for the diagnosis of leukemia over recent years and some of the works related to the detection of leukemia cells are discussed in brief in this section.

AimiSalihah et al [1] proposed colour image enhancement techniques and morphological features for leukemia cell detection. Contrast stretching, Bright stretching and Dark stretching are used to identify the blast cell as either Acute lymphoblastic leukemia cell or Acute Myeloid leukemia cell. Histogram technique is proposed for validating the results. The contrast stretching gives more accurate results than other methods. In [6], Himali Vaghela et al recommended histogram equalization and linear contrast stretching methods to detect the leukemia cells either as acute or chronic. Watershed transform method was adopted for nucleus segmentation to detect cancer cells. K- means clustering and shape based features are proposed to identify the blast cells with the accuracy of 72% and 73.5% respectively.

Luis Vogado et al [9] proposed leukemia cells segmentation based on the multi-space color channel. K-means clustering and morphological operations were used for cancer cell detection. The results are validated by different performance measures of accuracy and kappa index with the normalized values of 0.912 and 0.93 respectively. In [10], Silva and Kelson Aires proposed method for classification of leukemia cells using convolutional neural network. Alexnet, Caffenet and Vgg-f networks architectures are developed and images features are extracted according to the gain ratios for further classification. Transfer learning is adopted for the classification of images with different characteristics obtained from the different image databases.

Preeti Jagadev and Virani [14] proposed the method for the detection of leukemia and its types by using image processing techniques. Marker controlled water shed algorithm, k means clustering algorithm and HSV color based algorithms are used segmentation. SVM classifier is adopted for classification of different types of leukemia cells. Ruggero Donida Labati et al [15] discussed the details about acute lymphoblastic leukemia image data base where samples have been collected by experts of Tettamanti research Centre for childhood leukemia and hematological diseases, Monza, Italy and the morphological operations were used for identifying leukemia cells.

Subrajeet Mohapatra and Dipti Patra [16] proposed a method for leukemia detection using Hausdorff dimension and contour signature in blood microscopic images. K-means clustering is used to separate white blood cells from platelets and erythrocytes. Shape and texture features are used to detect the leukemia cells. Fractal features i.e. Hausdorff dimension is implemented to estimate the perimeter roughness and to classify lymphocytic cell nucleus. The extracted features are given as input for SVM classifier. The classification accuracy of 90% is obtained.

In [18], Shailesh Mishra and Deshmukh recommended the method for detecting leukemia using morphological features and image processing techniques. Dilation and border extraction are carried out to identify the blast cells by using morphological features. Thanh et al [19] proposed a system for identifying leukemia cells using convolutional neural network for clinical decision support system. Convolutional neural network with two convolutional layers of size 50x50x30, one max pooling layer of 25x25x30, one fully connected layer and one soft max layer is recommended and obtained the accuracy of 96.03%.

In [20], Worawut Srisukkham et al suggested an intelligent leukemia diagnosis with Bare-Bones PSO based Feature optimization technique. Marker –controlled watershed algorithm is adopted for lymphocytic membrane identification. Separation of nucleus and cytoplasm is achieved by stimulating discriminant measure (SDM) - based clustering algorithm. Initially 80 raw features are obtained. Feature optimization techniques are adapted to exhibit the optimized features. SVM classifier is used to classify lymphocytes and lymphoblasts using the identified optimized feature subsets.

III. PROPOSED METHODOLOGY

The proposed method for the automatic detection of acute lymphoblastic leukemia in microscopic blood smear images consists of four steps such as image pre-processing, image segmentation, feature extraction and classification. Fig.1 shows the system architecture for efficient detection and classification of acute lymphoblastic leukemia.

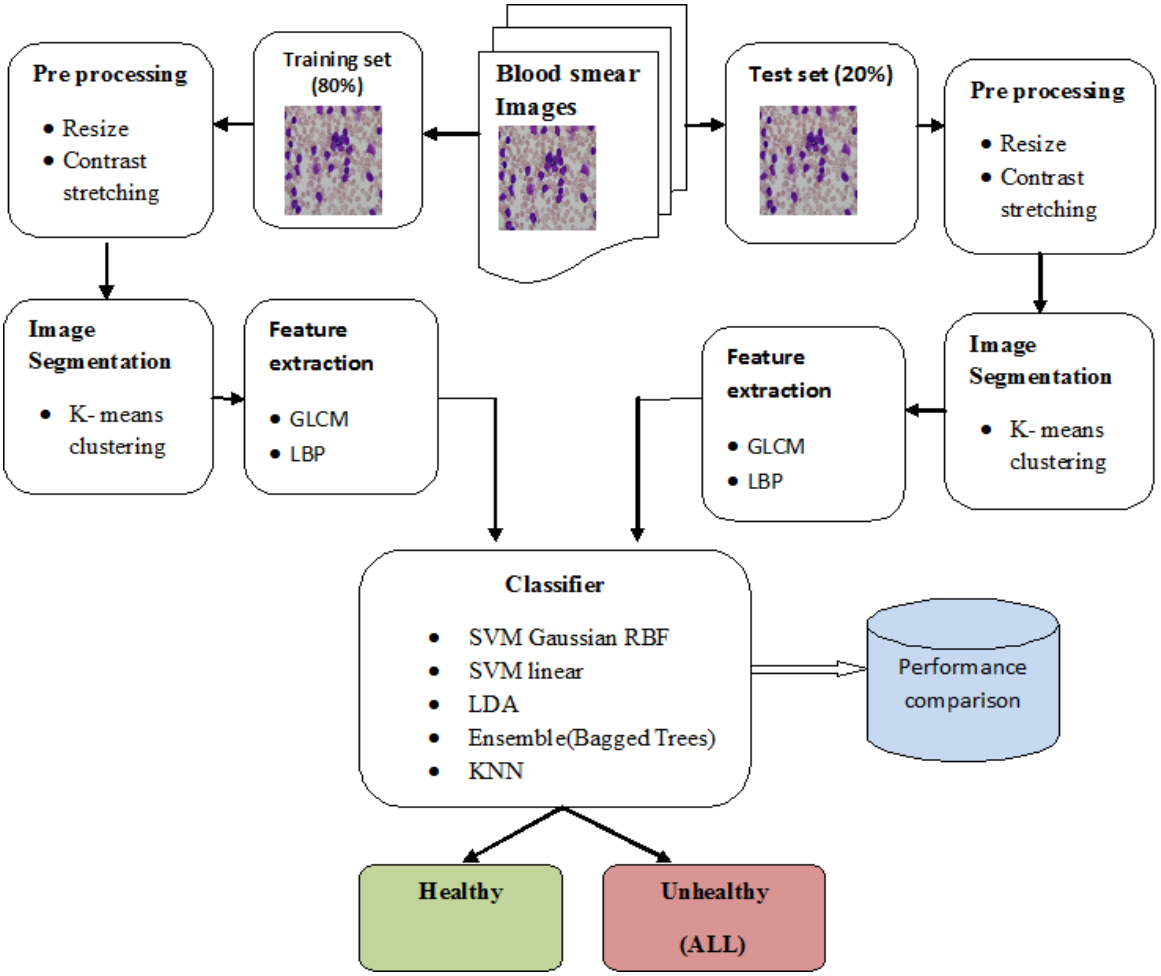


Fig.1: Proposed system architecture

A. Image Preprocessing

Image preprocessing is applied to the raw image in order to enhance the image features [1].Image resizing and contrast stretching are the two important preprocessing techniques adopted in the proposed work. The input images are resized to 256x256 sizes and the quality of the image is enhanced by contrast stretching technique.

B. Image segmentation

Image segmentation algorithms are applied to partition an image as regions which have more similarity based on some predefined criterion [3]. Region splitting and merging, region growing, thresholding and clustering are some of the techniques used for image segmentation [8]. K-means algorithm is adopted in the proposed work which is an iterative technique used to partition an image into K clusters. The selection of cluster centers are done by either randomly or based on some heuristic approach.

Each pixel in the image is assigned to the cluster that minimizes the distance between the pixel and the cluster centre. The cluster centers are re-computed by averaging all the pixels in the cluster. The steps are repeated until convergence is attained, that is until no pixels change clusters.

K-means clustering algorithm is proposed for image segmentation. The steps involved in the algorithm are given as steps. First, the microscopic blood images are read from database. The color image is

converted from RGB to L*a*b space. Colors are classified using k-Means clustering in a*b* space. Each pixel in the image is labeled from the results of k- means. The value of k is taken as three to obtain three clusters based on the shape, colour and texture.

C. Feature Extraction

The performance of a classifier depends on the features extracted. The features must be less sensitive to any actions in the image such as zooming the image, scaling and changing the orientation. The purpose of feature extraction is to represent the most relevant and important information from the image and present them in the lower dimensionality space (2). The prominent features extracted include shape, texture and color. In the proposed work, texture features are extracted from segmented image and used for classification.

D. Gray level co-occurrence matrix

A statistical method of examining texture that considers the spatial relationship of pixels is the gray-level co-occurrence matrix. The GLCM is used to characterize the texture of an image by calculating a spatial relationship between pairs of pixel with specific values as matrix.

E. Local Binary Pattern

Local Binary Pattern is a texture operator used to label the pixels of an image by thresholding the neighborhood of each pixel and considers the result as a binary number. Due to its discriminative power and computational simplicity, LBP texture operator has become a popular approach in various applications. The assigned label values are used as input for classification process.

F. Classification

Classification maps the pixels of an image into a particular class based upon the extracted features. The classification approach is implemented in two phases, i.e. training phase and testing phase [14]. The training data composed of features extracted from the images using gray level co-occurrence matrix and local binary pattern techniques. SVM is used for classification. The classification is performed by finding a hyper-plane that differentiates input data into two classes. Nonlinear transformation is also possible in SVM by adopting kernel trick. Linear, quadratic, polynomial, Gaussian radial basis function kernels are available in SVM classifier. The default kernel function is linear. SVM classifier with Gaussian radial basis function is adopted in the proposed work to obtain better classification accuracy.

G. Evaluation

The performance of the classifier is measured using various parameters such as accuracy, sensitivity, specificity and precision. A confusion matrix is generally used to describe the performance of the classifier [10]. It clearly shows the number of instances correctly classified, the number of instances which have been wrongly classified. If Class1 is a positive class and Class2 is a negative class, then the classifier's confusion matrix could be represented as in table 1. In table 1, TP stands for true positives i.e., the number of positive instances correctly classified; FP stands for false positives i.e., the number of negative instances classified as positive by the classifier. FN stands for false negatives i.e., the number of positive instances classified as negative; TN stands for true negatives i.e., the number of negative instances classified as negative.

	Predicted class 1	**Predicted class 2**
Action class 1	TP	TN
Action class 2	FP	FN

Table 1: Confusion Matrix Structure

Accuracy is the most common measure to check the performance of the classifier. Accuracy is the ratio of the number of correctly classified images to that of the total number of images.

$$\text{Accuracy} = (TP + TN)/ (TP + FP + FN + TN) \quad (1)$$

Sensitivity is also known as the true positive rate or recall. It states the rate at which the positive instances are correctly classified.

$$\text{Sensitivity} = TP / (TP + FN) \quad (2)$$

Specificity is known as the true negative rate. This parameter states the rate at which the negative class label are misclassified as correct label values.

$$\text{Specificity} = TN / (FP + TN) \quad (3)$$

Precision is the ratio of correctly classified positive instances to that of the total number of instances that have been classified as positive by the classifier.

$$\text{Precision} = TP / (TP + FP) \quad (4)$$

F-measure is known as F1 score or F score. F-measure is the harmonic mean of precision and sensitivity.

$$\text{F-measure} = 2*TP / (2*TP + FP + FN) \quad (5)$$

IV. EXPERIMENTAL RESULTS AND DISCUSSION

A. Image Acquisition

Leukemia is a blood cancer. The diagnosis of acute lymphoblastic leukemia is carried out in the proposed work. Totally 367 images of the blood smears of leukemia patients and non- leukemia patients obtained from Acute Lymphoblastic Leukemia-Image Database (ALL-IDB).The details of the ALL-IDB is shown in the table 2.

Images	**ALL-IDB1**			**ALL-IDB2**		
	HEALTHY	BLAST	**TOTAL**	HEALTHY	BLAST	**TOTAL**
	92	15	107	130	130	260

Table 2: ALL-IDB Data set details

ALL-IDB1 and ALL-IDB2 are the two image sets available in ALL-IDB. ALL-IDB2 images have similar grey level properties to the images of ALL-IDB1. The ALL-IBD2 is a collection of cropped area of interest of normal and blast cells that belongs to the ALL-IDB1 dataset. The description of ALL-IDB1 is as follows. The ALL-IDB1 image files are named with the notation ‘Im XXX_Y.jpg” were XXX is a 3- digit Integer counter and Y is a Boolean digit equal to 0 if no blast cells are present, and equal to 1 if at least one blast cells is present in the images. The image with labelled with Y=0 are from healthy persons,

and all images with label Y=1 are from acute lymphoblastic leukemia patients. Blood smear images of healthy and leukemia patients were obtained from acute lymphoblastic leukemia-Image Database (ALL-IDB). ALL- IDB1 contains 92 healthy blood cells samples and 15 blast cells whereas in ALLIDB2, healthy and leukemia cells images are 130 each. The normal and leukemia blast cells are shown in figure 2.

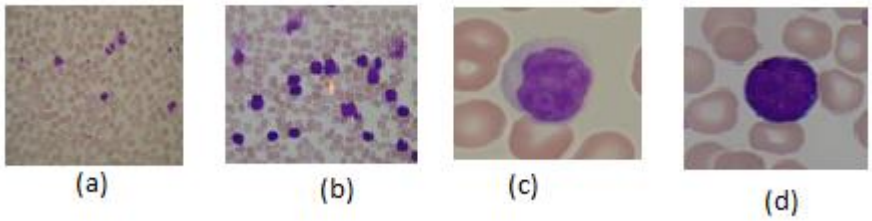


Fig. 2: Blood smear images (a) Healthy cells_ALL-IDB1 (b) Blast cells_ALL-IDB1 (c) Healthy cell_ALL-IDB2 (d) Blast cell_ALL-IDB2

ALL-IDB1 dataset images with healthy cells and blast cells are shown figures 2. (a) and (b). ALL-IDB2 dataset consists of single cell images. Figure 2. (c) Shows the healthy cell consists of nucleus and cytoplasm as separable one. In figure 2. (d), cytoplasm area are completely covered by nucleus indicates that the cell is affected by an acute lymphoblastic leukemia.

B. Image Pre-processing results

The blood smear image quality is improved by adopting image preprocessing techniques. Image resizing and contrast stretching are used in the proposed work. The images are resized to 256x256 sizes. Contrast stretching is done to preserve the edges of the blood cells as shown in figure 3.

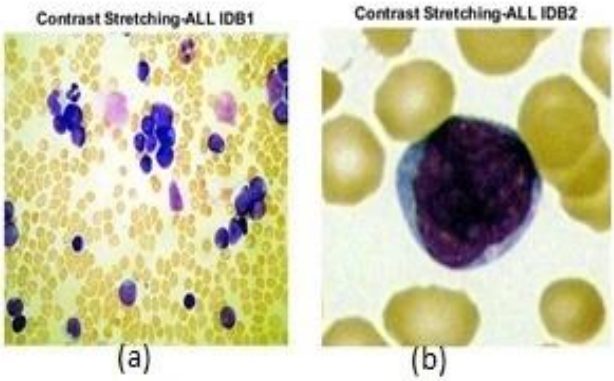


Fig. 3: Contrast stretching (a) ALL-IDB1 image (b) ALL-IDB2 image

C. Image Segmentation outputs and Feature extraction

The image is segmented after preprocessing by using k- means clustering algorithm with a cluster size of 3.Based on the colour, texture and shape the image pixels are grouped as three clusters as shown in figure 4. The texture features only separated by using the texture operators such as gray level co-occurrence matrix and local binary pattern techniques. The extracted texture features are used for further classification.

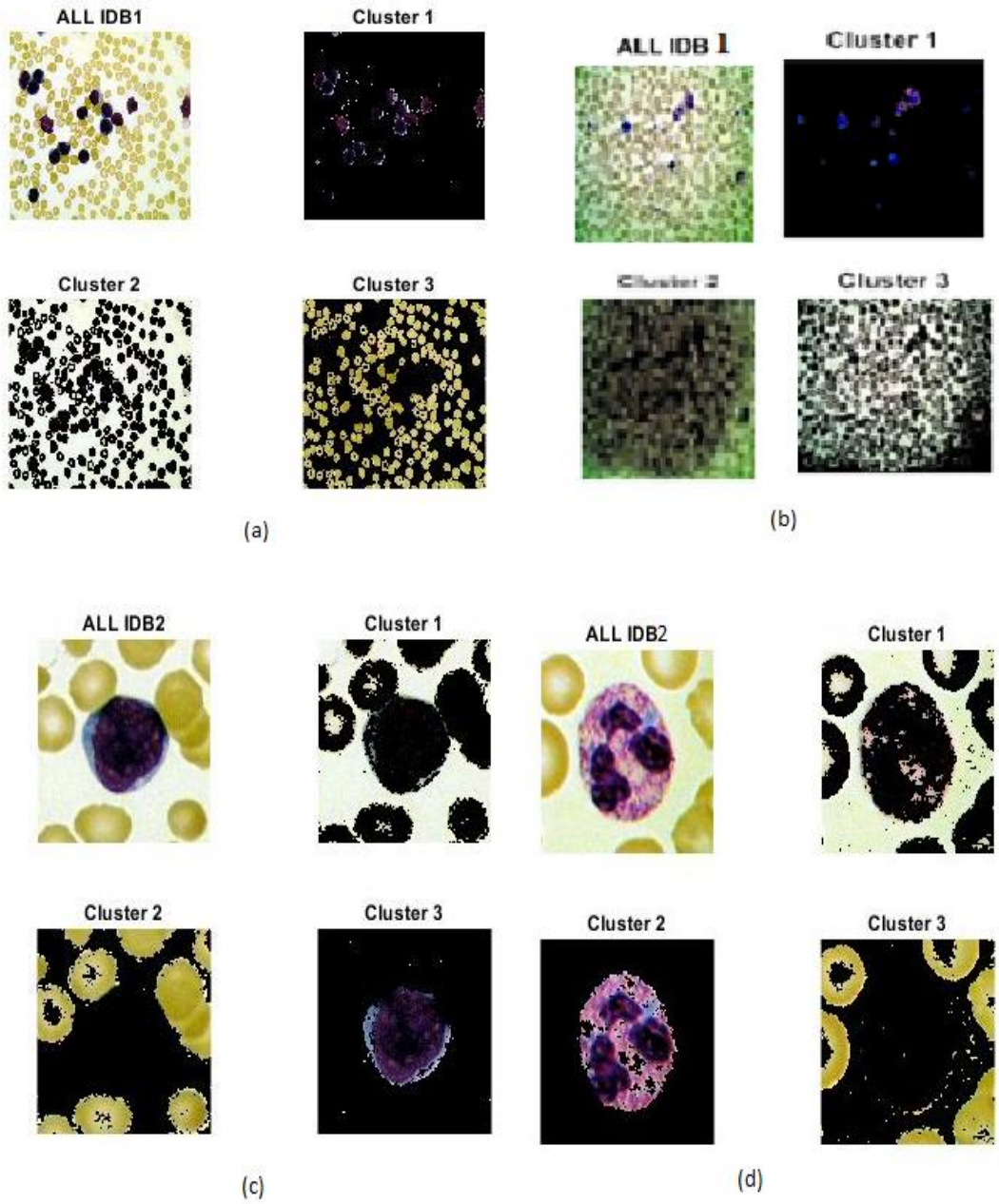


Fig. 4: Segmentation of blood smear image using K means clustering

(a) ALL-IDB1 healthy cells clusters (b) ALL-IDB1blastcells clusters

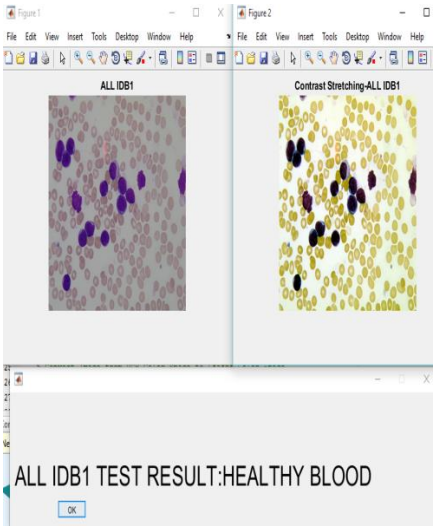
(c) ALL-IDB2 healthy cell clusters (d) ALL-IDB2 blast cell clusters

D. SVM classification results

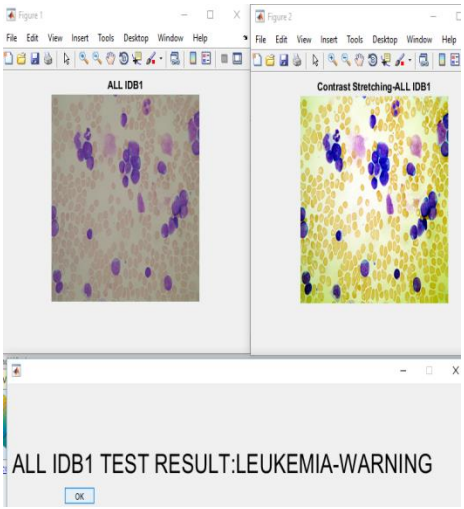
Binary SVM is a classifier which discriminates data points into two categories. Each data point is represented by an n-dimensional vector. Maximum separation between the two classes is achieved by selecting the hyper-plane with the largest margin. The margin is the summation of the shortest distance from the separating hyper-plane to the nearest data point of both categories.

The hyper-plane correctly classifies testing data points. SVM does the mapping from input space to feature space in order to support nonlinear classification problems. The kernel trick is helpful to make a linear classification in the feature space in to nonlinear classification in the input space.

In the proposed work, SVM with Gaussian radial basis function (RBF) is employed for classification. SVM classifier defines an optimal hyper-plane that separates the data into two different classes. The SVM classification output of the test image is either as healthy blood cell or leukemia blast cell. The classification results either diseased or normal. The results are displayed with the help of MATLAB GUI commands as shown in figure 5.

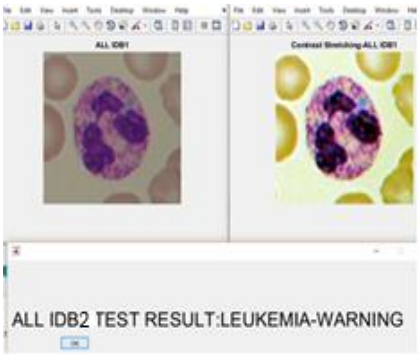


(a)



(b)

(c)



(d)

Fig. 5: SVM classifier with Gaussian radial basis function outputs

(a) Healthy person–ALLIDB1 image (b) Leukemia patient- ALLIDB1 image

(c) Healthy person–ALLIDB2 image (d) Leukemia patient- ALLIDB2 image

E. Metrics for performance evaluation

The accuracy, precision, specificity, sensitivity and F-measure of the proposed method are given in the table 3. SVM classifier with Gaussian radial function as a kernel gives the accuracy of 90.5% for ALL IDB1 and 95.30 for ALL IDB2.This is comparatively higher than SVM with linear kernel. The results are obtained by generating confusion matrix in Matlab version 2018a.The error rate is very low and the value of sensitivity is a measure that indicate the number of times the classification perfectly done. The specificity and F-measure are also calculated for the proposed SVM classifier with Gaussian radial basis function.

Metrics for evaluation	**ALL IDB1**	**ALL IDB2**
Accuracy	90.5%	95.30%
Sensitivity	87.83%	94.59%
Specificity	92.06%	95.23%
F-measure	87.25%	93.33%

Table 3: Evaluation metrics

The measured accuracy is compared with other existing classifiers such as Linear Discriminant (LD), Ensemble (Bagged trees) and KNN classifiers as shown in figure 6.The SVM classifier with Gaussian Radial Basis Function gives the accuracy of 90.5% for ALL-IDB1 and 95.3% for ALL-IDB2 .

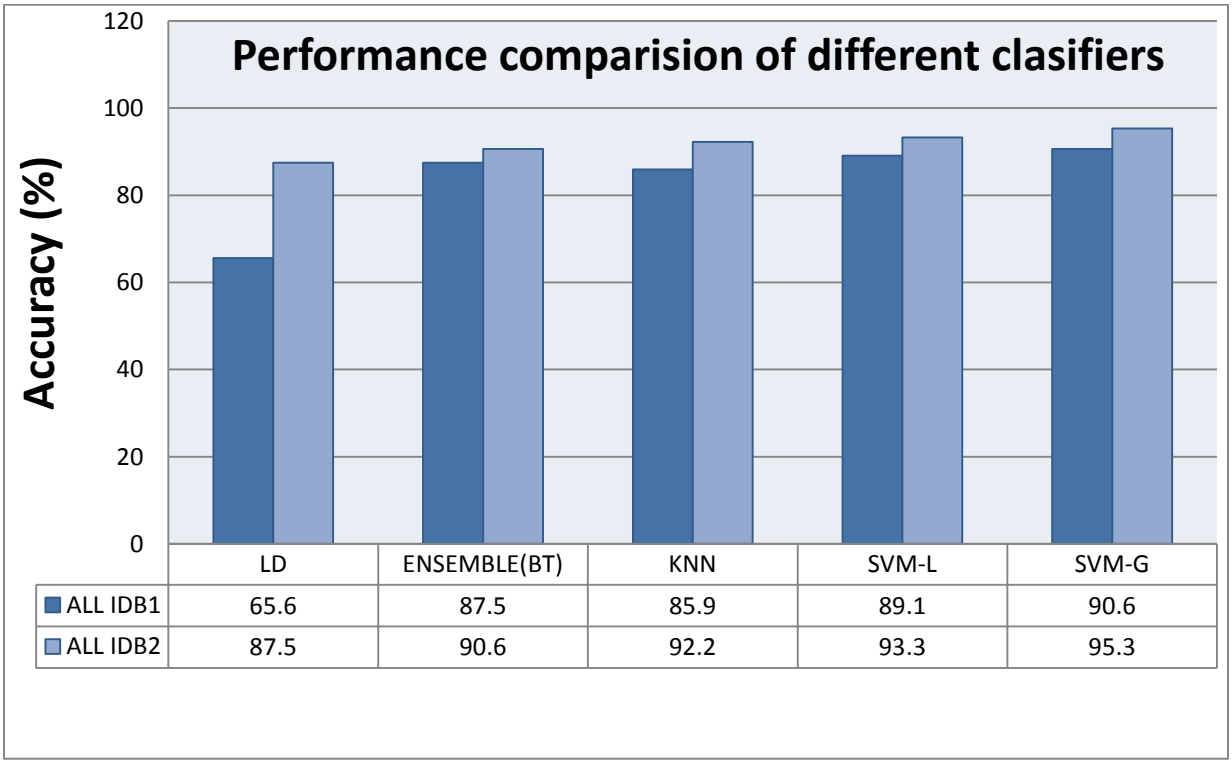


Fig. 6: Performance comparison of different classifiers

V. Conclusion

The diagnosis of acute lymphoblastic leukemia using texture features and SVM classifier is carried out and results are discussed in this paper. The proposed methodology is tested for 367 images from ALL-IDB. The accuracy of 90.5% forALL-IDB1 and 95.3% for ALL-IDB2 are obtained using SVM classifier with Gaussian radial basis function and the results are compared with other standard classifiers such as Linear Discriminant (LD), Ensemble (Bagged trees) and KNN. Deep learning algorithms, convolutional neural network (CNN) architectures will be used in future to ensure the unsupervised learning mechanisms and to obtain better accuracy. The proposed system will be validated with large amount of data and used in daily life, helping physicians and patients to diagnose the disease at the earliest.

ACKNOWLEDGMENT

The author would like to thank Labati et al for sharing the Acute lymphoblastic leukemia image database (ALL-IDB).

REFERENCES

[1] AimiSalihah, A.N, M. Y. Mashor, and Nor HazlynaHarun “Colour Image Enhancement Techniques for Acute Leukemia Blood Cell Morphological Features,” IEEE EMBS Conference on Biomedical Engineering and Sciences (IECBES), Volume 10 – No.1, 246-251, 2010.

[2] Ben Athiwaratkun and Keegan Kang,” Feature Representation in Convolutional Neural Networks”, arXiv: 1507.02313v1 [cs.CV] 8 Jul 2015.

[3] EmadA.Mohammed, Mostafa M.A.Mohamed, Christopher Nauglerand, C. P. Bean, Behrouz H. Far ,’Chronic lymphocytic leukemia cell segmentation from microscopic blood images using watershed algorithm and optimum thresholding“, IEEE Canadian Conference on Electrical and Computer Engineering (CCECE).,2013.

[4] Florence Yellin, Benjamin D. Haeffele, and Ren´eVidal,”Blood cell detection and counting in holographic lens free imaging by convolutional parse dictionary learning and coding”, IEEE,2017

[5] Geert Litjens, ThijsKooi , Babak Ehteshami Bejnordi ,Arnaud Arindra Adiyoso Setio , Francesco Ciompi, Mohsen Ghafoorian, JeroenA.W.M. van der Laak, Bram van Ginneken, and Clara I. Sánchez,” A survey on deep learning in medical image analysis”, Elsevier, Medical Image Analysis 42 (2017)60–88.

[6] Himali P. Vaghela, Hardik Mod, Manoj Pandya and M.B. Potdar “Leukemia Detection using Digital Image Processing Techniques,” International Journal of Applied Information Systems, Volume 10 – No.1, November2015.

[7] J. M. H. Du Buf, M. Kardan and M. Spann “Texture feature performance for image segmentation,” Pattern Recognition, Vol. 23, No. 3/4, pp. 291- 309, 1990.

[8] Keh-Shih Chuang, Hong-Long Tzeng, Sharon Chen, Jay Wu and Tzong- Jer Chen “Fuzzy c-means clustering with spatial information for image segmentation,” Elsevier2005.

[9] Luis H. S. Vogado, Rodrigo de M. S. Veras, Alan R. Andrade and Flavio H. D. de Araujo, “Unsupervised Leukemia Cells Segmentation Based on Multi-space Color Channels”, IEEE International Symposium on Multimedia, 2016.

[10]R.V. Silva and Kelson R.T. Aires.” Leukemia diagnosis in blood slides using transfer learning in

CNNs and SVM for classification, Engineering Applications of Artificial Intelligence 72 (2018) 415–422.

[11]Lokman Faivdullah, FaridAzahar, ZawZawHtike, and Wei Yan NyeinNaing," Leukemia Detection from Blood Smears", Journal of Medical and Bioengineering Vol. 4, No. 6, December 2015.

[12]Nikhil R Paul and Sankar K Paul," A review on image segmentation techniques", Pattern Recognition, Vol. 26, No. 9, pp. 1277 1294. 1993

[13]Patil Tejashri G and V. B. Raskar," A Robust Segmentation Method for Acute Lymphoblastic Leukemia Detection", International Journal of Advanced Research in Computer Science and Software Engineering, Volume 5, Issue 1, January2015.

[14]Preeti Jagadev and Dr. H.G. Virani," Detection of Leukemia and its Types using Image Processing and Machine Learning", International Conference on Trends in Electronics and Informatics ICEI 2017.

[15]Ruggero Donida Labati, Vincenzo Piuri and Fabio Scotti," ALL-IDB: The acute lymphoblastic leukemia image database for image processing", 18th IEEE conference on Image processing, 2011.

[16]Subrajeet Mohapatra and Dipti Patra," Automated Leukemia Detection using Hausdorff Dimension in Blood Microscopic Images", IEEE, 2010.

[17]Sos Agaian, Monica Madhukarand Anthony T. Chronopoulos, Automated Screening System for Acute Myelogenous Leukemia Detection in Blood Microscopic Images, IEEE. Trans., 2014.

[18]Shailesh J. Mishra and A.P.Deshmukh ,"Detection of leukemia using matlab ",International Journal of Advanced Research in Electronics and Communication Engineering, Volume 4, Issue 2,February 2015.

[19]Thanh. TTP, Giao N. Pham, Jin-Hyeok Park, Kwang-Seok Moon, Suk- Hwan Lee, and Ki-Ryong Kwo, "Acute leukemia classification using convolutional neural network in clinical decision support system", pp. 49– 53, 2017. © CS & IT-CSCP 2017.

[20]Worawut Srisukkham, Li Zhang and Siew Chin Neoh, "Intelligent Leukemia Diagnosis with Bare-Bones PSO based Feature Optimization", Journal on Applied soft computing, S1568-4946(17)30148-5, 2017.

IOT VIDEO ECOSYSTEM FOR VIDEO STORAGE AND ANALYSIS

Radha R*, Dr. Selvarani R*
Computer Science and Engineering,
Alliance University, Bangalore, Karnataka, India

Abstract- Cloud computing, needs that all things be associated to the central information storage, where enormous capacities of information are processed to find optimization solutions or make business decisions. IoT video data are collected through the sensors before it could be stored into cloud/storage server. We can apply few analyses to classify data which are useful to store and which can be discarded local to IOT device. The IOT video data consolidated from the devices to the Cloud is voluminous, and the quality of results is not significantly compromised, then the Edge analytics can be performed for the aggregation/transformation process to reduce the size of data streams before sending them to Cloud. If the control action needs to be relatively real time, then the action latency and efficiency are determining metrics to co*n*sider edge analytics. Our proposed framework collects IOT video from bigger and smaller ecosystem, do edge analytics to classify the useful data and draw an automatic prediction or an insight to categorize IoT videos and give them a semantic link to IoT videos. This categorization allows coming up with new mode or cloud ecosystem (smart city cloud, smart traffic cloud, education, energy) which allows group and store specific information for specific purpose.

Index Terms- **Edge analytics, video streaming, classification, and Prediction**

I. INTRODUCTION

The internet of thing is the network of devices that contain electronics, software, sensors, actuators, and connectivity which lets these things to link, act together and exchange data. The networking capabilities and miniature size of sensors have added to the proliferation of Internet of Things (IoT) and continuous sensing environments. Video data from such sensors must be analyzed real-time with low latency, Edge computing reduces latency because data does not have to traverse over a network to a data center or cloud for further process. Edge computing allows information produced by internet of things (IoT) devices to be administered closer to where it is created instead of sending it across long routes to data centers or clouds. This is ideal for situations where delays of milliseconds can be unacceptable, such as in financial services or manufacturing. Challenges in achieving this include: high data arrival rates, buffer overflows, context-switches, and object creation overheads.

However this ‘cool’ data system with its high-tech varied sensors are nonentity without being analyzed real-time by Stream Processing. Stream processing is a technology acting while the data is being produced, enabling its applications to act-on (collect, integrate, visualize, and analyze) real-time streaming data. In other words, Stream Processing enables us the facility to quickly process large amounts of data from multiple sources, in real-time. The cutting edge

Technique in Internet of Things (IoT) applications involve video analytics—a technology that applies machine-learning algorithms to video feeds, enabling sensor cameras to recognize people, objects, and situations automatically. These applications are new, but several factors are

encouraging their growth, including the increased sophistication of analytical algorithms and lower costs for hardware, software, and storage. Live video from sensors gives many advantages in relation to other sensing modalities. Most significant is its rigidity and open-endedness: new image and video processing algorithms can be developed to enhance the data pull out from an existing video stream. Furthermore, video offers high resolution, wide exposure, and low cost relative to other sensing modalities. The inactive nature of video sensing is especially striking for public spaces. A participant does not have to attire a special device, install an app, or do anything special. He or she merely has to be noticeable to a camera.

IoT video is collected form bigger and smaller ecosystem. By collecting, analyzing, applying various methodology/statistically analysis, this research aims discard 10% of the data which is not useful to store in cloud/server. Also, an automatic prediction or an insight can be drawn to categorize IoT videos and give them a semantic link to IoT videos. This categorization allows coming up with new mode or cloud ecosystem (smart city cloud, smart traffic cloud, education, energy) which allows group and store specific information for specific purpose, examples.- All smart cities based IoT videos will be moved smart city cloud storage, which later can used to analyze indent like theft analysis, power failure, and traffic.

II. IDENTIFY, RESEARCH AND COLLECT IDEA

1.Thilina Buddhika et.al(2016): stated in there paper achieving real time stream processing in IoT and sensing environments requires a holistic framework that accounts for the CPU, memory, network, and kernel issues that arise. Efficient scheduling of workloads through the use of thread pools and minimizing context-switches by processing streams in batches reduces the number of context switches during stream processing [1]

2.R. Pereira et.al(2016): stated in their research paper, some developments related to video streaming and the specific requirements associated with Low Power Personal Area Networks, the scenario relevant to the Internet of Things. They have considered alternative versions of the H.264 encoder standard, namely the AVC and its extension SVC. Evaluation is currently being conducted to assess their relative merits. Future work will include specific forms of streaming more suitable to devices of different capability, and considerations for adaptation of SVC into the DASH framework, following completion of the evaluation as indicated above [2]

3. Byungseok Kang et.al(2017): Gateways are emerging as a key element of bringing legacy and next generation devices to the Internet of Things (IoT). They integrate protocols for networking, help manage storage and edge analytics on the data, and facilitate data flow securely between edge devices and the cloud. Current IoT gateways solve the communication gap between field control/sensor nodes and customer cloud, enabling field data to be harnessed for manufacturing process optimization, remote management, and preventive maintenance [3]

4. Mr.Prabhu R et.al(2013): The mobile phones grow to be an essential part of our everyday life, with smart phone sales at present greater than before very much and also user demands to run lots of applications have enhanced. The victory of next invention mobile phone communication based on the capability of service suppliers to engineer innovative added worth to video service [4]

5. Yantao Li et.al(2016): Wireless video streaming on smartphones drains a significantly large fraction of battery energy, which is primarily consumed by wireless network interfaces for downloading unused data and repeatedly switching radio interface. In this paper, we propose an energy-efficient download scheduling algorithm for video streaming based on an aggregate

model that utilizes user's video viewing history to predict user behavior when watching a new video, thereby minimizing wasted energy when streaming over wireless network interfaces [5]

6. Jiyan Wu et.al(2017): Delivering high-definition (HD) wireless video under stringent delay constraint is challenging with regard to the limited network resources and high transmission rate. Concurrent multipath transfer (CMT) using stream control transmission protocol (SCTP) exploits the multihoming feature of mobile devices to establish associations with different access networks. In this paper, we study the multihomed HD video communication with SCTP over heterogeneous wireless networks[6]

7. Nalini Bagal et.al(2015): Real-time audio-visual communication has become the need of this era. Video conferencing may be one solution to saving both time and money. In fact, video conferencing may be a more effective way of communicating to clients and customers. Integrated Network Systems can install video conferencing that will allow you to share anything that is on your computer monitor and meet people face to face. In last few years, work is being done for real time transmission of audio and video. This paper presents detailed study of audio and video transmission through various channels. Most popular is using wireless [7].

8. Hitendra Patil et,al(2014): The standards like IETF as well as W3C are used to define the framework, protocols, and application programming interfaces. These interfaces provide further real-time interactive voice, video, and data in web browsers as well as other applications. This is explaining how media as well as data transfer in a peer-to-peer style directly between two web browsers. It's showing the protocols handled to transport & its secure the encrypted media, traverse NATs & firewalls, negotiate media capabilities, and provide identity for the media. Web Real-Time Communication (Web RTC) is an upcoming standard that aims to enable real-time communication among Web browsers in a peer-to-peer fashion [8].

III. PROPOSED SYSTEM

Live video analysis gives many advantages compared to other sensor data. In the proposed system the videos are collected from IOT devices such as video cameras from various locations and the analytics of video also called as video streaming is performed using edge analytics rather than cloud analytics. In edge analytics, small multi-tenant data centers named cloudlets are placed close to IOT devices, these cloudlets give vital privacy to public videos. Before the videos are uploaded to the cloud eco system, the video should be analyzed, recognized and perform face detection for privacy concern. Once the bounding boxes of faces have been detected, blurring those pixels is performed in cloudlets. Figure 1 shows the overall architecture of the proposed system, where the denaturing, the term used for preserving the privacy of video is performed in cloudlets, close to IOT devices.

Using longitude and latitude coordinates, we can recognize the location of the video, which lets us understand the nature of video. For example, If the crowd is more, from a particular geographic location which is usually crowded on every working days at the busy traffic signal, will give us conclusion note that the video is from the traffic camera. Likewise taking this as one of the parameter to recognize and categorize the video, we can well detect and denature the videos from public cameras. In future keeping the longitude and latitude coordinates for recognizing the nature of video, efficient algorithm will be applied for face detection and denaturing of the data in cloudlets and uploaded to cloud ecosystem for storage and distributed accessibility.

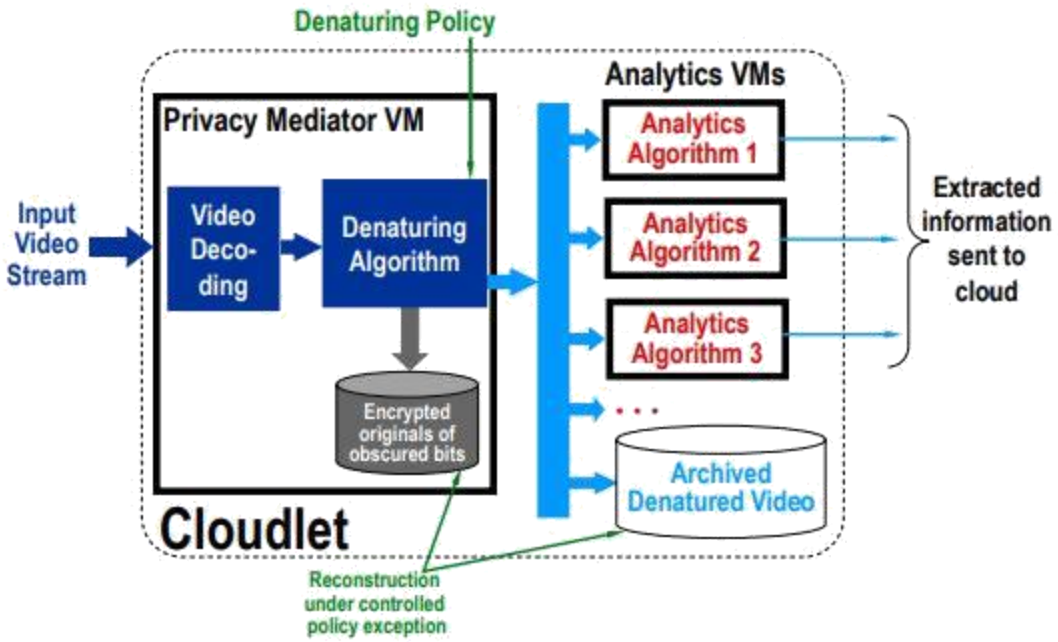


Fig 1: Architecture for live video streaming

REFERENCES

[1] Thilina BuddhikaNEPTUNE:RealTime Stream Processing forInternetofThings and Sensing Environments

[2] R. Pereira, 2015 29th International Conference on Advanced Information Networking and Applications Workshops

[3] Byungseok Kang, Daecheon Kim and Hyunseung Choo, Member, IEEE "Internet of Everything: A large-scale autonomic IoT gateway"

4] Mr.Prabhu R, Mr.Gautham K, Mrs.Nagajothi A "Adaptive Mobile Video Streaming and Efficient Social Video Sharing in Cloud"

5] Yantao Li ·Du Shen·Gang Zhou "Energy optimization for mobile video streaming via an aggregate model"

[6] Jiyan Wu, Member, IEEE, Chau Yuen, Senior Member, IEEE "Ming Wang and Junliang ChenContent-Aware Concurrent Multipath Transfer for High-Definition Video Streaming over Heterogeneous Wireless Networks"

[7] Nalini Baga, Prof. Shivani Pandita "A Review: Real-Time Wireless Audio-Video Transmission"

[8] Hitendra Patil, Amar Buchade "Review and Study of Real Time Video Collaboration Framework WEBRTC"

[9] Soumya Kanti Datta, Christian Bonnet, Navid Nikaein "An IoT Gateway Centric Architecture to Provide Novel M2M Services"

[10] Byungseok Kanga, Hyunseung Chooa "A Deep-learning-based Emergency Alert System"

[11] Monowar Hasan and Ekram Hossain, University of Manitoba "Random Access for Machine-to-Machine Communication in LTE-Advanced Networks: Issues and Approaches"

[12] Byungseok Kang1, Namyeong Kwon2, And Hyunseung Choo1, (Member, IEEE) "Developing Route Optimization-Based PMIPv6 Testbed for Reliable Packet Transmission"

[13] LilacAl-Safadi "The Effects of Real-Time Interactive Multimedia Teleradiology Syste" Aoxiang Xu Wieslaw Woszczyk "Real-Time Streaming of Multichannel Audio Data over Internet"

Siamese Triple Ranking Convolution Network in Signature Forgery Detection

Ojaswini Chhabra [*], Souradip Chakraborty [*]

[*] Data & Analytics, Walmart Labs, Bangalore

Abstract- Identifying a credible signature match based on a *base* signature of a person is an age-old problem. Despite recent automation and advances in this field using image recognition, a lot remains to be explored. In this paper, we develop an intelligent framework which can automatically detect a forged signature even if it is highly skilled, based on the developed feature embeddings and the corresponding algorithm. Siamese Triplet Convolution Neural Network is used to generate the feature embeddings for the signature images followed by a generalized Logistic Regression model to detect forgery. On the widely used SigComp dataset, our system achieves an accuracy of 96% in detecting forged signatures. Once the model is trained, it requires just one *base* image to determine whether another signature image is genuine or fraudulent with one shot learning. This algorithmic framework can be used in multiple commercial settings. One such example is testing customer or employee signatures on documents against a corresponding base signature saved beforehand.

Index Terms- Active and real-time vision, Fraud detection, Off-line signature recognition, Triplet loss.

I. INTRODUCTION

Signatures are widely relied upon for identity verification by business, financial organizations and governments to authorize transactions and documents. Accurate signature verification is imperative since forgery and fraud can cost organizations money, time, and their reputation. In the last few years, a lot of progress has been made in the field of automating signature forgery detection using machine learning and image recognition-based concepts.

Signature forgery can be broadly of two types:

- Blind Forgery: Where the forger has no idea what the signature to be forged looks like. This is easy to detect by machine because it is usually not very close to the appearance of a genuine signature.
- Skilled Forgery: Either simulation or tracing, in which the forger has a sample of the signature to be forged. In this case, detecting fraud requires more sophisticated tools to differentiate minute but critical details between genuine and forged signatures.

In this paper, an automatic off-line signature verification and forgery detection system using image processing and Deep Convolutional Siamese networks is proposed wherein a deep triplet ranking network is used to calculate the image embeddings. This is coupled with generalized linear model architecture with logistic loss functions and cross validation to arrive at the final model to label images as authentic or forged.

Training the model requires significant computation resources, but once the model is trained, it requires only one base image to determine whether another signature image is genuine or not with one shot learning. This process is instantaneous and can be carried out in real time. The main contribution of this paper is to enhance the robustness of the signature image embeddings using a FaceNet [1] based triplet network architecture with transfer learning using MobileNet CNN architecture [2]. The triplet loss

architecture with fine- tuned MobileNet CNN embeddings not only takes care of the separation between the genuine and forged signatures but also ensures the relative positioning of the genuine (positive) signatures in the embedding dimension.

An overview of the rest of the paper is as follows: in section II we review the literature in this area; section III delves into the method and framework developed in selecting triplets, learning the embeddings and the logistic loss parameters. Finally, in section IV we present some quantitative results of our embeddings.

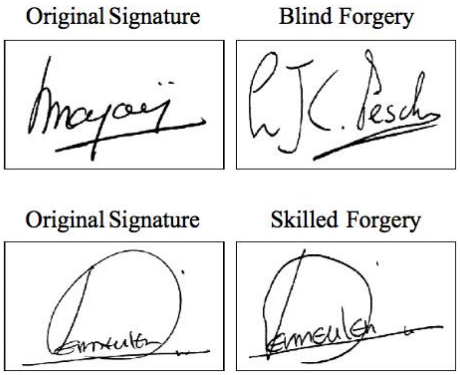


Fig. 1. Examples of blind and skilled forgeries from SigComp dataset

II. RELATED WORK

Offline signature verification is one of the most challenging tasks in the field of forgery detection and lot of work has been done in this field. Earlier methods to handle the problem include creating hand crafted features like block codes, wavelet, Fourier transformation etc. [3]. There is also a literature of work considering the geometrical and topological characteristics of local attributes such as position, tangent direction, blob structure, curvature as in Munich et al. [4].

Projection and contour based methods as shown in Dimauro *et al.* [5] gained a lot of popularity in this field. Other interesting approaches on direction profile [6], surroundedness features [7], grid-based features [8] etc. have also gained a lot of momentum in the past. There is also some literature related to structural methods where relationship between the local features is explored using graph-based matching [9].

Srinivasan *et al.* [10] explores the person-dependent and person-independent learning tasks which tend to capture both type of variances and thereby giving greater accuracy, but the method is highly sensitive to the number of samples per individual for consistency in estimation of the distribution.

Justino *et al*. [11] presents a very interesting learning process based on HMM where the objective is to get the best model that is able to represent each writer's signature while differentiating the intra-personal variation and interpersonal variation.

Drouhard *et al.* [12] proposes a neural network-based approach which uses a directional probability density function as a global shape factor wherein its discriminating power is enhanced by reducing its cardinality via filtering.

Alvarez *et al.* [13] focuses on building systems trained using VGG Convolutional neural network architecture with varying degrees of information, as well as experimenting with different objective functions to obtain optimal error rates.

Dey *et al*. [14] emphasizes on an offline writer independent signature verification task using a Siamese twin architecture which has the ability to learn complicated features to detect forgery which can't be achieved through hand-crafted features.

The architecture proposed in this paper uses depth-wise separable convolutions which significantly reduces the number of parameters when compared to normal convolutions with the same depth in the networks resulting in light weight deep neural networks. The triplet loss minimizes the distance between an anchor and a positive, both of which have the same identity, and maximizes the distance between the anchor and a negative of a different identity as explained in Schroff *et al.* [1], thereby giving state of the art results in face recognition problems.

We use this triplet loss function in conjunction with logistic regression model to efficiently identify the complex patterns in the signatures to detect blind as well as skilled forgery.

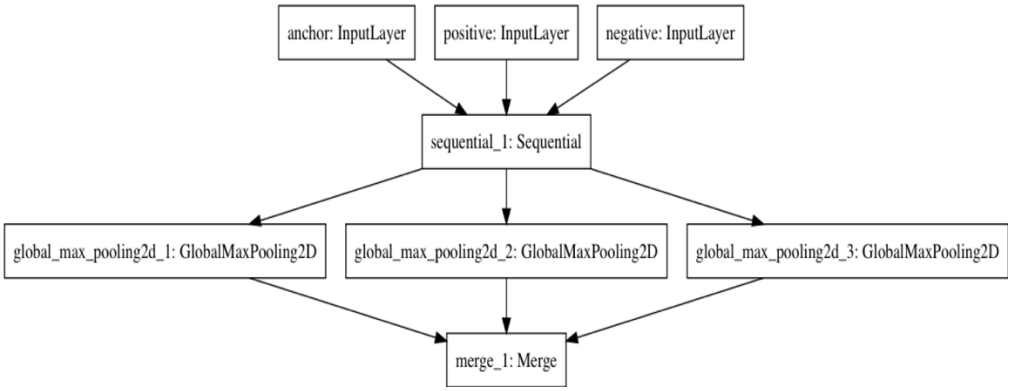


Fig. 2. Triplet loss architecture based on FaceNet [1]

III. METHOD

A. CNN Network Architecture

We use a deep convolutional Siamese network. Siamese convolution networks are twin networks with shared weights, which can be trained to learn the feature embed- dings where similar observations are placed in proximity and dissimilar are placed apart.

Triplet Loss function is used, wherein we contrive the data set in a triplet formation. This is done by taking an *anchor* image (genuine signature of a person) and placing it in conjunction with both a *positive* sample (another genuine signature of the same person) and a *negative* sample (a forged signature by someone else of the same person). This kind of framework ensures that the squared distance between two genuine signatures of the same individual is small, whereas the squared distance between a genuine and forged signature of an individual is large.

In implementing the CNN architecture, we have modified the pre-trained *MobileNet* CNN model [2] with additional layers and have used *Transfer Learning* in building the same. We have trained the last few layers and built dense layers on top of it to extract the embeddings with triplet loss function. Figure 3 shows the structure of convolution neural network based on [2] and the corresponding *trainable* and *non- trainable* parameters.

Layer (type)	Output Shape	Param #	Connected to
anchor (InputLayer)	(None, 128, 128, 3)	0	
positive (InputLayer)	(None, 128, 128, 3)	0	
negative (InputLayer)	(None, 128, 128, 3)	0	
sequential_1 (Sequential)	multiple	3228864	anchor[0][0] positive[0][0] negative[0][0]
global_max_pooling2d_1 (GlobalM	(None, 1024)	0	sequential_1[1][0]
global_max_pooling2d_2 (GlobalM	(None, 1024)	0	sequential_1[2][0]
global_max_pooling2d_3 (GlobalM	(None, 1024)	0	sequential_1[3][0]
merge_1 (Merge)	(None, 1)	0	global_max_pooling2d_1[0][0] global_max_pooling2d_2[0][0] global_max_pooling2d_3[0][0]

Total params: 3,228,864
Trainable params: 1,052,672
Non-trainable params: 2,176,192

Fig. 3. Structure and layer configuration of triplet CNN architecture based on MobileNet [2]

We use the Adam Optimizer with mean absolute error for back propagation to get to the final encodings. The image embeddings so obtained are such that the dissimilarity between the anchor image and positive image must be low and the dissimilarity between the anchor image and the negative image must be high for every triplet.

This kind of architecture ensures that even small differences in signatures can be captured in order to flag a skilled forgery.

The loss that is being minimized is then

$$Loss = \sum_{i=1}^{N} [\|f(x_{i=1}^{a}) - f(x_{i=1}^{p})\|_2^2 - \|f(x_{i=1}^{a}) - f(x_{i=1}^{n})\|_2^2 + \alpha] \qquad (1)$$

where,

f(a) refers to the image encoding of the anchor a

f(p) refers to the image encoding of the positive p

f(n) refers to the image encoding of the negative n

α is a constant used to make sure that the network does not try to optimize towards

f(a)- f(p) = f(a)- f(n) = 0

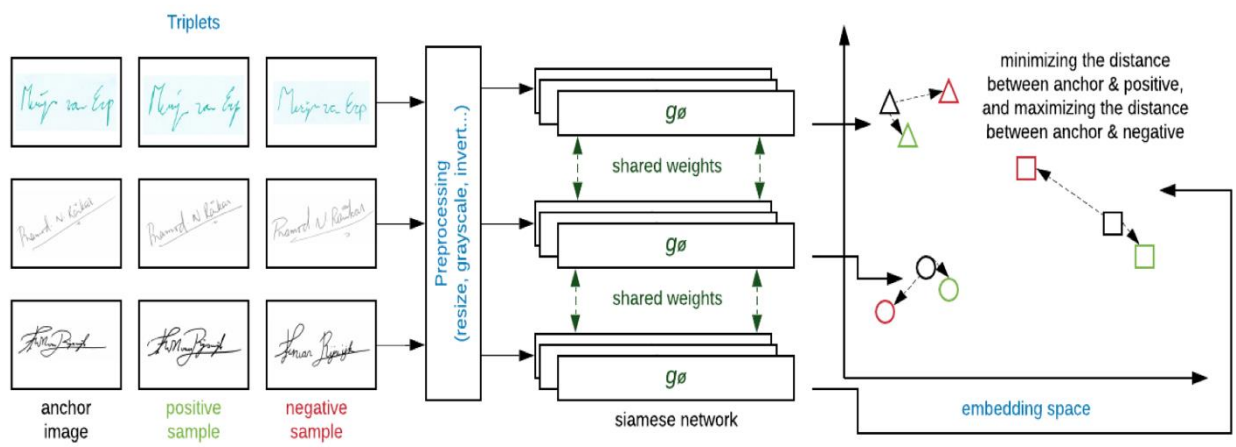


Fig. 4. Training the model using Deep Triplet Ranking CNN Network

B. Logistic Regression Model

Once the final image embeddings for all the training images is attained through the triplet loss architecture, we train a generalized linear model with logistic loss function to get the final model that declares any signature as genuine or forged against a base image signature.
For training this logistic function, we arrange the images in a pairwise manner where each observation is pair of images, either both of a person's genuine signature, or one of person's genuine signature, and the other as person's forged signature.
These will have labels (class) genuine or fraud assigned to them respectively. We use cross validation to get to the final logistic model taking the corresponding differences between the embeddings of each of the pairs (1024 length difference vector of embeddings) as the feature set and the class labels (genuine/ fraud) as the dependent variable y.

The loss function pertaining to this is

$$Loss = g(\sum_{i=1}^{N}(\sigma(W^T X_i(x_i^a, x_i^b) + b)), y_i) \quad (2)$$

where,
$g()$ is the Logistic Loss Function
y_i is the response class (0/1) for image pair i
x_i^a, x_i^b are the image embeddings from the trained CNN network of images in the image pair i, where x_i^a is a person's genuine signature, and x_i^b is either genuine or forged
$X_i(x_i^a, x_i^b)$ represents the entire set of difference based features computed from the corresponding embeddings of image pair i in space IR^m.
W and b represent the final weights obtained from generalized linear logistic loss function.

C. Final Framework Architecture

Once the training process is completed, all we need is the trained model outputs (encodings and logistic model weights). Next, we create a database framework to save the original signatures of every new individual against a unique ID. We pass these base images through the encodings we obtained in the training process and get the corresponding image embeddings. This is in the format of a vector of length 1024. We precompute these embeddings and save them against the individuals' unique ID in the database. Now, whenever we acquire a new signature image against one of the unique ID's that needs to be accessed to determine whether it is genuine or fraud, the framework passes that image through the encoding once again to get its image embeddings.

This new image embedding is then compared to its corresponding embedding of the base image of that individual to determine whether it is genuine or forged. This is done by taking the difference vector of the two embeddings and passing it through the logistic model to get the final prediction. If the resultant probability from the logistic model is low, then the framework declares the new image as a genuine, otherwise it is considered a forgery.
The test workflow is depicted below in Figure 5 and the framework flowchart is depicted in Figure 6.

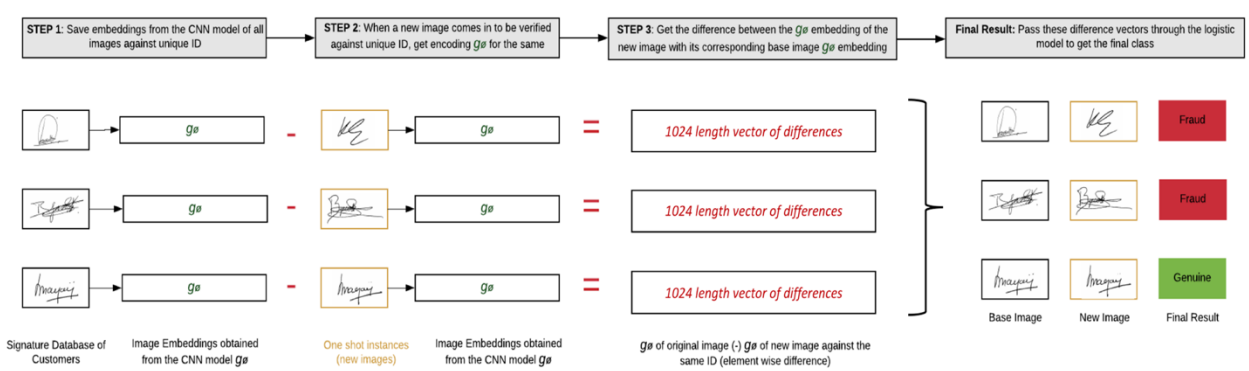


Fig. 5. Testing process of determining whether signature is genuine or fraud

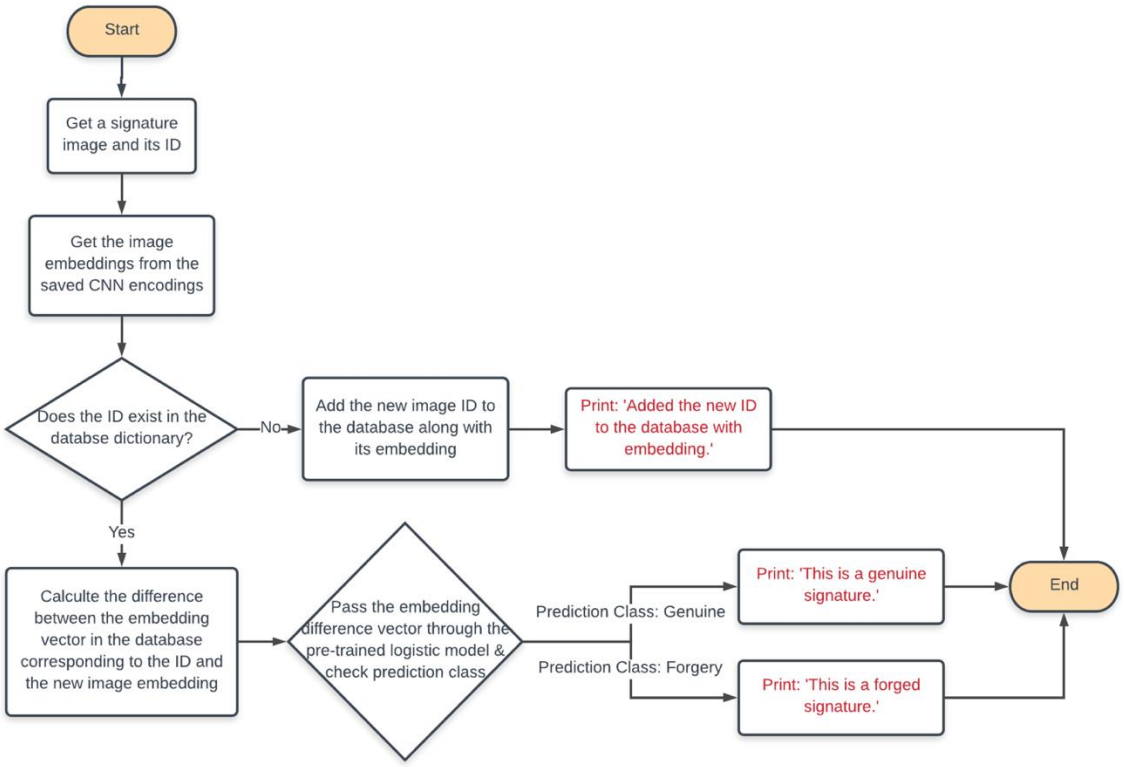


Fig. 6. Execution Framework Flowchart

IV. DATASETS AND EVALUATION

We evaluate our method on the SigComp datasets which are widely used in this field. The entirety of SigComp 2009, SigComp 2010, and SigComp 2012 was used. For SigComp 2011, we only used the Dutch offline dataset.

We create all combinations of triplets from this data: this is done by taking an anchor image (genuine signature of a person) and placing it in conjunction with both a positive sample (another genuine signature of the same person) and a negative sample (a forged signature by someone else of the same person). A total of 120k such triplet combinations was obtained. We resize all the images and convert them to arrays to be passed in the CNN triplet model.

A. Logistic Regression Model

We keep a holdout set of around 20%, that has the same distribution as our training set, but disjoint identities. For training the logistic model, we use 10-fold cross validation on the image pairs obtained.

[1] These datasets are from ICDAR Signature Verification Competition from years 2009-2012

TABLE I
RESULTS OF ACCURACY, PRECISION AND RECALL

Data	**Accuracy**	**Precision**	**Recall**
Training	98.49%		
20% Hold Out Test Set	96.54%	95.28%	96.75%
Average 10-fold CV	95.93%	94.74%	96.01%
Cross Validation fold 1	96.68%	95.83%	96.60%
Cross Validation Fold 2	95.37%	94.35%	95.10%
Cross Validation Fold 3	95.55%	94.37%	95.50%
Cross Validation Fold 4	95.76%	94.75%	95.60%
Cross Validation Fold 5	96.24%	95.07%	96.40%
Cross Validation Fold 6	96.24%	95.70%	95.70%
Cross Validation Fold 7	95.85%	94.66%	95.90%
Cross Validation Fold 8	96.37%	95.08%	96.70%
Cross Validation Fold 9	95.63%	94.20%	95.90%
Cross Validation Fold 10	95.59%	93.42%	96.70%

B. Personal Signature Images

We also curated a database of our personal signatures (both genuine pairs and forged) with clean labels to test the model output. This consisted of more than 50 signature pairs.

C. Results

Our model achieved a training accuracy of 98.5%. On the 20% hold-out set, it attained a test accuracy of 96.5%. The precision and recall for the same was 95.2% and 96.7% respectively.

The results from the 10-fold cross validation are presented in Table I. This gave an overall accuracy of above 96% which denotes that the model has high accuracy and generalizability, and is significant as well.

V. CONCLUSION

The differences between the images of a genuine signature and its skilled forgery is at times very minute and is challenging to detect by even a trained eye. Deep Triplet loss function is a very powerful loss function used in the industry for face recognition [1]. We have created our custom triplet model architecture with modified MobileNet CNN and dense layers with triplet loss function. Based on this loss, the image embeddings are created in such a way that the dissimilarity between the anchor image and positive image must be low and the dissimilarity between the anchor image and the negative image must be high for every triplet. This kind of architecture ensures that even small differences in signatures can be captured in order to flag a skilled forgery effectively.

REFERENCES

[1] F. Schroff, D. Kalenichenko, and J. Philbin, “Facenet: A unified embedding for face recognition and clustering,” *2015* ***IEEE*** *Conference on Computer Vision and Pattern Recognition (CVPR),* pp. 815-823, 2015.

[2] A. G. Howard, M. Zhu, B. Chen, D. Kalenichenko, W. Wang, T. Weyand, M. Andreetto, and H. Adam, "Mo- bilenets: Efficient convolutional neural networks for mobile vision applications," *CoRR,* vol. abs/1704.04861, 2017.

[3] M. K. Kalera, S. Srihari, and A. Xu, "Off-line signature verification and identification using distance statistics," *International Journal of Pattern Recognition and Artificial Intelligence,* pp. 228-232, 2003.

[4] M. Munich and P. Perona, "Visual identification by signature tracking," *Pattern Analysis and Machine Intelligence, IEEE Transactions on,* vo1. 25, pp. 200— 21 7, 03 2003.

[5] G. Dimauro, S. Impedovo, G. Pirlo, and A. Salzo, "A multi-expert signature verification system for bankcheck processing," *International Journal of Pattern Recognition and Artificial Intelligence,* vo1. 11, pp. 827-843, 08 1997.

[6] M. A. Ferrer, J. B. Alonso, and C. M. Travieso, "Offline geometric parameters for automatic signature verification using fixed-point arithmetic," *IEEE Transactions on Pattern Analysis and Machine Intelligence,* vol. 27, no. 6, pp. 993-997, June 2005.

[7] R. Kumar, J. D. Sharma, and B. Chanda, "Writer-independent off-line signature verification using surroundedness feature," *Pattern Recogn. Lett.,* vo1. 33, no. 3, pp. 301-308, Feb. 2012. [Online]. Available: http://dx.doi.org/10.1016/j.patrec.2011. 10.009 .

[8] R. Chandra and S. Maheskar, "Offline signature verification based on geometric feature extraction using artificial neural network," *2016 3rd International Conference on Recent Advances in Information Technology (RAIT1,* pp. 410—414, March 2016).

[9] S. Chen and S. Srihari, "A new off-line signature verification method based on graph" vo1. 2, 01 2006, pp. 869–872.

[10] H. Srinivasan, S. Srihari, and M. J. Beal, "Machine learning for signature verification," vol. 4338, 01 2006, pp. 761–775.

[11] E. Justino, M. El Yacoubi, F. Bortolozzi, P. Pontifcia, and R. Imaculada Conceio, "An off-line signature verification system using hmm and graphometric features," 08 2002.

[12] J.-P. Drouhard, R. Sabourin, and M. Godbout, "A neural network approach to off-line signature verification using directional pdf," *Pattern Recognition,* vol. 29, pp. 415–424, 03 1996.

[13] G. Alvarez, B. Sheffer, and M. Bryant, "Offline signature verification with convolutional neural networks," Tech. Rep.

[14] S. Dey, A. Dutta,J. IgnacioToledo, S. Ghosh, J. Lladós, and U. Pal, "Signet: Convolutional Siamese network for writer independent offline signature verification," 07 2017.

Power Reduction In Logic Circuits Using Power Gating For Deep Sub-Micron CMOS VLSI

Reeba Korah, **Neethu Vijayan**
Dept. of ECE, Alliance College of Engineering and Design

Abstract: Technology scaling leads to subthreshold leakages in deep submicron regime. There is a need for effective leakage reduction techniques to minimize leakage currents. This paper presents basic gates circuits and a full adder circuit with low power consumption. This crucial leakage reductions are achieved by turning off the logic cells from ground and supply rails by power gating technique. Power gating technique reduces static power consumption but it increases the delay of the logic cells badly in deep submicron CMOS circuits and the performance is besmirched to a great extent. In order to enhance performance additional transistors are added in the circuit which reduces both dynamic and leakage consumption enormously. Simulations are done with 45nm CMOS technology in Cadence Virtuoso. Circuits are working in 0.5V supply voltage. It reduces the static power without much affecting dynamic power consumption compared to those of equivalents. For the 10T full adder, the proposed design accomplish percent reduction in dynamic power consumption and percent reduction in static power consumption. The proposed designs are apt for low power and highly efficient adders.

Keywords: Power gating, retention, transistor stacking.

I. Introduction

Electronic devices such as laptop, cell phone, digital camera etc. are part of our day today life. Its battery life span is of great importance. When the mobile phone is operating in talk mode, some components of the mobile phone are turned off, but this doesn't stop the battery from getting depleted. This is because circuits which are de-activated by turning off certain component parts still have leakage currents flowing through them. It decreases the battery life over relatively long standby time although the magnitude of leakage current is less than the normal operating current. The normal operating current reduces the battery life over relatively short talk time. Thus low power circuits for different applications are of great interest.

Power dissipation in CMOS circuits [1] comes from two components. Dynamic dissipation may be due to charging and discharging of load capacitances as gates switch, or “short-circuit” current while both PMOS and NMOS stacks are partially ON, whereas Static dissipation may be due to Sub-threshold leakage through OFF transistors, Gate leakage through gate dielectric, Junction leakage from source/drain diffusions and Contention current in ratioed circuits. Power can also be considered in active, standby, and sleep modes. Active power is the power consumed while the chip is in the working condition. It is usually dominated by switching. Standby power is the power consumed while the chip is in idle condition. In sleep

mode, the supplies to unneeded circuits are turned off to eliminate leakage [2]. This drastically reduces the sleep power required, but the chip requires time and energy to wake up so sleeping is only viable if the chip will idle for long enough. There are many proposed technique for power and energy reduction. Voltage supply (Vdd) scaling is considered one of the most effective elements in the process of reducing power dissipation in CMOS circuits. In order to maintain the required current drive, Threshold Voltage (Vth) has also be reduced. Reducing threshold voltage Vth results in an exponential increase in leakage power. The ratio Vdd /Vth tend to decrease with technology scaling to keep the leakage power under control. Another technique is power gating. In this, the power supply and the ground line are separated from the circuit block while the chip is in sleep mode.

In this paper, we synthesized NOT, NAND and NOR gates and proposed novel techniques to minimize power and to improve the performance of these gates in nanoscale technology [3]. Also, 1-bit 10 T full adder has been realized with power gating for the leakage power reduction in 45nm technology [4]. The rest of the paper is organized as follows. Section II describes the related work. Section III shows the design considerations and section IV presents the proposed design. Section V describes the implementation and simulation together with the simulation of a 1bit 10 T full adder. Finally, the conclusions are presented in section VI.

II. Related Work

Ankita Nagar [5] proposed the reduction of power dissipation in basic gates using transistor stacking. It is found that when the number of low input increases in case of NAND gate the power dissipation decreases and for NOR gate power dissipation decreases with the increase in high input vector combination. Jatin Mistry, James Myers and Basher M [6] presents a technique called sub clock power gating. Amit Bakshi [7] gives an idea on reducing subthreshold leakage power consumption and ground bounce noise during the sleep to active mode transition. Shota Ishihara, Masanori Hariyama [8] gives the classification of power gating technique as fine grain power gating and coarse grain power gating. In coarse grain power gating a single sleep transistor is shared among all the components whereas in finegrain power gating each component will have sleep transistors. In order to reduce subthreshold leakage in sleep mode Seta et al applied reverse body bias [9] for deep submicron circuits. T Kuroda et al proposed variable threshold CMOS [10]using TWIN well or triple well technology. An extra circuitry to apply bias to the footer was proposed by Kawaguchi et al [11] and this technique is called super cut off CMOS power gating. Large cores power deduction is possible through multi-mode power switches.It is proposed by Zhaobo Zhang et al by using three transistors attached to footers to the main core to operate in four modes such as active mode, snore mode, dream mode and sleep mode. In this paper we proposed a design to reduce the power consumption using power gating with reverse body bias and extra devices to improve performance.

III. Design Considerations

Supply voltage scaling is an important step in the technology scaling process. Supply voltage scaling helps in maintaining the power density of an integrated circuit below a

limit dictated by available cost effective cooling techniques and it also guarantees the long term reliability of the devices fabricated in a scaled semiconductor technology. In current CMOS technologies, dynamic power dissipation due to switching is the main component of the total energy consumption of an integrated circuit. The dynamic switching energy is proportional to the square of the supply voltage in a full voltage swing CMOS circuit. Moreover the leakage and short circuit energy components also depend super linearly on the supply voltage. Reducing the supply voltage, therefore, is an effective way to lower the power dissipation.

Power gating is one of the technique to reduce the leakage power. Power gating uses PMOS or NMOS as sleep transistors that are having low leakages. Power gating consists on switching off the power supply from blocks that are in standby mode and switching the power back on when their functionality is required. In order to switch power off, high Vt transistors are used as switches and placed between the block power and ground pins and the rails. Header Switch which is a PMOS transistor is placed between the power pins and the power rails. Footer switch which is a NMOS transistor placed in between the ground pins and ground rails. This way the controlled block, is no longer powered by the main power rails (always-on rails), but by a switched/virtual power rail. The control of the power switches is achieved through an enable or sleep signal. For a header switch, the enable signal takes the logic value 1 in order to switch power off. In the footer switch case, the opposite happens. As important as the power switches themselves is to distribute and deliver power in the best way possible across the whole design. The power network should be designed to minimize the voltage drop and to correctly power all blocks and standard cells in the design. Additionally, a state retention strategy can be implemented. Depending on the application, it can be necessary to save the state of the block. This way, when power is switched back on, the block can return to the exact same functioning state [14]. This is achieved by using state retention transistors. It may be also important to isolate the powered-off block output signals. These signals, if floating, can induce a crow-bar current in an adjacent block, to which represent input signals. Sub-threshold leakage current that is flowing through a stack of series-connected transistors decreases when more than one transistor in the stack is turned off. This effect is known as stacking effect. It is also known as self-reverse bias. Leakage currents in NMOS or PMOS transistors depend exponentially on the voltage at the four terminals of transistors [15]. The gate voltage Vg is “0” and this will increase source voltage Vs of NMOS transistor. This reduces sub-threshold leakage current exponentially. In our design we consider transistor stacking, retention transistor and novel hybrid technique to reduce the power consumption.

IV. Proposed Work

NOT Gate

Novel hybrid NOT gate

The design of a hybrid NOT gate is shown in fig 1.The circuit contain an inverter design with sleep transistors. In the static mode, the sleep transistors P0 and N1 are off. Therefore the logic circuit is isolated from supply and ground rails. So the static power is almost low. In the dynamic mode the sleep transistors are on and when the input is zero, PMOS P3 will

be ON and the little leakage current in the NMOS N0 is fedback to the input through the PMOS P3 and thereby the dynamic power is reduced. When the input is high, PMOS P3 will be OFF and the NMOS N0 would be ON and it is connected to ground through high threshold voltage transistor N1. This low voltage threshold PMOS P3 added to the input of the inverter increases the throughput. The source and gate of this PMOS is shorted. Its drain is connected to the ground and also fed to the virtual ground. This low threshold voltage transistor is used to minimize the standby leakage.

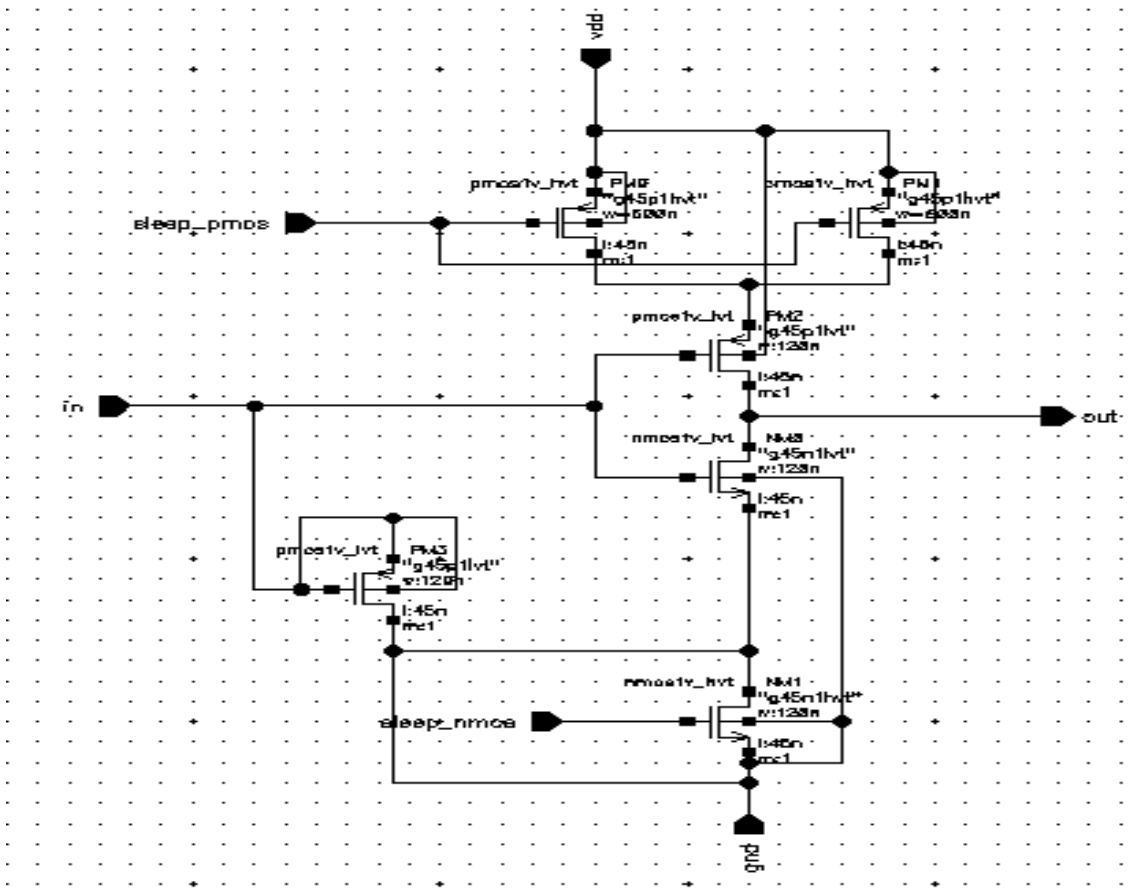


Fig.1 Novel Hybrid NOT gate

Novel Hybrid NOT gate with transistor stacking

In this design transistor stacking is employed. Here instead of one transistor in the logic circuit, two PMOS/NMOS transistors (P2, P3 and N0, N1) are placed one over the other. It will not affect the total width or length of the transistors. But the standby power is reduced further because leakage current through a stack of two off transistors will be less than that of a single off transistor. The dynamic power consumed will be more as extra transistors are used in this technique. When such series connected NMOS transistors N0 and N1 are turned off, the internal node makes the gate to source voltage of upper transistor N0 to be negative. Hence the subthreshold leakage current is reduced. Care should be taken such that addition of extra transistors should not impact adversely the performance of the logic circuit. In order to improve

the performance of a static CMOS inverter a low threshold voltage transistor PMOS P4 is connected. Also by applying a body bias using an extra circuitry standby leakage can be further suppressed.

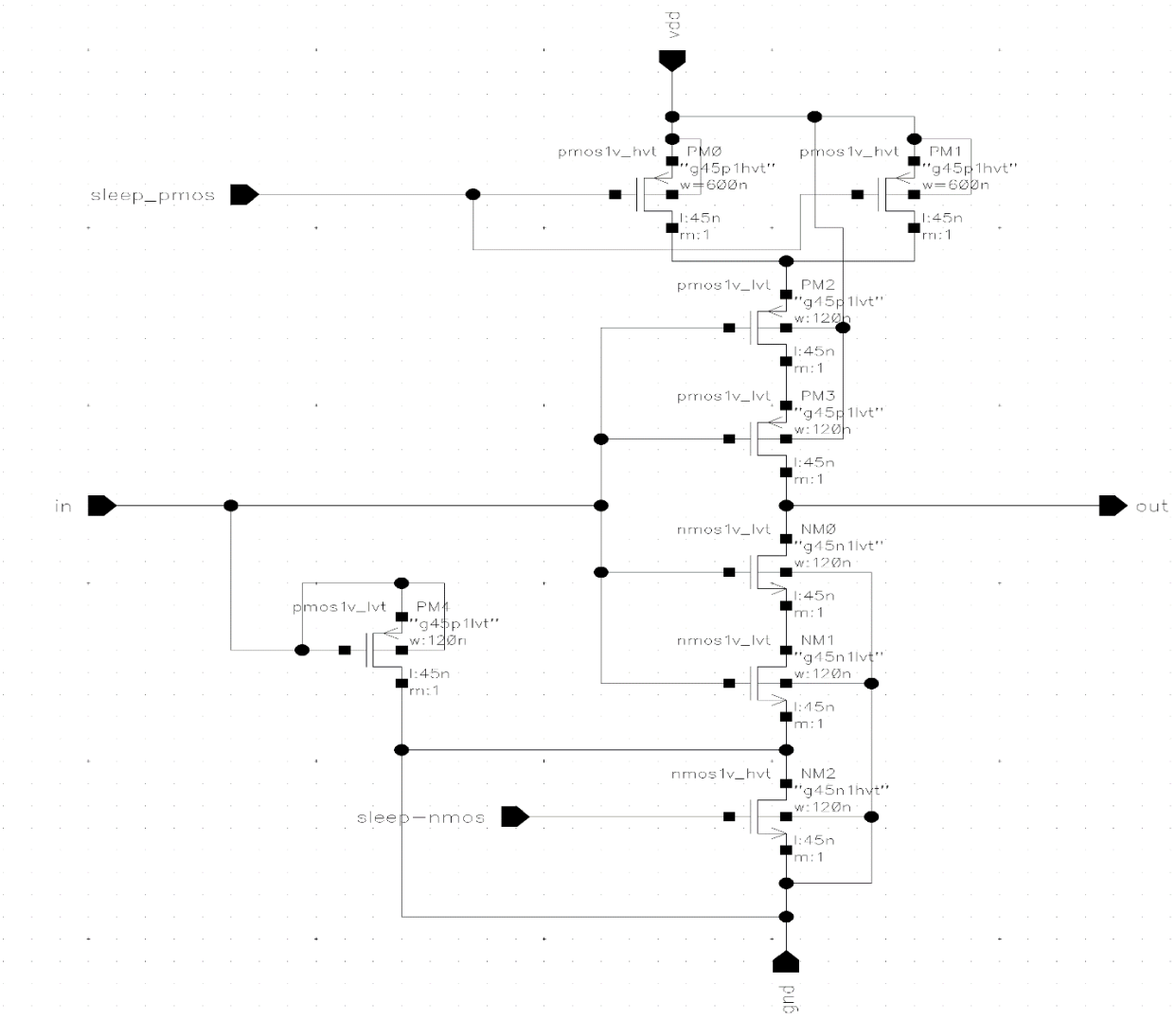


Fig 2. Novel Hybrid NOT gate with transistor stacking.

Novel Hybrid NOT gate with retention transistor

Here the circuit is designed with retention transistors is shown in fig 3. Here symmetric virtual rail clamping technique is employed. This technique work by reducing the virtual supply to less than Vth rather than shutting down completely as in the case of conventional power gating. Symmetrical Virtual rail clamping is used to achieve reduced voltage. For this a pair of NMOS and PMOS transistors are used at the head and foot of the power gated logic. A high threshold NMOS transistor N2 is connected parallel to high threshold PMOS transistors P0 and P1 in the pullup network and also a high threshold PMOS transistor P4 is connected parallel to high threshold NMOS transistor N1 in the pulldown network.

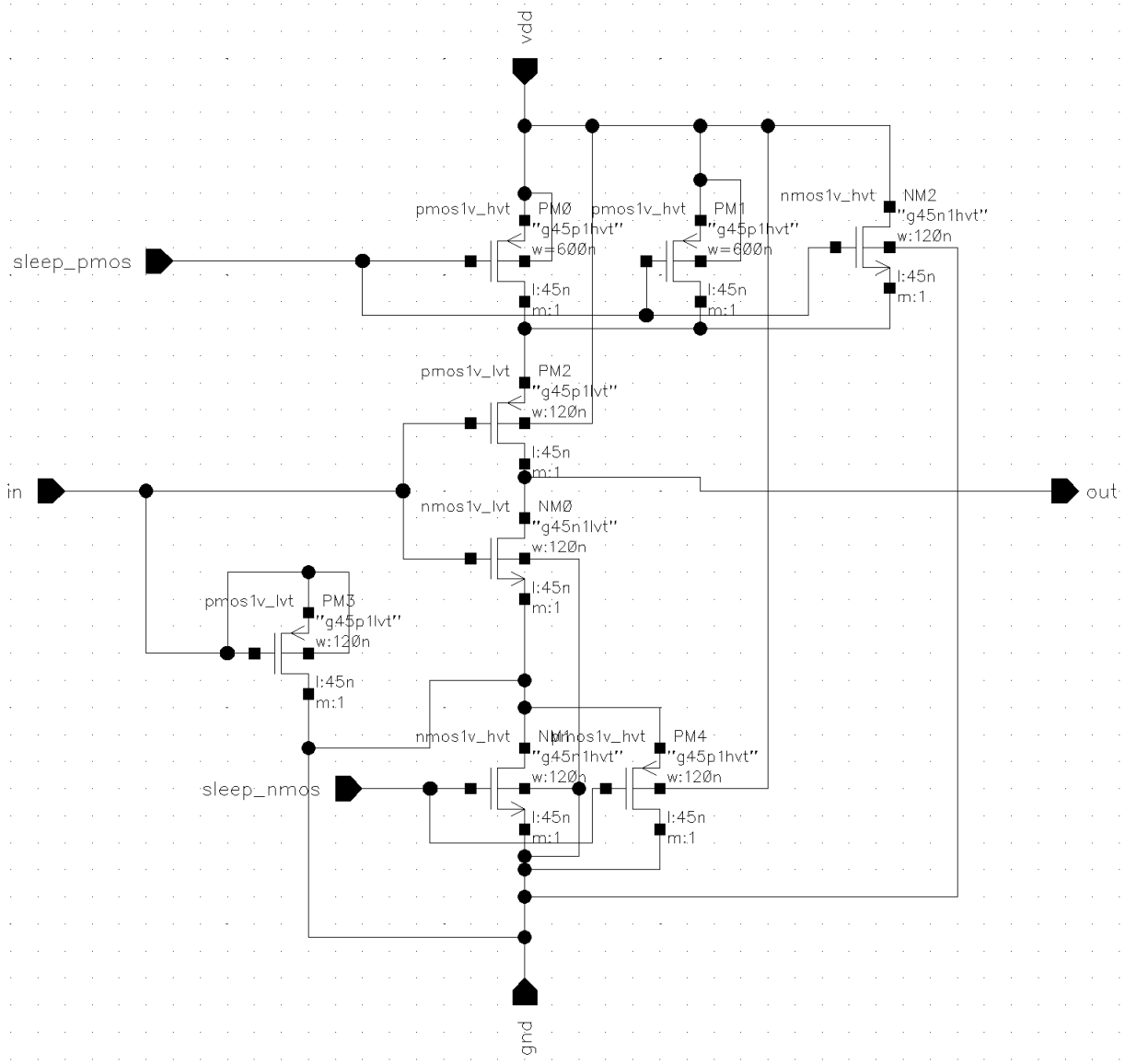


Fig 3.Novel Hybrid NOT gate with retention transistor

In the static mode when sleeppmos and retention transistors are given logic high, virtual supply rail is clamped to V_{dd} -V_{thn} and virtual ground rail is clamped to V_{dd}+V_{thp}, where V_{thn} and V_{thp} are the threshold voltages of the NMOS N2 and PMOS P4 transistors. This results in enormous reduction in voltage across power gated logic due to the following reasons. The first reason is that the charge stored in the supply rails is recycled back to charge the virtual ground rail in sleep mode through the PMOS P4.The second reason is that by connecting the PMOS body and NMOS body in the power gated circuit to the actual supplies achieves a new threshold voltage due to reverse body bias on all the transistors in the logic circuit which reduces leakage further. The third reason is that virtual rail clamping helps gates to keep its correct output for minimum supply voltages. Also in the static mode, when the input is low, PMOS P3 would be ON and thereby voltage in the virtual ground is fed back to the input and thereby static power is reduced.

On the other hand, when the input is high, PMOS P3 would be OFF and voltage in the virtual ground could not be fedback to the input. Hence static power reduction is more only when the input is low. In the dynamic mode when sleeppmos (P0, P1) and retention (N2) are given logic low, dynamic power is reduced more compared to the previous circuits as fewer number of transistors are ON in this technique.

NAND Gate

Novel Hybrid NAND gate

In the dynamic mode, when both inputs are high, both the NMOS, N4 and N3 in the pull down network are ON and the high threshold PMOS transistor P2 is OFF and the output goes low as all the transistors in the pullup network are OFF and current from the pull down network flows back to the input through this transistor P2 thereby reducing the dynamic power consumption. Here the PMOS transistor P2 whose gate and source is tied together and anyone of the inputs is given to the gate/source and the drain is connected to both ground and virtual ground as shown in fig 4.This extra HVT transistor supports in the reduction of dynamic and standby power consumption to a large extend without degrading the circuit performance. The output of CMOS NAND gate is high when any one of the input is low. On the other hand, when both inputs are high, the output is low. When both inputs are low, the outputs goes high as both the PMOS transistors in the pull up network are ON and both the NMOS transistor in the pulldown network are OFF.A part of the current in the transistor N2 flows through high threshold PMOS P2 as it is ON and thereby reducing dynamic power consumption.

When the inputs IN1 and IN2 are High and low then high threshold PMOS transistor P2 is ON and the output goes High as the PMOS transistor P3 is ON and the current in the NMOS transistor N4 flows back to the input through the high threshold PMOS P2 thereby reducing dynamic power. When the inputs are low and high then the high threshold transistor P2 is OFF and the output goes High as the PMOS transistor P4 is on and the current flowing in the NMOS transistor N1 drains to the ground through reverse body biasing. In the active mode, pmossleep and nmossleep are set Low and High respectively and the sleep transistors P0, P1 and N2 are turned on. As these sleep transistors ON resistances are small, the virtual supply rail more or less functions as the actual power rail. In the standby mode, power gating technique reduces the leakage power by using pmossleep and nmossleep. In the standby mode. pmossleep and nmossleep are set High and Low respectively and the sleep transistors are turned OFF, detaching the NAND gate circuit from the power supply and ground lines. This scheme reduces leakage power to a large extent and also the dynamic power consumption of the circuit is reduced but it increases area and delay of the circuit. Since the NMOS ON-resistance is smaller for the same width compared to PMOS, NMOS high threshold transistors are sized smaller compared to high threshold PMOS transistors.

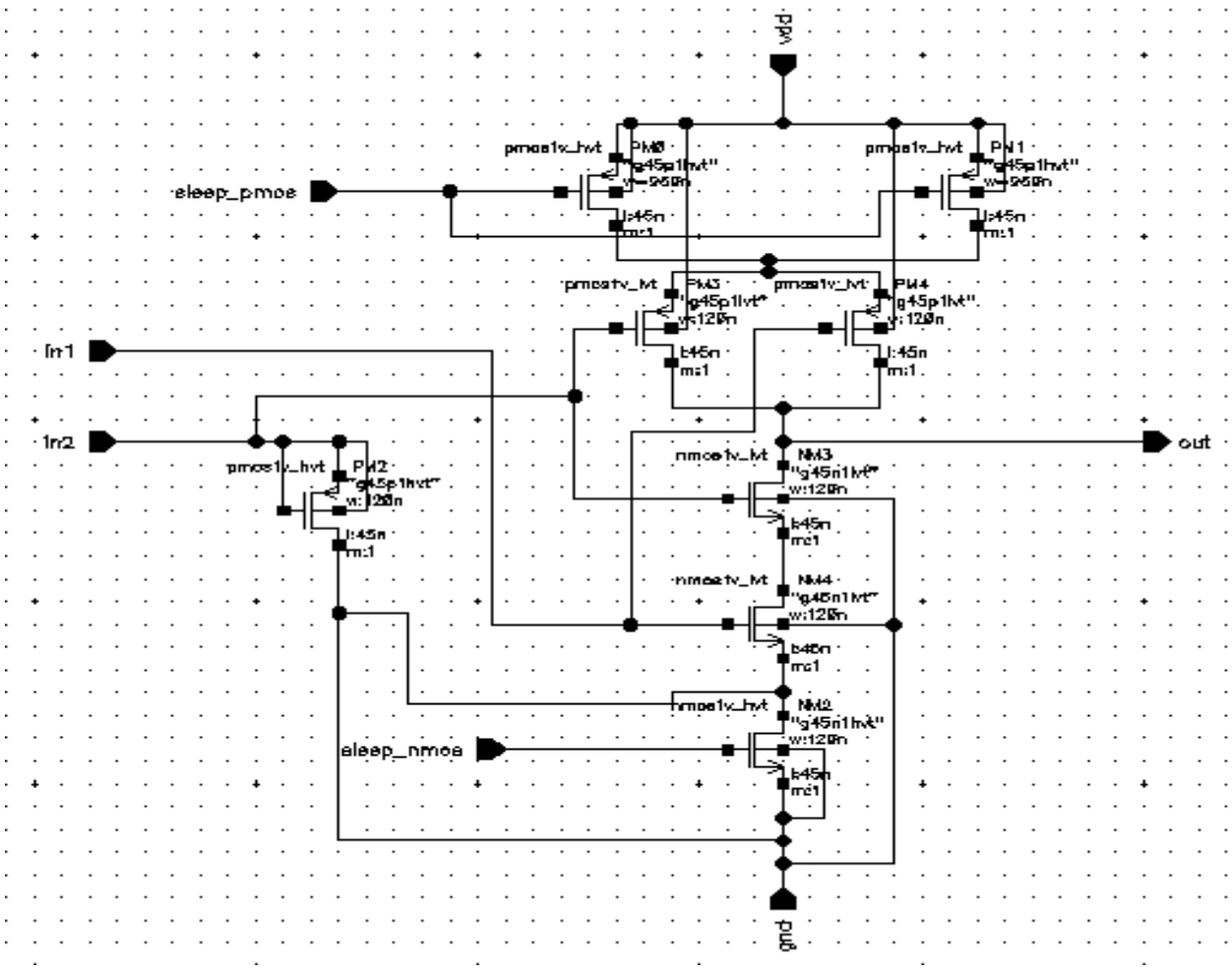


Fig 4.Novel Hybrid NAND gate.

Novel hybrid NAND gate with transistor stacking

One way of influencing the throughput of the static NAND circuit is by adding an extra device, a high threshold voltage PMOS P6. Here the PMOS transistor P6, whose gate and source is tied together and any one of the inputs is given to the gate/source and the drain is connected to both ground and virtual ground as shown in Fig. 5. This extra HVT transistor supports in the reduction of dynamic and standby power consumption to a large extent without degrading the circuit performance. The CMOS NAND gate output goes high when any one of the input is Low. Alternatively, when both inputs are High, the output is Low. Moreover reverse body bias and transistor stacking techniques are applied to reduce standby leakage power. In the active mode, pmossleep and nmossleep are set Low and High respectively and the sleep transistors P0, P1 and N4 are turned ON. As these sleep transistors ON resistances are small, the virtual supply rail more or less functions as the actual power rail. When both inputs are high the output is low as all the PMOS transistor in the pull up network are OFF and all the NMOS transistors (N0, N1, N2 and N3) in the pull down network are ON and the current flows to the

ground through NMOS transistor N4 and thereby reducing dynamic power consumption. Likewise when both the inputs are low all the NMOS transistor in the pull down network are ON and the current in the HVT transistor N4 flows through the HVT PMOS transistor P6 as it is ON and thereby reducing dynamic power consumption.

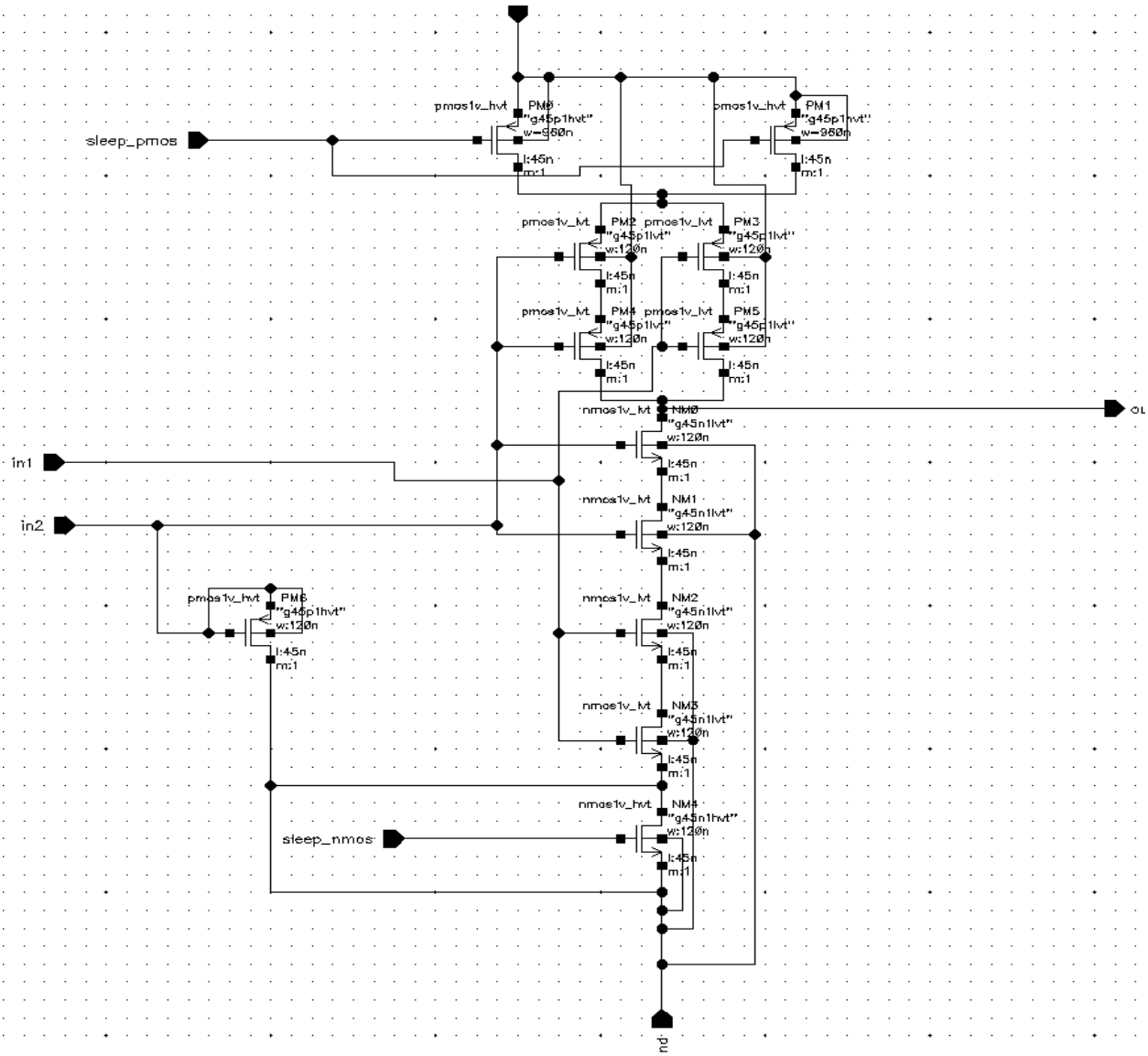


Fig 5.Novel Hybrid NAND gate with transistor stacking

When the IN1 and IN2 are high and low respectively, the output is high as PMOS transistors P2 andP4 are ON and HVT PMOS transistor P6 is ON and the current in the NMOS transistors (N2 and N3) flows back to the input through the PMOS P6 thereby reducing dynamic power. When the inputs IN1 and IN2 are low and high respectively, then the PMOS P6 is OFF and the output is high as the transistors P3 and P5 are ON and N2 and N3 are OFF and the leakage current flowing through the NMOS Transistor N0 and N1/ drains to the ground through reverse body

biasing. In case of standby mode power gating technique reduces the leakage power by using pmossleep and nmossleep.in the standby mode, pmossleep and nmossleep are set High and Low respectively and the sleep transistors are turned off, detaching the NAND gate circuit from the power supply and ground lines. This scheme reduces leakage power to a large extent and also the dynamic power consumption of the circuit is also reduced but it increases area and delay of the circuit.

HVT NMOS transistors are sized smaller compared to HVT PMOS transistors. Using transistor stacking in the NAND gate circuit increases the resistance and results in maximum reduction of sub threshold leakage current flowing through a stack of series connected transistors when more than one transistor is turned off in the stack. Here due to stacking effect, the subthreshold leakage current through a logic gate dependent on the states of the primary inputs. But the dynamic power of this circuit is increased as more number of transistors is utilized in this circuit.

Novel hybrid NAND gate with retention transistor

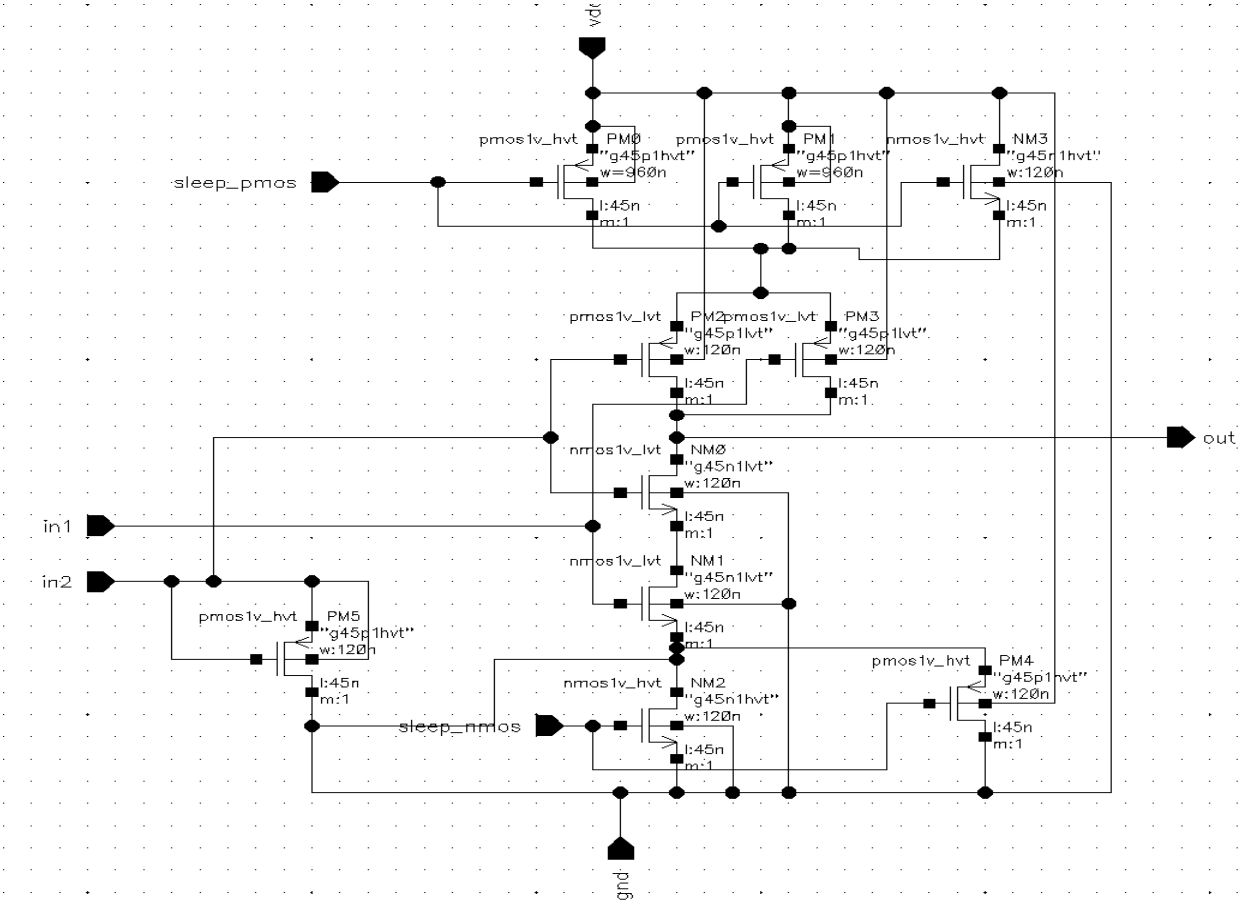


Fig 6.Novel Hybrid NAND gate with retention transistor

Here Virtual rail clamping has been introduced as a way to maintain a voltage across the power gated logic to retain register state. This technique reduces the recharge, glitching and wake up cost associated with power gating, as valid logic outputs are maintained. In general virtual rail clamping ensures a single voltage drop reduction across the power gated logic. But in order to increase the leakage power savings of power gated logic, symmetric virtual rail clamping is used to reduce the clamped voltage by two threshold voltages. Using this technique static power consumption is reduced to a large extent. Also symmetric virtual rail clamping causes the NAND gate to hold its output for very low supply voltages also. Moreover the dynamic power consumption is also deceased as the number of transistors used here is less compared to the previous circuit. Fig 6 shows the circuit of NAND gate with data retention transistors.

NOR Gate

Novel Hybrid NOR gate

In the dynamic mode, when both inputs are High, these two high threshold voltage NMOS transistors N0 and N1 are ON and also the two NMOS N2 and N3 in the logic circuit are ON. As the output is pulled to zero, the current through these two transistors N2 and N3 are fed back to the input through these HVT NMOS transistors N0 and N1 and thereby reducing the dynamic power consumption. When input In1 is Low, the transistor P2, connected next to Vvdd in the logic circuit is ON and the same input is fed to the high threshold transistor P4 whose gate and drain are shorted. This switches ON the high threshold voltage PMOS P4 whose source is connected to Vvdd trying to pull up the output High. If the input In2 is high, then the PMOS P3 is OFF and thereby pulling the output to Low. On the other hand, if the input In2 is also Low, then the PMOS P3 is ON and the NMOS N2 and N3 are OFF and thereby the output is pushed high. When the input In1 is High andIn2 is low, then the PMOS P2 is OFF and the PMOS P3 is ON and the high threshold PMOS P4 is also OFF. Even though this PMOS P4 is ON, this transistor is no more connected to supply voltage as the PMOS P3 is OFF and hence the output goes Low. Fig 7 shows the circuit of novel hybrid nor gate.

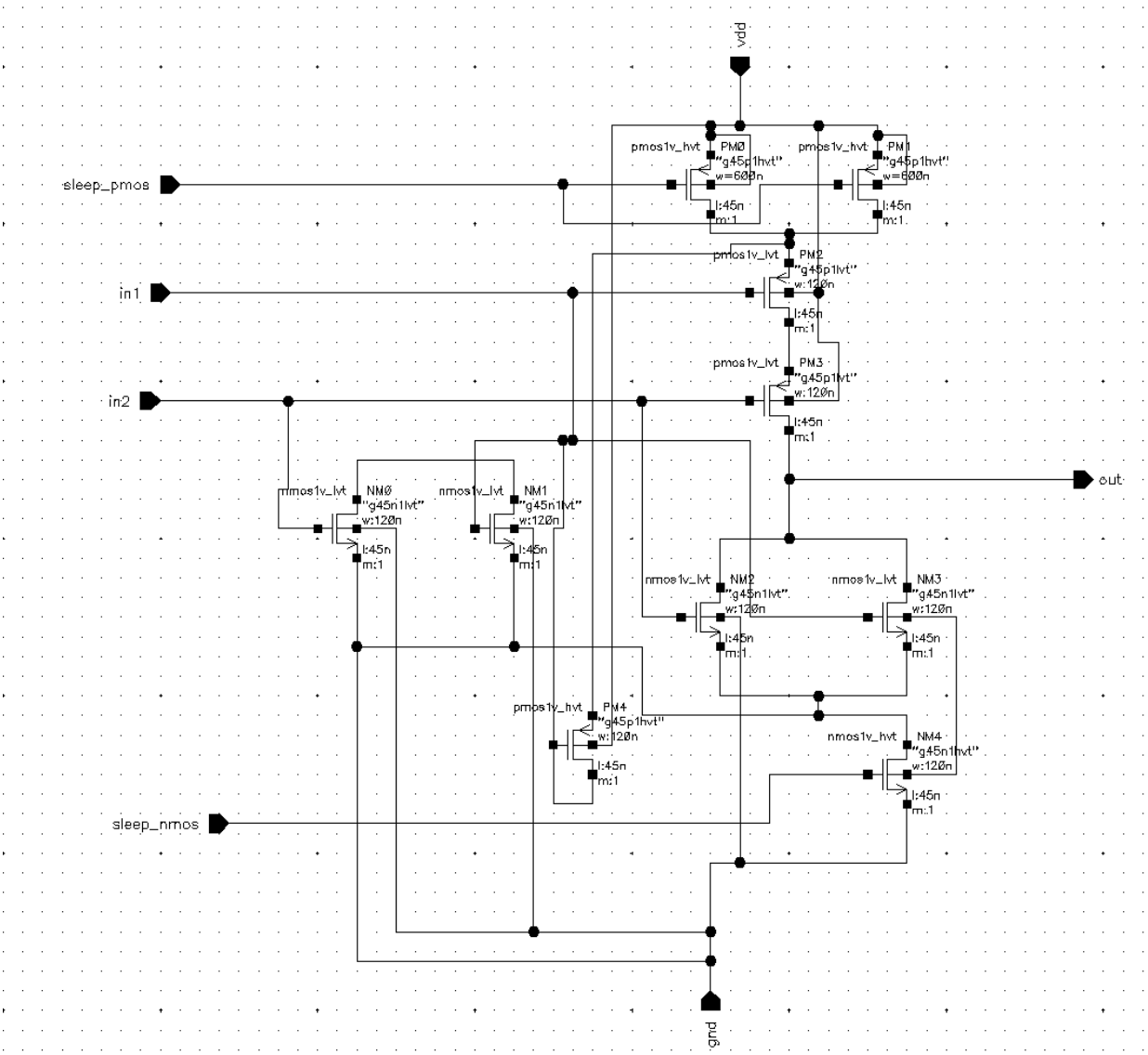


Fig 7.Novel Hybrid NOR gate.

Novel hybrid NOR gate with transistor stacking

In this circuit two high threshold voltage NMOS transistors N0 and N1 are connected as shown in fig 8.and their source are tied together to the virtual ground VGND and another high threshold voltage PMOS P6 is connected with gate and drain shorted and its source is connected to virtual supply Vdd. This technique employs transistor stacking. This technique helps in the maximum minimization of leakage power. In general the output of the NOR gate is low if any one of the input is High and the Output is high if both the inputs are low. In this transistor stacking technique, the resistance between the power supply and actual ground increases and thereby large percentage of the leakage power gets reduced. Addition of high threshold voltage transistors over the transistor stacking of logic circuit helps in the further

reduction of leakage power compared to simple power gating technique. Also supply voltage is reduced to 0.5V which helps to achieve definite reduction in active and standby power.

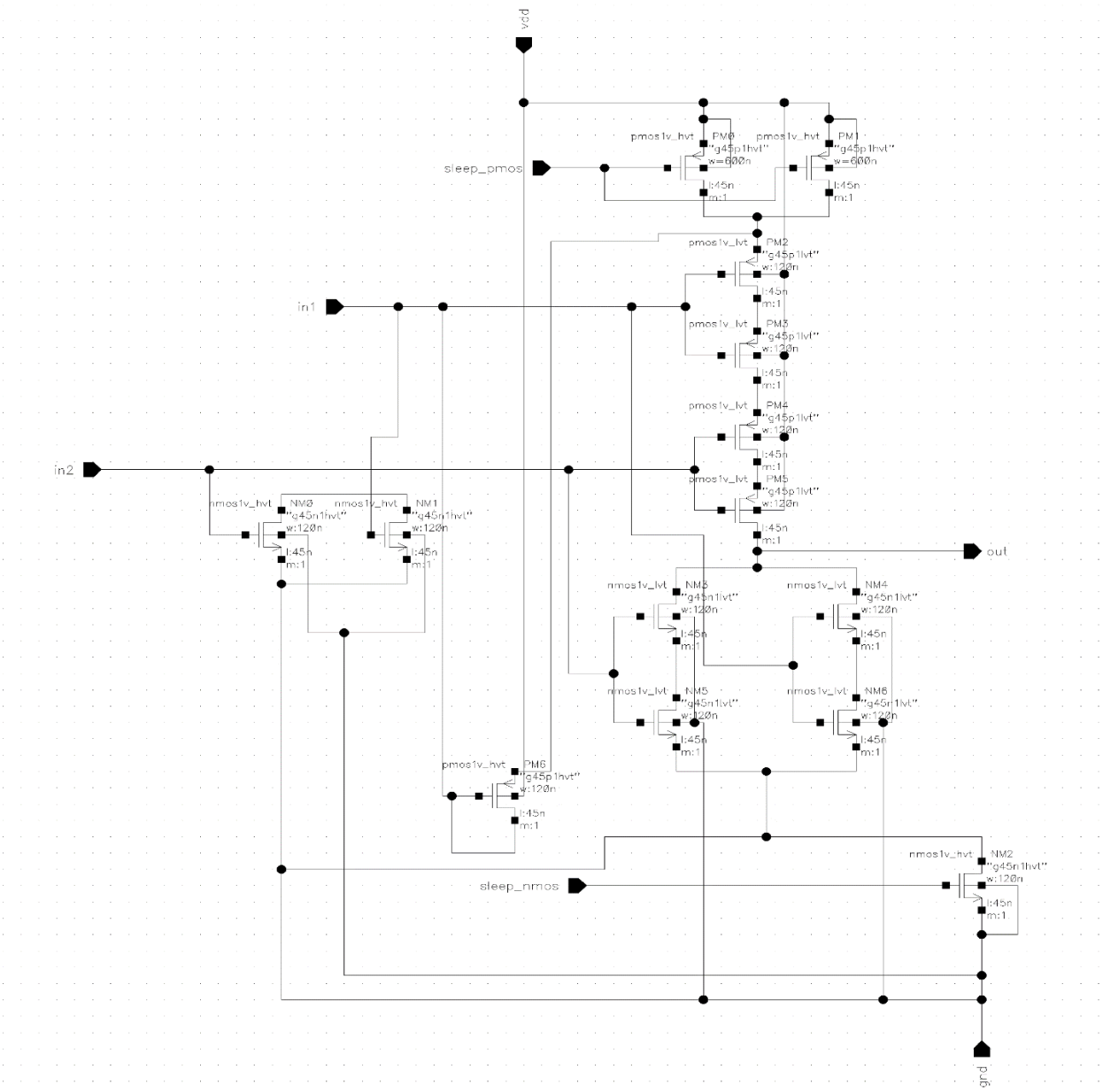


Fig 8.Novel Hybrid NOR gate with transistor stacking

In the dynamic mode, when both the inputs are high these two HVT NMOS N0 and N1 are ON and also all NMOS transistors in the logic circuit are also ON.As the output is pulled to zero, the current through these transistors N3, N5, N4 and N6 are feedback to the input through these HVT NMOS transistor N0 and N1 and thereby reducing dynamic power consumption. When the inputs IN1 is low, the transistors P2 and P3, connected next to Vdd in the logic circuit is on and also the same input is fed to HVT PMOS transistor P6.This switches on the HVT PMOS P6 whose source is connected to Vdd, trying to pull up the output high. If the input IN2 is high, then

the PMOS transistors P4 and P5 are off and thereby pulling the output LOW. On the other hand, if the other input IN2 is also low, then the PMOS transistors P4 and P5 are on and thereby the output is pushed high. When the input IN1 is high and IN2 is low then the PMOS P2 and P3 are off and high threshold voltage PMOS P6 is also off and the PMOS transistors P4 and P5 are on. Even though these PMOS transistors P5 and P6 are on, these transistors are no more connected to supply voltage as the PMOS P2 and P3 are off and hence the output is low.

Novel Hybrid NOR gate with retention transistors

In the second technique, retention transistors N5 and P5 are employed to preserve the output values. Using feedback connection in the circuit, in general preserves the output of the circuit. But here addition of HVT NMOS transistor N5 in parallel to HVT PMOS transistors and the addition of HVT PMOS transistor P5 in parallel to HVT NMOS transistor N4 help in retaining the output. This technique helps in reducing the leakage power to some extent and also the dynamic power of the circuit. In the static mode, when both the inputs IN1 and IN2 are high, the transistors P2 and P3 are off and hence no more attached to supply rail. Hence in these two cases, static power is greatly reduced. Moreover reverse body bias is employed in all the techniques to increase the threshold of the transistor in the standby state. Reverse body biasing is done by connecting the PMOS substrate to the supply rail and thereby increasing N well voltages of the PMOS. Hence the effective threshold of the PMOS transistor are increased which subsequently helps in the leakage current reduction in PMOS transistors. Similarly, by connecting the NMOS substrate to the ground rail increases the P-well voltages of the NMOS, which helps to raise the NMOS threshold and thereby leakage reduction occurs in the NMOS transistors. But a large voltage is required to achieve a mall rise in the threshold voltage. Hence this method is less effective.

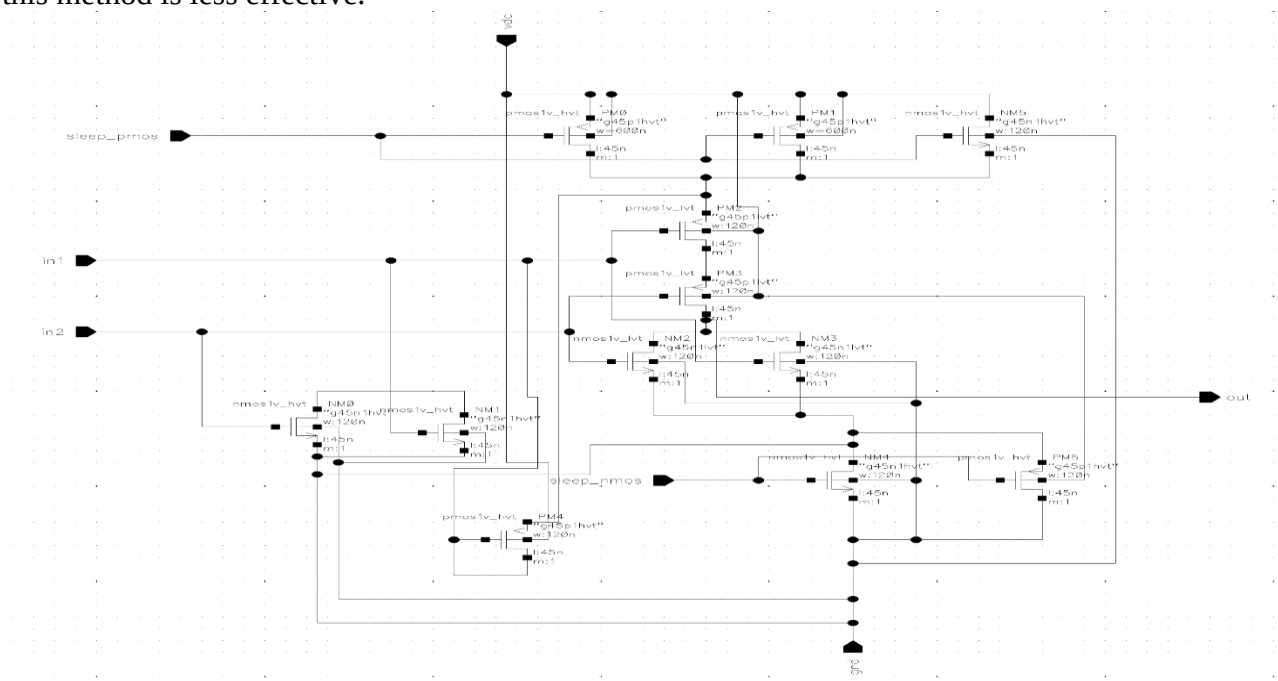


Fig 9.Novel Hybrid NOR gate with retention transistor.

For the proposed techniques, the width of the High Threshold Voltage transistor (HVT) and the list of high voltage transistors used in the circuit are tabulated in table I, II and III.

Table I
HVT TRANSISTORS AND WIDTH OF HVT TRANSISTORS UTILIZED IN NOT GATE

Technique	HVT Transistors	Width of power gating PMOS transistors(nm)	Width of power gating NMOS transistor (nm)
Novel Technique	P0,P1,N1	600	120
Novel Technique with transistor stacking	P0,P1,N2	600	120
Novel Technique with retention transistor	P0,P1,P4,N1,N2	P0,P1-600 P4-120	120

Table II
HVT TRANSISTORS AND WIDTH OF HVT TRANSISTORS UTILIZED IN NAND GATE

Technique	HVT Transistors	Width of power gating PMOS transistors(nm)	Width of power gating NMOS transistor (nm)
Novel Technique	P0,P1,P2,N2	P0,P1-960 P2-120	120
Novel Technique with transistor stacking	P0,P1,P6,N4	P0,P1-960 P6-120	120
Novel Technique with retention transistor	P0,P1,P5,P4,N2,N3	PI,P2-960 P5,P6-120	120

Table III
HVT TRANSISTORS AND WIDTH OF HVT TRANSISTORS UTILIZED IN NOR GATE

Technique	HVT Transistors	Width of power gating PMOS transistors(nm)	Width of power gating NMOS transistor (nm)
Novel Technique	P0,P1,P4,N4	P0,P1-600 P4-120	120
Novel Technique with transistor stacking	P0,P1,P6,N0,N1,N2	P0,P1-600 P6-120	120
Novel Technique with retention transistor	P0,P1,P4,P5,N0,N1,N4,N5	P0,P1-600 P4,P5-120	120

V. SIMULATION RESULTS

We applied different techniques to reduce the dynamic and static power consumption of NOT, NOR and NAND gate. All the design are done and simulation is performed in Cadence design environment using 45 nm technology. The power dissipation is measured during active mode and standby mode of operation for all circuits. In the static mode, power gating PMOS transistor are constantly maintained at 0.5V and power gating NMOS transistor are connected to ground potential. In the active mode, power gating PMOS transistor are connected to ground and NMOS transistors are connected to 0.5V supply. The period and the pulse width of the input are set as 1.5ns and 750ps respectively. The rise time and fall time are set within 50ns for NOT gate and 82ns for NAND and NOR gate.
The static and dynamic power dissipation during all operating modes are measured by using CADENCE result browser and calculator. Table IV and V summarizes the simulated results of standard NOT gate with novel hybrid technique, transistor stacking technique and retention transistor technique. Figures 10 shows the simulated output waveform of hybrid NOT gate.

Table IV
DYNAMIC POWER CONSUMPTION OF NOT GATE

INPUTS	Standard CMOS (nw)	Novel Hybrid technique (nw)	Novel Technique with transistor stacking (nw)	Novel Technique with retention transistor (nw)
0	131.6	41.64	11.87	41.72
1	131.1	41.46	13.78	41.39

From the dynamic power consumption result obtained, we can clearly see that the dynamic power consumption is reduced compared to the standard CMOS NOT gate and the transistor stacking technique have the least power consumption.

Table V
STATIC POWER CONSUMPTION OF NOT GATE

INPUTS	Standard CMOS (pw)	Novel Hybrid technique (pw)	Novel Technique with transistor stacking (pw)	Novel Technique with retention transistor (pw)
0	4.14	13.71	2.389	11.87
1	3.94	2.69	13.68	13.78

From the above result we can see that the leakage power consumption is reduced with the novel techniques.

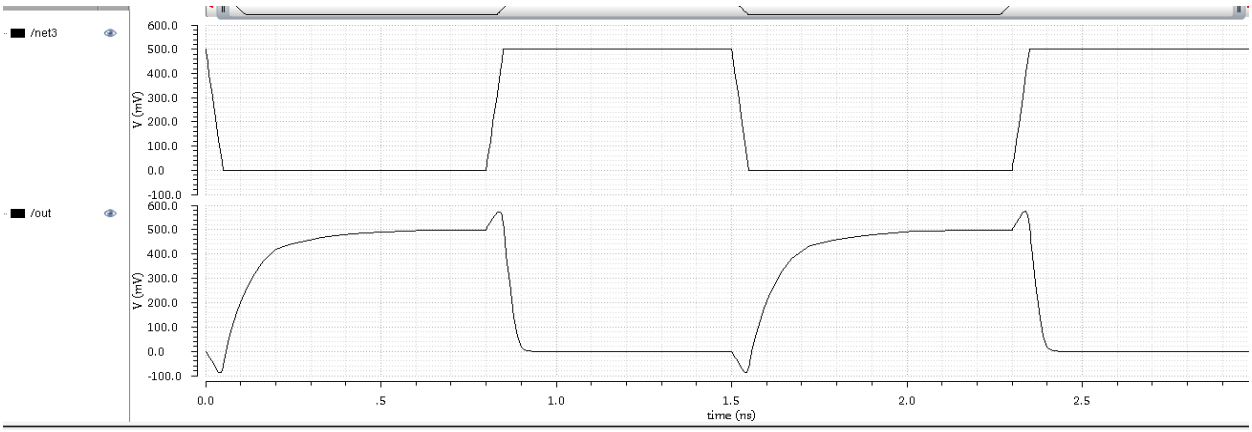


Fig 10.Novel Hybrid NOT gate

Delay of NOT gate with different techniques are given in table VI. Delay is less in retention transistor technique.

Table VI
Delay of NOT gate

INPUT	Standard CMOS (ps)	Novel Hybrid Technique (ps)	Novel Technique with transistor stacking (ps)	Novel technique with retention transistor (ps)
0	120	44	81	44.4
1	174	90	141	85

Table VI and VII summarizes the simulated results of standard NAND gate with novel hybrid technique, transistor stacking technique and retention transistor technique. Figures 11 shows the simulated output waveform of NAND gate with transistor stacking.

Table VI
DYNAMIC POWER CONSUMPTION OF NAND GATE

INPUTS	Standard CMOS (nw)	Novel Hybrid technique (nw)	Novel Technique with transistor stacking (nw)	Novel Technique with retention transistor(nw)
00	204	66.94	78.54	66.88
01	93.85	20.29	26.69	20.09
10	38.61	7.213	6.653	17.1
11	217	66.13	69.57	66.15

Table VII
STATIC POWER CONSUMPTION OF NAND GATE

INPUTS	Standard CMOS (pw)	Novel Hybrid technique (pw)	Novel Technique with transistor stacking (pw)	Novel Technique with retention transistor(pw)
00	1.44	2.951	1.621	4.486
01	4.14	4.423	3.443	12.37
10	3.83	3.922	2.909	11.78
11	7.89	1.499	1.495	1.622

Even in all these techniques, novel technique with transistor stacking consumes less static power. But the dynamic power consumption is less in the case of novel technique with hybrid NAND and retention transistors.

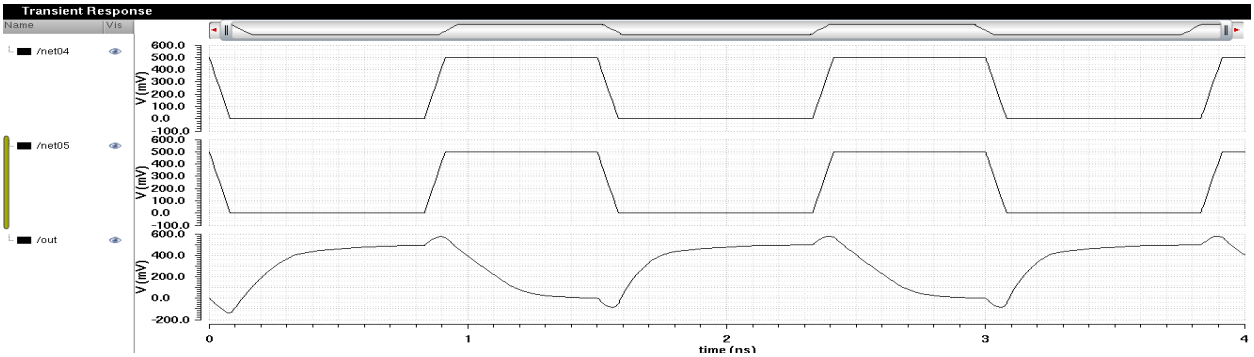


Fig 11 .Novel hybrid NAND gate with transistor stacking

Table VII
DELAY OF NAND GATE

INPUT	Standard CMOS(ps)	Novel Hybrid Technique (ps)	Novel Technique with transistor stacking (ps)	Novel technique with retention transistor (ps)
0	102	122	212	123
1	168	89	189	86.69

Delay is minimum in novel hybrid technique with retention transistor. The simulated result of NOR gate is shown in Table VIII and IX and the output waveform is shown in figure 12.

Table VIII
DYNAMIC POWER CONSUMPTION OF NOR GATE

INPUTS	Standard CMOS (nw)	Novel Hybrid technique (nw)	Novel Technique with transistor stacking (nw)	Novel Technique with retention transistor(nw)
00	198.3	75.91	87.71	76.58

01	48.88	41.85	36.77	37.39
10	103.5	26.88	25.97	22.09
11	188	60.81	49.29	62.25

Here novel technique with retention transistors consumes less dynamic power compared to other technique. Static power consumption is less in case of novel technique with transistor stacking and hybrid technique.

Table IX
STATIC POWER CONSUMPTION OF NOR GATE

INPUTS	Standard CMOS (pw)	Novel Hybrid technique (pw)	Novel Technique with transistor stacking (pw)	Novel Technique with retention transistor(pw)
00	8.289	2.636	2.595	46.06
01	3.946	1.88	2.167	34.37
10	3.686	0.669	0.653	0.703
11	1.951	0.6209	0.63	0.576

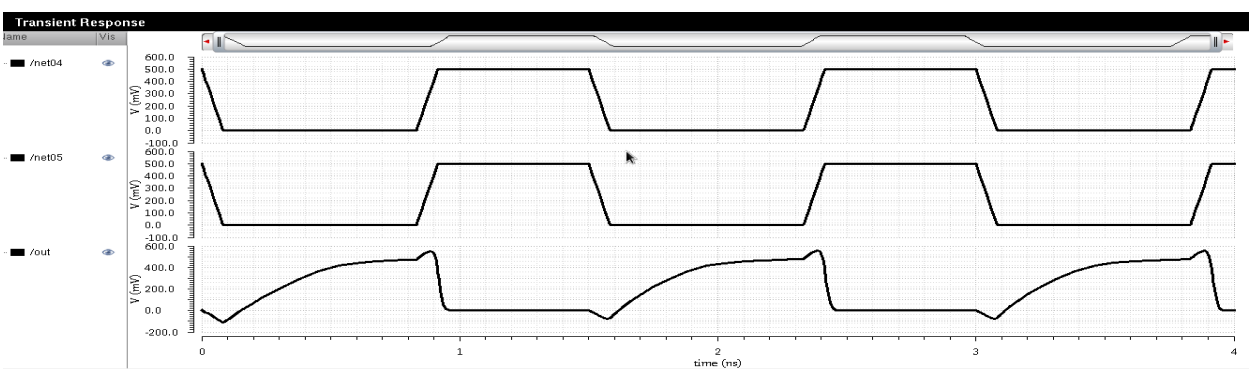


Fig 12. Novel Hybrid NOR gate with retention transistor

Table X
DELAY OF NOR GATE

INPUT	Standard CMOS (ps)	Novel Hybrid Technique(ps)	Novel Technique with transistor stacking(ps)	Novel technique with retention transistor(ps)
0	116.69	46.47	58.32	48
1	334	297.2	646	299

Delay is less in novel technique with retention transistors as extra transistors are added for state retention, helps in quick low to high transition and thereby the delay is minimized as shown in table X

From the ideas obtained we tried to reduce the power consumption of a full adder. For that a 10T full adder is considered. Figure 13 shows schematic of 10T full adder[15]

cell designed using CMOS 45 nm technology consist energy recovering logic reuses charge and therefore consumes less power than non-energy recovering logic. It also uses the sleep transistors to reduce the leakage power. The circuit consists of two xors realized by 4 transistors. Sum is obtained from the output of the second stage xor circuit. The cout can be obtained by multiplexing a and c which is controlled by the output of a xor b. Let us consider that there is a capacitor at the output node of the first XOR module. The charge stored at the load capacitance is reapplied to the control gates. The combination of not having a direct path to ground (depends on input) and the re-application of the load charge to the control gate makes the 10T full adder an energy efficient design. The circuit produces full-swing at the output nodes. But somewhat less to provide so for the internal nodes. A

Table XI
DYNAMIC POWER CONSUMPTION OF 10T FULL ADDER

10T full adder circuit (nw)	10T fulladder circuit using sleep transistor(nw)
62.57	53.9

Table XII
STATIC POWER CONSUMPTION OF 10T FULL ADDER

10T full adder circuit (pw)	10T full adder circuit using sleep transistor(pw)
4.436	0.383

From the result it is clear that the power is reduced by four times in the proposed design.

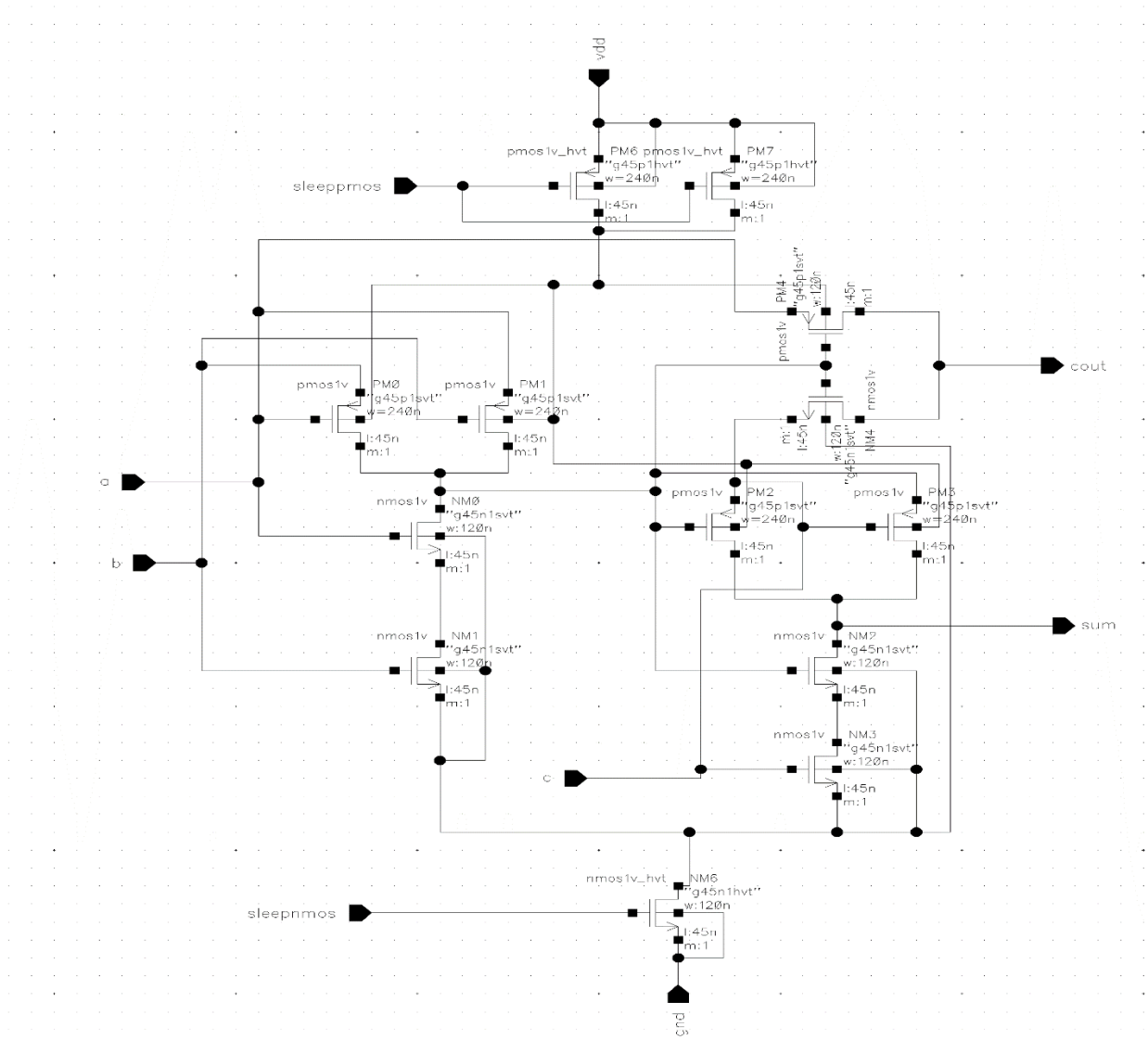


Fig 13.Schematic diagram of 10T full adder with sleep transistors

Conclusion

We have proposed three techniques to reduce the power consumption of logic gates and the design is extended for the performance analysis of 10T full adder. Voltage scaling is one of the most efficient ways for reducing power and energy. For ultra-low voltage operation, techniques which allows bulk CMOS circuits to work in the sub-0.5V supply territory is presented. We are also reducing the leakage power with the help of power gating techniques. The hybrid multiple mode power gating reduces the static power, delay as well as power delay product.

References

[1] Neil H E Weste, David Harris, Ayan Banerjee, "CMOS VLSI Design: A Circuits and Systems Perspective", Third edition, Pearson Education Inc., 2012

[2] Sandip B. Rahane, A.K. Kureshi , Sachin D. Pable "A Review Of Conventional And Emerging Power Gating Techniques For Leakage Power Reduction", International Journal of Innovative Research in Science, Engineering and Technology Volume 3, Special Issue 4, April 2014

[3] Zhaobo Zhang,Xrysovalantis Kavousianos, Krishnendu Chakrabarty and Yiorgos Tsiatouhas, "Static Power Reduction Using Variation –Tolerant and Reconfigurable Multi-Mode Power Switches", IEEE Trans. on VLSI systems,vol.22,No.1,January 2014.

[4] Brajmohan Baghel ,Shipra Mishra , " Analysis & Implementation of Low Power MTCMOS 10T Full Adder Circuit in Nano Scale" , ISSN No: 2309-4893 International Journal of Advanced Engineering and Global Technology I Vol-05, Issue-02, March 2017

[5] Ankita Nagar, Vidhu Parmer, "Implementation of Transistor Stacking Technique in Combinational Circuits" IOSR Journal of VLSI and Signal Processing (IOSR-JVSP) Volume 4, Issue 5, Ver. I, Sep-Oct. 2

[6] Jatin N.Mistry, James Myers, Bashir M AL-Hashimi, David Flynn, John Biggs and Geoff V. Merrett "Active Mode sub-clock power gating", IEEE trans on VLSI systems, September 2013

[7] Amit Bakshi , "Implementation of power gating technique in CMOS full adder cell to reduce leakage power and ground bounce noise for mobile application, " International Journal of Electronics, Communication & Instrumentation Engineering Research and Development (IJECIERD)ISSN 2249-684X Vol.2, Issue 3 Sep 2012 97-108

[8] Shota Ishihara, Masanori Hariyama and Michitaka Kameyama, "A low power FPGA Based on Autonomous Fine grain Power gating", IEEE trans on VLSI sysytems, vol19, N0.8, August 2011.

[9] Seta k,Hara H,Kuroda T,Kakumu.M and Sakurai T, "50% active power saving without sped degradation using standby power reduction(SPR)circuit", In Proceedings of the IEEE International Solid State Circuits conference 318-319,1995.

[10] Kuroda T et al, "A 0.9V 150MHz 10mW 4mm 2-D Discrete Cosine Transform Core Processor with Variable Threshold Scheme", in Digital Technical Papers IEEE International Solid state circuits Conference 88-89, 1998.

[11] Kawaguchi H, Nose K and Saurai T, "A super Cut-off CMOS(SCCMOS)Scheme for 0.5V Supply Voltage with Picoampere Standby Current", IEEE J Solid State Circuits Vol.35,no.10,1498-1501,October 2000.

[12] Balraj Singh 2Mukesh Kumar 3J. S. Ubhi "Analysis of CMOS based NAND and NOR Gates at 45 nm Technology" , International Journal of Electronics, Electrical and Computational System IJEECS ISSN 2348-117X Volume 6, Issue 4 April 2017

[13] Pramod Kumar M P,A S Augustine Fletcher , "A Survey on leakage Power Reduction Techniques by using power gating Methodology" ,International journal of engineering Trends and technology, IJETT volume 9 number 11-March 2014.

[14] S .Karunakaran 1 and B. Poonguzharselvi "Exploration on Power Delay Product of Basic Logic Gates for various CMOS Logic Styles", International Journal of Engineering Studies. ISSN 0975-6469 Volume 9, Number 2 (2017), pp. 111-120

[15] K .Dhanunjaya, Dr.M N.Giri Prasad, Dr.K.Padmaraju , "Performance analysis of low power full adder cells using 45nm CMOS Technology", International Journal of Microelectronics Engineering (IJME), Vol. 1, No.1 , 2015

Clustering with modified mutation strategy in Differential Evolution

Seema Patil[*], R. J Anandhi [**]

[*] *Department of Computer Science & Engineering, The Oxford College of Engineering*
[**] *Department of Information Science & Engineering, New Horizon College of Engineering*

Abstract- In this paper, a clustering approach based on modified mutation strategy in the Differential Evolution has been proposed. The objectives of modification are to achieve high rate of convergence and to obtain better cluster efficiency. The proposed form of modification has been applied on probabilistic environment to define the differential vector through randomly selected members and the best solution has been obtained. Over number of benchmark dataset, clustering efficiency have been estimated and compared with Conventional Differential Evolution as well as Particle Swarm Optimization. The proposed solution has delivered the superior and consistent performance over the considered benchmark.

Index Terms-Clustering, Convergence, Differential evolution, Mutation, Particle swarm optimization

I. INTRODUCTION

The tremendous growth of data-based knowledge in scientific studies has presented lot of challenges before the researchers to extract useful information from them using traditional data base techniques. Hence effective mining methods are essential to discover the implicit knowledge from huge data warehouses. Data based knowledge offer numerous opportunities in various practical applications like bioinformatics, engineering, biology, healthcare, medicine, prediction analysis, forecasting the crime and various computing techniques.

To perform this, knowledge extraction is done with the help of data mining techniques such as classification and clustering. The important task of combining various population or data points into clusters is clustering which performs similarity of points. It is one of iterative process of discovery of knowledge which involves major trial and failure. The clustering process does not require any kind of feedback to perform similarity of data points, it is self-organized [1]. Clustering defines a new swarm intelligence (SI) for partitioning any datasets into an optimal number of groups through one run of optimization. SI is an innovative distributed intelligent paradigm for solving optimization problems that originally took its inspiration from the biological examples by swarming, flocking and herding phenomena in vertebrates.

Data clustering is a popular approach of automatically finding classes, concepts, or groups of patterns. Particle Swarm Optimization (PSO) incorporates swarming behaviors observed in flocks of birds, schools of fish, or swarms of bees, and even human social behavior, from which the idea is emerged. Data clustering using PSO can be used to find the centroids of a user specified number of clusters. For automatic clustering of large unlabeled data sets, Differential Evolution (DE) is used. [2]

This work proposed the method for clustering, based on differential evolution. Even though DE is very efficient, but sometimes it suffers from the issue of slow convergence and difficulties in achieving the global solution. To overcome these, balance between exploration and exploitation has been maintained by adding the two modules in the conventional DE. To increase the level of exploitation, under the probabilistic mode, selection between best and randomly selected member takes place. The Differential vector made by best solution, deliver the fast change in the solution and results in faster convergence. The

multi-culture approach helps in exploration of new and efficient solution. Gathering and selection of solution from different environments will maintain the diversity in the population.

II. RELATED WORK

The author Gupta [3] et al., has proposed a new efficient clustering approach which was applied on k harmonic means (KHM) by using PSO. The local optimum problem of KHM was overcome by PSO. Also, fuzzy logic was used to control the various parameters of PSO. The author Pranav [4] et al., has achieved the global optima on clustering by making use of two validation indices criteria. These indices were simple and robust against other outliers and shown best clustering which has lower computation cost and parallel execution and faster convergence. The author Wang [5] et al., combines PSO and DE approach by taking velocity update of PSO and mutation parameter of DE to generate the new population. The DE re-mutation, crossover and selection are performed throughout the optimization process to get the good results. This approach gives the best result compared to inertia weight PSO and comprehensive learning PSO and basic DE. The author Zhu et al., [6] has discussed complications associated with K-means clustering algorithm and centroid all rank distance concept has been presented. To overcome the difficulties associated with density and delta-distance clustering (DDC) when data derived from the two indicators are large, an efficient and intelligent DDC algorithm has been discussed by author Liu et al [7]. A robust recommendation algorithm based on kernel principal component analysis and fuzzy c-means clustering has been presented by author Huawei et al., [8]. The author has presented a variation of differential evolution (DE) algorithm to solve an automatic clustering problem [9]. The author [10] describes the new improved approach of PSO by improving the diversity mechanism and mutation operator to employ new neighborhood search strategy. These new approaches were tested on well-defined benchmark data sets. Based on matrix partitioning a hierarchical clustering algorithm has been presented in [11].

III. PROPOSED WORK

A. Modified Mutated DE (MMDE)

To increase the convergence speed of DE, a new approach in mutation operation has been presented. It has two possibilities of differential change under the probabilistic environment. In the first case, differential change is defined through best member and random selected member while in second case three random members are selected to define the differential change. A threshold value is defined to determine the selection of differential change type. Best member based differential change generate the faster change, while the random member-based selection tries to prevent from suboptimal convergence. The pseudo code for applied mutation strategy has been shown below.

- Define the Threshold value (*Thr*)
- r = U [0, 1]; a random number generated through uniform distribution in range of [0 1];
- *if* $r < Thr$
  - Select two members' m1 & m2 randomly from population
  - Select best member BM from population
  - Mutation vector defined as: Mv = m1+ mf* [BM- m2];

  Else
  - Select three members m1, m2 & m3 randomly from population
  - Mutation vector defined as: Mv= m1+ mf*[m2-m3]
- *End*

B. Multi-domain-based DE

A multi-culture concept called "Multi-culture modified mutation Differential Evolution" has been developed to evolve the individual population independently and later exploit to form a better community to search the solution space efficiently. This approach is inspired very much by present human society, where at fundamental level two things happen (i) the independent existence of a number of separate population, and they get their progress under the same environment up to a certain period of time. (ii) with respect to objectives, a number of individuals are selected from the different population and form a new population to achieve the objectives. Rather than working under monoculture formed by one population as in conventional PSO, multiculture environment has been proposed, where a number of different environments created by a different set of population independently. Each population has evolved socially, independently to generate the multiculture and later among all, best individuals are selected to finish the task. This is a dual stage process where first stage finds some potential solution discovered from different regions of solution space, and later in the second phase, each individual contributes more efficiently to find a global solution. Even with the small size of the population, the proposed method has achieved better quality solution with the very high value of consistency.

In the working principle of MMDE, population (POP) are the initial random population, which is evolved by the DE process individually and independently for a fewer number of iterations and creates the multi-culture new population (NPOP). Even though the process of creating the NPOP is same for all POP, because of difference in leadership and different community surrounding, each NPOP has different characteristics. Through the fitness-based selection process, among all members from all NPOP, better members are selected to form a new population (SPOP), which has the same size as initial POP. In SPOP, there are a number of good candidates, which are different and have higher fitness value, hence the high level of diversity exists. Finally, over SPOP, MMDE has been applied till terminating criteria has not meet, to obtain the Final Population (FPOP).

IV. EXPERIMENTAL RESULTS AND ANALYSIS

For the data set namely "Wine data"," Iris", and "Glass" data set which are available in UCI repository[12] have been considered to analyze the work. In the first part, only the MMDE has been applied and performances have been obtained for 5 independent trials. Comparison has been made with conventional DE(CDE) and dynamic weighted PSO(DYPSO). For all the cases, the size of population has been considered as 100, mutation rate and crossover rate as 0.4 and 0.5. The allowed number of iterations were 600.The performances have been represented in terms of correctly placed data samples in the clusters, number of data samples placed wrongly, cluster efficiency and total intra cluster distance value. In second part, multidomain based experiment has been included with MMDE and performances have been estimated over "Glass" data set. Experimental process has been developed in the MATLAB environment.

A. Dataset: Wine Data

There are total 178 set of data carrying 3 clusters. Each data contains 13 attributes.

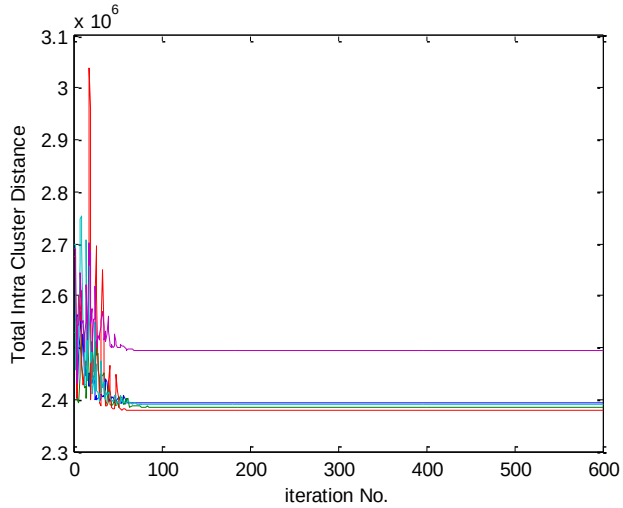


Fig.1 DYPSO based convergence in 5 trials for winedata set

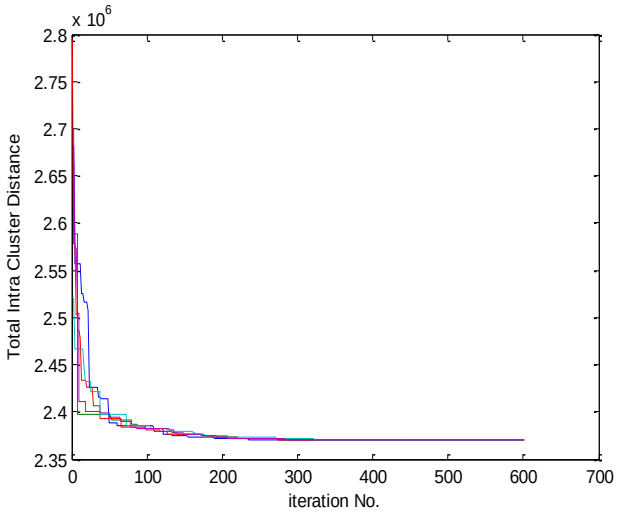


Fig.2 CDE based convergence in 5 trials for winedata set

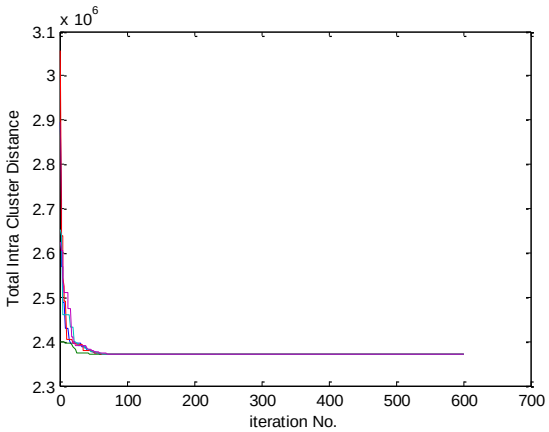


Fig.3: MMDE based convergence in 5 trials for winedata set

Table1: Mean Performance over 5 trials by different algorithm over winedata set

	Correctly clustered data samples	Wrongly clustered data samples	Clustered efficiency	Total Intra Cluster Distance value 1.0e+006 *
DWPSO	125	53	70.22	2.4088e+006
CDV	125	53	70.22	2.3707e+006
MMDV	125	53	70.22	2.3707e+006

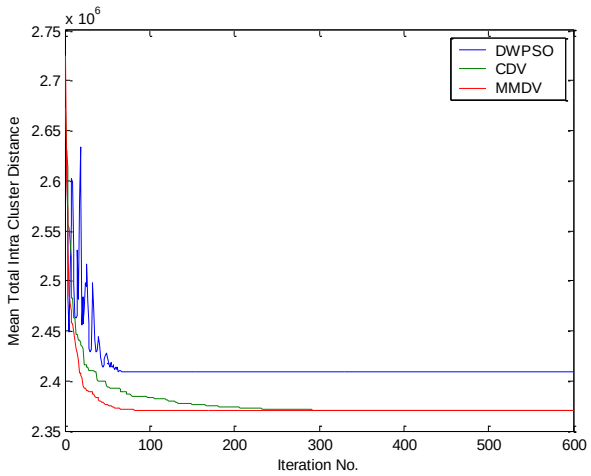
Table 2: Centroid position for winedata

C1	3.0351	3.0067	3.0065	3.0541	3.2816	3.0057	3.0043	3.0108	3.0041	3.0154	3.0024	3.0067	4.9797
C2	3.0375	3.0051	3.0065	3.0462	3.2867	3.0078	3.0081	3.0008	3.0051	3.0154	3.0029	3.0084	6.2486
C3	3.0339	3.0067	3.0062	3.0565	3.2508	3.0057	3.0048	3.0010	3.0040	3.0111	3.0024	3.0067	4.2455

The performances obtained under 5 independent trials by different algorithms have been shown in Table1.It can be observed that all the three algorithms have nearly the same performances, while there is little more distance measure appeared for the DYPSO. The obtained centroid value by MMDE for 1st trail have been shown in Table2.The convergence characteristics for DYPSO, CDE and MMDE have been shown in Fig.1 to Fig.3.To get the relative convergence speed, Fig.4 has plotted the mean convergence characteristics. Proposed MMDE has shown the fastest rate of convergence while DYPSO was the poorest.

B Dataset: IRIS Data

Contain total 150 data set and each data has 4 attributes. Three different global clusters exist in dataset. The convergence performances of DYPSO, CDE and MMDE have been shown in Fig. 5 to Fig.7, while the statistical performances have been shown in Table 3 to Table 5. It can be observed that MMDE has shown very consistent performance in all trials and in Fig.8 comparative convergence has been shown. The obtained best value of centroid has been shown in Table 6.



Fig,4: Mean convergence comparison for Iris data set

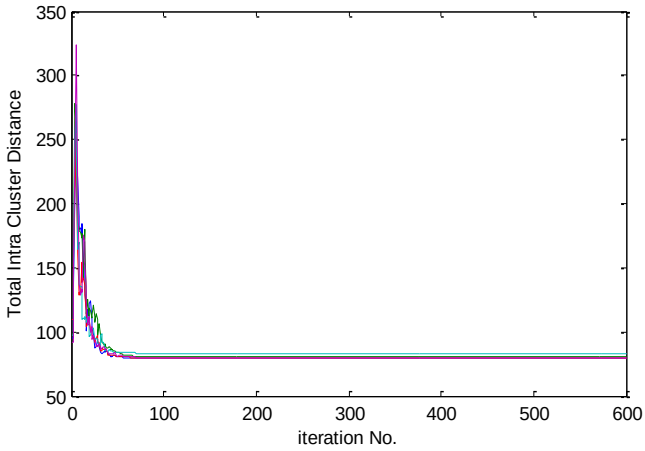


Fig.5: DYPSO based convergence in 5 trials for Iris data set

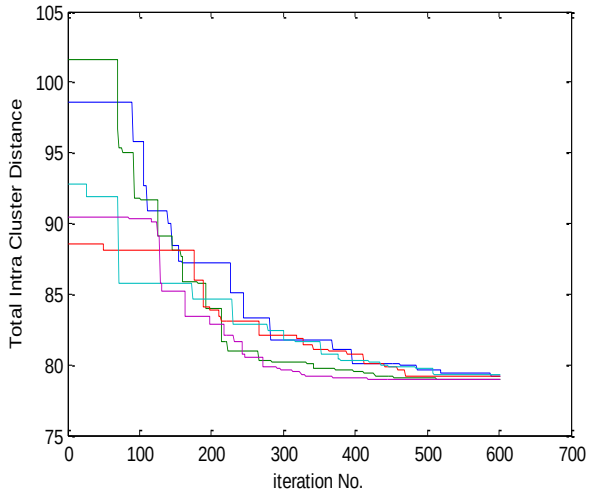


Fig.6: CDE based convergence in 5 trials for Iris data set

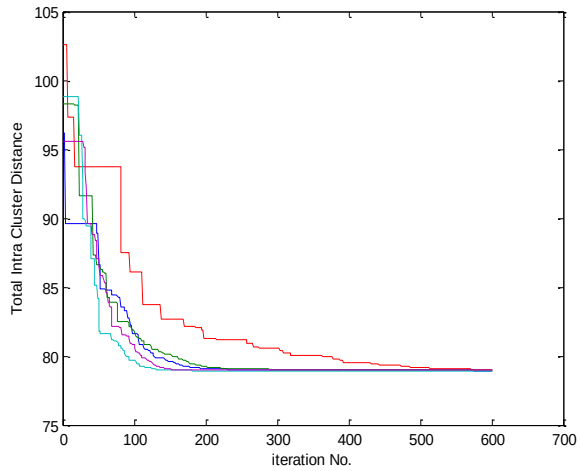


Fig.7: MMDE based convergence in 5 trials for Iris data set

Table 3: DYPSO performance over Iris data

Trial No. IRIS(PSO)	Correctly clustered data samples	Wrongly clustered data samples	Clustered efficiency	Total Intra Cluster Distance value
1	134	16	89.33	79.3157
2	134	16	89.33	80.2949
3	133	17	88.67	79.4755
4	136	14	90.67	83.2333
5	133	17	88.67	79.7068
Mean	**134**	**16**	**89.33**	**80.4052**

Table 4: CDE performance over Iris data

Trial No. IRIS(CDV)	Correctly clustered data samples	Wrongly clustered data samples	Clustered efficiency	Total Intra Cluster Distance value
1	134	16	89.33	79.2028
2	134	16	89.33	78.9563
3	133	17	88.67	79.1462
4	134	16	89.33	79.2389
5	134	16	89.33	78.9430
Mean	**133.8**	**16.2**	**89.2**	**79.0974**

Table 5: MMDE performance over Iris data

Trial No. IRIS(MMDV)	Correctly clustered data samples	Wrongly clustered data samples	Clustered efficiency	Total Intra Cluster Distance value
1	134	16	89.33	78.9471
2	134	16	89.33	78.9631
3	134	16	89.33	79.0133
4	134	16	89.33	78.9454
5	134	16	89.33	78.9494
Mean	**134**	**16**	**89.33**	**78.9637**

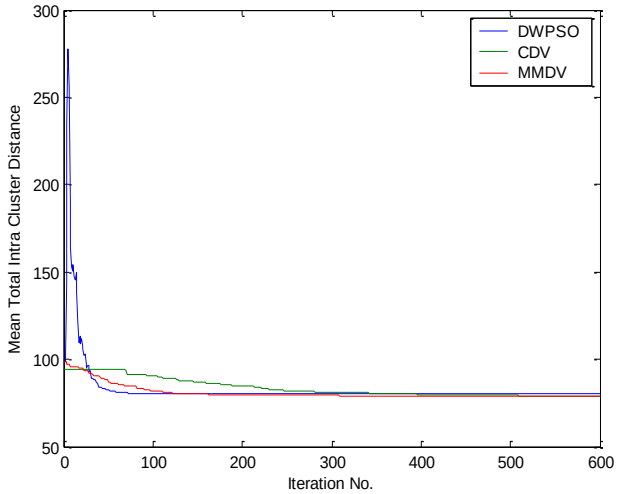


Fig.8: Mean convergence comparison for Iris data set

Centroids of IRIS Dataset				
C1	5.8863	2.7456	4.3731	1.4115
C2	5.0173	3.4385	1.4452	0.2704
C3	6.8326	3.1128	5.7640	2.0469

Table 6: Centroids value for Iris data set

C. Dataset: Glass Data

This data set contains total 214 data set. Each data set carried 10 attributes and 6 clusters exists.

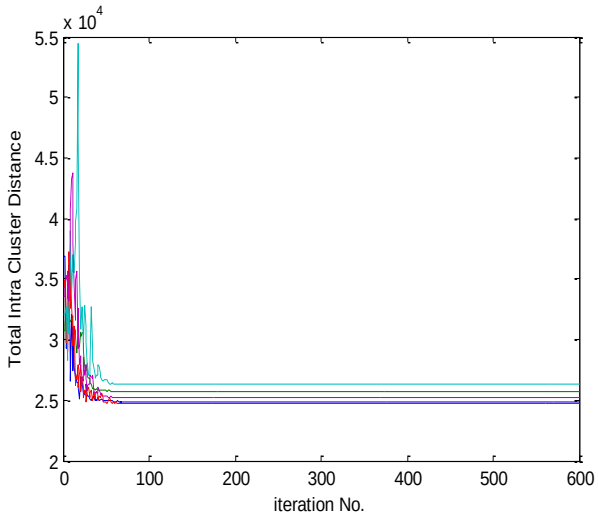


Fig.9: DYPSO based convergence in 5 trials for Glass data set

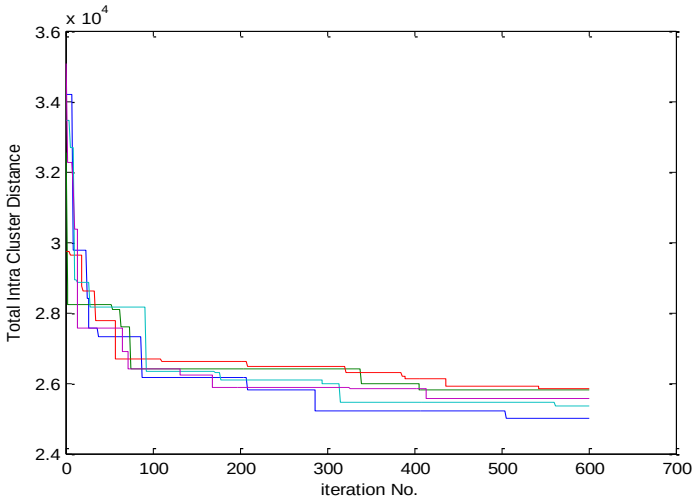


Fig.10: CDE based convergence in 5 trials for Glass data set

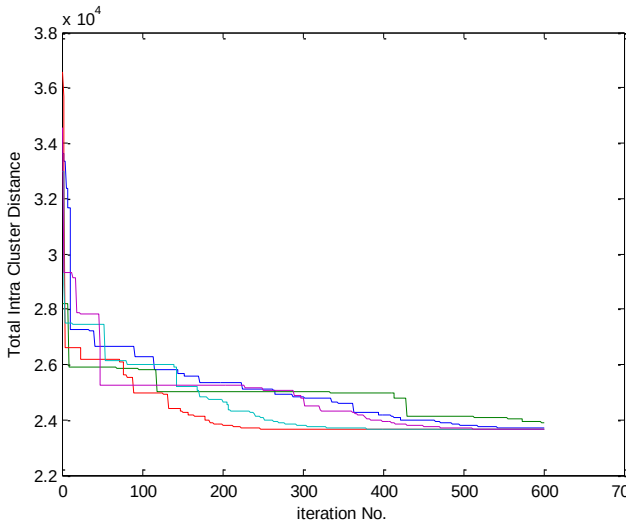


Fig.11: MMDE based convergence in 5 trials for Glass data set

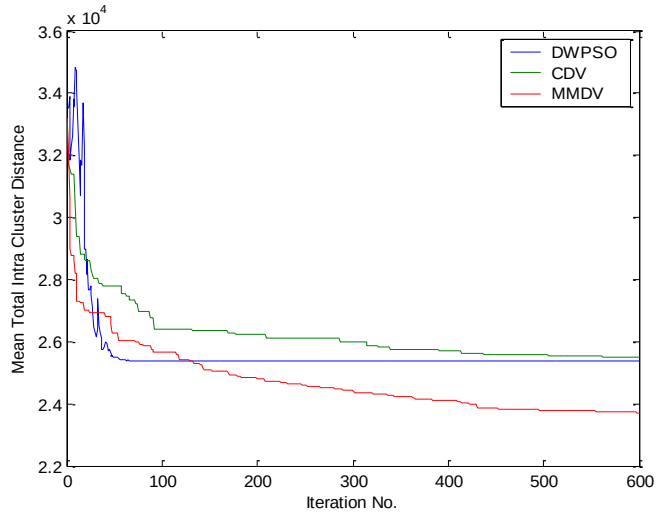


Fig.12: Mean convergence comparison for Glass data set

Table 7: DYPSO performance over Glass data

Trial No. Glass (PSO)	Correctly clustered data samples	Wrongly clustered data samples	Clustered efficiency	Total Intra Cluster Distance value
1	183	31	85.51	2.4897 e+004
2	189	25	88.32	2.5737 e+004
3	178	36	83.18	2.4721 e+004
4	184	30	85.98	2.6271 e+004
5	188	26	87.85	2.5209 e+004
Mean	**184.4**	**29.6**	**86.17**	**2.5367e+004**

Table 8: CDE performance over Glass data

Trial No. Glass (CDE)	Correctly clustered data samples	Wrongly clustered data samples	Clustered efficiency	Total Intra Cluster Distance value
1	183	31	85.51	2.4990 e+004
2	189	25	88.32	2.5797 e+004
3	178	36	83.18	2.5850e+004
4	184	30	85.98	2.5368 e+004
5	188	26	87.85	2.5546 e+004
Mean	**184.4000**	**29.6000**	**86.17**	**2.5510e+004**

Table9: MMDE performance over Glass data

Trial No. Glass (MMDE)	Correctly clustered data samples	Wrongly clustered data samples	Clustered efficiency	Total Intra Cluster Distance value
1	187	27	87.38	2.4990 e+004
2	187	27	87.38	2.5797 e+004
3	187	27	87.38	2.5850 e+004
4	189	25	88.32	2.5368 e+004
5	184	30	85.98	2.5546 e+004
Mean	**186.8**	**27.2**	**87.29**	**2.5510e+004**

Table10: Centroids value for Glass data set

C1	166.0782	2.4471	13.7061	3.5266	2.2563	73.3031	2.4611	10.7421	-0.1976	0.5747
C2	198.4844	2.5638	16.2827	3.2212	2.7751	73.5565	1.7972	9.9803	1.6024	-0.1853
C3	54.2369	2.1344	14.2542	4.4666	1.9043	72.6730	1.0457	9.7003	1.4352	0.2335
C4	18.5031	2.1863	13.2582	4.4278	1.5191	74.4194	1.3567	10.2181	0.4565	10.1096
C5	129.9205	0.8875	13.9521	4.3390	2.7228	75.5818	0.9168	8.7067	1.4468	1.4522
C6	91.0957	2.8459	14.1901	3.6017	2.9122	72.2789	0.9257	10.0617	0.7071	1.1787

For the Glass data set the obtained convergence characteristics have been shown in Fig.9 to Fig.11. Comparative mean convergence has been shown in Fig.12. It can be observed that, in spite of more number of clusters, superior convergence has appeared. The obtained statistical performance has been shown in Table7 to Table9. For MMDE, maximum cluster efficiency has been obtained. The obtained best centroid value has also been shown in Table10.

D. Multidomain based MMDE

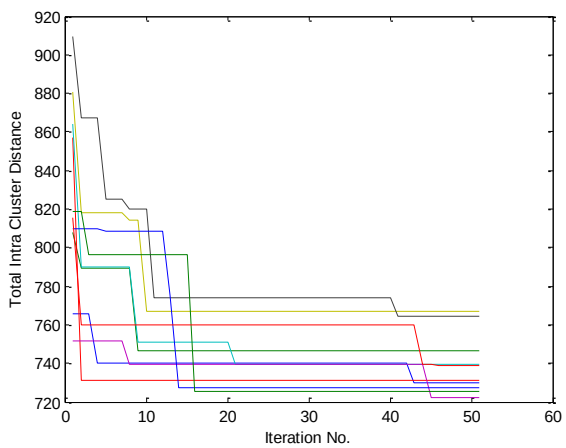


Fig.13 Convergence characteristics in 1st Stage for multidomain MMDE

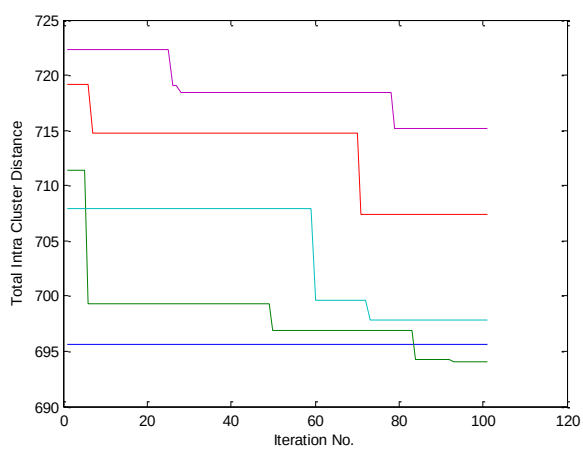


Fig.14 Convergence characteristics in 2nd stage for multidomain MMDE

Convergence characteristics over Glass data set for multidomain MMDE has been shown in Fig.13, for the 1st stage and in Fig.14 for the 2nd stage. The obtained performances have been shown in Table11. It can be observed that maximum efficiency 87.48% has been obtained. The corresponding centroid value has also been shown in Table 12.

Table11: Multidomain MMDE performance over Glass data

Trial No. (MMDE) GLASS	Correctly clustered data samples	Wrongly clustered data samples	Clustered efficiency	Total Intra Cluster Distance value
1	188	26	87.85	695.5811
2	188	26	87.85	694.0454
3	189	25	88.32	707.4350
4	190	24	88.79	697.8723
5	181	33	84.58	715.1624
Mean (Std.Dev)	**187.2 (3.5637)**	**26.8 (3.5637)**	**87.48 (0.1252)**	**702.0192 (9.042)**

Table12: Centroid values by Multidomain MMDE

C1	16.0000	1.5165	13.4754	3.3530	2.4072	74.6342	0.0100	8.7993	0.0894	0.2050
C2	201.3622	1.5122	14.7074	0.1029	1.2528	72.3216	0.1859	8.6580	1.3473	0.0031
C3	165.4855	1.5189	12.7370	2.3479	2.1774	71.8032	0.7419	7.7070	0.2396	0.0068
C4	48.0214	1.5246	11.9324	4.4900	1.1781	72.9279	0.7290	9.8281	0.0987	0.0876
C5	88.8809	1.5116	13.4721	3.3903	1.0875	72.9210	0.3255	7.9812	0.0100	0.1157
C6	127.1936	1.5134	13.9751	3.8544	1.4775	73.6876	0.2323	9.0625	0.0100	0.1454

E. Comparative study of MMDE with K-Means

Comparative performance between Multi-Domain MMDE and K-Means over all the three different data sets have been shown in Table13-15. For each data set 5 independent trials have been applied. It can be observed with outcomes that the problems with K-Means algorithm are twofold. First it may not deliver the optimal performances, second, there is high level of variations in the performances over trails which is really a serious issue from the practical point of view. This happens because of sensitivity of K-Means algorithm towards initialization. Whereas the proposed method Multi-domain MMDE has delivered not only better performance because of exploration but also variation level is very less.

Table 13: Comparative Performance of MMDE and K-means for Wine Data

WineData	Multi-Domain		K-Means	
Trial	MMDE Samples		K means Samples	
	Correctly clustered	Wrongly Clustered	Correctly clustered	Wrongly Clustered
1	125	53	125	53
2	125	53	120	58
3	125	53	120	58
4	125	53	120	58
5	125	53	120	58
Mean	**125**	**53**	**123.75**	**54.28**
Efficiency	**70.22**		**67.98**	

Table 14: Comparative Performance of MMDE and K-means for Iris Data

Iris Data	Multi-Domain		K-Means	
Trial	MMDE Samples		K means Samples	
	Correctly clustered	Wrongly Clustered	Correctly clustered	Wrongly Clustered
1	135	15	134	16
2	134	16	134	16
3	137	13	100	50
4	133	17	134	16
5	134	16	100	50
Mean	**134.6**	**15.4**	**120.4**	**29.6**
Efficiency	**89.73**		**80.27**	

Table 15: Comparative Performance of MMDE and K-means for Glass Data

Glass Data	Multi-Domain		K-Means	
Trial	MMDE Samples		K means Samples	
	Correctly clustered	Wrongly Clustered	Correctly clustered	Wrongly Clustered
1	188	26	187	27
2	188	26	187	27
3	189	25	187	27
4	190	24	187	26
5	191	33	187	27
Mean	**187.2**	**26.8**	**187**	**26.8**
Efficiency	**87.48**		**87.38**	

V. CONCLUSION

In this paper, a modified mutation strategy for differential evolution has been proposed to facilitate the clustering requirement of data. This modification increases the convergence rate and deliver the cluster efficiency up to the mark. To increase the level of exploration, two stage based a multimodal structure has also been proposed. With this structure, the bias variation sensitivity of cluster activity decreased. Number of benchmarks have been tested which had the number of clusters from 2 to 6 to ensure the generalize capability. Proposed solution has outperformed the conventional form of DE as well as dynamic weighted form of PSO. Proposed work has been evaluated only using datasets of UCI Repository, further it can be applied on application oriented dataset to evaluate performance.

ACKNOWLEDGMENT

I would like to offer my special thanks to Dr. R J Anandhi, Professor and Head, Department of Information Science and Engineering, New Horizon College of Engineering, Bengaluru, my research supervisor, for her patient guidance, enthusiastic encouragement and useful critiques of this research work. I would also like to thank Dr. R V Praveena Gowda, Principal, The Oxford College of Engineering, Bengaluru, for his advice and assistance in keeping my progress on schedule. Finally, I wish to thank my husband and parents for their support and encouragement throughout my research work.

REFERENCES

[1] "Nuria Gómez Blas, Octavio López Tolic, "Clustering using Particle Swarm Optimization", International Journal Information Theories and Applications. 23, pp24-33, 2016

[2] Swagatam Das, Ajith Abraham, "Automatic Clustering Using an Improved Differential Evolution Algorithm", IEEE Transactions on Systems, Man, And Cybernetics—Part A: Systems And Humans, Vol. 38, NO. 1, pp 218-237, JANUARY 2008

[3] Yogesh Gupta, Ashish Saini, "A new swarm-based efficient data clustering approach using KHM and fuzzy-logic", Soft Computing, Springer, pp 145-162, 2019.

[4] Pranav Nerurkar, Aruna Pavate, Mansi Shah and Samuel Jacob, "Performance of Internal Cluster Validations Measures for Evolutionary Clustering", Computing, Communication and Signal Processing, Springer pp 305-312 2019

[5] H. Wang, L. L. Zuo, J. Liu, W. J. Yi, B. Niu1, "Ensemble particle swarm optimization and differential evolution with alternative mutation method", Natural Computing, Springer Nature pp 1-14,2018

[6] Yehang Zhu,Mingjie Zhang,Feng Shi, "Application of Algorithm CARDBK in Document Clustering", Wuhan University Journal of Natural Sciences,December 2018, Volume 23, Issue 6, pp 514–524

[7] Xuejuan Liu,Jiabin Yuan,Hanchi Zhao, "Efficient and Intelligent Density and Delta-Distance Clustering Algorithm", Arabian Journal for Science and Engineering,December 2018, Volume 43, Issue 12, pp 7177–7187

[8] Huawei Yi "Robust Recommendation Algorithm Based on Kernel Principal Component Analysis and Fuzzy C-means Clustering", Wuhan University Journal of Natural Sciences,April 2018, Volume 23, Issue 2, pp 111–119.

[9] Kuo, R.J. & Zulvia, F.E. "An improved differential evolution with cluster decomposition algorithm for automatic clustering", Soft Computing, Springer, pp 1–17,2018.

[10] Dang Cong Trana,c,, Zhijian Wua and Changshou Dengb, "An improved approach of particle swarm optimization and application in data clustering", Intelligent Data Analysis 2015 pp 1049–1070

[11] Hewijin Christine Jiau,"MPM: a hierarchical clustering algorithm using matrix partitioning method for non-numeric data", Journal of Intelligent Information Systems,May 2006, Volume 26, Issue 3, pp 303–306.

[12] https://archive.ics.uci.edu/ml/index.php

Automated Catalog Management and Image Quality Assessment using Convolution Neural Networks and Transfer Learning

Souradip Chakraborty*, Mani Kanteswara Rao Garlapati*

** Data & Analytics, Walmart Labs, Bangalore*

Abstract- Catalogue management is a very important aspect in the field of ecommerce as it helps the visitors in efficiently selecting the necessary interest items. In an online store, customers are unable to touch the product before buying it and this can only be compensated by providing a good sensory experience through image catalogue and efficient management of the same. In every retail website, all the items in the catalogue are in a particular order of different categories. In this work, we have developed an entire pipeline where the first task is to automatically classify the various orientations (front view, side view, top view etc.) of the images sent by the vendor using CNN and Transfer learning. In the second part of our pipeline, we have eased the process of catalogue management with the image quality assessment of the vendor images using No reference image quality assessment and finally the automatic ordering of items are done as per thresholding. Good quality images from all orientations plays a critical role in making a customer-friendly online store leading to customer satisfaction.

Index Terms- Convolution Neural Networks, Transfer Learning, Image quality assessment, Structural similarity index, Quality Embedding.

I. INTRODUCTION

Efficient Catalogue management is very important and vital for ecommerce retailers since it helps online visitors in selecting the necessary items and if the catalogues are well organized it serves as a great aid for the customers which help them in turning to loyal customers. Many research works have been done in the field of image classification using convolution neural network [1] and Transfer learning [2], but very few works have been done using a combination of both in classification of various orientations (different views like side view, front view etc..) of images of items sent by vendors which is being done as a part of catalogue management in this work and hyperparameter tuning has been done using Bayesian optimization [3] which gave much superior results when compared to the baseline model. Since manual/decision rule based ordering of the images sent by vendors are being done in majority of industries currently which is extremely time-consuming and hence it can be improved vastly by the our methodology . Secondly, quality of the images sent by the vendor plays a crucial part as improper image quality in an online platform might directly lead to customer dissatisfaction. The way human perceives image quality is very unique and to make

the machine understand and learn that way makes image quality assessment a very difficult task to perform[4]. Hence, Structural similarity index [5] has been considered as a metric in this case for assessment of the quality of images of items sent by vendor which gives the human-perceived notion of quality . The major challenge faced with respect to quality is the blurring effect in images sent by vendors which also is one of the primary cause of customer dissatisfaction as understood from various customer feedback and surveys .Hence the second phase of the pipeline deals with assessing the quality of the image automatically once it falls below a certain quality threshold. The main contribution of the paper lies in the development of quality embeddings which projects each and every image in some latent dimensions which represents various quality attributes and using the same ,the human perceived quality metric has been predicted for every image. The concept of quality embeddings have not been used before and it helps in no-reference image quality assessment task efficiently. Another contribution of the paper lies in the synthetically generated noisy datasets which eliminated the manual annotation process very effectively and helps in the no-referencing quality assessment. The key idea lies in the concept that human beings while detecting if an image is of poor quality or not doesn't need the true reference superior quality version of the image. If an image is a lit blurred, human beings are well adept in detecting the same and hence for machines to reflect the same intuition , the above methodology has been implemented. Bayesian optimization has been leveraged in the process of hyperparameter tuning which reduces the time complexity of the pipeline significantly and provides an intelligent approach to search the best hyperparameters in the given space.

II. IDENTIFY,RESEARCH AND COLLECT IDEA

There is lot of research work that has happened over the years in the field of image classification and orientation detection, but in majority of the models developed there is a requirement that the dataset size should be large enough since deep convolution based models will have a lot of parameters. The complexity lies in this case since there will be multiple new items for which dataset size won't be large and the model thus developed should be robust enough to work in such scenarios as well. It also includes other constraints such as time complexity, simplicity and light weight models for the pipeline to work optimally. The creation of robust features using light weight MobileNet CNN helps in achieving the objective of orientation classification.

There has been work done in the field of image quality assessment but in many of the cases manually annotation of datasets have been used. In our case , we have synthetically generated noisy datasets which reduces the manual efforts of annotating. The quality embeddings developed in our architecture has never been used/developed till date for image quality assessment.

The dataset that has been used for image orientation classification consists of 3 classes- Front view, Side view and Top view and the size has been kept low to meet the constraints mentioned earlier. The dataset consists of 312 images in total out of which 95 of back view, 108 of front view and 109 of side view images have been used to train. The challenge was to show good accuracy even with small datasets.

Figure 1: Back, Front and Side view of the images trained

III. STUDIES AND FINDINGS

Image Orientation Classification using Convolution Neural Network and Transfer Learning

Histogram of Oriented Gradients(HOG) as Baseline Model:

The HOG features are mainly used in image processing object detection tasks The key idea behind the histogram of oriented gradients descriptor is that local object appearance within an image can be described by the distribution of edge directions. The image is divided into small connected regions and for the pixels within the regions, a histogram of gradient directions are computed. The final feature vector is the combination of all these histogram features.

For implementation of the task of classification of image orientation into one of the 3 categories, the baseline model that has been used is with the histogram of oriented gradients features as it has been widely used in many places where image orientation classification is the prime objective [7].

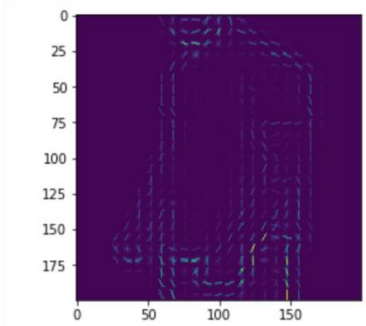


Figure 2: Histogram of oriented gradient features of Image side view

Using Histogram of oriented gradient features as predictors, 5 different classification models were fitted to the training data and for each of the models, the ideal hyper parameters were computed using Bayesian Optimization of hyper parameters [3], the convergence plot of the same (sample) is shown in Figure 3.

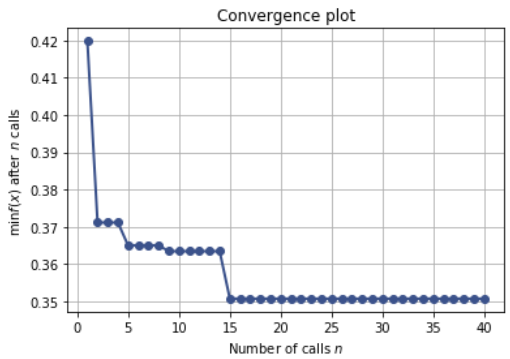


Figure 3: Convergence plot of model hyper parameters in Bayesian Optimization

The cross-validation accuracy of each of the models thus computed is shown below in Table 1.

Table1: Cross Validation accuracy of various classification models with Histogram of oriented gradient features

Classifiers	Cross-Validation Accuracy
SVM	62.22%
Multinomial Logistic	71.23%
Naïve Bayes	62.12%
Decision Tree	55%
Random Forest	70%

Bayesian Optimization helps in reducing the time complexity associated with grid search for the hyperparameters significantly as it implements an intelligent way of searching the space using Gaussian process. So, at each iteration it implements a trade-off between exploration and exploitation and thus forms an utility function and optimizing the same it chooses the next best hyperparameter.

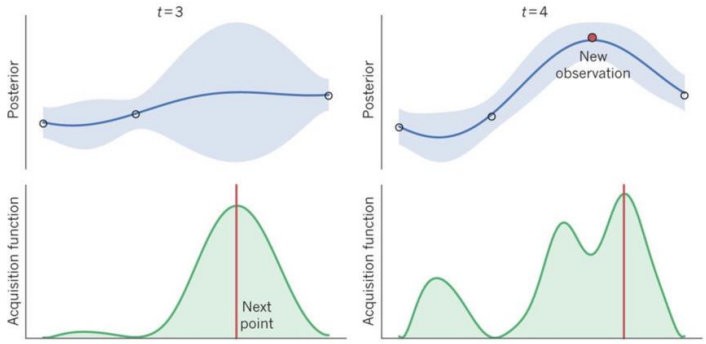


Figure 4: Bayesian Optimization and Gaussian Process

But as can be seen in Table1, HOG features fail to give a good accuracy in orientation classification problem and thus we use our CNN and Transfer Learning based approach to implement the same.

Convolution Neural Networks and Transfer Learning features based model:

Recently image classification task using Convolution Neural Networks (pre-trained on ImageNet dataset) and Transfer Learning has gained huge success [1], [2]. So, to solve the image orientation classification problem (front, side and back view) three pre-trained Convolution Neural Network model features have been extracted. The three models are Mobile net [8], VGG16 [9] and Inception [10] from which the last layer features have been extracted which consists of the most important and specific features for the classification task. Each of the pre-trained features has been finally trained on our dataset. The pre-trained features act as the predictors and all the 5 models mentioned previously which consists of SVM, Multinomial Logistic, Naive Bayes, Decision Tree and Random Forest with the response variable having 3 classes' i.e. Front view, Side view and Back view.

The cross-validation accuracy for each of the pre-trained features and each model has been shown in Table 2.

Table 2: Cross Validation accuracy of Pre-trained CNN features for each of the Classification models

	Cross Validation Accuracy		
Classifiers	Mobile net	VGG16	Inception
SVM	94.2%	83%	82.9%
Multinomial Logistic	94.69%	89%	90%
Naïve Bayes	85.6%	80.2%	69%
Decision Tree	90%	76%	65%
Random Forest	93%	88%	81%

As shown in Table 2, the MobileNet features clearly outperform each of the other pre-trained CNN features even for a relatively small dataset and hence the same has been chosen for the process of image orientation classification.

The major advantage of MobileNet is that it uses depth wise separable convolutions to build light weight deep neural networks[8]. Another advantage being it has only two global hyperparameters which can be tuned very easily for the trade-off between latency and accuracy.

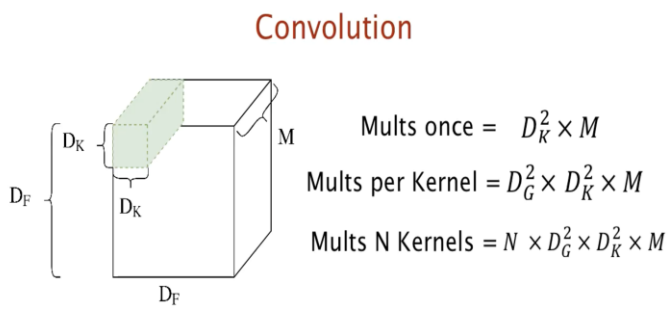


Figure 5: Number of Computation in Vanilla Convolution

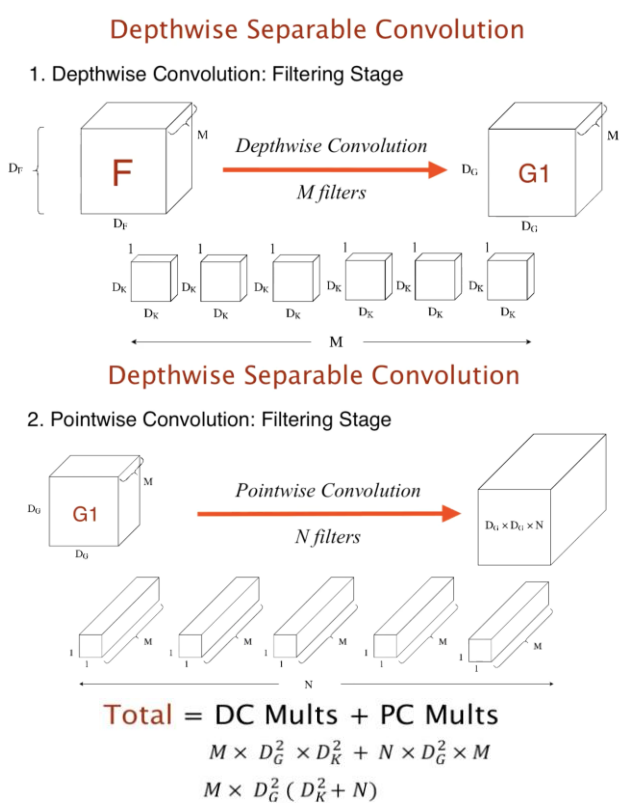


Figure 6: Number of Computations in Depth wise Convolution

As discussed, the number of computations has been significantly reduced in case of depth wise convolution and thereby decreasing the time complexity and making the model light.

It can be clearly seen from both Table 1 and 2 that amongst all the classifiers used, Multinomial logistic regression clearly outperforms all the other classifiers for both pre-trained CNN features as well as Histogram of oriented gradient based features. Hence, Multinomial Logistic Regression model with Mobile net features has been selected as the final model of classification of image orientation which gives an accuracy of approximately 95%.

Comparison of final model with baseline model based on Cross validation accuracy: Statistical Significance

A 10 fold cross validation was performed for both the MobileNet feature based multinomial logistic regression model and Histogram of gradient feature based multinomial logistic model and a Student's t-test [11] was performed to show that the accuracy in the former is significantly better than the later as shown in Table 3. The p-value<0.05 which indicates the statistical significance.

Table 3: Student's t-test for comparison of Cross-validation accuracy of models

t-Test: Two-Sample Assuming Equal Variances		
	Multinomial logistic _Hog	*Multinomial Logistic Mobilenet CNN*
Mean	0.68747	0.93525
Variance	0.012784722	0.002838069
Observations	10	10
Pooled Variance	0.007811396	
Hypothesized Mean Difference	0	
df	18	
t Stat	-6.26883626	
P(T<=t) one-tail	3.26456E-06	
t Critical one-tail	1.734063607	
P(T<=t) two-tail	6.52912E-06	
t Critical two-tail	2.10092204	

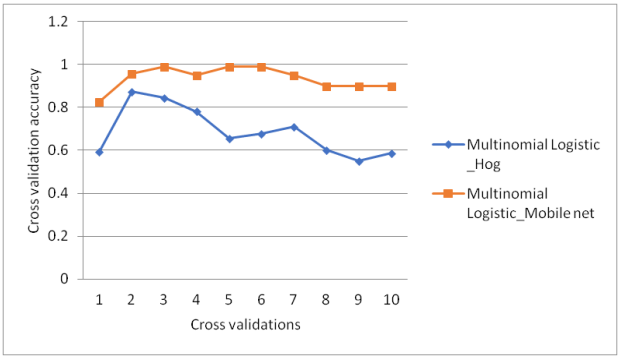


Figure 7: Cross-validation accuracy for both the models

As it can be seen from both Table 3 and Figure 7, MobileNet CNN model features with Multinomial Logistic Regression classifier trained on our dataset outperform significantly our

baseline model and hence that has been selected for the image orientation classification. This constitutes the first part of our pipeline.

Image Quality Assessment using Structural Similarity Index and Transfer Learning

For the task of quality assessment of images sent by vendor automatically, structural similarity has been used as the desired index as mentioned in [5]. The conventional metrics such as the peak signal-to-noise ratio (PSNR) and the mean squared error (MSE) which operate directly on the intensity of the image don't qualify as human visual system-based quality metric. But in our case, it is very important to use a quality index which is very similar to human perception and hence Structural similarity index which considers the impact of changes in luminance, contrast and structure in an image has been considered as shown in Figure 8.

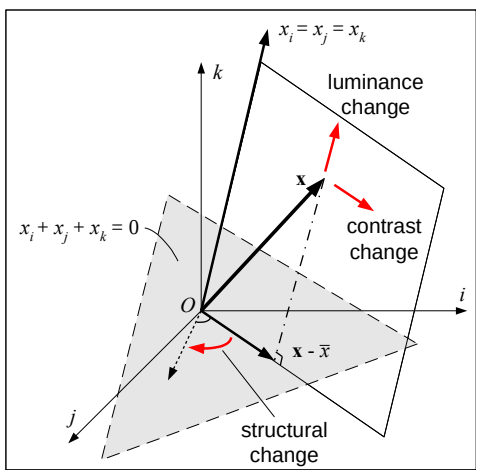


Figure 8: Structural Similarity Index

$$l(\mathbf{x},\mathbf{y}) = \frac{2\mu_x\mu_y + C_1}{\mu_x^2 + \mu_y^2 + C_1} \tag{1}$$

$$c(\mathbf{x},\mathbf{y}) = \frac{2\sigma_x\sigma_y + C_2}{\sigma_x^2 + \sigma_y^2 + C_2} \tag{2}$$

$$s(\mathbf{x},\mathbf{y}) = \frac{\sigma_{xy} + C_3}{\sigma_x\sigma_y + C_3} \tag{3}$$

$$SSIM(\boldsymbol{x},\boldsymbol{y}) = l(\boldsymbol{x},\boldsymbol{y}) \cdot c(\boldsymbol{x},\boldsymbol{y}) \cdot s(\boldsymbol{x},\boldsymbol{y}) \tag{4}$$

As shown in Figure8 and Equation 4, the structural similarity metric incorporates the illuminance, contrast and structural components of an image and hence is likely to capture the human perception whereas other metrics like MSE and PSNR etc. only captures the pixel wise difference between the two images which is not the way human perceives quality.

Introduction of Noise to the images of our dataset: Synthetic Data Generation

The way human perceives quality is very unique and every time some image of poor quality comes, it is a very easy task for human to detect that the quality is not adequate may be some blurring, other noise factors are there in the image. Humans won't need any reference image of superior quality for that poor quality image to tell that which motivates us to the concept of no-reference image quality assessment.

The main challenge in the field of image quality assessment is that we won't have the perfect image of an item every time with its corresponding imperfect/poor quality version for assessing the quality of the images. Hence, we need a methodology where quality of the image can be assessed without reference image [6]and which can work for small datasets as well. The idea is to make the machine learn the way human perceives quality in such cases.

The first step is to add distortion to the reference images of the datasets with different noise signals and artificially create our own datasets of good images and distorted images. There can be various types of noise signals which can be given to the image but for our case we have considered blurring as the noise factor with various factors and kernels of the same. The noise signals considered are different types of blurring since that is one major area of concern for the images sent by vendor which is shown in Table 4.(Here reference image is only for the training set, for test set there won't be any).

Table 4: Different distortion types added to reference images

Type of Noise added	Kernels and Parameters
Mean Blur	(5,5),(25,25),(55,55),(75,75)
Gaussian Blur	(5,5),(25,25),(55,55),(95,95)
Bilateral Blur	(9,50,50),(9,125,125)
Median Blur	5,27

The operation has been done for all the 312 images and each type of distortion has been considered as a separate class/category which makes a total of 13 categories including the reference good images. Since each parameter induces blurring of different types and each type has been considered as a separate class for the supervised framework that we have created.

Image quality based classification using Mobile net CNN features and Deep Learning Classification algorithm: **Quality Embeddings**

In the second step of the process of image quality assessment, the pre-trained MobileNet[8] last layer features have been extracted for all the images of 13 different classes mentioned above which includes the good/reference class images, Mean blurred images (4 different classes),Gaussian Blurred images (4 different classes),Bilateral Blurred images(2 different classes) and Median Blur(2 different classes) . The MobileNet [8] final layer features of the images contain all the important features and information about them. As discussed earlier as well the benefits of having a light weight model with depth wise convolution, MobileNet CNN captures the most relevant features from the image . Then we have built deep layers on top of it which basically projects the features into different dimensions. Finally using a SoftMax layer, we have classified them into the 13 different classes and the model is trained on the same. The deep learning architecture after extraction of the MobileNet embeddings have been shown below.

Table 5: Deep Learning Model Description

Layers	No. of Neurons	Activation Function
Input Layer	1024	-
Hidden Layer1	512	Relu
Hidden Layer 2	256	Relu
Hidden Layer 3	112	Relu
Output Layer	13	Softmax

The key idea and innovation of the work lies in the concept of creating quality embeddings for each of the images. The last but one layer before the SoftMax layer of the model described above projects the images into quality dimensions. The main intuition behind that is if the model is generating such features in the final layer such that it is being able to classify images which are similar otherwise and the only difference lies in quality of the images, then the features that are generated are quality based features.

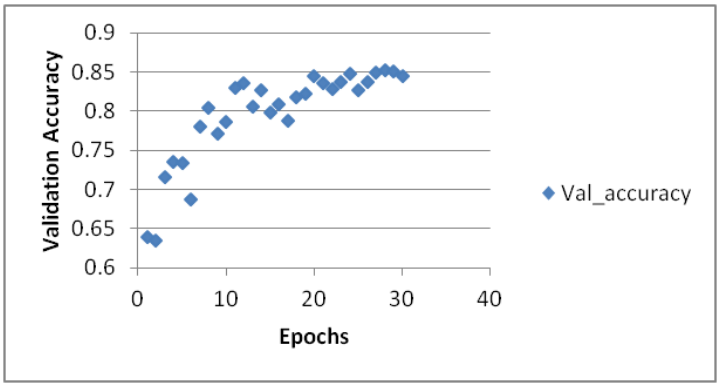


Figure 9: Validation accuracy of the Deep Classification model

As shown in figure 9, the cross-validation accuracy obtained by the model was 84.5% which is quite high considering the amount of data used. The final layer of the deep model is extracted as these features are the quality embeddings or quality-related features for these images. The main idea as mentioned above as well behind the statement is that in these image classes (13) the only difference is the image quality and all other things are same for all the classes and hence if a model is differentiating between these images it clearly indicates the features will be those features which are related to quality characteristics of the images. The diagrammatic workflow has been shown below in Figure 10.

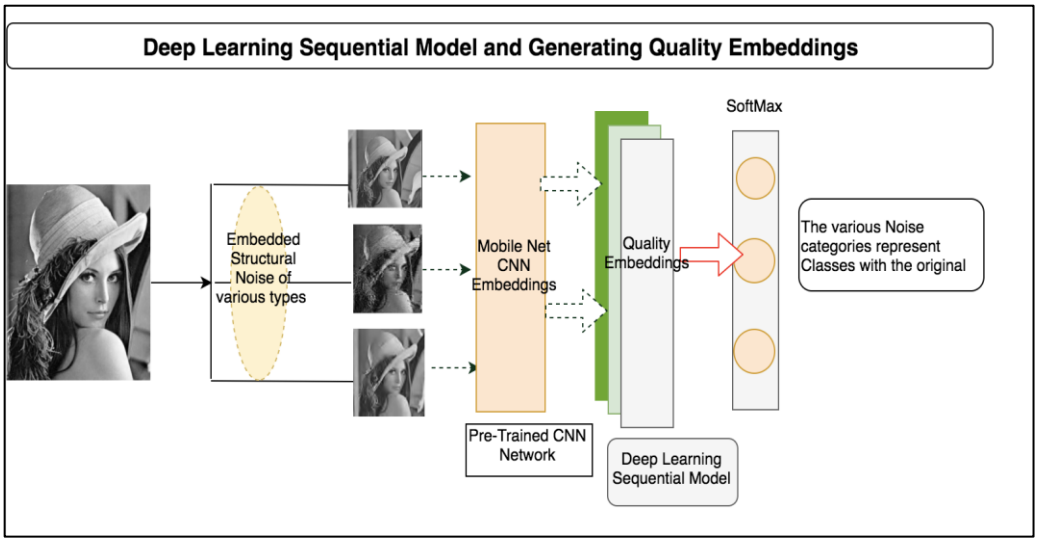


Figure 10: Generating Quality Embeddings using Deep architecture

Computation of Structural Similarity Scores and prediction using Ridge Regression model

Once the quality related features have been extracted for the images, the Structural similarity scores for all the known synthetically generated distorted images and original images are computed from the reference images. So, for the true reference images the structural similarity will be 1 and as the distortion in the images increasing the structural similarity metric value decreases.

Then the quality embeddings for all the images have been extracted using the deep learning model described in Table5 and extracting the last but one layer weights.

Once the images are projected into quality dimensions, the quality embeddings have been taken as predictor variables and the Structural similarity scores computed for the same images as the response variables and a Ridge regression is fitted with an 80-20 validation and a validation accuracy of 83% is achieved by this methodology. So, now whenever a new image is there, the quality embeddings are extracted from the images by projecting them in the latent quality dimensions and then considering the same as a test feature for our Ridge regression model, the Structural similarity score for that image will be predicted using the model and based on which and a business decided threshold value, necessary actions will be taken. The ridge regression model equation is shown in Equation 5.

$$\begin{aligned}\hat{\beta}^{\text{ridge}} &= \underset{\beta \in \mathbb{R}^p}{\operatorname{argmin}} \sum_{i=1}^{n} (y_i - x_i^T \beta)^2 + \lambda \sum_{j=1}^{p} \beta_j^2 \\ &= \underset{\beta \in \mathbb{R}^p}{\operatorname{argmin}} \underbrace{\|y - X\beta\|_2^2}_{\text{Loss}} + \lambda \underbrace{\|\beta\|_2^2}_{\text{Penalty}}\end{aligned} \tag{5}$$

Finally the ordering is done as per business requirements which complete the pipeline of our process and the flow has been shown below in Figure 11.

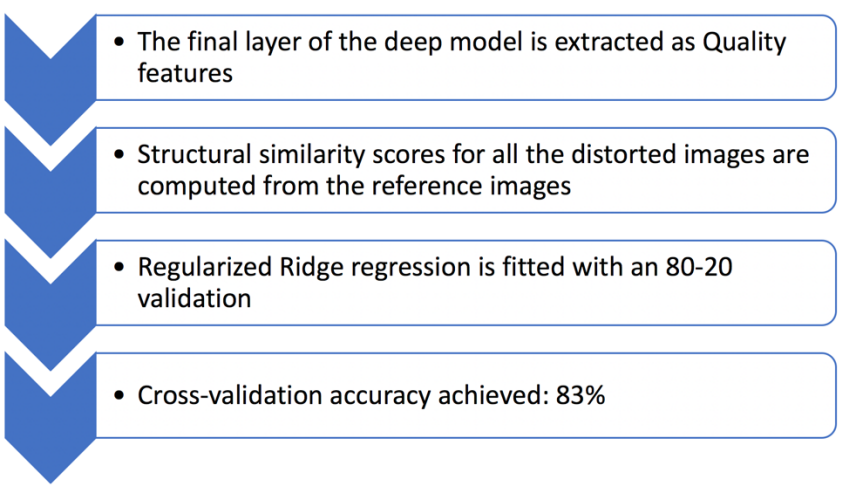


Figure 11: Directional Flow for predicting Structure Similarity metric using Ridge regression

Evaluation:

As shown in Table 1 and Table 2, our MobileNet CNN features with Multinomial logistic regression performs much better than the baseline model and other pre-trained CNN features compiled in Table 6.

Table 6: Cross validation accuracy of various models

	Cross Validation Accuracy			
Classifiers	Hog	MobileNet	VGG16	Inception
SVM	0.6222	0.942	0.83	0.829
Multinomial Logistic	0.7123	0.9469	0.89	0.9
Naïve Bayes	0.6212	0.856	0.802	0.69
Decision Tree	0.55	0.9	0.76	0.65
Random Forest	0.7	0.93	0.88	0.81

The statistical significance test has been performed to check if the increase in accuracy is statistical significant or not and hence a paired t-test has been done to do the same.

Table 7: Student's paired t-test for comparison of Cross-validation accuracy of models

t-Test: Two-Sample Assuming Equal Variances		
	Multinomial logistic _Hog	*Multinomial Logistic Mobilenet CNN*
Mean	0.68747	0.93525
Variance	0.012784722	0.002838069
Observations	10	10
Pooled Variance	0.007811396	
Hypothesized Mean Difference	0	
df	18	
t Stat	-6.26883626	
P(T<=t) one-tail	3.26456E-06	
t Critical one-tail	1.734063607	
P(T<=t) two-tail	6.52912E-06	
t Critical two-tail	2.10092204	

As it can be seen that MobileNet CNN features with Multinomial logistic regression performance is much superior and that has been tested in Table 7 via paired t-test.

Table 8: Validation accuracy of Deep Learning Model for Image quality Classification

Epochs	Validation Accuracy
1	0.6394
7	0.78
15	0.7982
20	0.8445
25	0.8263
28	0.8528
30	0.8453

An accuracy of 85% was achieved by the deep learning quality classification model and finally the Ridge regression model had an accuracy of 83%.

The final workflow of the pipeline has been shown below in Figure 12.

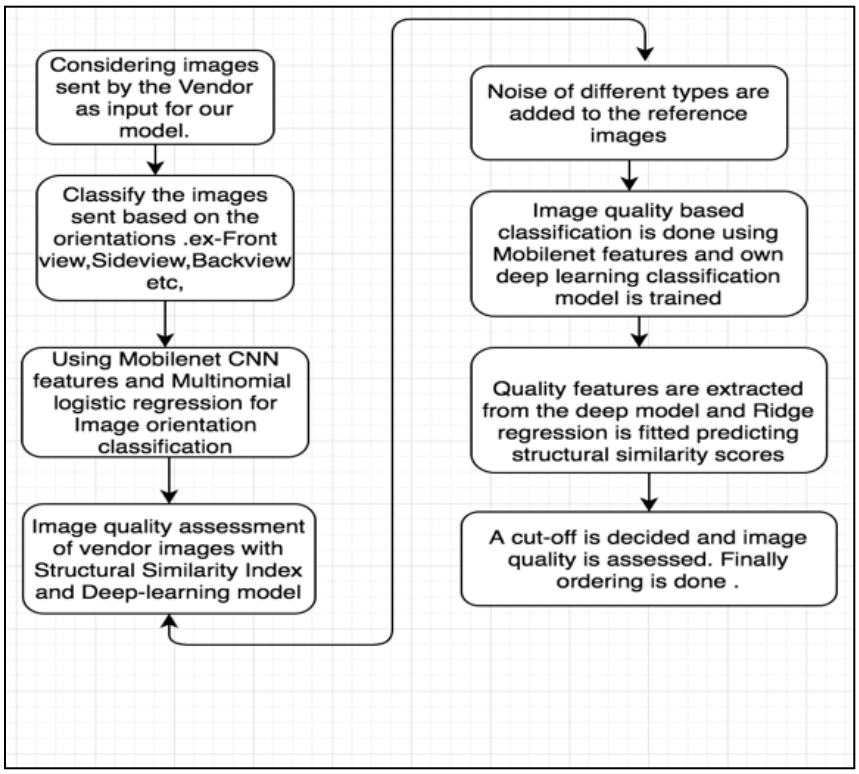


Figure 12: Final Workflow of the Pipeline of Catalogue Management.

IV. CONCLUSION

In this work, we have successfully build a pipeline where in the first step we have classified the image orientations, in this case Front-view, Side-view & Back-view with a cross validation accuracy of 94% with pre-trained Mobile Net features and Multinomial Logistic Regression approach and that too with small dataset which was one of the challenge for our work. This process actually reduces the manual labor and helps in easing the process of catalogue management.

The next most important part of our pipeline of automated catalogue management was to successfully implement image quality assessment with no-reference image. This is a very important area since many of the images of items sent by the vendor are not as per required quality which causes the customer to move to different industries. Moreover, this is a reasonably challenging task to assess image quality when the reference image is not present.

In the methodology developed to solve this problem, the first step is to add distortions/noise to our reference images and then extract MobileNet features and finally a deep learning model is trained in such a way that it can uniquely identify the different classes of images. The last layer

features from this deep learning model has been extracted since it consists of the quality characteristics of the images.

The structural similarity index has been used as the index to measure the structural similarity between the reference and distorted images as it is almost similar to the way human perceives image quality. Using the structural similarity scores as the response and the features of the deep model as predictor, a Ridge regression model is being fitted with an accuracy of 83% which is quite good considering the complexity of the problem. So, now whenever a new image comes, first the MobileNet features will be extracted from it and its structural similarity score will be predicted from the Ridge regression model.

Finally, the ordering is done as per Business requirements and this wraps up the pipeline built for automated catalogue management.

Further scope of research is there to classify more orientations of images for image orientation classification. In image quality classification task, ensemble models can be used to make the accuracy better and many different types of noise signals can also be added to make the model much better.

ACKNOWLEDGMENT

We want to grasp this opportunity to convey our deepest gratitude to all of them who aided us in completing our paper. Our heartfelt gratitude to AICAAM team, for this opportunity. Our great thankfulness to Data and Analytics Team, Walmart International, for introducing us to this topic and guiding us throughout our paper. A special thanks to Ashish Gupta, Manager Walmart Labs as our internal guide of the paper, for his persistent encouragement and motivation towards us. We are thankful to all of them who directly and indirectly helped us in accomplishing this paper.

REFERENCES

[1]. Alex Krizhevsky, Ilya Sutskever, and Geoffrey E. Hinton. 2012. ImageNet classification with deep convolutional neural networks. In Proceedings of the 25th International Conference on Neural Information Processing Systems - Volume 1 (NIPS'12), F. Pereira, C. J. C. Burges, L. Bottou, and K. Q. Weinberger (Eds.), Vol. 1. Curran Associates Inc., USA, 1097-1105.

[2]. Murinto, Murinto & Prahara, Adhi & Winiari, Sri & Pramudi Ismi, Dewi. (2018). Pre-Trained Convolutional Neural Network for Classification of Tanning Leather Image. International Journal of Advanced Computer Science and Applications. 9. 10.14569/IJACSA.2018.090129.

[3]. Jasper Snoek, Hugo Larochelle, and Ryan P. Adams. 2012. Practical Bayesian optimization of machine learning algorithms. In Proceedings of the 25th International Conference on Neural Information Processing Systems - Volume 2 (NIPS'12), F. Pereira, C. J. C. Burges, L. Bottou, and K. Q. Weinberger (Eds.), Vol. 2. Curran Associates Inc., USA, 2951-2959.

[4]. Wang, Zhou & Bovik, Alan & Lu, Ligang. (2002). Why is image quality assessment so difficult? Proc IEEE Int Conf Acoustics Speech Signal Process. 4. IV-3313. 10.1109/ICASSP.2002.1004620.

[5]. Zhou Wang, A. C. Bovik, H. R. Sheikh and E. P. Simoncelli, "Image quality assessment: from error visibility to structural similarity," in *IEEE Transactions on Image Processing*, vol. 13, no. 4, pp. 600-612, April 2004.

[6]. S. Bosse, D. Maniry, K. Müller, T. Wiegand and W. Samek, "Deep Neural Networks for No-Reference and Full-Reference Image Quality Assessment," in IEEE Transactions on Image Processing, vol. 27, no. 1, pp. 206-219, Jan. 2018.

[7]. P. E. Rybski, D. Huber, D. D. Morris and R. Hoffman, "Visual classification of coarse vehicle orientation using Histogram of Oriented Gradients features," *2010 IEEE Intelligent Vehicles Symposium*, San Diego, CA, 2010, pp. 921-928.

[8]. Howard, A.G., Zhu, M., Chen, B., Kalenichenko, D., Wang, W., Weyand, T., Andreetto, M., & Adam, H. (2017). MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications. *CoRR, abs/1704.04861*.

[9]. Simonyan, Karen & Zisserman, Andrew. (2014). Very Deep Convolutional Networks for Large-Scale Image Recognition. arXiv 1409.1556.

[10]. C. Szegedy *et al*., "Going deeper with convolutions," *2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Boston, MA, 2015, pp. 1-9.

[11]. Douglas C. Montgomery, George C. Runger, *Applied Statistics and Probability for engineers*. 6th edition, Wiley India.

Relationship between Financial Ratios and Systematic Risk in Steel Industry: A study

Rohith Varma*, Selvarani.R**

*Department of Computer Science, Alliance College of Engineering and Design,
**Program Director, Department of Computer Science, Alliance College of Engineering and Design.

Abstract- The objective of this study is to determine the relationship between financial ratios and systematic risk of Steel industry. The financial variables that have been utilized for the study are Quick ratio, return on assets excluding Revaluations, Operating Profit Margin, Net Profit Margin, Debt equity ratio. Four years data from 2015 to 2018 have been collected from Yahoo Finance, National Stock Exchange, Moneycontrol.com. The consequences of 15 Steel Companies have been studied through the application of Pearson algorithm and the results are discussed elaborately in this paper. This work supports investors to decide on ventures for investment.

Index Terms- Systematic Risk, Quick Ratio, Net profit margin, Operational profit margin, Pearson Correlation, Debt equity ratio

I. INTRODUCTION

Systematic Risk is a part of total risk which is beyond the control of an individual or a company. Every investment or security is subjected to systematic risk therefore it is not a type of risk which can be diversified. Systematic Risk is a crucial factor for any investor. Investors tend to invest in companies which have a high return. To measure the Systematic Risk, beta is used. Beta varies from industry to industry. Financial Ratios are good indicators of a company's performance and its financial situation. Therefore relationship between Systematic Risk and the financial ratios can be used to find if we could invest in the company or not. Therefore, in this paper we will find the relationship between the financial ratios and Systematic Risk in companies related to the Steel Industries of India.

II. BACKGROUND

The total risk of a company includes Systematic Risk and Unsystematic Risk. Unsystematic Risk is specific to the firm, Systematic Risk is the market related risk. The Unsystematic risk can be reduced by diversifying the stocks therefore the main concern to the investors is the systematic risk. Gu and Kim(1998) suggested that the systematic risk (Beta) of each firm can be estimated by the equation or the characteristic line. The slope of the characteristic line of each company which is estimated against S&P 500 return represent the stock's return and its beta. Previous studies on financial ratios: The financial ratios in quick service and full service Restaurants were studied by Kim, Ryan and Ceschini (2007), the data from year 1999 to 2003 was utilized. According to the research the return on investment is found to be negatively correlated to beta. Debt to equity ratio was not very significant in the restaurant services. Kim et al (2007) estimated that there is a positive relation between quick ratio and risk(beta). Gu and Kim (1998) evaluated the current ratio, leverage ratio (total liabilities to total assets), asset turnover, and profit margin of 35 casino firms and found that only asset turnover was significant and negatively correlated

with Beta at the .10 level. All other variables were found to be statistically non significant. Further they also suggested that it is better to use quick ratio instead of current ratio as it represents the liquidity better. Till date, there has been no study which explored the relationship between Steel Industry in India and the financial Ratios.

III. PROPOSED SYSTEM

The ratios that are selected for consideration for our research are: -

Quick Ratio: The quick ratio is an indicator of a company's short-term liquidity position and measures a company's ability to meet its short-term obligations with its most liquid assets. [1]

Quick Ratio = (Cash + Marketable Securities + Accounts Receivable) / CurrentLiabilities

Return on assets: Return on total assets (ROTA) is a ratio that measures a company's earnings before interest and taxes (EBIT) relative to its total net assets.[2]

Return on assets = Net Income/Total Number of Assets.

Operating profit margin: Operating Profit Margin is a profitability or performance ratio used to calculate the percentage of profit a company produces from its operations, prior to subtracting taxes and interest charges. [3]

Operating profit margin = Operating profit/Total Revenue.

Net Profit Margin: Net profit margin is the percentage of revenue remaining after all operating expenses, interest, taxes and preferred stock dividends (but not common stock dividends) have been deducted from a company's total revenue. [4]

Net Profit Margin = Net Profit/Total Revenue.

Debt Equity ratio: The debt-to-equity (D/E) ratio is calculated by dividing a company's total liabilities by its shareholder equity.[5]

Debt Equity ratio = Total Liabilities/Total Shareholder's Equity.

The relation between each of the financial ratio and the Systematic Risk is analyzed.

The companies that were used for the research are :-
Large scale companies in steel industry:-

- JSW Steel Ltd.

- Tata Steel.
- Steel Authority of India Ltd.
- Manaksia Steels Ltd.
- Visa Steel Ltd.

Medium and small scale :-

- Jindal stainless and Power Ltd.
- Usha Martin Ltd.
- Mukand Ltd.
- Beekay Steel Industries Ltd.
- Shivalik Bimetal Controls Ltd.
- Kamdhenu Ltd.
- Vardhman Special Steels Ltd.
- Balasore Alloys Ltd.
- Panchmahal Steels Ltd.
- SAL Steel Ltd.

IV. ANALYSIS ON SYSTEMATIC RISK(β)

Systematic risk can be calculated by the sensitivity of a security's return with respect to market return. This sensitivity is calculated by using the β (Beta) Coefficient.
The value of Systematic Risk (β) can be calculated using the following formula:
β =Slope (Rp, Rm)
Rm: Market Return
Rp: Portfolio Return
The beta has been calculated by the taking the slope of the Market Return and the Portfolio Return. The tool that was used for the calculation of beta is Excel. The data for the Market Return and the Portfolio return has been extracted from Yahoo Finance.

V. RESULTS AND FINDINGS

Financial Ratios (Large Scale Industry) /Financial Years	FY18	FY17	FY16	FY15
JSW Steel				
Quick Ratio	0.68	0.56	0.41	0.67
Return on Assets Excluding Revaluations	115.98	100.28	84.44	1,032.60
Operating Profit Margin(%)	21.14	22.07	17.35	19.24

Net Profit Margin(%)	7.11	6.84	-9.61	4.7
Debt Equity Ratio	1.14	1.38	1.58	1.06
Tata Steel				
Quick Ratio	0.34	0.28	0.32	0.27
Return on Assets Excluding Revaluations	510.87	511.31	725.65	686.4
Operating Profit Margin(%)	26.46	24.74	18.87	23.95
Net Profit Margin(%)	6.99	7.17	12.82	15.41
Debt Equity Ratio	0.45	0.61	0.44	0.39

Steel Authority of India Ltd				
Quick Ratio	0.4	0.38	0.42	0.55
Return on Assets Excluding Revaluations	86.46	87.18	94.89	105.33
Operating Profit Margin(%)	8.02	0.08	-7.42	10.18
Net Profit Margin(%)	-0.83	-6.37	-10.29	4.57
Debt Equity Ratio	1.18	1.08	0.84	0.65
Manaksia Steels Ltd.				
Quick Ratio	1.69	4.04	3.25	2.78
Return on Assets Excluding Revaluations	26.88	24.35	23.4	22.64
Operating Profit Margin(%)	6.99	7.49	5.87	4.81
Net Profit Margin(%)	3.37	2.73	1.72	2.48
Debt Equity Ratio	0.53	0.36	0.35	0.3
Visa Steel Ltd.				

Quick Ratio	0.1	0.19	0.24	0.26
Return on Assets Excluding Revaluations	-68.52	-59.43	-41.92	11.13
Operating Profit Margin(%)	0.02	1.16	0.58	-0.22
Net Profit Margin(%)	-9.08	-10.21	-115.49	-26.18
Debt Equity Ratio	--	--	--	22.59

*The data has been extracted from moneycontrol.com

Table 1 : Financial Ratios of Large Scale Steel Industry for the period:- FY2015 – FY2018

Financial Ratios (Small and Medium Scale)/Financial year	FY18	FY17	FY16	FY15
Jindal Stainless				
Quick Ratio	0.4	0.44	1.64	1.42
Return on Assets Excluding Revaluations	49.07	42.96	-25.46	-7.64
Operating Profit Margin(%)	11.87	13.31	7.9	5.05
Net Profit Margin(%)	2.95	0.7	-5.9	3.71
Debt Equity Ratio	1.84	2.98	--	--
Usha Martin Ltd.				
Quick Ratio	0.38	0.37	0.39	0.31
Return on Assets Excluding Revaluations	5.69	14.95	26.66	39.26
Operating Profit Margin(%)	11.77	10.65	8.47	16.22
Net Profit Margin(%)	-6.99	-10.93	-12.22	-7.8
Debt Equity Ratio	20.06	7.78	4.6	2.8
Mukand Ltd.				
Quick Ratio	0.79	1.39	1.28	1.19

Return on Assets Excluding Revaluations	80.2	25.56	32.61	32.49
Operating Profit Margin(%)	2.02	13.13	12.2	11.99
Net Profit Margin(%)	1.43	0.42	0.05	0.05
Debt Equity Ratio	1.46	7.19	5.36	5.09
Beekay Steel Industries Ltd.				
Quick Ratio	1.81	1.42	1.14	0.94
Return on Assets Excluding Revaluations	157.53	121.59	102.05	93.13
Operating Profit Margin(%)	13.64	11.68	12.57	11.59
Net Profit Margin(%)	7.23	4.78	3.71	3
Debt Equity Ratio	0.59	0.78	0.82	0.91
Shivalik Bimetal Controls Ltd.				
Quick Ratio	1.49	1.64	1.99	1.52
Return on Assets Excluding Revaluations	22.2	37.1	33.85	31.7
Operating Profit Margin(%)	17.2	15.98	13.67	16.11
Net Profit Margin(%)	10.01	7.25	4.16	5.19
Debt Equity Ratio	0.36	0.36	0.57	0.64
Kamdhenu Ltd.				
Quick Ratio	1.6	1.62	2.07	1.93
Return on Assets Excluding Revaluations	54.96	49.11	45.16	42.52
Operating Profit Margin(%)	3.7	3.94	3.75	3.09
Net Profit Margin(%)	1.32	0.98	0.96	0.82
Debt Equity Ratio	0.68	0.92	1	0.98

Vardhman Special Steels Ltd.				
Quick Ratio	2.16	1.42	1.73	2.14
Return on Assets Excluding Revaluations	94.9	106.75	96.56	91.47
Operating Profit Margin(%)	7.2	8.88	6.37	1.74
Net Profit Margin(%)	2.91	2.84	0.64	-2.28
Debt Equity Ratio	0.65	1.28	1.65	2.02
Balasore Alloys				
Quick Ratio	0.9	0.87	0.71	0.56
Return on Assets Excluding Revaluations	103.57	101.47	52.9	55.12
Operating Profit Margin(%)	12.43	17.7	9.78	12.75
Net Profit Margin(%)	5.41	8.26	2.23	3.44
Debt Equity Ratio	0.18	0.2	0.32	0.3
Panchmahal Steels Ltd.				
Quick Ratio	0.58	0.55	0.52	0.52
Return on Assets Excluding Revaluations	63.39	63.67	63.51	70.11
Operating Profit Margin(%)	4.14	4.94	-0.63	2.7
Net Profit Margin(%)	0.13	0.08	-4.33	-0.9
Debt Equity Ratio	0.58	0.57	0.51	0.42
SAL Steel Ltd.				
Quick Ratio	0.48	0.43	0.27	0.26
Return on Assets Excluding Revaluations	0.45	-4.56	-4.13	-0.74

Operating Profit Margin(%)	1.27	2.65	6.87	7.14
Net Profit Margin(%)	8.49	0.24	-8.52	-20.75
Debt Equity Ratio	44.12	--	--	--

*The data has been extracted from moneycontrol.com

Table 2: Financial Ratios of Medium and Small scale industry for the period:- FY2015 – FY2018

Analysis on Systematic Risk (Beta)

Systematic Risk for Large Scale Industry /Financial Years	FY18	FY17	FY16	FY15
Tata Steel Ltd	-0.028	0.293	0.251	0.213
JSW Steel Ltd	-0.148	1.073	0.511	0.835
SAIL Ltd	-0.521	1.291	0.821	0.0195
Manaksia Steels Ltd	-0.021	0.081	0.552	-0.046
Visa Steel Ltd	0.088	0.468	1.503	0.180

*The following data has been calculated from Yahoo Finance website

Table 3: Systematic Risk associated with Large scale industry for the period:- FY2015 – FY2018

Systematic Risk for Medium and small scale Industry /Financial Years	FY18	FY17	FY16	FY15
Jindal Stainless	0.400	1.658	1.391	1.266
Usha Martin Ltd	0.619	0.111	1.105	0.722
Mukand Ltd	-0.180	-0.173	0.562	0.428
Beekay steel Ltd	0.583	-0.724	-0.148	
Shivalik Bimetal Controls Ltd	1.118	1.547	0.831	-0.289
Kamdhenu Ltd	0.248	-0.591	0.694	0.702
Vardhman Special Steel Ltd	0.405	-0.343	0.850	-0.483
Balasore Alloys	-0.284	1.295	1.396	0.741
Panchmahal Steels Ltd	0.153	0.504	-0.942	1.165
SAL Steel Ltd	-0.402	-0.209	0.691	1.325

*The following data has been calculated from Yahoo Finance website

Table 4: Systematic Risk associated with Medium and Small scale industry for the period:- FY2015 – FY2018

Correlation Analysis between Systematic risk and Financial ratios

This study was calculated by using Pearson's correlation coefficient on the 15 steel companies mentioned above.

Financial Ratios	**Correlation**
Quick Ratio	Negative Correlation
Return On Assets Excluding Revaluations	Positive Correlation
Operating Profit Margin	Negative Correlation
Net Profit Margin	Negative Correlation
Debt/Equity Ratio	Positive Correlation

Table 5: Correlation between Systematic Risk and Financial Variables

II. CONCLUSION

The main goal of an organization is to improve shareholders' value. The total risk of the investment is calculated and measured by the variance or, most commonly the Standard deviation of the return. To comprehend the parameters related with Systematic Risk is valuable for financial investors and company managers. This study finds the relation between systematic risk and quick ratio, return on assets excluding evaluations, operating profit margin, Net profit margin and Debt-Equity Ratio. Correlation analysis has been used for estimation. Study included 15 financial companies listed in NSE for the period of 2015 to 2018. Quick Ratio is negatively correlated with beta (β), Return on assets excluding revaluations is positively correlated with beta(β), Operating profit margin is negatively correlated with beta(β), Net profit margin is negatively correlated with beta**(β),** Debt/Equity ratio is positively correlated with beta(β). Managers can estimate these parameters to control systematic risk and to improve financial performance of a firm. Financial ratios do play an important role in finding systematic risk. This study is essential for investors and finance managers to understand what kind of relationship exist between financial ratios and systematic risk. The future research will involve larger number of ventures and more number of parameters for better analysis and results, which will support the investors for better decision on their investments.

REFERENCES

1. https://www.investopedia.com/terms/q/quickratio.asp
2. https://www.investopedia.com/terms/r/returnonassets.asp
3. https://corporatefinanceinstitute.com/resources/knowledge/finance/operating-profit- margin/
4. https://investinganswers.com/financial-dictionary/financial-statement-analysis/net-profit- margin-2233
5. https://www.investopedia.com/terms/d/debtequityratio.asp
6. https://in.finance.yahoo.com/
7. https://www.moneycontrol.com/
8. Rowe, T., & Kim, J. (2010). Analyzing the Relationship Between Systematic Risk and Financial Variables in the Casino Industry. UNLV Gamming & Research Journal Vol. 12 No.2.
9. Sharpe, W. F. (1963). A simplified model of portfolio analysis. Management Science, 9(2), 277-

293.

10. Bowman, R. (1979). The Theoretical Relationship Between Systematic Risk and Financial (Accounting) Variables.
11. Gupta, N., & Gurjar, A. S. (2014). A Study of Systematic Risk with Reference of Selected Companies.
12. qbal, M., & Ali Sha, S. Z. (2012). Determinants of Systematic Risk.
13. Logue, D. E., & Merville, L. J. (1972). Financial policy and market expectations. Financial Management, 37-44.
14. Lau, S. C., Quay, S. R., & Ramsey, C. M. (1974). The Tokyo stock exchange and the capital asset pricing model.
15. Mnzava, I. D. (2009). The significance of corporation tax as a determinant of systematic risk: Evidence using United Kingdom (UK) data. KCA Journal of Business Management, 2(1).
16. Gu, Z., & Kim, H. (1998). Casino Firm's risk Features and Their Beta Determinants. Prosress in Tourism and Hospitality Research 4, 357-365.
17. Gu, Z., & Kim, H. (2002). Determinants of Restaurant Systematic Risk: A Re-examination. The Journal of Hospitality Financial Management, 10(1), 1-13.
18. Logue and Mervile. (1972). Financial Policy and Market Expectations. Financial Management, 1(3), 37-44. Iqbal, M., & Ali Sha, S. Z. (2012).
19. Determinants of Systematic Risk. Moyer, R., & Chatfield, R. (1983). Market Power and Systematic Risk. Journal of Economics and Business, 35 (I), 123-130. Rowe, T., & Kim, J. (2010).
20. Analyzing the Relationship Between Systematic Risk and Financial Variables in the Casino Industry. UNLV Gaming & Research Journal Vol. 12 No.2.

Retail Based Cost Reverse Engineering and Cost comparison within Item Similarity Clusters for Cost Negotiations

Mani Kanteswara Rao Garlapati*, Souradip Chakraborty*

* Data & Analytics, Walmart Labs, Bangalore

Abstract- Detection of negotiable items and performing various cost negotiation strategies is an age-old problem in the field of retail but despite of that, there are lot of areas that are unexplored in the above field.

This paper talks about detection of negotiable items and how cost negotiations can be performed by leveraging retail price-based cost reverse engineering negotiation methods. Retail price change of an item over time is analyzed along with the cost trend of the item to determine potential for cost negotiations. These comparisons are performed within different item similarity clusters, so that we can see how different peer items within the same similarity cluster perform with respect to retail price and cost price trends. Hence our methodology efficiently captures and detects the most robust set of negotiable items considering major attributes including margin percent, sales volume, substitutability etc.

Index Terms- Item substitutes, Community detection, Parent supplier, Dis-intermediation, Joint buy.

I. INTRODUCTION

This paper guides us on Cost negotiations of different items using different negotiation levers to be identified during the process. In this paper we would focus on cost negotiation using one such lever which is based on retail price-based cost reverse engineering. Retail price of any item usually consists of both promotional and non-promotional sales. For performing cost negotiations, we should be able to identify the trend of regular retail price of the item over time which would be obtained after removing all the promotional events. The regular retail price of an item can be computed by removing different promotional effects like clearance, rollback, price adjustments etc. By understanding the trend of change in regular retail price we can identify if there is an opportunity for cost negotiations by comparing it with the trend of average unit cost for each item using correlation coefficient-based metric. These comparisons can also be performed by comparing the trends among similar item description clusters. These item similarity clusters are obtained by leveraging a combination of Lucene elastic search implementation and community detection algorithm.

Retail price of an item changes constantly based on the demand, competitor pricing and various other factors. Price change of an item for rollback is set centrally for any retail and it gets applicable across all items within certain departments. Clearance and Price adjustments are made by store manager when items are not getting cleared and which remain in the shelf for a long period of time. These are tagged with certain report codes for different types of promotions along with the quantity of units sold at the promotional price that was offered. Use these report codes we will be able to separate regular sales and units using which we can then calculate Regular Average Unit Retail (AUR), which is the sell price of unit quantity. The change in trend of the

regular AUR and its comparison with the trend in the Average Unit Cost (AUC) would help in and efficient negotiation of the item cost.

II. IDENTIFY, RESEARCH AND COLLECT IDEA

There is lot of work happening in this space around how various negotiations can be performed. Economic, behavioural, and software agent perspectives based integrative negotiation gained quite a lot of popularity in the field of electronic ecommerce [1].

There has been work around leveraging retail price trend for negotiation cost of an item. Research has also been done on how to identify similar items based on their item details. Community detection and elastic search has not been leveraged till date for detection of similar items based on their descriptions. But both of these together have not been leveraged for cost negotiations.

We have utilized Walmart items details like descriptions, signage etc. for identifying item similarity clusters. Along with these descriptions we have utilized retail price, promotional activities, cost information, warehouse data, external data like holidays, events, climatic conditions etc. for obtaining the regular retail trend and cost price trends for different items across departments.

III. STUDIES AND FINDINGS

Regular retail price of an item refers to non-promotional price over time. We can build out data at week level in order to avoid day level nuances in the data and avoid data sparsity for low selling items. Also, we are averaging the prices of different items across different stores. This help in generalizing the retail and cost data trend over time and across stores. We will need to compute actual sales, units across different weeks. Flag different promotional events basis their report codes from the data to identify the promotional contribution of various events for their sales and units. Regular sales and units can be computed using the below formula.

Regular Sales = Actual Sales – Rollback Sales – Clearance Sales – Price Adjustment Sales

Regular Units = Actual Units – Rollback Units – Clearance Units – Price Adjustment Units

Regular Average Unit Retail (AUR) can be computed as a ratio of Regular Sales and Regular units over different weeks. Similarly, Average Unit Cost (AUC) can be computed as a ratio of Net ship cost to the Net ship quantity over different weeks.

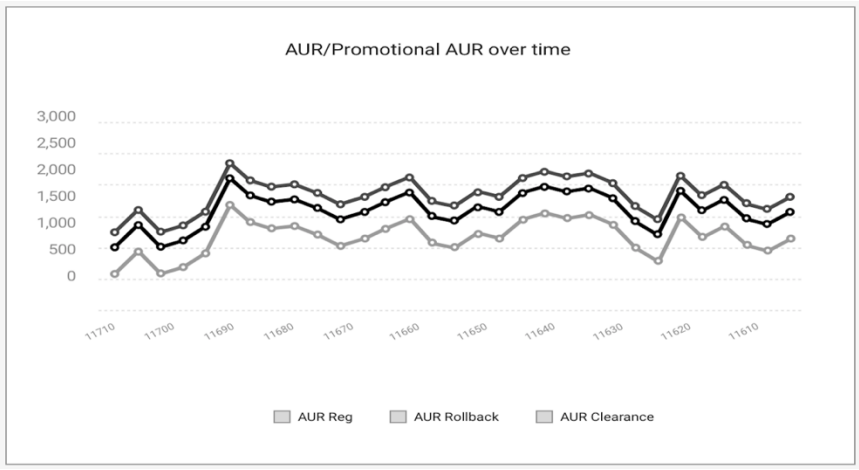


Figure 1: Promotional Breakdown of Retail Price

On computation of AUR and AUC separating out the promotional effects, the trend of the variables has been compared using Pearson's correlation coefficient ratio metric.

$$r\_ratio = (r_{AUR,time}) / (r_{AUR,time} + \varepsilon) \quad (1)$$

where,

r_ratio - correlation coefficient ratio metric

$r_{AUR,time}$ – correlation coefficient of AUR over time

$r_{AUR,time}$ - correlation coefficient of AUC over time

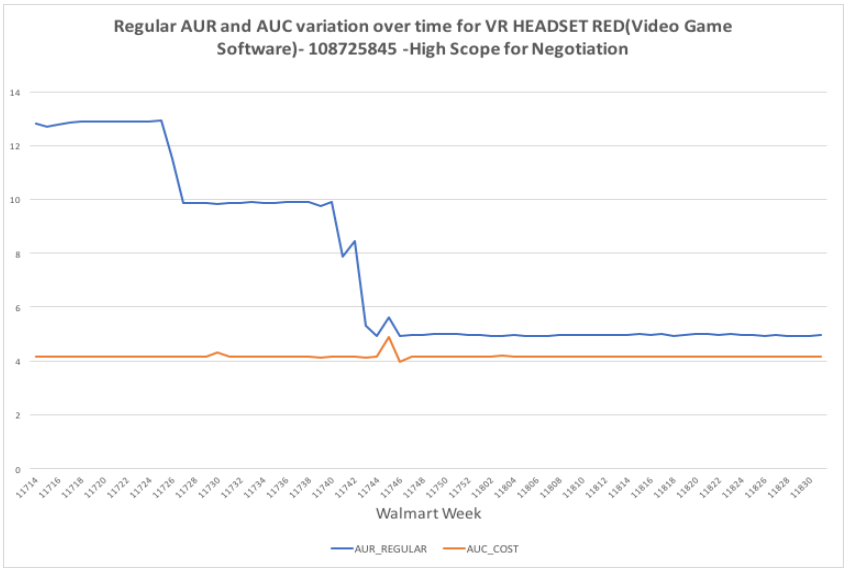


Figure 2: Trend based analysis of AUC and AUR

The correlation coefficient gives an adept understanding of the variation of the one of the variables with respect to another and hence in this case it has proven to be efficient in detecting negotiable items efficiently.

The entire scenario can be broken down into the following cases:

1. Sale Price increasing, Cost Price increasing
    i. r_ratio > 1 - No action needed
    ii. r_ratio <1 - Scope for Negotiation
2. Sale Price decreasing, Cost Price decreasing
    i. r_ratio > 1 - Scope for Negotiation
    ii. r_ratio <1 - No action needed
3. Sale Price decreasing, Cost Price decreasing
    i. No actions needed
4. Sale Price decreasing, Cost Price increasing
    i. Scope for Negotiation

As shown in Figure2, the AUR retail price decreased with time for that item whereas the unit cost price remains almost constant giving an indication for potential negotiation.
Comparison of Average unit cost (Cost per unit) across different items within similar item clusters can also aid to cost negotiation. Our aim is to find items with similar description and create clusters, so that we can perform comparison of AUC among the items within the same cluster. Also, these clusters can be leveraged for additional cost negotiation opportunities such as Join Buy, Dis-intermediation and Parent-Supplier Connection.

For creation of item similarity clusters, we need first obtain similar description-based items. It is computationally expensive as we need to find similarity score between all the possible item pairs. Product description of that item includes the most relevant information of the same including its textual details, category, subcategory, fineline information,color, texture, brand etc.
To compute the pairwise item similarity, we have extracted various natural language based features including n-grams, percentage match, words share etc. The metrics thus formed have been used to compute various similarity scores like Jaccard index, Cosine similarity etc. whose weighted average gives the separation between the two items which is an indication of item substitutability.

To make the number of item pairs for computing similarity metrics, we have leveraged Lucene search library [9]. Lucene is an inverted full-text index. It takes all the documents, splits them into words and then builds an index for each word. We would take help of the item descriptions like signage, upc, supplier inputs etc. to create the index for the items. Lucene library is utilized for building the index on the concatenated descriptions of the item. Once we have the index built, we parse each item through the index to obtain top 50 most relevant similar items based on their descriptions. The resultant set would be up to 50 most relevant items for each such item that has been parsed based on the similarity score computed by Lucene. This process continues for all the items that have been considered. The output is in the form of similar item pairs with their scores. The flow for computing these similar cluster items can be shown in Figure 3.

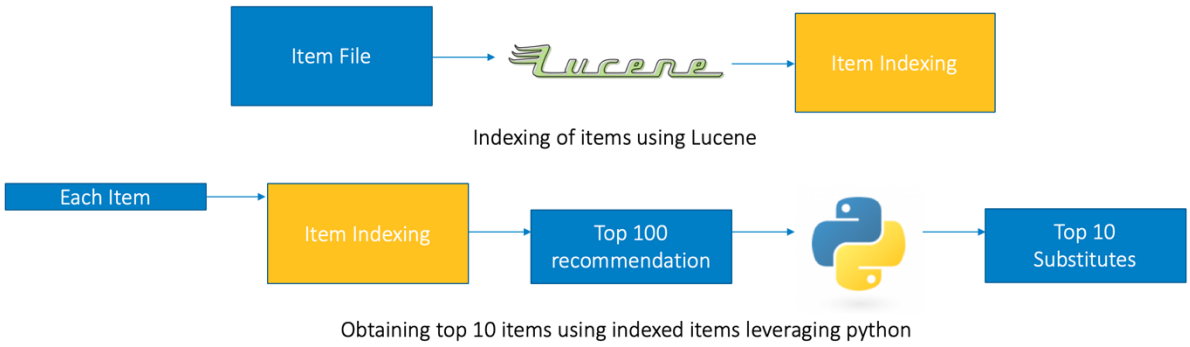


Figure 3: Flow for Computing similar cluster items

The similar item pairs obtained using the Lucene search functionality can be visualized in the form of a network. Each item would resemble a node and every item pair resembles an edge connecting the nodes as shown in Figure3. The scores between each item pair is utilized as edge weight in the graph. We need to identify communities of similar items among these item pairs..

There are numerous algorithms to achieve this, but we need to choose an algorithm which can follow top-down approach. Firstly, lets understand how different algorithms operates.

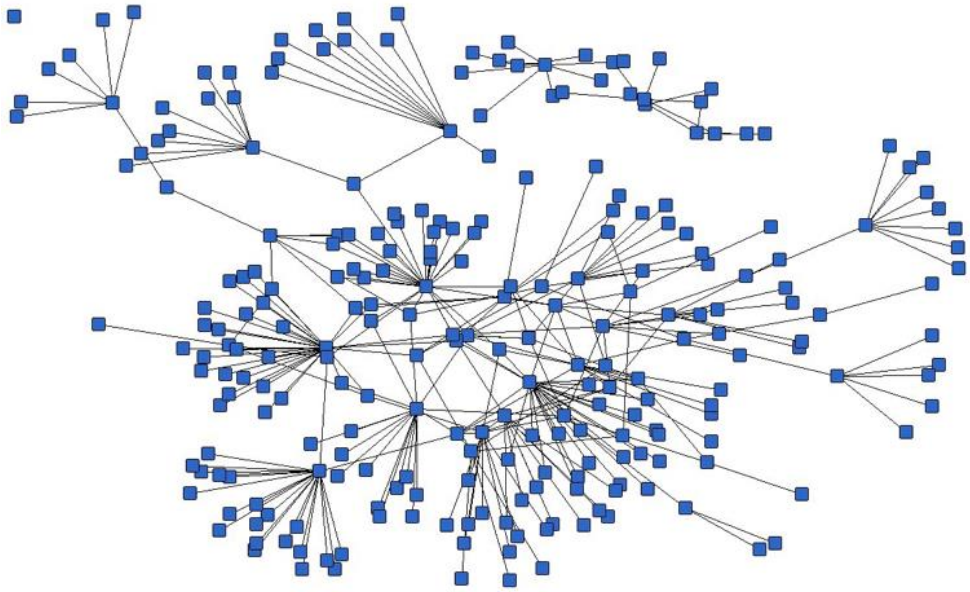


Figure 4: Community Detection in Graphical Networks

Edge-betweenness-community is a hierarchical decomposition method where edges are removed in the decreasing order of their edge betweenness. This is done assuming that edges connecting different groups mostly contain multiple shortest paths. It yields good results but is slow since its computation intensive due to calculations of edge betweenness. Also, they need to be re-calculated after every edge removal. It's ideal for few thousands of vertices in the data. It builds a full dendrogram and does not give the final cut off point to obtain the final groups.

Fast greedy-community is another hierarchical approach, but it is bottom-up instead of top-down. Modularity gets optimized in a greedy manner. Communities are merged iteratively considering every vertex belongs to a different community so that each merge is locally optimal. The algorithm stops when modularity has reached its saturation. The method is fast and hence most sorted as there are no parameters that needs to be fine-tuned. One drawback is that communities below a given size gets merged with neighboring communities.

Walk trap-community performs random walks on the network. These walks tend to stay within the same group because there would be few edges that lead to outside group. Walk trap runs short random walks of 3-5 steps and results are used to merge separate communities in a bottom-up fashion. Modularity score can be leveraged to select where to cut the dendrogram. It is a bit slower than the fast-greedy approach but is more accurate.

Spinglass-community is from statistical physics based on Potts model. Each vertex of the network can be in one of c spin states, and the edges of the specify which pairs of vertices would prefer to stay in the same spin state. Communities get defined after simulating for a given number of steps based on the spin states of the vertices in the network. This method is not fast and not deterministic due to the inherent simulation, but has a tunable parameter to determine the cluster size.

Leading-eigenvector-community is a top-down hierarchical method which optimizes the modularity function. At each iteration the network is split into two parts such that the separation yields a significant increase in modularity. Split is determined by evaluating leading eigenvector of modularity matrix. Due to the eigenvector calculations, it might not work on degenerate networks.

Label-propagation-community is a simple method where every node is assigned one of k labels. During different iterations it re-assigns labels to nodes in a way that each node takes the most frequent label of its neighbors. The method stops when the label of each node is one of the most frequent labels in its neighborhood. It is very fast but yields different results based on the initial seeds. Hence, one should run the algorithm a large number of times and then build the final communities which is tedious.

For this paper we have utilized customized constraint-based Louvain community detection [10] algorithm to determine the optimal clusters. This algorithm also optimizes modularity which is a scale value between -1 and 1 that measures the density of edges inside communities to edges outside communities. First small communities are identified by optimizing modularity locally on all nodes, then they are grouped into one node for the next iteration.

$$Q = \frac{1}{2m} \sum_{ij} \left[A_{ij} - \frac{k_i k_j}{2m} \right] \delta(c_i, c_j), \tag{2}$$

where,

- A_{ij} represents the edge weight between nodes i and j .
- k_i and k_j are sum of weights of the edges attached to nodes i and j
- $2m$ is the sum of edge weights in the graph
- c_i and c_j are communities of the nodes
- δ is a delta function

The modularity (2) has been maximized using constrained optimization on the UPC information.

To identify how granular the communities needs to be formed is a critical task. For solving this problem, we have leveraged the UPC information. The same UPC might be supplied by different suppliers and would have different item numbers. Since we are grouping different item numbers under the communities, the UPC information would help us to determine if the communities have been formed correctly. We need to ensure that most of the communities that have been formed contains all the items of the same UPC number falling into the same community. If this fails for 90% of the items then we re-run the algorithm by fine tune the edges and breaking the communities further to achieve the above constraint.
On removing weak edges from the network, the Louvain algorithm would create more granular clusters leading for most of the items having same UPC to fall into the same community. Several iterations are performed and under each iteration a new network is created with the fine-tuned edges and the communities are identified ensuring that 90% of the items having same UPC fall into the same community.

Once communities are identified and all the items are tagged to each of the community. We would compute the volume and margin percent for each of the item. Basis the distribution of the volume and margin percent we are going to classify each of the cluster into four groups as shown in below table. Each group would help every retailer to take a strategic decision for improved quality and sales as shown in Table 1.

Table 1	
Group	**Strategic Decision for Retailers**
High Selling High Margin	Strategic Items
High Selling Low Margin	Negotiation Opportunity
Low Selling High Margin	Marketing/Promotional Opportunity
Low Selling Low Margin	Alternate Sourcing Opportunity

Table 1: Shows how retailers can take strategic decisions using the Group information different items get assigned

Also, retail-based cost negotiations can be performed within each cluster that has been formed. For example, in the below cluster we can see that there are 4 similar items. But one can observe that we have a good margin percent and good volume percent for the first two items on comparison with the bottom two items. Customers are tending to buy smaller packs comparative to the larger packs in this variety of items. As a retailer we can improve the profits by procuring more quantities of first two items as they have been liked by the customers and they have been having maximum margin percentage as shown in Table 2.

Table 2		
Item Name	**Units%**	**Margin%**
GV NFC ORANGE JUICE 59OZ CARAFE	0.69	0.29
GV NFC ORANGE JUICE WITH PULP 59OZ CARAF	0.36	0.29
GV NFC ORANGE JUICE 109OZ CARAFE	0.01	0.04
GV NFC ORANGE JUICE WITH PULP 109OZ CARAF	0.01	0.07

Table 2: Shows items falling in same cluster for cost negotiations

All the item clusters that have been formed would be classified to one of the groups as shown in the table 3. Each group would help to take a strategic decision by the retailer which would eventual help in increased customer satisfaction, increased sales and more footfalls.

Table 3		
Cluster No	**Item Name**	**Group**
1	MYBIOME WOMENS PROBIOTIC	Low Selling High Margin
1	MYBIOME MENS PROBIOTIC	Low Selling High Margin
1	MYBIOME 50+ PROBIOTIC	Low Selling High Margin
2	OLE CREMA SQ 15OZ	High Selling High Margin
2	OLE CREMA SQUEEZE	High Selling High Margin
3	SIMPLY GRAPEFRUIT JUICE 52OZ	High Selling Low Margin
3	SIMPLY CRANBERRY COCKTAIL 52OZ	High Selling Low Margin
3	SIMPLY GRAPEFRUIT JUICE	High Selling Low Margin
4	TROPICANA PP LEMONADE 52OZ	Low Selling Low Margin
4	TROPICANA PP RASPBERRY LEMONADE 52OZ	Low Selling Low Margin
4	TROPICANA PP PEACH LEMONADE 52OZ	Low Selling Low Margin

Table 3: Shows sample clusters number and the groups they have been assigned

IV. CONCLUSION

Analyzing just average unit retail and comparing with the average unit cost would help in identifying cost negotiation possibility in one dimension. On identification of item similarity clusters and comparing AUC with AUR across different items within the same item similarity cluster would help in leveraging cost negotiation opportunities in multiple dimensions, since item cost performance across different items is being compared. Tagging of each item into different groups in terms of their selling and margin performance helps the sourcing managers to better negotiate with different suppliers across the globe.

ACKNOWLEDGMENT

We want to grasp this opportunity to convey our deepest gratitude to all of them who aided us in completing our paper. Our heartfelt gratitude to AICAAM team, for this opportunity. Our great thankfulness to Data and Analytics Team, Walmart International, for introducing us to this topic and guiding us throughout our paper. A special thanks to Ashish Gupta, Manager Walmart Labs as our internal guide of the paper, for his persistent encouragement and motivation towards us. We are thankful to all of them who directly and indirectly helped us in accomplishing this paper.

REFERENCES

[1] Robert H. Guttman,Pattie Maes, "Agent-Mediated Integrative Negotiation for Retail Electronic Commerce", in *AMET*, 1998, pp 70-90.

[2] Lewicki, D. Saunders, and J. Minton. Essentials of Negotiation. Irwin, 1997.

[3] Rosenschein and G. Zlotkin. Rules of Encounter: Designing Conventions for Automated Negotiation among Computers. MIT Press, 1994.

[4] C. Beam, A. Segev, and J. G. Shanthikumar. „Electronic Negotiation through Internet-based Auctions." CITM Working Paper 96-WP-1019, December 1996.

[5] E. H. Miller, "A note on reflector arrays (Periodical style—Accepted for publication)," *IEEE Trans. Antennas Propagat.*, to be published

[6] C. Sierra, P. Faratin, and N. Jennings. "A Service-Oriented Negotiation Model Between Autonomous Agents." *Proceedings of the Eighth European Workshop on Modeling Autonomous Agents in a Multi-Agent World (MAAMAW'97)*. Ronneby, Sweden, May 1997.

[7] Kraus, S.: Negotiation and Cooperation in Multi-Agent Environments. Artificial Intelligence journal, Special Issue on Economic Principles of Multi-Agent Systems, 94(1–2):79–98, (1997).

[8] De Paula, G. E., Ramos, F. S., Ramalho, G. L.. Electronic Commerce Negotiation Model Formalization. Technical report UFPE-DI 99-20. 1999

[9] Balipa, Mamatha & Ramasamy, Balasubramani. (2015). Search Engine using Apache Lucene. International Journal of Computer Applications. 127. 27-30. 10.5120/ijca2015906476.

[10] De Meo, Pasquale & Ferrara, Emilio & Fiumara, Giacomo & Provetti, Alessandro. (2011). Generalized Louvain Method for Community Detection in Large Networks. International Conference on Intelligent Systems Design and Applications, ISDA. 10.1109/ISDA.2011.6121636.

[11] Blondel, Vincent D; Guillaume, Jean-Loup; Lambiotte, Renaud; Lefebvre, Etienne (9 October 2008). "Fast unfolding of communities in large networks". *Journal of Statistical Mechanics*: Theory and Experiment. 2008

Diagnosis of Neurodegenerative Disorders in Brain MRI Using Tissue Variation and SVM Classifier

N.Ahanapriyanka and G.Kavitha'

Department of Electronics Engineering, MIT Campus, Anna University, Chennai, ahanachellian@gmail.com

***Abstract*—**Dementia is a neurological impairment, which results in loss of mental ability to perform a regular task. Alzheimer disease (AD) and Mild Cognitive impairment (MCI) are most common forms of dementia. The tissue variation in brain helps to identify the pathological changes in these disease progressions. Thus, an attempt is made to diagnose the severity of such neurodegenerative disorders based on pattern changes in brain tissues. Initially, MR images are pre-processed using robust brain extraction (ROBEX) tool. The skull stripped images are subjected to Tsallis entropy based on multilevel thresholding to segment the brain tissues such as white matter (WM), gray matter (GM) and cerebrospinal fluid (CSF). Pyramid histogram of gradient (PHOG) and Zernike moments (ZM) features are extracted from different segmented brain tissues for normal, MCI and AD. The significant features selected based on principal component analysis (PCA) are subjected to least square SVM (LSSVM) and SVM classifiers. Result shows that prominent tissue variations are observed in PHOG features compared with ZM. It is noticed that SVM classifier is able to classify the normal, MCI and AD images better than LSSVM. Higher classification accuracy is obtained for GM. These finding suggest that GM is able to show significant difference between normal, MCI and AD subjects. Thus, this study could aid to analyse the brain tissue interior variation and connection between neurodegenerative disorders. Hence this approach could be used as a supplement in the investigation of dementia disorder.

***Keywords*—**Alzheimer disease, mild cognitive impairment, brain tissue, white matter, gray matter, cerebrospinal fluid, pyramids histogram of gradient(PHOG), Zernike moment(ZM)

I. INTRODUCTION

Dementia is the most common chronic degenerative symptom. This leads to various brain disorders such as Alzheimer (AD), Mild cognitive impairment (MCI), Vascular Dementia, Dementia with Lewy Bodies (DLB), Frontotemporal Dementia (FTD) and mixed dementia. AD and MCI is recognized as important cognitive problem as a dementia risk [1]. Various researches suggest that subjects with MCI have greater chance to convert into AD due to prominent memory disturbance. Approximately 50 million people worldwide live with dementia [2]. It causes major impact in cognitive skill and cause difficultly to lead a social economic life. The identification of cognitive feature is of greater importance to provide an advent treatment. Currently, diagnosis of dementia relies on clinical assessment which limits to highlight the substantial changes in the brain regions. However, many studies suggest that tissue variations of these disorders incorporate the potential information about the diseases progression [3].

The diagnostic strategy for these degenerative disorders can be developed with the combination of clinical assessment and analysis of pathological changes [4]. These changes are generally observed by Magnetic Resonance (MR) imaging. It is a noninvasive, nondestructive and flexible imaging technique that does not require ionizing radiation.

The MR brain images relatively contain extra-cranial tissue. The delineation of these extra-cranial tissues requires preprocessing which is commonly referred as skull stripping. Various techniques [5] such as morphological operation, elastic registration and tissue segmentation are used to separate the brain from extra-cranial tissues. Based on various report Robust learning based Brain Extraction System (ROBEX) tool results better delineation of T1 weighted MR brain image [6]. It is based on point distribution model adjusted by using voxel classification with random forest algorithm. This helps to identify discrimination in tissue effectively.

Recently, a variety of data analysis and machine learning methods are used to discriminate among normal, MCI and AD subjects. Based on various study deterioration of brain widely affect the brain tissue regions such as White matter (WM), Gray matter (GM) and cerebrospinal fluid (CSF) [7]. Diagnosis of tissue variation in these disorders is challenging due to heterogeneous pattern that is analogous to other brain disorder. A precise and accurate diagnosis of these tissues variation would facilitate the diagnosis and understanding of this disease. Hence, segmentation is carried out.

Segmentation of brain tissue is a tough task due to the irregular intensity between tissues and its boundaries [8]. Many techniques such as voxel based method, vertex based method, thresholding, clustering, graph based method and Region of Interest (ROI) have been attempted for segmentation of brain tissues [9]. Among them thresholding is commonly used. In an image, pixel with gray level values higher than a certain threshold value ‘T’ are categorized as object of the image and remaining gray level values are lesser than ‘T’ are categorized as background image [10] . However, in case of medical image, bi-level thresholding does not give appropriate performance. As a result, there is strong requirement for multilevel thresholding. Tsallis entropy is simple to implement and easily extended to multilevel thresholding problems compared to other entropy due to its non extensive property [11]. The moment-preserving principle used to select the thresholds of the gray-level image. The segmentation result requires quantification study to understand its pathology. Hence quantification is carried out by feature extraction techniques.

Now-a-days, feature based analyses are utilized in the study of complicated disorders associated to brain of human for the reliable diagnosis. Zernike moment (ZM) [12] and Pyramid Histogram of gradient (PHOG) [13] descriptors are considered for this study. These features have been effectively used in variety of medical image analysis including detection of tumors and degenerative disorders. However, identifying discriminative features is challenging owing to the large set feature.

Feature selection method help to identify the prominent feature set. In this study, feature selection is performed using principal component analysis (PCA). This method effectively represents the intrinsic pattern of the given features [14]. This methodology has been applied to face recognition and image denoising techniques. Neuroanatomical pattern classification has recently facilitated the identification of imaging biomarkers. Recently various machine learning techniques are applied for classification. Among them, SVM [15] and least square SVM (LS-SVM) [16] classifier is widely used in EEG signal classification, schizophrenia, breast cancer and breast microscopic images. Hence in this work SVM and LS-SVM are adopted for classification of normal, MCI and AD images.

The contribution of this work is to identify the variations of normal, MCI and AD subjects in their neuroanatomical region of brain MR images. This study preciously made to interpret the

tissue pattern changes in WM, GM and CSF effectively. Initially T1 weighted MR brain images are skull stripped using ROBEX tool. Then the images are subjected to segmentation using multilevel Tsallis entropy based segmentation for delineation of brain tissue. Different features such as ZM and PHOG are extracted from segmented brain tissue. Then, the effective feature set is identified using PCA. These feature sets are help to discriminate and analyse the tissue pattern changes. Finally, SVM and LS-SVM is used to classify normal, MCI and AD images.

The paper is structured as follows. In Section 2 brief description of methods that includes dataset details, skull-stripping, feature extraction, feature selection and classifier are focused. Section 3 presents the results and discussions of framework. Finally, the conclusions are given in Section 4.

II. METHODS AND MATERIALS

The flow diagram of the work is depicted in Fig. 1. The input images, obtained from publically available database are exposed to skull-stripping process using ROBEX tool. Segmentation of various brain tissues are carried out using multilevel Tsallis entropy. The pattern variations features such as ZM and PHOG are extracted from segmented brain tissue. The significant features are selected using PCA, and are classified as normal, MCI and AD using SVM and LS-SVM classifiers.

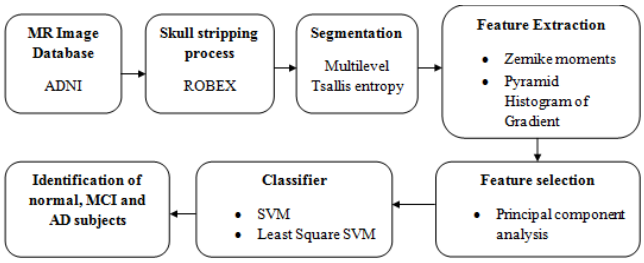


Fig. 1. Flow diagram of the proposed work

IMAGE DATABASE

In this study, total of 750 images (N=250, MCI=250 and AD=250) are considered. The images are acquired from Alzheimer's Disease Neuroimaging Initiative (ADNI) database (http://adni.loni.usc.edu/data-samples/). The T1-weighted transaxial view MR images are used for the analysis. The MCI and AD images are selected based on the Mini mental state examination (MMSE) and clinical dementia rating (CDR) score which is available in the database. All the MR scans are atlas registered and bias field corrected images. These MR images are subjected to further processing.

SKULL STRIPPING

The normal, MCI and AD T1 weighted MR Images are skull stripped using ROBEX tool. It is based on the combination of generative and discriminative model to delineate the non brain

tissue [6]. The brain boundary is detected with the help of discriminative model using random forest classifier. The generative model finds the contour of brain tissue with highest like hood in accordance with discriminative model. The contours are obtained using graph cut method to attain the target shape. Thus the tool produces the result with improved intensity standardization, higher boundary probability volume and better final segmentation.

MULTILEVEL TSALLIS ENTROPY BASED SEGMENTATION

Multilevel thresholding uses more than one threshold value and creates an output image with multiple groups [17]. Tsallis entropy is used to measure the improbability of the system. Tsallis entropy can be extended to non-extensive system based on a general entropic formula which is represented in Eq. (1) [18].

$$s_q = \frac{1-\sum_{i=1}^{k}(p_i)q}{q-1} \tag{1}$$

where k is the total number of possibilities of the system and q is the measure of degree of non-extensivity of the system called Tsallis parameter or entropic index . This entropic form can be extended for a statistical independent system by a pseudo-additive entropic rule which is expressed in Eq. (2) [11].

$$s_q^{A+B}(t) = s_q^A(t) + s_q^B(t) + (1-q)s_q^A(t)s_q^B(t) \tag{2}$$

This method is used to obtain threshold value for segmentation of an image. The information measure between required classes is maximized and the corresponding gray level is considered to be the desired threshold values of particular region. Tsallis entropy of an image I depends on the desired threshold t. Here, $s_q^A(t)$ and $s_q^B(t)$ represent the entropy of each class, and the third term $s_q^A(t)s_q^B(t)$ represents the interaction between the threshold classes. Parameter q can be lesser or greater than 1. The value of q is less than one as the entropy becomes sub-extensive and greater than one as the entropy become super-extensive. In this work, non-extensiveness parameter ‘q’ is considered to be 0.5 in order to improve the correlation and effectiveness of the threshold values.

ZERNIKE MOMENT (ZM)

ZM is a global descriptor with the class of orthogonal moments. ZMs belong to the class of orthogonal rotation invariant moments (ORIMs). It is obtained by projecting the input image onto the complex orthogonal Zernike polynomials. The discrete form of the ZM for an image with the size N × N is expressed as follows [12],

$$z_{n,m} = \frac{n+1}{\lambda_N}\sum_{C=0}^{N-1}\sum_{r=0}^{N-1} f(x,y)V_{n,m}(x,y) \quad \text{Where}$$

$$V_{n,m}(x,y) = R_{n,m}(\rho_{xy})e^{-jm\theta_{xy}} \tag{3}$$

In Eq. (3) $V_{n,m}(x,y)$ is the Zernike basis functions, $R_{n,m}$ is a radial polynomial and ρ is the length of vector from the origin to (x, y) with the range of 0 ≤ xy ≤ 1, where N is a normalization factor. This enables the contribution of each moment to be unique and independent of information in an image. To analyze the performance and effectiveness of ZM a group of high-order and low-order moments have been extracted from normal, MCI and AD images. The first, second and third group includes 36, 44 and 25 moments which is based on Eq. (4).

$$Group1 = \begin{cases} 1 \le n \le 10 \\ |m| \le n \\ n - |m| = 2k \\ k \in N \end{cases} \text{ and } Group2 \ \& \ 3 = \begin{cases} 11 \le n \le 20 \\ |m| \le n \\ n - |m| = 4k \text{ for group2} \\ n - |m| = 8k \text{ for group3} \\ k \in N \end{cases} \quad (4)$$

PYRAMID HISTOGRAM OF GRADIENT (PHOG)

PHOG descriptors are used to represent the shape and spatial patterns of an image. The objective of the PHOG [19] is to take the spatial property of the local variation into account. In PHOG edge contours are extracted using the Canny edge detector for an image. The grid at resolution level (L) had desires cells along each dimension. The orientation gradients are then computed using a 3*3 Sobel mask without Gaussian smoothing [20]. Here shape descriptors are used in the orientation range [0-180]. PHOG bin size is fixed throughout as k bins (here k = 8). Each PHOG feature for the entire image is a vector with the dimensionality $k\sum_{l \in L} 4^{l}$. Finally, based on the levels L, the images are divided and HOG is computed. Finally, the histograms of all the blocks consist of a whole HOG descriptor.

PRINICIPAL COMPONENT ANALYSIS (PCA)

PCA is able to transform samples into a new space and to use lower-dimensional representation from the new space to denote the sample. Consider m be the numbers of features are taken, with respect to n as m × n matrix A. Let k < n be the dimensionality of the space that seeks to embed features. Assume that the columns (features) of A are mean-centered [21]. Then, PCA returns the top k left singular vectors of A as m ×k matrix P_k. After that projects the data on the k-dimensional subspace spanned by the columns of P_k. Let $M_{U_K} = U_K U_K^T$ be the projector matrix on the resultant subspace. It is well-known that the resulting projection is optimal in its residual which is represent in Eq. (5),

$$\| A - P_{U_K} A \| \xi \quad (5)$$

Thus this process minimized over all possible k-dimensional subspaces.

CLASSIFIER

Support vector machine (SVM) maps the input points into a high-dimensional feature space and finds a separating hyperplane that maximizes the margin between two classes in this space. Without any knowledge of the mapping, the SVM finds the optimal hyperplane by using the dot product functions in feature space that are called kernels [22]. The solution of the optimal hyperplane can be written as a combination of a few input points that are called support vectors. In the feature space, the decision function that separates classes is given as Eq. (6)

$$D(x) = w^T \phi(x) + b \tag{6}$$

Where w and b is the one dimensional vector and bias term respectively. Then the SVM is formulated in the primal form as,

$$Q = \frac{1}{2} w^T w + c \sum_{i=1}^{M} \xi_i \ \ subject \text{ to } y_i D(x_i) \geq 1 - \xi_i \tag{7}$$

In Eq. (7), ξ_i are artificial slack variables representing classifier errors and C is a constant. It used to avoid parameter sensitivity.

The Least Square SVM is the Least Square interpretation of Support Vector Machine. It finds the pattern in the data which can be used for classification and regression analysis [23]. This approach minimizes the sum of squares of the errors made in the results of every single equation to determine the line of best fit to the model. In LS-SVM for function estimation, the optimization problem is formulated as [24]

$$\min_{a,b,c} J(a,e) = \frac{1}{2} \| a \|^2 + \frac{1}{2} \gamma \sum_{k=1}^{N} e_k^2 \tag{8}$$

Where, $e_k \in R$ are error variables; and $\gamma \geq 0$ is regularization constant. To solve this optimization problem, Lagrange function is constructed as

$$\frac{1}{2} \| a \|^2 + \frac{1}{2} \gamma \sum_{k=1}^{N} e_k^2 - \sum_{k=1}^{N} \alpha_k \left\{ ek + (w, \phi(x_k)) + b - y_k \right\} \tag{9}$$

where, $\alpha_k \in R$ are Lagrange multipliers. Solving by differentiating the above Eq. (8) and Eq. (9) least square model can be solved in kernel space. This will improve the robustness of the classifier.

III. RESULT AND DISCUSSION

The image dataset used in this work includes normal, MCI and AD subjects. The MR brain images are skull stripped using ROBEX tools. The T1 weighted axial view of MR images are skull stripped. Fig.2 (a) shows the original image in the database. The corresponding skull stripped images are shown in Fig. 2(b). Visual results show that the ROBEX tool is able to extract the whole brain tissues better for both normal, MCI and AD subjects. It is able to observed the there exist a tissue variation in delineated images as the disease progresses. Finally, the obtained skull strip images are used for segmentation. This automated skull stripping methods help to improve the accuracy of prognostic and diagnostic procedures in brain image segmentation and analysis.

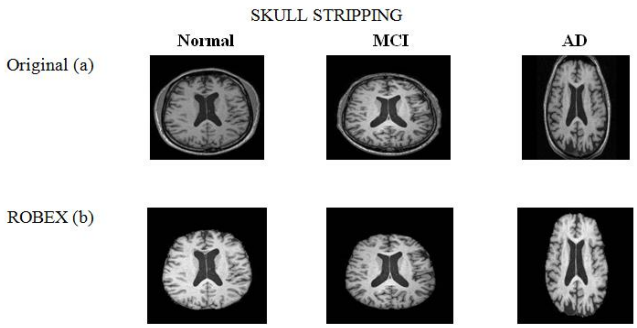


Fig. 2. Skull Stripping using ROBEX (a) Typical MR image and (b) Segmented Whole brain image

Fig. 3(a) represents T1 weighted MR image of Normal, MCI and AD subjects. Fig. 3 (b), (c), (d) and (e) represent the segmented image, white matter, grey matter and CSF. The visual representation of the images indicates there exist the tissue variation among normal, MCI and

AD subjects respectively. In the Fig 3(a) original image, it is evident the visibility of different tissues is strongly dependent on the different threshold values.

In Segmentation Tsallis entropy is used to segment the tissue. Based on global and objective property of the image histogram in Tsallis entropy it is suitable to implement in multilevel thresholding case. Parameters are tuned for Tsallis to improve the image thresholding values to segment the brain tissue. It is observed from segmented image in Fig.3 (b) that brighter, darker and background regions are well separated by different threshold values to obtain different tissues.

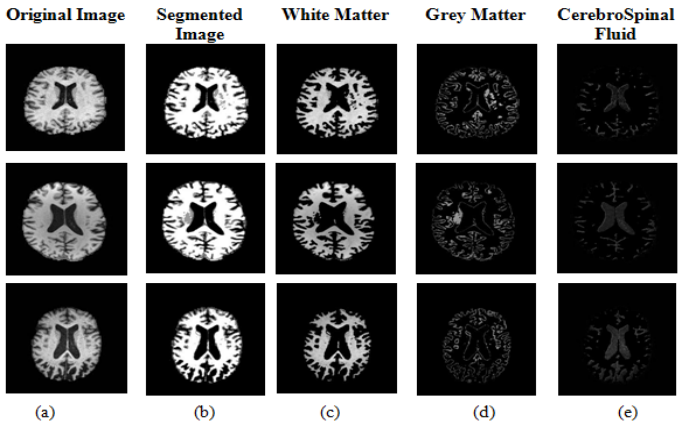


Fig. 3. Segmentation of Brian Tissue (a) Typical MR image (b) Segmented image(c) White matter (d) Grey matter (e) Cerebrospinal fluid (1st, 2nd and 3rd row represent the normal,MCI and AD respectively)

Then from the segmented WM, GM and CSF- PHOG and ZM features are extracted for normal, MCI and AD images. PHOG features is extracted for L=1, 2, 3 levels which indicate the number of pyramids. Fig. 4 (a), (b) and (c) depicts PHOG descriptor at each level respectively.

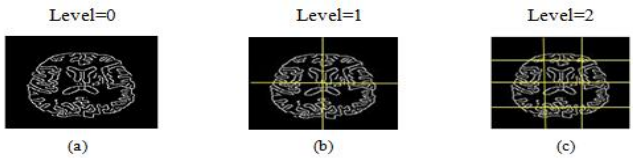


Fig. 4. Typical Representation of PHOG feature at each level (a) Level=0 (b) Level=1 and (c) Level=2

From the Table I . It shows each level corresponding to different set of features.

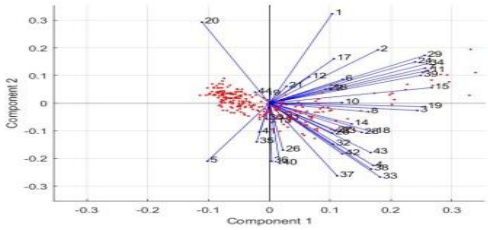
TABLE I. FEATURES EXTRACTED FROM THE SEGMENTED IMAGE

Pyramid Histogram of Gradient		
Level	***Bin***	***Number of Features***
1	8	40
2	8	168
3	8	680
Zernike moment		
Group	***order (n)***	***Number of moments***
1	1-10	36
2	11-20	44
3	11-20	25

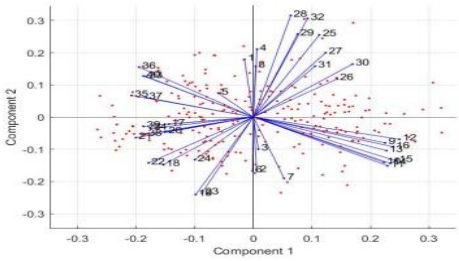
Zernike moment is based on orthogonal property where moment represents the descriptor information of an image. Zernike feature require moment order (n) and repetition moments (m) to extract the features. Here Zernike amplitude features for three groups such as (G=1, 2, 3) are extracted. To analyze the effect of orders of Zernike moments on the performance of the

overall system, a group of high-order and low-order Zernike moments have been extracted. These three groups of Zernike moments are analysed to determine the various detail of the tissue.

The extracted features from PHOG and ZM for different order and various levels are given to feature selection method. PCA perform selection based on the highest eigen values. The obtain eigen value is used to select the significant features set.



(a)



(b)

Fig. 5. Biplot representation of (a) Zernike moments and (b) PHOG features

The multivariate features of PCA can be visually represents by means of biplot. In biplot the features in original space are more representatives in the new generated space. In Fig. 5, PCA biplots allow visualization of the magnitude and sign of each features contribution to the largest two principal components. The x-axis is represented by principal component 1 and the y-axis by Principal component 2. In the above biplots, PCA scores are represented in red points and the original features are represented in blue vectors. The distance between the scores and the vector shows how much the score influences the vector or vice versa [25]. If the score are situated on the opposite side of the PC coordinate compared to the vector then the scores have less effect on features. This factor determines the influence of significant features. The size of the angle between vectors determines the correlation of the features. A longer vector means the variable is well represented by the plot and vice versa. A small angle indicates a strong positive correlation, 90 degrees represents no correlation and 180 degrees represents a negative correlation.

A visual analysis of the biplots indicates that Zernike and PHOG methods seem to capture the information in different ways, as evidenced both by the different nature of the radial dependence of features with respect to the first two principal components as well as by the respective distribution of the scores. In ZM feature the score are less influence to the vector when compare to PHOG. The angle of projection for PHOG and ZM comparison of actual vector lengths in the biplots is difficult, largely due to the matrix-specific PCA variable scaling.

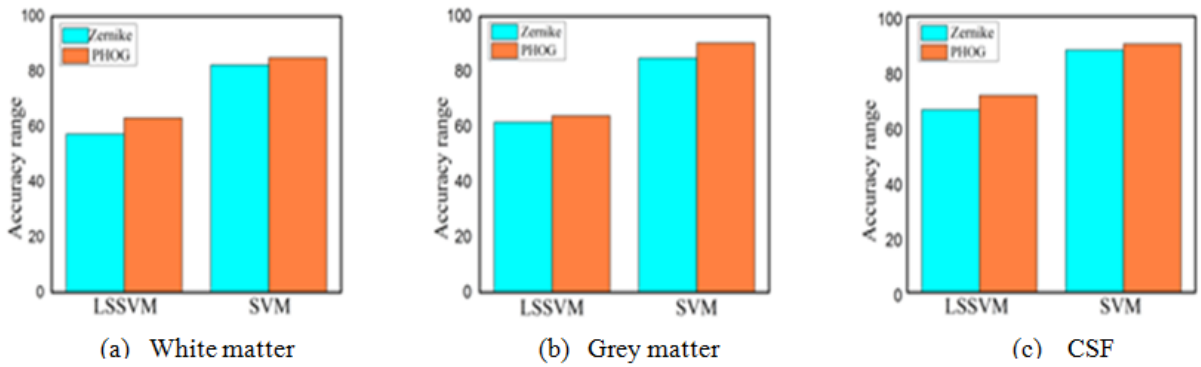


Fig. 6. Comparison of Zernike and PHOG features based on classifier accuracy for different brain tissues (a) White matter (b) Grey matter and (c) CSF

The considered groups and levels of extracted feature classification were carried out for each brain tissues. The classification measure accuracy is consider to evaluate the performance of LS-SVM and SVM. The average classification accuracy of ZM and PHOG feature for various brain tissues are shown in Fig. 6. It is observed that SVM gives more accuracy when compare to LS-SVM. It is notice that ZM give lesser accuracy and variation in brain tissues when compare to PHOG. This shows that PHOG that are extracted from WM, GM and CSF seems to have important discriminating features between normal, MCI and AD images.

TABLE II. CLASSIFICATION ACCURACY OF PHOG FEATURES USING LS-SVM AND SVM

PHOG (% Accuracy)	***LS-SVM***			***SVM***		
	WM	***GM***	***CSF***	***WM***	***GM***	***CSF***
LEVEL 3	94.8	98	97.6	98	99.2	98.8
LEVEL 2	94	96.4	95.6	96	97.6	96.8
LEVEL 1	58.4	60.4	70.4	60.8	74.4	74.4
Average	**82.4**	**84.9**	**87.8**	**85**	**90.4**	**90**

The classification result of PHOG features for each level is represent in Table 2.The results shows a that there exist a reliable tissue variation in every level in PHOG between the brain tissues. SVM perform better in each levels of PHOG than LS-SVM. In SVM it is observed that at each level there exist a difference in tissue variation in WM, GM and CSF. At each level GM appear to have high discrimination than WM and CSF. From the above result, it is clear that level 3 in SVM shows higher classification accuracy when compared to other levels. Hence SVM in level 3 gives maximum accuracy of 99.2% for grey matter than WM (98%) and CSF (98.8%).Thus the significant feature PHOG and its corresponding level is identified. The significant difference in GM is able to classify the normal, MCI and AD subjects. Another critical observation shows that changes in GM [26] cause shrinkage in hippocampus, cortical and sub-cortical along the discriminating directions. Such patterns of changes are well known to characterize the disease progression in AD and related dementia.

TABLE 3 COMPARISON OF ACCURACY OBTAINED BY THE PROPOSED METHODOLOGY AND CONVENTIONAL METHOD

Database	***Approach***	***Accuracy***	***Reference***
ADNI	*SPM tissue segmentation +PCA+SVM*	83.48%	*27*
ADNI	*Fuzzy possibilistic tissue segmentation + SVM Classification*	71%	*28*
ADNI	***Proposed Method***	90.4%	-

Table 3 compares the accuracy of the proposed method with other state of art methods. It shows that, the proposed framework gives improved accuracy than other methods. These findings suggest that these features can be a useful representation for characterizing brain tissue variationscomplementary to volumetric analysis for better diagnosis.

IV. CONCLUSION

In this work an attempt is made to analyse the tissue variation in normal, MCI and AD subjects from MR brain images. At the beginning delineation of non brain tissue was carried out using ROBEX tool. It is observed that it properly segment the non brain tissue using hybrid approach. Segmentation of a brain tissue from the skull stripped is carried out using multilevel Tsallis entropy. It is also identified from the results better segmentation is carried out based on the improved threshold values. The ZM and PHOG features are extracted for further evaluation. The significant feature set obtains using PCA. It is noticed that compare to ZM feature PHOG performs well. Results show that accuracy of SVM is higher than LS-SVM. Similarly in SVM, GM pattern variation is well diagnosed for normal, MCI and AD images using PHOG feature. These observation shows that GM shows a discriminate variation to classify normal, MCI and AD subjects effectively. These findings suggest that tissue variation can be a useful representation for characterizing dissimilarities in brain structure that is complementary to volumetric analysis. Thus, this framework could be used to study the neuropsychiatric disorder such as dementia. In future, this work could be extended by considering other regions such as caudate, hippocampus and thalamus from different views with more number of samples.

REFERENCES

[1] Sarah Tomaszewski Farias, Dan Mungas, Bruce R. Reed, Danielle Harvey, and Charles DeCarli, MDet, “Progression of Mild Cognitive Impairment to Dementia in Clinicvs Community-Based Cohorts”, Archives of Neurology, pp.1151-1157, Vol 66(9), 2009.

[2] Nina B. Silverberg, Laurie M. Ryan and Maria C. Carrillo et al., “Assessment of cognition in early dementia”, Alzheimer's & dementia: the journal of the Alzheimer's Association, pp.60-76, Vol7 (3), 2011.

[3] https://www.alz.co.uk/research/statistics.

[4] P. Kalavathi and V. B. Surya Prasath,“Methods on skull stripping of MRI head scan images-a review”, Journal of Digital Imaging, Vol 29(3), pp.365-79, 2015.

[5] Francisco J. Galdames, Fabrice Jaillet and Claudio A. Pereza,“An Accurate Skull Stripping Method Based on Simplex Meshes and Histogram Analysis in Magnetic Resonance Images”,Rapport de recherche RR-LIRIS-2011-019.

[6] Juan Eugenio Iglesias, Cheng-Yi Liu, Paul Thompson and Zhuowen Tu,“Robust Brain Extraction across Datasets and Comparison with Publicly Available Methods”,IEEE Trans. on Medical Imaging, 2011.

[7] Lilia Lazli1, Mounir Boukadoum and Otmane Aït-Mohamed, “Brain Tissue Classification of Alzheimer Disease Using Partial Volume Possibilistic Modeling: Application to ADNI Phantom Images”, 2017.

[8] Sanjay Agrawal , Rutuparna Panda, Leena Samantaray and Ajith Abraham,“A novel automated absolute intensity difference based technique for optimal MR brain image thresholding”,Journal of King Saud University – Computer and Information Sciences,2017.

[9] K. R. Anandh, C. M. Sujatha2 and S. Ramakrishnan, "A Method to Differentiate Mild Cognitive Impairment and Alzheimer in MR Images using Eigen Value Descriptors", Journal of Medical System, Vol 40 (1), pp.1-8, 2016.

[10] Arora, S., Acharya, J., Verma, A., & Panigrahi, P. K., "Multilevel thresholding for image segmentation through a fast statistical recursive algorithm", Pattern Recognition Letters, Vol. 29(2), pp. 119–125, 2008.

[11] Qianqian Lin and Congjie Ou n, "Tsallis entropy and the long-range correlation in image thresholding", Signal Processing, Vol.92, pp.2931–2939, 2012.

[12] Amir Tahmasbi n, Fatemeh Saki and Shahriar B. Shokouhi,"Classification of benign and malignant masses based on Zernike moments", Computers in Biology and Medicine, Vol 41(8), pp.726-35, 2011.

[13] U. Rajendra Acharya, Yuki Hagiwaraa, Joel E.W. Koha, Jen Hong Tana, Sulatha V. Bhandary , A. Krishna Rao and U. Raghavendra ,"Automated screening tool for dry and wet age-related macular degeneration using pyramid of histogram of oriented gradients and nonlinear features",Journal of Computational Science, Vol 20,pp.41-51,2017.

[14] G.Nagarajana, R.I.Minub, B Muthukumar, V.Vedanarayanan and S.D.Sundarsinghe, "Hybrid Genetic Algorithm for Medical Image Feature Extraction and selection",International Conference on Computational Modeling and Security, Vol85, pp. 455-462, 2016.

[15] Mahmoodi NM, Arabloo M and Abdi JLaccase, "Immobilized manganese ferrite nanoparticle: synthesis and LSSVM intelligent modeling of decolorization", Water Res, Vol 67, pp. 216-226, 2014.

[16] Takuya Inoue and Shigeo Abe, "Fuzzy support vector machines for multilabel classification",Pattern Recognition, Vol 48(6), pp.2110–2117, 2015.

[17] K.P. Baby Resma and Madhu S. Nair,"Multilevel Thresholding for Image Segmentation using Krill Herd Optimization Algorithm",Journal of King Saud University - Computer and Information Sciences,2018.

[18] Liying Zheng and Kai Tian, "Detection of small objects in sidescan sonar images based on POHMT and Tsallis entropy", Signal Processing, Vol. 142, pp.168–177, 2018.

[19] Shanwen Zhang , Yihai Zhu, Zhuhong You and Xiaowei Wu, "Fusion of superpixel, expectation maximization and PHOG for recognizing cucumber diseases",Computers and Electronics in Agriculture, Vol 140,pp. 338-347, 2017.

[20] N. Dalal and B. Triggs, "Histograms of oriented gradients for human detection", Proc. of IEEE Conference on Computer Vision and Pattern Recgnition, pp. 886-893, 2005.

[21] Christos Boutsidi, Michael W. Mahoney and Petros Drineas,"Unsupervised Feature Selection for Principal Components Analysis", Proceedings of the 14th International conference on Knowledge discovery and data mining,pp.61-69,2008.

[22] Takuya Inoue and Shigeo Abe, "Fuzzy support vector machines for multilabel classification",Pattern Recognition, Vol 48(6), pp.2110–2117, 2015.

[23] Suykens JAK, Van,Gestel T, DeBrabanter J, DeMoor B, Vandewalle JLeast squares support vector machines,World Scientific Pub. Co., Singapor e (2002).

[24] Mahmoodi NM, Arabloo M and Abdi JLaccase, "Immobilized manganese ferrite nanoparticle: synthesis and LSSVM intelligent modeling of decolorization", Water Res, Vol 67, pp. 216-226, 2014.

[25] Mehrdad Hesampour et al.,"Nanofiltration of single and mixed salt solutions: Analysis of results using principal component analysis", Chemical Engineering Research and Design,Vol 8,pp. 1569–1579,2010.

[26] Christiane Möller, Hugo Vrenken , Lize Jiskoot , Adriaan Versteeg, Frederik Barkhof, Philip Schelten and Wiesje M. van der Flier, " Different patterns of gray matter atrophy in early- and late-onset Alzheimer's disease", Neurobiology of Aging , Vol 34(8),pp. 2014-22,2013.

[27] A. Domínguez , J. Ramírez , J. M. Górriz , F. Segovia , D. Salas-Gonzalez , F. J. Martínez-Murcia, I. A. Illán , "Statistical feature selection and classification models for Alzheimer's disease progression assessment", IEEE Nuclear Science Symposium, Medical Imaging Conference and Room-Temperature Semiconductor Detector Workshop (NSS/MIC/RTSD), 2016.

[28] Lilia Lazli ; Mounir Boukadoum ; Otmane Ait Mohamed, "Computer-Aided Diagnosis System for Alzheimer's Disease Using Fuzzy-Possibilistic Tissue Segmentation and SVM Classification", IEEE Life Sciences Conference (LSC), 2018.

Analysis and Tactics of Online Merchandise

Prof. Srinatha D K*, Dr. R. Selvarani*, Devanshi Hiteshbhai Dave*
Department of CSE, Alliance University

,

Abstract- There are so many n number of e-commerce websites available in market which is the reason for the need of having an analysis tools with the help of which an organization can determine whether they are meeting their sales goals or not. Due to which we have aim to develop "Sales Analytic Tools" which is describe briefly, in our paper we have explained about which system do we need to analyses the database transaction of any e-commerce website with help of various data mining algorithms and techniques like k-means algorithm, linear regression, logistic regression. This research paper is all about how to develop a system which takes an input from transaction stored in database. Transaction means details of product sold. After taking the input from database, the input data will be segmented, based on which we will analyse a graph and obtain the products list which are now a day's trending in market most and will get to know the pattern in which we should sale our products. With the help of which our system will provide us tactics of how we can improve our sales and online sales planning to overall increase productivity and profits of an enterprise.

Keywords – Sales analytic tool, data mining, k-means algorithm, logistic regression, linear regression, online sales planning.

I. INTRODUCTION

E-commerce websites are most used in this world of internet as it provides almost all the products which are useful for their customers which make other e-commerce websites difficulty and challenges. In e-commerce websites the very crucial thing is data, data of an enterprise which can be a really big data [2]. There were 35 million online shoppers in India in 2014 and is 100 million by 2016 and is expected to cross $100 billion by 2020 in India out of which $35 billion can be through fashion e-commerce[5]. This means that online sales will increase four times in coming years.

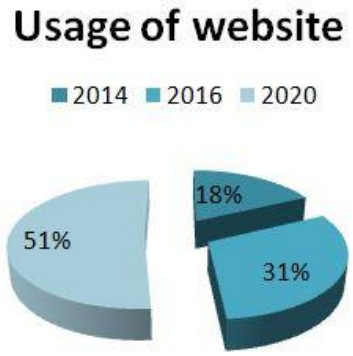


Fig. 1.1 Increase in website usage according year

II. Literature Survey

E-commerce websites are the most trending in this busy world. Since it is very easy and fast way to purchase products in just few clicks. There is very strong competition in e-commerce enterprise for the best and fast sales of products. Therefore it is very important for every enterprise to stay trending with trending fashion. In our system we are performing an algorithm in order to analyse the production and sales of products with various data mining techniques. Data mining helps to analyse database of our system which help us to find different tactics to increase sales of available products. After obtaining tactics it will work on sales optimizations. Optimization helps to manage products by grouping it as fast or slow selling products. This makes need of this systems for all enterprise as it provides profit and increases productivity.

Data mining [1]: Data mining is abstraction of secret information from big database as it has very good potential to help enterprise for data warehouse. Enterprise makes knowledge driven decisions (KDD) with help of data mining tools which helps predicting future trends. It also helps in solving problems related to business with less time consuming.

The customer data is stored in form of transaction in database which is classified with some important factors, which interns helps in decision making. The pattern of product sale tells us about new trends and also helps in strategizing business goals and needs. It leads to give us best strategy in competition [6].

Product review and rating plays very important role, since it helps us to convince other customers to purchase that particular product. Also if the rating or review of the product is less then it helps organization to understand that the product is not so good either in quality wise or trend wise.

Apriori Algorithm: is a data mining algorithm. It is used when transactions are stored as frequent item set and are mined and used association rule on them. Those item sets which are most frequently used are determined by Apriori can be used for the determined of association rules which the highlight general trends.

K-means: is a clustering and is also called nearest centroid classifier. It is popular for cluster analysis in data mining. K-means clustering is used to create k groups from a set of objects or items, so that the members of a group are more similar. It is a well-known and best cluster analysis technique used for clustering. The best value of k usually used is between 5 to 6. Logistic Regression [4]: It is a powerful algorithm for binomial outcome (output is in form on yes / no or 1 / 0). It is used to predict whether a particular product will be sold or not in sense if future sales based on past sales.

Linear Regression [4]: It is an algorithm which gives the result in form of graph when there is a relationship between dependent and independent variables. This technique is used for prediction. It uses research and analysis to predict the future outcomes. Enterprise uses logistic regression after restructure of budget.

III. EXISTING SYSTEMS

Existing systems helps in classifying various systems which are existing. The classifications of existing systems can be as follow:

Mind tree:

- It provides complete view of customers.
- Since planning, tracking, customer interaction etc is also required it consists of a module called manager module.

Tableau:

- Excellent user interface: Since it provides best convenient, straightforward and manageable user interface due to which more customers likes to use it.
- It integrates with many big data platforms such as Hadoop.

Micro Strategy:

- It uses visual data exploration interface.
- Multiple source data is combined and used.
- Advance analytic tools are used for trend and financial analysis.

Drawbacks of Existing Systems:

- Big data is used.
- Data can be breach due to big analytic tool.
- Tools are not exact, since they use big data sets.
- Existing systems have very complex environment.

IV. OUR APPROACH

The Sales Analytic Tool takes input as transaction data of sales by using data mining techniques. Then the data is analysed. The raw data is then segmented based on product rating, review, sold etc important factors [3]. Then the profiling is done which is in terms of graph, which helps to see which products are sold more and which are sold less. With the help of which an enterprise to do planning for more sales examples if the product is not getting sales we can keep it for offers and if product is getting sold more than we can increase the rate of product. This provides high profit organizations. As shown in below figure 1 and 2:

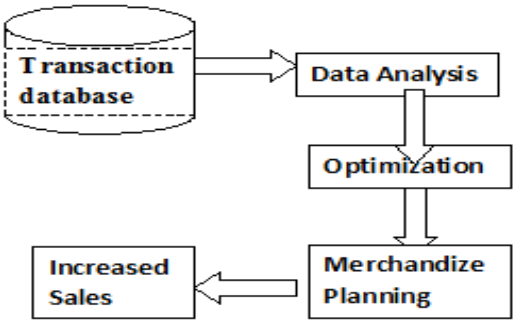


Fig.4.1. Block Diagram of Analytic Tool

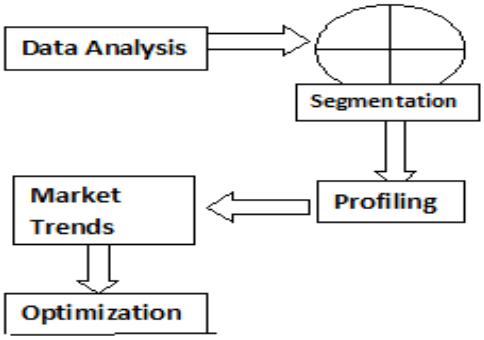


Fig.4.2. Overview of process between data analysis and optimization

Features of our system:

- Database: contains details of all transaction of customers, also contains details of products available and sold.
- Product Sold: Analysis of sold/ unsold products.
- Product Rating and Reviews: Based on customer review on particular product, enterprise can suggest other customers also to purchase those products.
- Market Trends: Those products which are trending more should be available so that enterprise can have more profit.
- Product Selling Patterns: Products are generated based on fast selling or slow selling.
- Sales Strategies: Based on profiling product's price, offer, discount etc is been done.
- Current Stock and Price: Those products which are sold more should be available to increase profit of company and availability to customers, therefore stock of product should be always there. The prices of products are revised as per profiling is done.

V. IMPLEMENTATION

Since it is all about e-commerce website, we need to work on both front-end as well as back-end.

Front End: The website with the help of which our enterprise's business will run. Our website consists of all those products which are supposed to be sold by our company. It consists of home page which consists of various categories which are available in our enterprise.

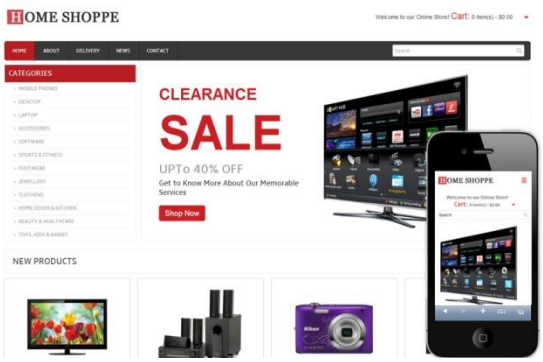


Fig.5.3. Home Page of an e-commerce website

Back End: It is where we will store our website's transaction done by customers. And those data is segmented based on fast and less selling. After which our tool performs optimization of products. After optimization our tool will work on pricing of products for increasing enterprise's profit.

Product_id	Product_name	Product_type
1234	Jeans	Cloths
5678	Heels	Shoes

Table.5.1. Transaction of products stored in database.

Product_id	Product_name	Fast_selling	Slow_selling
1234	**Jeans**	**No**	**Yes**
5678	**Heels**	**Yes**	**No**

Table.5.2. Segmentation of products.

Since we are using data mining it is necessary for us to store our data in a warehouse which is database in our case. The data or transactions are than taken by our tool in terms of input, which further proceeds for segmentation which means what kind of similar data are there are clustered as one segment, and then profiling is done where the output is shown in form of graph which interns helps entrepreneur to understand in which way the products are getting sold, which also provides different strategies to gain more profit in business. Since our tool works on k-means algorithm which is a machine learning algorithm, which means that it works on cluster formation of similar type of data and different clusters are dissimilar to each other.

Product name	**Product sold**	**Original cost**	**New cost**
Jeans	15	500/-	700/-
Heels	4	800/-	600/-

Table.5.3. Output strategy of products based on their selling

Product name	**Product sold**	**Original cost**	**New cost**
Pen	27	50/-	80/-
Cover	0	300/-	Free

Table.5.4. Output strategy of products based on their selling

VI. TESTING

This tool "Sales Analytic Tool" has been developed to find whether requirements are reached or not.

Feature of our tool-

Optimization of sales: Our tool helps them to increase their profit of enterprise with help of merchandize planning. Merchandize planning means change in cost of product based on their sale.
Analysis of Evaluation: Our tool meets user's requirements and satisfies them as it provides better strategy to increase profit of an organization.
Analysis of usability: Our tool is very easy for users to use it and understand it.
Analysis of stock: Our tool helps user to know about trends and also to update stock of those products which are fast selling. As shown below:

Fig 6.1 Product's selling rate

VII. RESULTS

Sales analytic tool provides the output for stock left in warehouse. As shown in figure below:

Transaction id	Product id	Product name	Product stock
4	1	Jeans	15
8	1	Jeans	14
9	1	Jeans	13

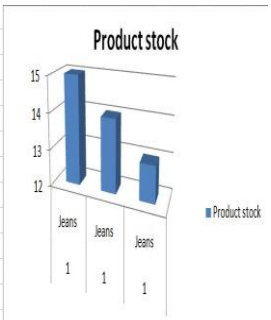


Fig. 7.1 Stock left in warehouse

Sales analytic tool also helps entrepreneur to know which products are getting sold more, which are trending products and all, as shown in figure below:

Product id	Product name	Product frequency per day
1	Jeans	3
2	T-shirt	7
3	Shirt	8
4	Kurti	6

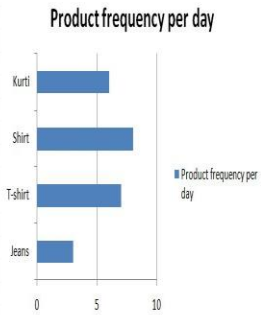


Fig. 7.2 Product frequency per day

If we compare our tool with normal e-commerce websites than our tool provides output in form for enterprise to how to increase sales, where else other e-commerce website does not provide any business strategies to increase sales. As shown in figures below:

Product id	Product name	Original cost	Changed cost
1	Jeans	1000	700
2	T-shirt	400	450
3	Shirt	550	700
4	Kurti	700	750

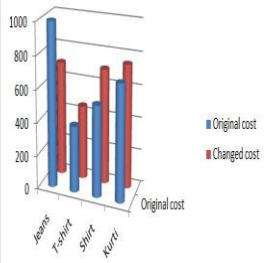


Fig. 7.3 Sales analytic tool's output for better merchandise planning.

Product id	Product name	Original cost	Changed cost
1	Jeans	1000	1000
2	T-shirt	400	400
3	Shirt	550	550
4	Kurti	700	700

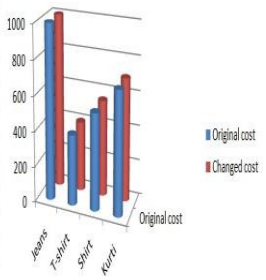


Fig 7.4 Other websites output.

VIII. CONCLUSION

Due to rapidly increasing in growth of e-commerce websites it is very important for any enterprise to use such tools which provides all the required information of resources and sales transactions. That product whose sales are more is known as frequently sold and are found using data mining algorithms. Therefore we use so many different data mining algorithms such as Market Basket Analysis, Apriori etc. By using such algorithms we can increase the product selling and profit margin of an enterprise. After which we can categorize our products as fast and slow selling products, and accordingly we can update them and apply some business strategies on products.

REFERENCES

[1] "Rating of product using data mining of sales data", International Journal for research in applied Science and Engineering and Technology [February 2016].

[2] "Mining frequent pattern of stock data", India Conference, 2011 Annual IEEE [December 2011].

[3] "A method for online retail sales estimation based on semantic features of web pages" IEEE [2014].

[4] "Analysis and Optimization of online sales of products", IEEE [2017].

[5] "Use data mining techniques to improve the effectiveness of sales and marketing" [April 2015].

[6] "Implementation of data mining techniques to perform market analysis", [November 2014].

RepoAI: A Novel Approach Towards Automated Project Reporting

Manikiran P[1], Abhay Subramanian K[2], Kshithij R. Kikkeri[3]
[1]Department of Information Technology, Alliance University, Bangalore
[2]Department of Computer Science & Engineering, Dayananda Sagar Academy of Technology and Management, Bangalore
[3]Department of Computer Science & Engineering, BNM Institute of Technology, Bangalore

***Abstract*— A report is fundamentally a document that gives out information on a given topic in a methodical fashion. Writing the report is not as overwhelming as some may initially think, more so, for school students. A student's interaction with prevailing literature often requires perusing a substantial amount of text. In this paper, we have proposed a novel method to summarize project reports automatically so that students can perform their tasks by consulting shorter summaries instead of entire literature. This model produces summaries that are statistically better than summaries produced by existing models. We found that summaries helped the students save time and provide them a guiding light, that there was no evidence that accuracy degraded when summaries were used and that most students preferred working with the proposed model to working by themselves browsing the internet.**

Keywords: Android, Data Mining, Web Scrapping, Google NLP

I. INTRODUCTION

School education poses no lack of hurdles to students. Some of the issues faced by students can appear to be so frustrating that the students might find it impossible to cope with. Most of the times, it is the case that the students are barely expected to pull out their part of the work. However, the task might look straightforward at hindsight, a multitude of students feel that it is a laborious errand to classify the topics and writing relevant content under it in a specified format as they utterly fail to throw some light on the bird's eye view of the topic and make some generic points on the given topic before dealing with the problem statement. They become unsuccessful to give an apt problem statement; they fail to furnish a relationship between the problem statement and all of the topic sentences in the research paper; they fail to reinforce the research paper with enough statistics or facts that are relevant only to the topic of the research paper under consideration. A great deal of students faces challenges with the formulation of the research paper as they do not completely scan through and understand all the data available at the student's disposal. As a consequence of this, the choice is left to the student whether to choose a generic topic or a specific topic for the amount of content required. For example, the topic of marine pollution can be a 200-page thesis because there is a cornucopia of data available on marine pollution. Notwithstanding this, the general topic of marine pollution by itself, is extremely humongous for a 10-page research paper. Drafting any report takes a lot of time, energy, and effective organization of contents in the specified format. To prevent pullulation of key wrongdoings and maintain the naturalistic exactitude, one must devote adequate time to do research in a proper manner and jot down the findings, support one's report with adequate data

and statistics, and ensure that the report is in the correct format, within the report and in the appendices. Most students, however, fall short in these metacognitive skills (Graesser & Person 1994). Afolabi (1992) identified some of the most common problems that students have when writing a literature review, including not being sufficiently critical, lacking synthesis, and not discriminating between relevant and irrelevant materials. Hence the proposed model was developed to provide students a guiding light as to how to draft reports and the model is justified by extensive experimentation.

II. BACKGROUND

A. Artificial Intelligence (AI)

According to the father of Artificial Intelligence, John McCarthy, it is "*The science and engineering of making intelligent machines, especially intelligent computer programs*" [1]. AI is a way of instructing a computer, a computer-supervised bot, or a program to think logically, in the parallelly as the intellectual human processes his thoughts.

B. Frontend and Backend

Frontend and backend are terminologies that describe the user interfaces and program services that are work together to provide the complete final application. A frontend application is a piece of software that users interactive with directly. For example, considering login page, the form which user gets to enter, edit and submit is treated as frontend. Frontend applications typically comprise of HTML, CSS and JavaScript. A backend application or program is generally a script or piece of code that runs without the user knowledge to provide support of the front-end services, typically by residing closer to the key resources such as database and can directly communicate with the required resources. Backend is a program that maybe invoked directly by the frontend or indirectly with the help of intermediate programs. Backend is built using languages such as JSP, PHP or Python.

C. Hybrid Application

Hybrid applications are generally a set of web scripts or programs running within a a browser shell of the application which has privileges to utilize the native features of the mobile, such as camera, notifications and storage. Hybrid applications have a lot of advantages compared to pure native applications such as wide range of platform support, reduced development time, and support for using third party libraries or scripts.

D. Ionic framework

According to the Ionic Official website, "*Ionic Framework is an open source UI toolkit for building performant, high-quality mobile and desktop apps using web technologies (HTML, CSS, and JavaScript)*" [2]. It is a mobile application development framework targeted at building hybrid mobile apps.

E. Database

A database is an assortment of data that is systematized so that it can be effortlessly retrieved, managed and restructured. There are multiple types of databases, such as SQL, MySQL, NoSQL, MS Access, MongoDB, etc.

F. API

An application programming interface (API) is a set of communication rules and commands that facilitate in building applications. In general, an API is used to provide unified and simple access to resources, either belonging to the same project/organization or any other third party provided resources.

G. The Cloud Natural Language API

Google Cloud's Natural Language API provides users with the structure and context of original text using a powerful pretrained machine learning models through a simplified endpoint or REST API. The documents are classified into general categories such as news, technology and entertainment. It facilitates features like entity recognition, sentiment analysis, entity sentiment analysis and other text annotations to users of the API.

III. PREVIOUS WORK

Though several intricate and absorbing models of project reporting structure presently exist, a persistent issue has been how to identify the text present in a given website in the report of a particular model in a new piece of writing. In the field of discourse analysis, researchers have generally resorted to using trained people who rate or specialist informants from the target field; this is a time-consuming process [4], [5], [6]. inexperienced readers in the classroom, on the other hand, need to gain a more immediate view of the different structural moves used in a text. Similarly, if novice writers are able to monitor the structural changes in their own compositions, they can make corrections to the flow of a text where necessary [7] There are several research papers on automatic bug reporting and unsupervised bug reports categorization. Also, there are many research papers on automatic report generation for smart office/inventory management. Boris et al have come up with a model which automatically generates project reports which can be edited and stored in a SQL database. They are useful for project managers.[8] Zalte et al have proposed an automatic question paper generator which is streamlined and secure and is mostly inclined toward academia and hence found its mention here.[9] Most of the existing systems focus on the industries and none on the academia. Hence, we have developed a system that concentrates on the students need for report generation which will save students' time and energy to a large extent.

IV.IMPLEMENTATION

A. *Backend*

The development and deployment of backend part is done in a number of steps which is described in great detail

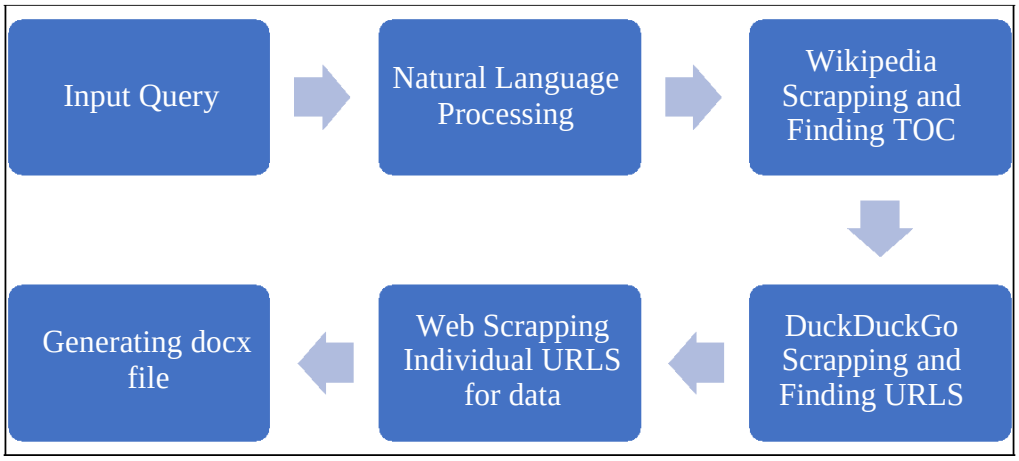
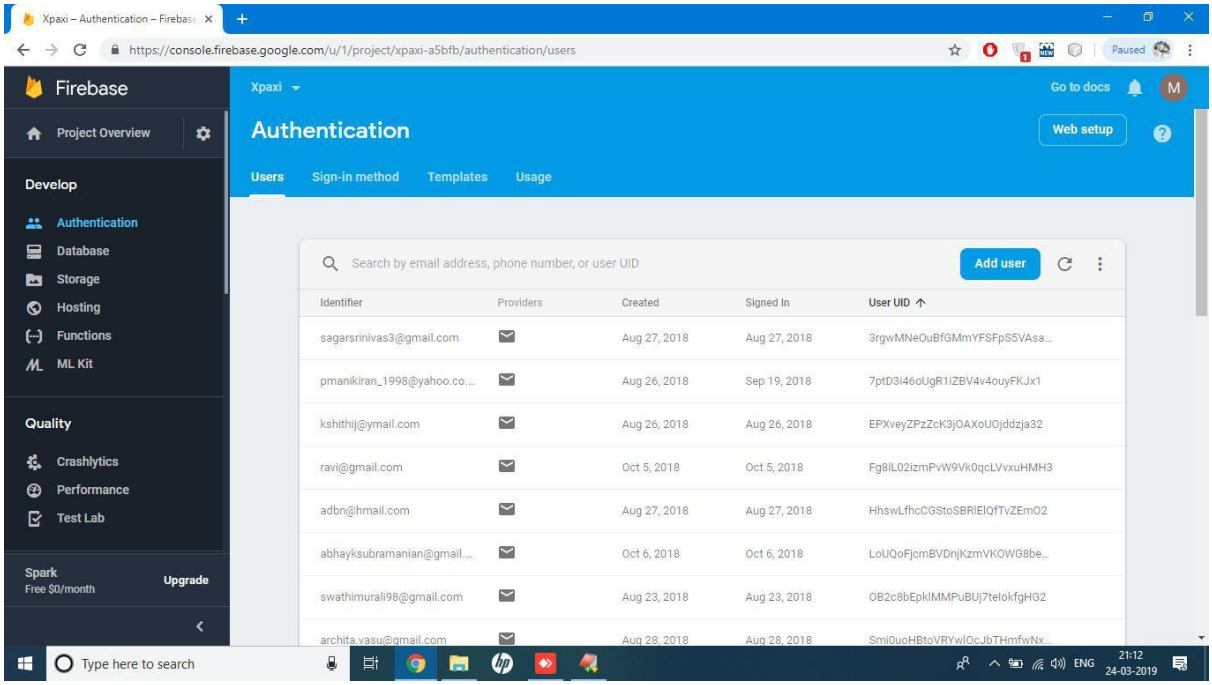


Fig 1. Backend Workflow

1. Login and Signup

The data entered by the user is filtered to check for SQL Injections and then passed on to the FireBase API. The response from the FireBase API is sent back to the user through our API.



2. Google Scrapping

The backend development starts with the usage of Python. Some background checks are done at this stage. If an article on the input query is present on Wikipedia, then it jumps to step 3. Else, a predefined set of topic headings are employed namely [“Introduction”, “Cause”, “Effect”, “"Precautions”, “Applications”, “Conclusion”]. Subsequent Google searches will be performed on the predefined headings. Taking an example of ‘Marine Pollution’ as the input string, Google searches on “Marine pollution introduction”, “Marine pollution causes” etc., will be performed.

3. NLP & Wikipedia Scrapping

Google Cloud Natural Language Processing from the “google.cloud” library is used to extract the main topic query from a set of long input query. Instead of having predefined topics, a Wikipedia search is performed to check if there is any related topic, and scrapped the TOC (Table of Contents) to get the topics for the particular query. Then the search is run similar to step 2.

The input query which is received by the API backend script is cleaned by using NLP for filtering out unwanted words and leaving behind only essential keywords. These keywords are used to find related Wikipedia pages, if any, and the Table of Contents of that particular Wikipedia page is scrapped and a list of necessary headings again after cleaning and filtering is returned back to the user.

4. DuckDuckGo Scrapping

Initially, Google Custom Search API was employed for fetching query result links, but due to the limitation of 100 query searches per day, we moved on to DuckDuckGo scraping . The searching efficiency was improved by using DuckDuckGo instead of Google, as Google blocks bots which try to search a lot of queries within a short period of time. The URLs from the scraped results were refined by filtering out the ads, and giving more priority to popular websites likes Wikipedia and other similar websites.

5. Individual URL Scrapping

Since each website has a different html DOM structure, it is really hard to scrape using same technique. Alternately, the route chosen was to grab the text from the entire webpage. Now a linear search algorithm is run to find the match of text from the DuckDuckGo result summary. From the initial position the matching text found to the required length of string/paragraphs, that is specified by the user in the app, is scrapped.

6. Using Flask & Generating .docx file

Once this functionality is complete, “Flask” library of Python is incorporated to implement a real-time backend server/API endpoint. Finally, the data is collected and organized which was scrapped from the WWW, and is exported into docx file format, using the Python library

"docx". This file is saved in the server, and also the path to this file is saved in database, for users to see their past generated documents.

B. *Frontend*

1. First Stage – Ionic

A hybrid app was built using Ionic framework. Ionic framework is built over Cordova to generate apps for both Android and iOS. The user gets a login and signup options, for storing previous search results for future purposes. The user has in his prerogative to do a part of the report on his phone and the other part on his desktop. This exhibits a macaronic charm and has an edge over other similar applications.

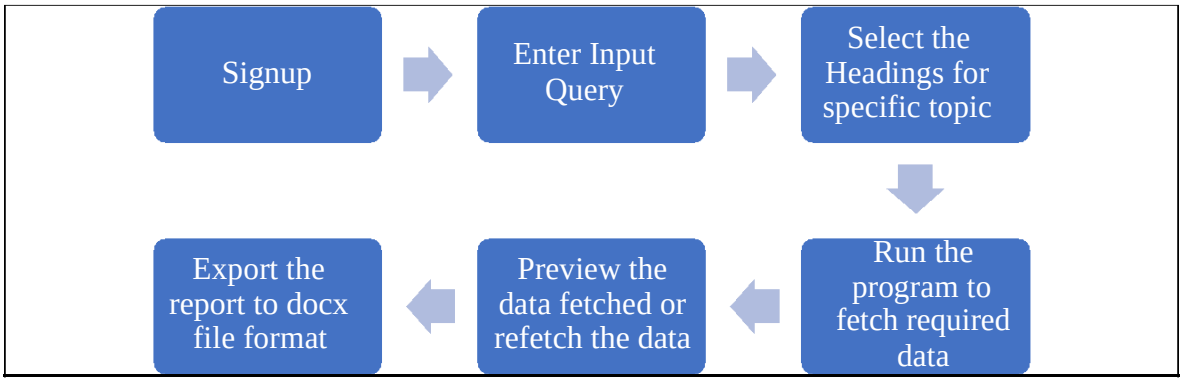


Fig 2. Frontend workflow

2. Input Query

Prima facie, the user enters any search query in the given text field. A new FormData object is created including the search query which is sent as a HTTP POST request to the API which is located at http://api.repoai.ml. The response thus obtained is a list of heading which is further filtered and processed using our cleaning techniques including NLP.

3. Selecting the Headings for specific topic

The user is displayed with all the topics fetched from the API request in a list format with option to select, add, edit and remove them. The user can add new heading(s) of his/her choice if they find it missing from the Wikipedia TOC and our predefined headings set. They can also edit the headings, if they find any mistakes or irrelevancy.

4. Generating Data and Reviewing it

Once a user selects any particular heading, the data gets fetched automatically and is shown to the user to check for mistakes or do any modifications. If he/she feels that the content is irrelevant, they can click regenerate option to fetch data from different sources. The same procedure is followed with rest of the headings. The user is also provided with the option to reorder the headings to match his/her requirements.

5. Exporting Report

The user can click on Generate Report button to select the file format he/she desires to export it to, either PDF or Docx file format. Once they click on a particular file format, the request is sent to the API along with the entire data, and a path to the generated file is returned back. The app uses this path to download using the default download manager.

6. Solving local storage issue

A plugin called “cordova-sqlite-storage” was used which helps in maintaining an SQLite database in a persistent storage location (phone storage).

7. Saving Recent Searches

An algorithm was written to save N recent searches on local storage, similar to that of LRU (Least Recently Used) cache which is a cache that, when low on memory, evicts least recently used items. LRU is an eviction policy that makes a lot of sense for the typical kind of cache we all deal with on a daily basis. [10]

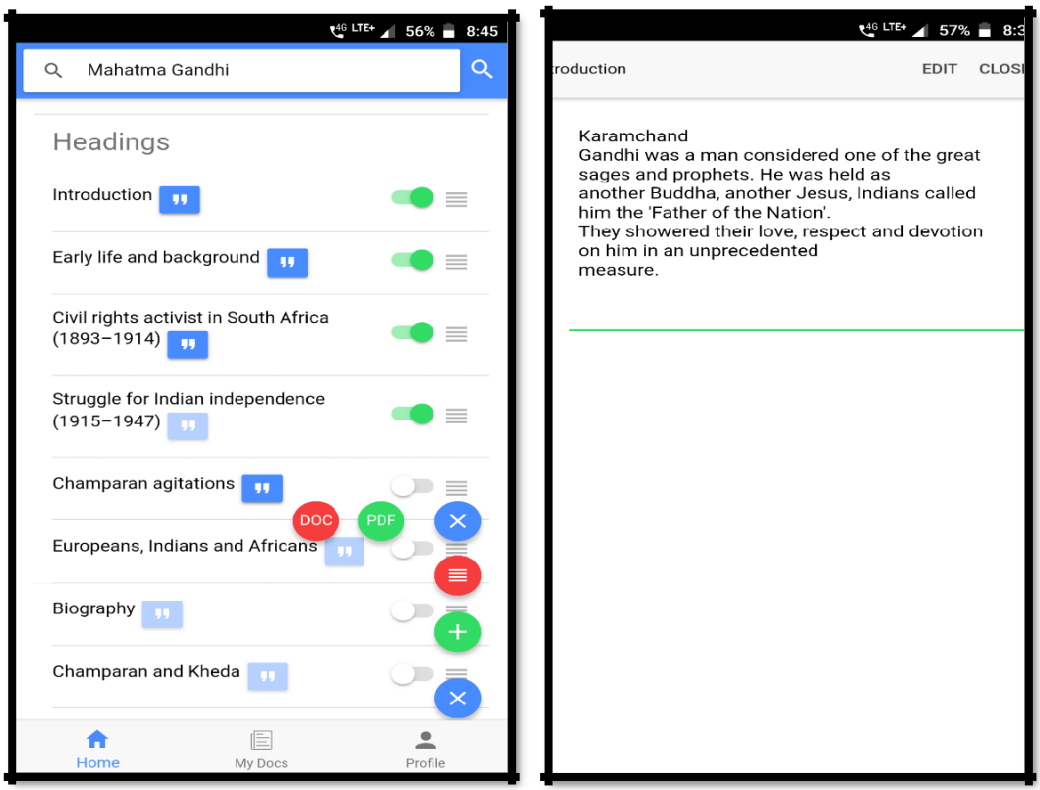


Fig 3. Screenshots

V. CONCLUSION AND FUTURE ENHANCEMENTS

This project has an immense scope in the coming days as there are a lot of functionalities and modules which can be incorporated into it, but the scope of the project is limited due to various constraints. Some of the areas in which there can be enhancements are listed below

1. Addition of word limit to the report so that the student can comply with the word limit given to him by the teacher.
2. Addition of images in the report so that it becomes more presentable and get a clearer picture of what the report is conveying.
3. Implementation of figures and statistics which help increase the readability and gives an idea about the past, present and future scenarios of the topic under consideration.
4. Support of multiple formats prescribed by different institutions/organizations like IEEE, IJERT etc, also formats prescribed by different universities for project reports.
5. Google is working hard to provide cloud services that facilitate human-computer interaction through tools that are able to consume human language like GCP (Google Cloud Platform), hence it is provides a path for wider enhancement of user experience thereby increasing the student's usage of the app.

ACKNOWLEDGMENT

Many professors have mentored and guided us during the course of developing our application. Their input was extremely useful in the subsequent development of the application and without their help it would have been an uphill task. Several people have given us constructive feedback on different modules of the application. Their help is hereby sincerely acknowledged. We now understand why family members are unfailingly mentioned in this segment. Without our families' love and support, this project wouldn't have been successful. Many thanks to them for being our pillars of support.

REFERENCES

[1] https://www.tutorialspoint.com/artificial_intelligence/artificial_intelligence_overview.htm

[2] https://ionicframework.com/docs/intro/concepts

[3] J. M. Swales, Genre Analysis: English In Academic and Research Settings. Cambridge, U.K.: Cambridge Univ. Press, 1990.

[4] L. Anthony, "Characteristic features of research article titles in computer science," IEEE Trans. Prof. Commun., vol. 44, pp. 187–194, Sept. 2001.

[5] L. Anthony, "Writing research article introductions in software engineering: How accurate is a standard model?," IEEE Trans. Prof. Commun., vol. 42, pp. 38–46, Mar. 1999

[6] Laurence Anthony AND George V. Lashkia, "Mover: A Machine Learning Tool to Assist in the Reading and Writing of Technical Papers" IEEE TRANSACTIONS ON PROFESSIONAL COMMUNICATION, VOL. 46, NO. 3, SEPTEMBER 2003

[7] Boris Veytsman, Maria Shmilevich: Automatic report generation with Web, TEX and SQL. TUGboat, Volume 28 (2007), No. 1 — Proceedings of the Practical TEX 2006 Conference

[8] Zalte S.V., Jadhav C.C, Mangire A.A, Hole A.D.4 and Tulshi A.R: Automatic Question Paper Generator System, International Journal of Advanced Research in Computer and Communication Engineering Vol. 7, Issue 3, March 2018

[9] https://www.openmymind.net/Writing-An-LRU-Cache/

AUTHORS

Manikiran P is currently a student at the department of Information Technology, Alliance University, Bangalore, India. His research interests include Natural Language Processing and Machine Learning. He can be contacted at pmanikiran1998@gmail.com

Abhay Subramanian K is currently a student at the department of Computer Science & Engineering, Dayananda Sagar Academy of Technology and Management, Bangalore, India. His research interests include Game Theory and Machine Learning. He can be contacted at abhaysubramaniank@gmail.com

Kshithij R. Kikkeri is currently a student at the department of Computer Science & Engineering, BNM Institute of Technology, Bangalore, India. His research interests include Natural Language Processing and Machine Learning. He can be contacted at kshithij@ymail.com

Ensemble Neural Network Classifier Design using Differential Evolution

Shobha T[*], R J Anandhi[**]

[*] Department of Computer Science and Engineering, The Oxford College of Engineering
[**] Department of Information Science and Engineering, New Horizon College of Engineering

Abstract- In this paper, ensemble neural network has been designed for the classification purpose using the Differential Evolution (DEEANN). The ensemble structure has been designed into two different stages of classifier, in the first stage four multilayer perceptron classifier have been applied which has same size of training data but some new data also has been included to embed some new information in the process of learning. All the four classifiers have been trained through the gradient descent algorithm. In the second stage, there is decision neural network, which considers the output of all the four first stage classifiers as inputs and develop the final decision with the help of differential evolution. Two other possibilities of integrating the classifier outputs is based on majority voting method and mean decision value method have also been considered to compare the performances. The proposed method has shown not only the high efficiency but also resistivity with trial variation.

Index Terms- Classification, Differential evolution, Ensemble architecture, Feed-forward neural network, Neural network

I. INTRODUCTION

Machine learning refers to a system that has the capability to automatically learn knowledge from experience and other ways. In machine learning classification is one of the prime important mechanism available, where Classifier predicts categorical labels that exist with applied input data to take the necessary action or place the data in the different, already existed groups. In many situations, where single classifier is used, face challenge of not having generalized knowledge of classification, ensemble approach in which many neural networks are integrated in proper manner can increase the reliability of classification decision. The main causes of problem with single classifier may be limited size of data, inappropriate learning, and noisy data etc. The major challenges with ensemble neural network are structure formation and determination of role of each classifier involved. These two properties have significant effect over ensemble network generalization capability. The most powerful way to design the classifier is to use artificial neural network, which has better ability of learning and handling the nonlinearity in optimum manner.

Computing system which is inspired by biological nervous systems are defined as Artificial Neural Network (ANN). ANN consists of large number of processing elements called neurons. Neurons of network work parallel to learn from the input data, to organize internal processing, and to obtain optimized output. It is the most widely used algorithm for solving real-world computational applications. The challenges faced by ANN are choosing appropriate initial value of connection weights, selecting number of hidden nodes, convergence of learning algorithm and training error. To overcome these problems, evolutionary algorithms (EA) have been used in recent days to evolve ANN for obtaining better and efficient network performance. EA can help in choosing best connection weights, reduced number of hidden nodes and better convergence.

In this paper feed forward architecture has been considered to design the classifiers and its ensemble has been created with differential evolution. Each classifier has obtained the learning through the gradient descent method. Diversity within individual classifier has been created by providing some variation over training data set. A next stage classifier which is again a feed forward architecture, weights has been evolved through the stochastic search method. The proposed learning through differential evolution has been applied over benchmark XOR classification problem to analyze the training capability and then later ensemble network has been developed to define the classification over other data sets.

II. RELATED WORK

Based on Genetic algorithm a ensemble approach has been presented in [1]. An ensemble neural network based on the Akaike information criterion (AIC), has been proposed in [2]. Proposed algorithm first searches for best configuration weights of each component and then it is used as an automating tool to find the best configuration weights of ensemble neural network. An ENN in [3], uses entropy theory to combine the components of network. Using ENN, the best structure of each component network is searched first, and then employed to determine the best combining weights of ensemble NN. To forecast complex time series, authors of [4] have designed ensemble neural network using hybrid method based on particle swarm optimization with fuzzy aggregation of responses. Based on feature selection, a multi-sided multi-granular neural network ensemble optimization method has been presented in [5], using different attribute granularity and the corresponding subsets, proposed algorithm divides attribute granularity of dataset from multi-side, and structures multi-granular individual neural networks. Zhiye et al., have proposed ensemble neural network framework for dependency parsing of natural sentences. Ensemble models using a convolutional neural network has been discussed in [7]. An ensemble of models, each of which is optimized for a limited variety of poses, is capable of modeling a large variety of human body configurations. An ensemble of over complete patch-based neural networks has been presented in [8] which segments accurate quantification of white matter hyper intensities (WMH) from Magnetic Resonance Imaging. Due to over complete nature of proposed algorithm, accurate and regular segmentations have been obtained, and by using a boosted ensemble of neural networks the segmentation error has been minimized. An ensemble DNN (EDNN) algorithm using Deep neural network and its applicability to metabolomics studies has been presented in [9]. To predict PPIs based on different representations of amino acid sequences, an EnsDNN (Ensemble Deep Neural Networks) algorithm has been proposed in [10]. In article [11], to identify and classify four types of blur images, an ensemble convolution neural network (CNN) has been designed. To enhance model discriminability without incurring additional computing burden, a two-stage pipeline, comprised of deep compression and ensemble technique has been used. In [12], Computational Modification Sites with Ensemble Neural Network (CMSENN), has been proposed to detect protein modification.

III. PROPOSED WORK

The proposed work has two stages: (i) In the first stage, number of individual feed-forward architecture has been considered for classifier. Four same size architecture have been considered which share some overlapping in the training data of one classifier with other. Gradient descent has been applied to train the individual classifier as shown in the Fig.1. (ii) In the second stage a feed forward architecture (called integrated classifier) shown in Fig.2 takes the output of each 1st stage classifier as the input and get the training over complete data set. In this stage, integrated classifier has learnt how to assign the weightage to output delivered by each classifier to decide on final decision. The learning is given through the differential evolution algorithm which has high level of exploration capability. Fig.3. depicts architecture of evolving learning with Differential Evolution. The weight of neural network is first transformed into one dimensional representation and the pre-processing has been applied as linear normalization as defined by Eq. (1), where x_{mn} and x_{mx} are the minimum and maximum value of the particular attribute. This pre-processed input evolve by differential evolution as given in next section.

$$x_n = \frac{x - x_{mn}}{x_{mx} - x_{mn}} \qquad \text{Eq. (1)}$$

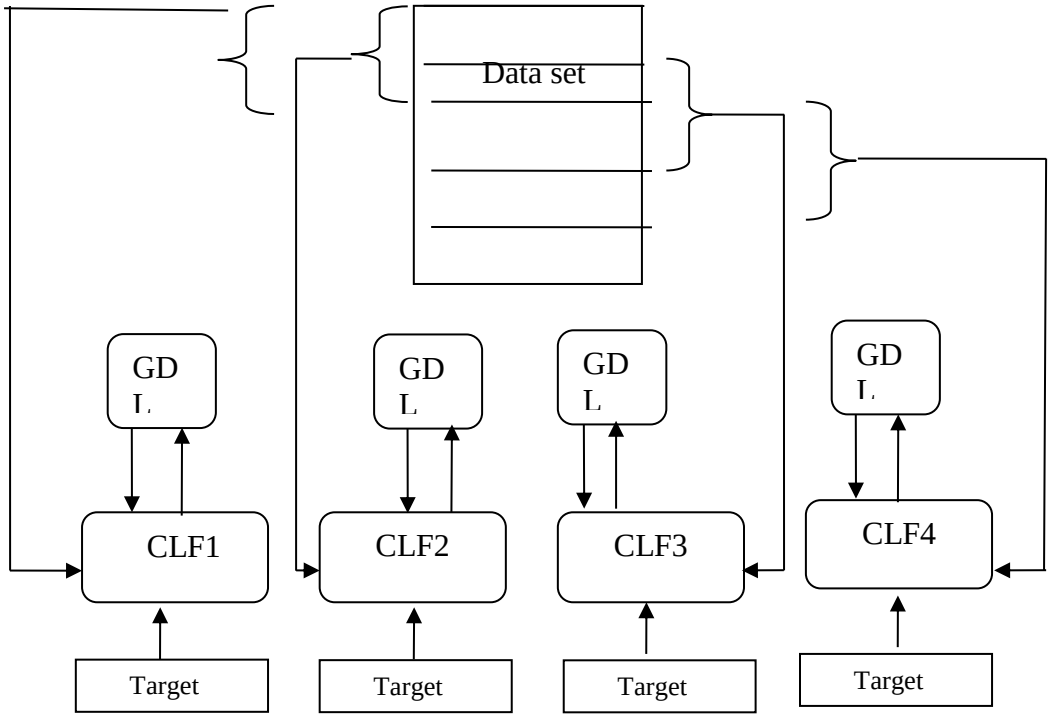


Fig.1: First stage training phase

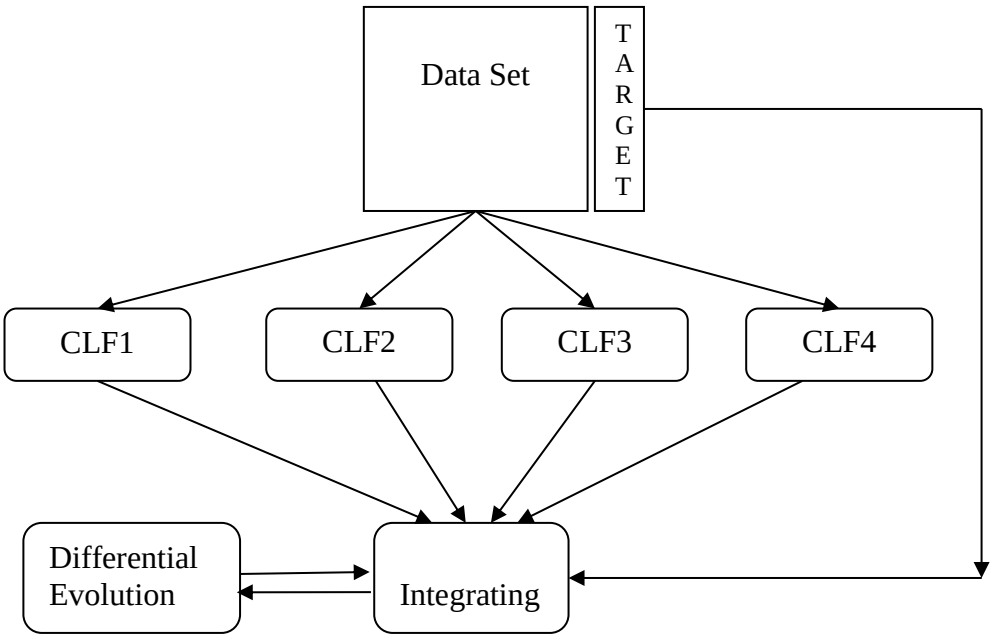


Fig.2: Functional module for Second stage development

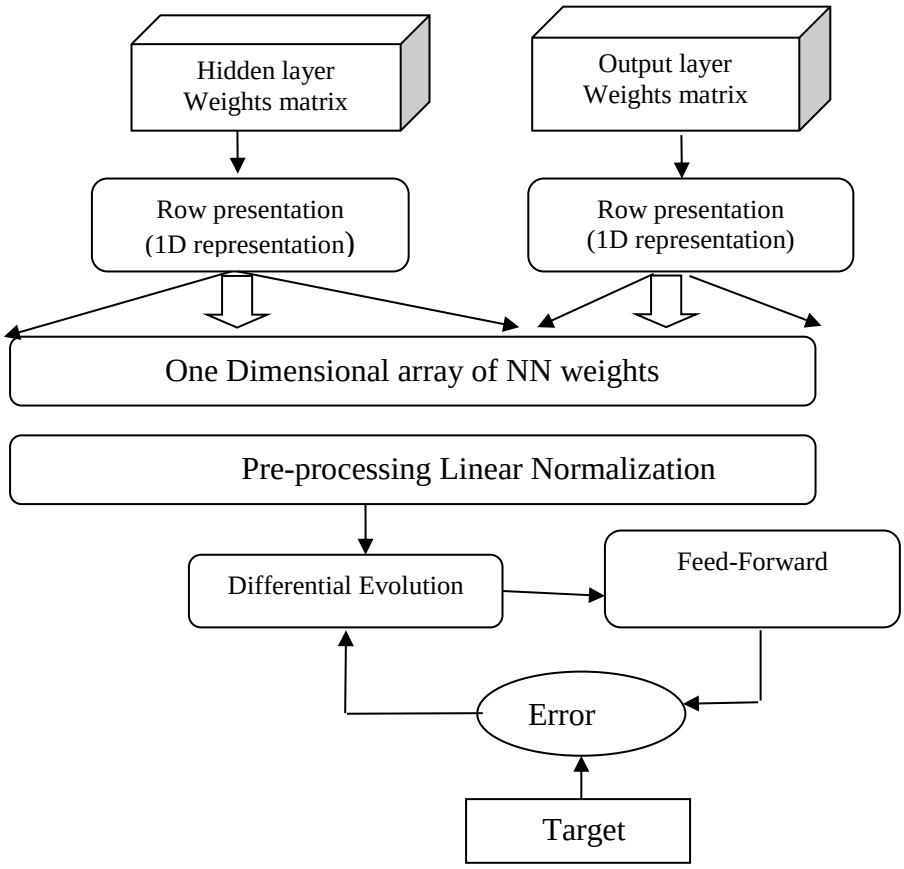


Fig.3. Evolving Learning with Differential Evolution

A. Differential Evolution

Differential evolution (DE) is one of the most powerful stochastic real-parameter optimization algorithms used to solve global optimization problems. It has 3 basic operations: mutation, crossover, and selection. During each generation, mutated vector is generated using linear combination of a base vector and one or more differential vector as defined in Eq. (2). Trial vectors have been developed by applying crossover operator under probabilistic environment as shown in Eq. (3). CR is a controlling parameter within the range [0,1]. jrand is a randomly selected index to make sure that at least 1-D of the mutated vector will enter into the newly generated individual. Then the selection process selects between target and trail vectors to choose vectors for the next generation according to Eq. (4).

$$V_i^{(G)} = X_{r1}^{(G)} + F * \left(X_{r2}^{(G)} - X_{r3}^{(G)}\right) \quad Eq.(2)$$

$$u_{\mathrm{ij}}^{(G)} = \begin{cases} v_{ij}^{(G)} & if\ rand(0,1) \le CR\ or\ j = j_{rand} \\ x_{ij}^{(G)} & otherwise \end{cases} \quad Eq.(3)$$

$$x_{\mathrm{ij}}^{(G)} = \begin{cases} u_i^{(G)} & if\ f\left(u_i^{(G)}\right) \le f\left(x_i^{(G)}\right) \\ x_i^{(G)} & otherwise \end{cases} \quad Eq.(4)$$

B. Learning algorithm with Gradient Decent

Weight update equations for Back Propagation, the rate of the updates are proportional to the derivative of the nonlinear activation functions. A typical activation function for neurons in multilayer perceptron neural network is of sigmoid type with bell-shaped derivatives. The change in hidden and output layer weights have been shown below. To make the learning faster, momentum has also been included as shown in Eq.7.

1. Initialize the weights in the network according to Gaussian distribution random number initialization process.
2. From data set, the set of training data, derive the network response.
3. Compare the preferred network responses with the definite output of the network and the local error is calculated according to

   For output layer: $\delta_i^s = \left(d_q - x_{out,i}^s\right) g(u_i^s)$ Eq. (5)

   For hidden layer: $\delta_i^s = \sum_{h=1}^{n2} \delta_h^{s+1} w_{hi}^{s+1}\, g(u_i^s)$ Eq. (6)

4. The weights of the network can be updated as

   $$w_{ij}^s(t+1) = w_{ij}^s(t) + \mu \delta_i^s x_{out,j}^s + \alpha\left[w_{ij}^s(t) - w_{ij}^s(t-1)\right] \quad \text{Eq. (7)}$$

5. Stop the iteration if network converged, else go back to step 2.

IV. EXPERIMENTAL RESULTS & ANALYSIS

Proposed algorithms have been implemented using MATLAB. To understand the proposed form of learning of neural network through differential evolution, a benchmark problem of classification XOR has been considered. A proper classification of XOR problem will ensure the classification capability of developed algorithm. A feed forward architecture having size [2 3 1] has been selected and weights have been evolved for number of iterations with population size 100 until mean square error is greater than 0.007. The obtained result for hidden layer weights and output layer weights have been shown in Table1 and in Table2 respectively. The learning error with iteration for best solution and whole population mean error have been shown in Fig.4. It can be observed that there is nearly close chase by population mean error which indicate nearly complete population is having similar type of convergence. The final obtained output by neural network has been shown in Table3 and it is observed that output is very close to the target.

Table1: Hidden layer weights obtained after learning completion

Input node	1st Hidden node	2nd Hidden node	3rd Hidden node
1	-4.3460	6.1492	12.0641
2	15.4912	9.5074	-5.7446

Table2: Output layer weights obtained after learning completion

Hidden node	Out put node
1	-17.5399
2	22.6341
3	-16.2271

Table3: Final obtained output by DE-ANN

Input Data		Target	DE-ANN O/P
0	0	0	0.0038
0	1	1	0.9936
1	0	1	0.9978
1	1	0	0.0000

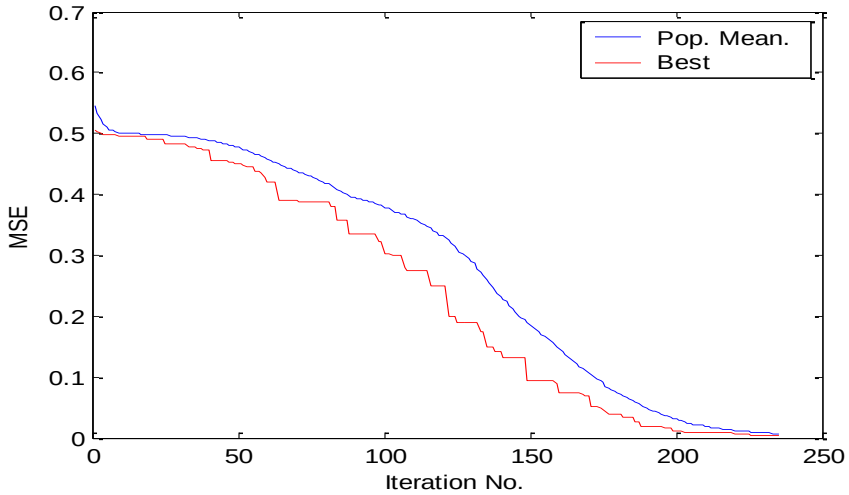


Fig.4: Learning error convergence characteristics

To get the practical benefit of proposed ensemble classifier," Heart-Disease" data set of UCI repository has been considered. It has total 270 data set and each data having 13 attributes and have two classes. In development of training data set for each classifier, 100 data set have been considered while remaining data set has been considered for the testing. (1st CLF: 1-100; 2nd CLF: 51-150; 3rd CLF: 101-200; 4th CLF: 151-250 data samples have been considered for training). The learning rate and momentum constant 0.2 and 0.1 respectively has been considered. The learning of each and every classifier has been allowed to update up to 2000 iterations. The obtained learning error convergence for all the 4 classifiers have been shown in Fig.5. It can be observed that for all the cases, proper convergence has been obtained. The classification efficiency for training and test data along with mean square error has been shown in Table4. It can be observed that, very fine learning happens for each classifier and performance is around [98.6%, 98.4%, 97.2% and 99.2%] on an average for training data. But performance is very poor over test data which is [76%, 77%, 83% and 79%]. Such kind of performance over test data may not be acceptable when considering the critical applications.

Table4. Mean square error/Training /test data performance for each classifier

Trail No.	CLF1 MSE / Tr. / Test	CLF2 MSE / Tr. / Test	CLF3 MSE / Tr. / Test	CLF4 MSE /Tr. / Test
1	0.108 / 99 / 79	0.197 / 98 / 76	0.324 / 97 / 81	0.0105 /99 /80
2	0.0109 / 99 / 78	0.0103 / 99 / 76	0.0156 / 98 / 76	0.0104 / 99 / 79
3	0.0123 /99 / 79	0.0203 / 98 / 77	0.0311 / 97 / 84	0.0014 / 100/ 78
4	0.0308 / 97 / 73	0.0203 / 98 / 77	0.0313 / 97 / 82	0.0102 / 99 / 79
5	0.0119 / 99 / 76	0.0103 / 99 / 77	0.0311 / 97 / 83	0.0104 / 99 / 79
Mean (Tr.	98.6000 /77.0000	98.4000 /76.6000	97.2000/81.2000	99.2000/79.0000

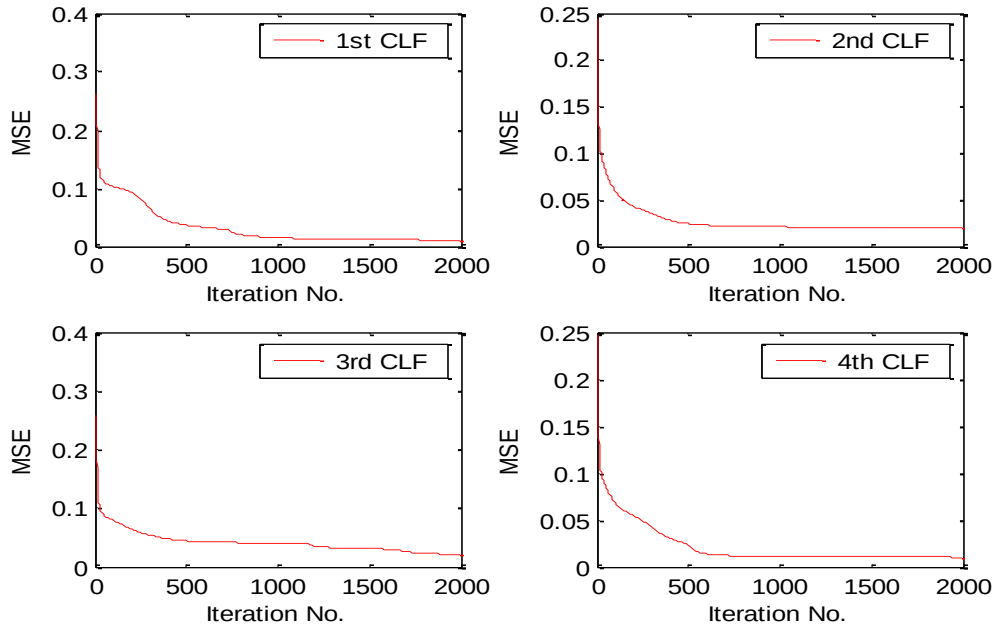
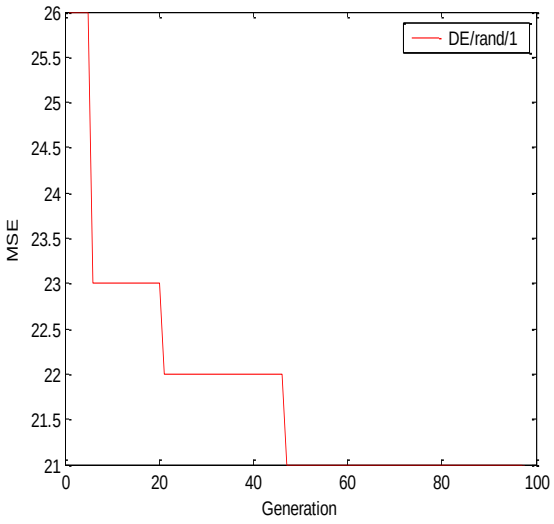


Fig.5.Individual classifier error convergence

Table5: Performance delivered by different types of ensemble network

Trial No.	Majority Decision Voting Efficiency %	Mean Decision Value Efficiency %	DEEANN				
			CLF1 WT	CLF2 WT	CLF3 WT	CLF4 WT	Efficiency %
1	88.5185	88.8889	0.3533	0.1890	0.1265	0.2744	92
2	87.4074	90	0.2522	0.3336	0.1976	0.3378	92
3	90	90.7407	0.3946	0.1857	0.2795	0.1000	93
4	88.8889	89.6296	0.1405	0.4412	0.2579	0.2868	92
5	88.6845	89.7253	0.2437	0.3349	0.1729	0.3448	92

Apart from differential evolution-based ensemble, two other different type of ensemble has been developed. The obtained result after 5 independent trials have been shown in Table5. In first case, final decision of each classifier has been considered depending upon the majority voting strategy. In second approach, rather than counting the final decision, numeric decision given for most favorable class has been considered and mean is estimated for overall decision values. In this case, there is advantage that, none of the classifier outcome is ignored completely. With this, process opinion of every expert is being utilized. These two processes are easy to implement, and it can be observed that there is betterment achieved by these two-ensemble compared to individual. But, it is also observed that ensemble network that used the mean value had better performance over majority voting method. When the differential evolution-based ensemble has been applied, the obtained weights for each classifier under five trials have been shown in Table5 along with achieved efficiency.



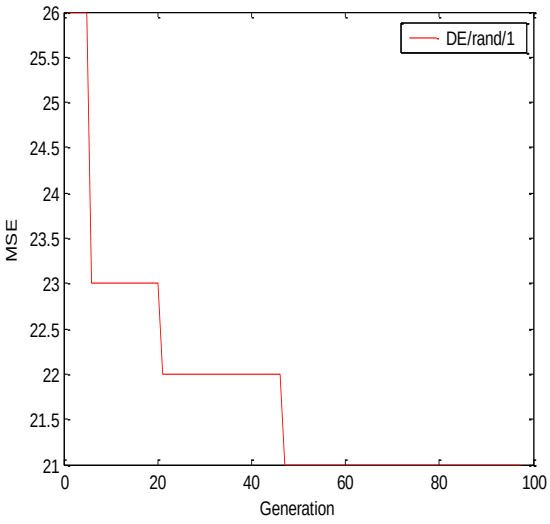


Fig.6. Weight evolution for ensemble by DE **Fig.7. Learning error minimization by DE for ensemble**

Table6: Comparative performance between all classifiers

CLF1	CLF2	CLF3	CLF4	MJVT	MNDS	DEEANN
77.0000	76.6000	81.2000	79.0000	88.6999	89.7969	92.2000

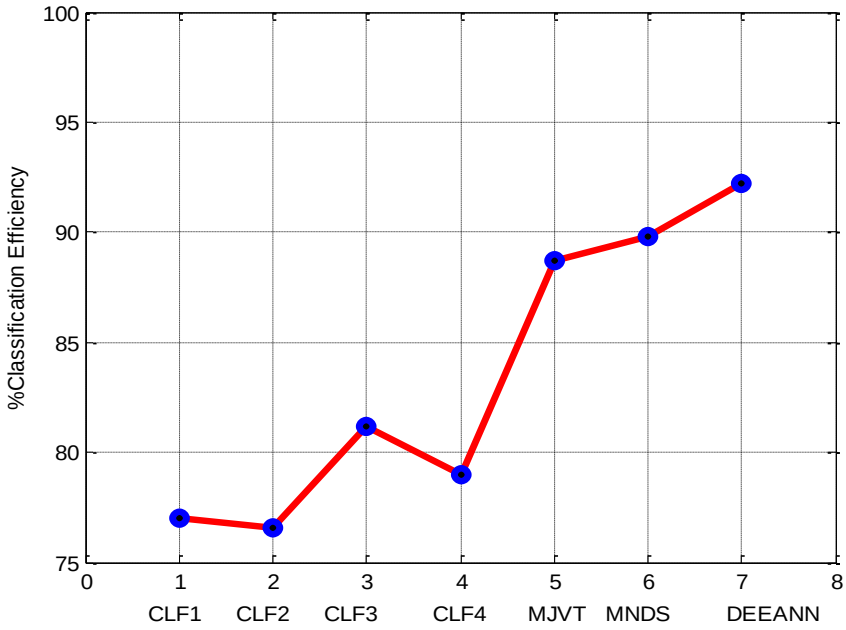


Fig.8.Comparative performance for different classifiers

It is very interesting to note that a very high efficiency around 92% has been achieved along with consistent performance. The obtained variation in weights for all the individual classifier of first stage has been shown in Fig.6. While the learning error minimization convergence by DE for ensemble has been shown in Fig.7.The final overall comparative performance of each classifiers of first stage of ensemble neural network (CLF1-CLF4), overall performance of ensemble neural network with majority voting (MJVT), ensemble neural network with mean decision (MNDS) and differential evolution based ensemble neural network(DEEANN) has been shown in Table 6. The comparative graphical plot for classification efficiency has been shown in Fig.8, which shows the benefit of ensemble as well as benefit of different structures of ensemble neural networks.

Table7: Sensitivity and Specificity performances by different classifier structures

Classifiers	Sensitivity (%)	Specitivity (%)
CLF1	85.83	83.33
CLF2	84.17	86.00
CLF3	82.50	92.67
CLF4	82.50	90.00
MJVT	78.33	97.33
MNDS	84.17	94.00
DEEANN	90.00	93.33

Ideally, classifier having 100% sensitivity and 100% specificity is absolutely best. To understand the predictability of individual classifier, in terms of sensitivity and specificity, performances obtained by a different form of classifiers have been evaluated and shown in Table7. It is clear that all the individual neural network classifier (CLF1-CLF4) have more or less nearly the same performances. A betterment in specificity and a remarkable drop in sensitivity has been observed in ensemble neural network with majority voting (MJVT). Performance has been improved further by ensemble neural network with mean decision (MNDS) when compared to individual classifiers and MJVT. But DEEANN has outperformed, when compared to the individual neural network classifiers (CF1-CF4), MJVT, and MNDS.

V. CONCLUSION

In this proposed work, the importance of ensemble neural network in the area of classification has been shown. The observations are very clear and appealing that instead of any single classifier, ensemble classifier is always beneficial. The structural form of ensemble has achieved with evolutionary process, which has delivered the outstanding efficiency as well as high level of consistency, which is very important in the critical applications. The proposed method is computation efficient and has shown quality improvement, compared to either individual classifier as well as conventional form of ensemble methods. Computational efficiency of DEEANN has been obtained at the cost of little extra computational cost. Here, proposed work has been evaluated only on Heart-Disease dataset. Further, the application-specific dataset can be used for evaluation of performance.

ACKNOWLEDGMENT

I would like to express my deep gratitude to Dr. R J Anandhi, Professor and Head, Department of Information Science and Engineering, New Horizon College of Engineering, Bengaluru, my research supervisor, for her patient guidance, enthusiastic encouragement and useful critiques of this research work. I would also like to thank Dr. R V Praveena Gowda, Principal, The Oxford College of Engineering,

Bengaluru, for his advice and assistance in keeping my progress on schedule. Finally, I wish to thank my husband, parents and family members for their support and encouragement throughout my study.

REFERENCES

[1] Zhi-HuaZhou, JianxinWu, ”Ensembling neural networks: Many could be better than all”, Artificial Intelligence, Volume 137, Issues 1–2, May 2002, Pages 239-263

[2] ZhiyeZhao Yun Zhang HongjianLiao,” Design of ensemble neural network using the Akaike information criterion”, Engineering Applications of Artificial Intelligence, Volume 21, Issue 8, December 2008, Pages 1182-1188

[3] ZhiyeZhao YunZhang,” Design of ensemble neural network using entropy theory”, Advances in Engineering Software, Volume 42, Issue 10, October 2011, Pages 838-845.

[4] Martha Pulido, Patricia Melin, Oscar Castillo, ”Particle swarm optimization of ensemble neural networks with fuzzy aggregation for time series prediction of the Mexican Stock Exchange”, Information Sciences, Volume 280, 1 October 2014, Pages 188-204

[5] HuiLi Xuesong Wang ShifeiDing,”Research of multi-sided multi-granular neural network ensemble optimization method”, Neurocomputing, Volume 197, 12 July 2016, Pages 78-85

[6] KangilKim, YunJin,”Center-shared sliding ensemble of neural networks for syntax analysis of natural language”, Expert Systems with Applications, Volume 83, 15 October 2017, Pages 215-225

[7] YukiKawana, NorimichiUkita, Jia-BinHuang,”Ensemble convolutional neural networks for pose estimation”, Computer Vision and Image Understanding, Volume 169, April’18, Pages 62-74

[8] Jose V.Manjón,”MRI white matter lesion segmentation using an ensemble of neural networks and overcomplete patch-based voting”, Computerized Medical Imaging and Graphics, Volume 69, November 2018, Pages 43-51

[9] TaigaAsakura, YasuhiroDate,” Application of ensemble deep neural network to metabolomics studies”, Analytica Chimica Acta, Volume 1037, 11 December 2018, Pages 230-236

[10] LongZhang GuoxianYu,” Protein–protein interactions prediction based on ensemble deep neural networks”, Neurocomputing, Volume 324, 9 January 2019, Pages 10-19

[11] WangRui , “Blur image identification with ensemble convolution neural networks”, Signal Processing, Volume 155, February 2019, Pages 73-82

[12] WenzhengBao, BinYang,” CMSENN: Computational Modification Sites with Ensemble Neural Network”, Chemometrics and Intelligent Laboratory Systems, Volume 185, 15 February 2019, Pages 65-72

IMPACT ON GROUNDWATER AND SOIL DUE TO SOLID WASTE DUMP

A CASE STUDY OF S. BINGIPUR IN BANGALORE

Dr.Basavaraj Paruti*, Dr. Santhaveeranagoud. B**

* Department of Civil Engineering
Alliance College of Engineering and Design,
Alliance University Bangalore, India

** Department of Civil Engineering
UVCE, Bangalore University
Bangalore, India

Abstract- Municipal Solid Waste Management has become one of the major problems in urban and semi-urban areas. Improper MSW disposal and management causes all types of pollution: air, soil, and water. Indiscriminate dumping of wastes contaminates surface and ground water supplies. Health and safety issues also arise from open dumping. The report starts with various approaches to manage municipal solid waste and a plan to implement an integrated solid waste management for a city. Solid wastes have potential for causing serious adverse impact on the environment. Ground water & Surface water Contamination, Land Pollution, and Air Quality Deterioration. Leachate is a toxic liquid that seeps through solid waste in a land fill. This process extracts soluble dissolved and suspended materials from the waste. It contains bacteria, toxic substances, heavy metals, etc . The impact assessment of the open dumping was assessed by collecting and analyzing ground water and soil (within 5 km of the site) around S Bingipur village dump yard in Bangalore city. The focus of this study is to assess the contribution of waste dumping in soil contamination and in groundwater pollution. Collected surface soil samples from the open waste dumping area and controlled site (away from dumping yard) were examined and found variation in the soil composition. On the other hand, ground water samples were collected from the nearby village bore wells and lake, were analyzed and observed contamination of groundwater up to certain limit. This paper presents the impact of open dumping of solid waste on surrounding water and soil.

Index Terms- Municipal Solid Waste Management, Soil & Groundwater pollution, open dumping and Landfill, Leachate

I. INTRODUCTION

The threat of environmental pollution has been remaining the human world and is still growing fast due to excessive population growth in developing countries. Municipal solid waste (MSW) normally termed as garbage or trash is an unavoidable consequence of human activity. Population growth and economic development lead to enormous amounts of solid waste generation by the dwellers of urban areas. Urban MSW is usually generated from human settlements, small industries and commercial activities .Solid waste from hospitals and clinics is an additional source of MSW. Most of the countries do not have any specific technique of managing hospital and clinical wastes. So, they are mixed with MSW and pose a threat to human population and surrounding environment. Unsuitable disposal of MSW causes all types of

pollution: air, soil, and water. Indiscriminate dumping of wastes contaminates surface and ground water supplies. In urban areas, MSW clogs drains, creating stagnant water for insect breeding and floods during rainy seasons. Open burning of MSW contributes significantly to urban air pollution. Open dumping is quite common in developing countries due to low budget available for waste disposal. It also poses serious threat to groundwater. Health and safety issues also arise from improper MSWM. Insect and rodent vectors are attracted to the waste and can spread diseases such as cholera and dengue fever. Using water polluted by MSW for bathing, food, irrigation and drinking water can also expose individuals to disease organisms and other contaminants. In India, dumping on land is the most common method of waste disposal, because it is the cheapest method of waste disposal. Still, this method requires large area and proper drainage. The land disposal of municipal and industrial solid waste is potential cause of groundwater contamination. Unscientifically managed dumping yards are prone to groundwater contamination because of leachate production. Leachate is the liquid that seeps from solid wastes or other medium and have extracts with dissolved or suspended materials from it.

The volume of leachate depends principally on the area of the landfill, the meteorological and hydro-geological factors and effectiveness of capping. It is essential that the volume of leachate generated be kept to a minimum and ensures that the access of groundwater and surface water is minimized and controlled. The volume of leachate generated is therefore expected to be very high in humid regions with high rainfall, or high run off and shallow water table. Leachate from the solid waste dump has a significant effect on the chemical properties as well as the geotechnical properties of the soil. Leachate can modify the soil properties and significantly alter the behavior of soil.

The present study has been focused to conduct a detailed analysis of S.Bingipura solid waste landfill site to fulfill the following objectives:

- Assessment of quality of water bodies surrounding S.Bingipura
- To determine the nature of soil around the landfill site.
- Also compared the soil characteristics for contaminated and uncontaminated soil in the study area.

II. MATERIALS AND METHODS

2.1. Description of the Study Area

Bangalore is also known as the silicon valley of India. Bangalore urban district is located on the Deccan Plateau in the south eastern part of Karnataka. Bangalore district lies between $12^0$39’ to $13^0$18’ North Latitude and $77^0$22’ to $77^0$52’ East Longitude. The temperature in the district is known to vary between 39^0C (Max.) to 11^0C (Min.). The average rainfall in the district is found to be 831mm. The district comprises of the following river: Shimsha, Kanva, Arkavathi, South Pennar and Vrishabhavathi. Total geographical area of the district is 2196 sq.km. The city is situated at an elevation of 920m above MSL.

The district is spread across four Taluks; Bangalore North, Bangalore East, Bangalore South and Anekal. Bangalore is a hub for Information Technology, Biotechnology, Aerospace, & key knowledge based industries.

As per provisional reports of Census India, population of Bangalore in 2011 is 96, 21, 551; of which male and female are 50,22,661 and 45,98,890 respectively. The sex ratio of Bangalore is 916 females per 1000

males. The population density of Bangalore is 4,381 per sq.km. The Population growth of the city as per Census 2011 was found to be 47.18%.

The study was carried out at S.Bingipura. village located in the state of Karnataka as shown in Figure 2.1. The village lies in Bangalore Urban district and the block/tehsil is Anekal. S.Bingipura is situated about 21.30 km from the city, with an average height of about 915m above MSL. The study started in the month of January 2016, but presently the site is being closed down and they are proposing a park at the site. The site is known to receive 1.45 lakh tons quantity of waste from Bommanahalli BBMP zone area.

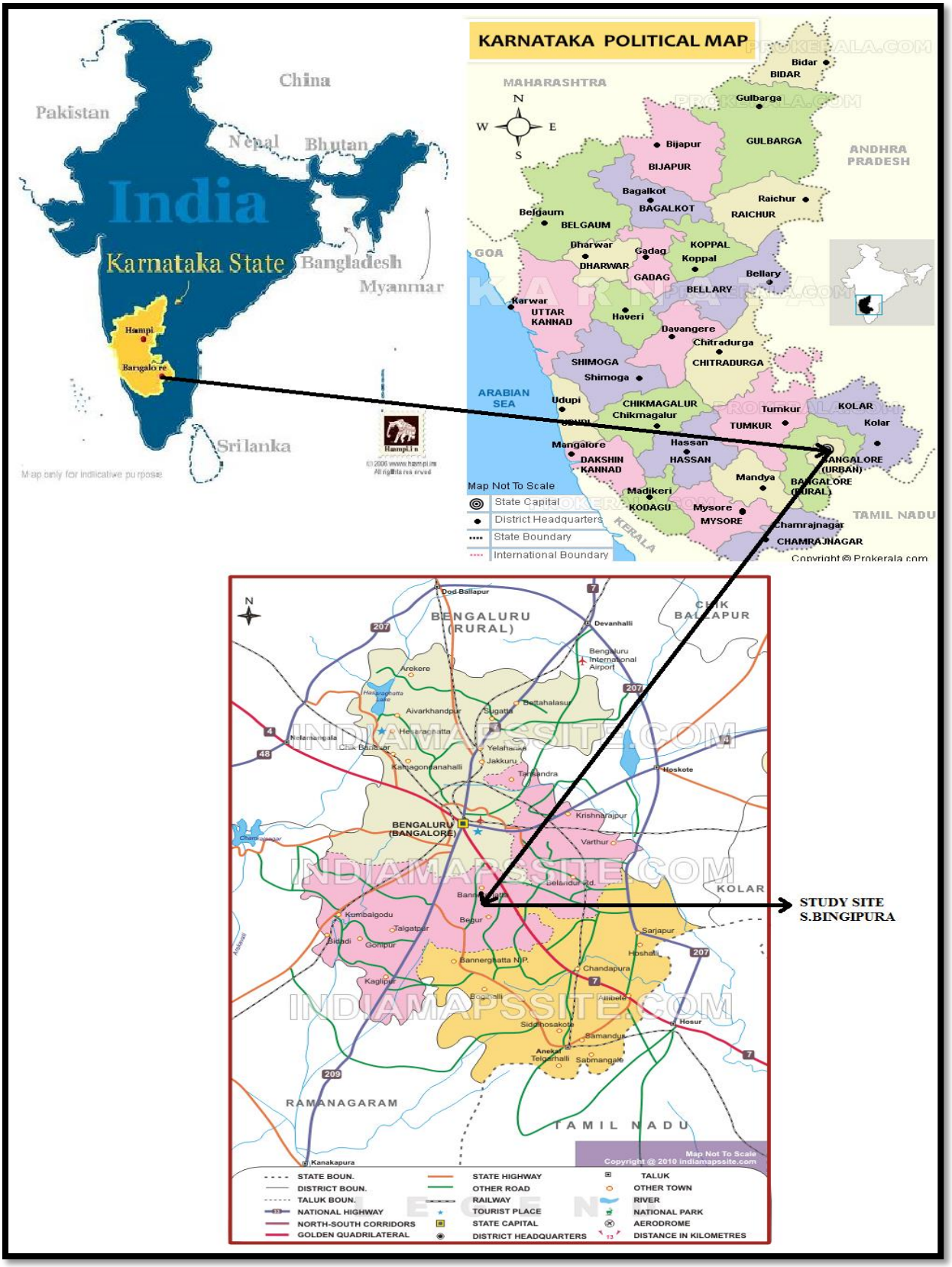


Fig.2.1. Index map of the study site

2.2 Sampling and Analytical Methods

Since there is no proper solid waste treatment and disposal, at the dump yard, there is a possibility of contamination to soil and groundwater in and around the site. So, a soil sample from the dump yard and soil away from the dump yard are collected for testing and comparison. Similarly, to check whether the ground water is being contaminated or not, the ground water samples were collected from a neighboring area (5 km) and tested. Soil samples were collected from the dumpsite, by removing the surface debris and subsurface soil dug to a depth of about 30cm and 1m with a hand auger. 5 Kg of soil sample was taken into the sterile containers and labeled. The samples were carried to laboratory and analyzed for water and soil chemical properties. The analysis was done as per the standard methods. Various Physico-chemical parameters examined in water samples include, pH, electrical conductivity (EC), total dissolved solids (TDS), total alkalinity (TA), total hardness (TH), calcium, magnesium, potassium, iron, chlorides, turbidity, Nitrates. Similarly soil samples were tested for moisture content, specific gravity, density of soil,gradation of soil properties, bulk density, electrical conductivity (EC) .The results were compared with BIS standard limits. The sampling locations were located on map(Figure 2.2 and 2.3) with help of GPS and detail of the site is given in Table.2.2. The methods adopted for the various parameters of soil and water analysis is mentione in the Table-2.3 and 2.4.

Table-2.2. Details of the Sampling locations

Location	Code	Latitude	Longitude	Environmental Attribute
S Bingipur	LT1	$77^0$37’43.57” E	$12^0$50’6.71” N	Leachate quality
	LP1	$77^0$37’30.41” E	$12^0$50’5.63” N	Leachate quality
	L 1	$77^0$37’54.4224” E	$12^0$49’59.214” N	Surface Water sampling
	BW 1	$77^0$37’17.462” E	$12^0$50’25.7064” N	Ground water sampling
	BW 2	$77^0$37’18.64” E	$12^0$50’21.27”N	Ground water sampling
	BW 3	$77^0$37’57.2736” E	$12^0$50’30.0696” N	Ground water sampling
	BW 4	$77^0$37’55.1172” E	12^0 50’25.1592” N	Ground water sampling
	BW 5	$77^0$38’0.2364” E	$12^0$49’59.4948” N	Ground water sampling
	BW 6	77^0’37’42.4956” E	$12^0$49’55.3836”N	Ground water sampling
	BW 7	$77^0$37’16.8924” E	$12^0$49’11.9136” N	Ground water sampling
	SS1	$77^0$37’40.6056” E	$12^0$50’13.1964” N	Soil quality Sampling location
	SS2	$77^0$37’44.1588” E	$12^0$50’10.2558” N	Soil quality Sampling location
	SS3	$77^0$37’43.8888” E	$12^0$50’6.8676” N	Soil quality Sampling location
	SS4	$77^0$37’53.9688” E	$12^0$49’13.2248”N	Soil quality Sampling location
	SS5	$77^0$37’47.4816” E	$12^0$49’5.2532” N	Soil quality Sampling location

Fig 2.2 Water Sampling locations

Fig 2.3 Soil Sampling locations

Table-2.3.The Methods of water and leachate Analysis

Sl.No	Parameter	Unit	Method adopted
1	Color	Hazens	Tintometer
2	Turbidity	NTU	Nephleometer
3	pH value	-	Digital pH meter
4	Conductivity	µS/cm	Conductivity meter
5	Total dissolved Solids	mg/l	Filter paper method
6	Suspended solids	mg/l	Filter paper method
7	Total solids	mg/l	Oven drying method
8	Total Hardness as CaCO3	mg/l	EDTA method
9	Calcium Hardness as CaCO3	mg/l	EDTA method
10	Magnesium Hardness as MgCO3	mg/l	EDTA method
11	Total Alkalinity as CaCO3	mg/l	Titration
12	Acidity	mg/l	Titration
13	Chlorides as Cl-	mg/l	Aginometric Titration
14	Sulphates as SO42-	mg/l	Flame Photometer
15	Nitrates as NO3-	mg/l	Titration
16	Fluorides as F-	mg/l	Ion Analyzer
17	Sodium	mg/l	Flame Photometer
18	Potassium	mg/l	Flame Photometer

Sl.No	Parameter	Unit	Method adopted
19	Ammonia	mg/l	Titration
20	Iron as Fe	mg/l	Spectro-photometer
21	DO	mg/l	Winkler's method
22	BOD	mg/l	Dilution method
23	COD	mg/l	Autoclave method
24	Lead	mg/l	Absorption Spectro-photometer
25	Nickel	mg/l	Absorption Spectro-photometer
26	Cadmium	mg/l	Absorption Spectro-photometer
27	Manganese	mg/l	Absorption Spectro-photometer
28	Zinc	mg/l	Absorption Spectro-photometer

Table-2.4. Tests on Soil

Sl.No.	Parameters	Method adopted
1	pH	Digital pH meter
2	Electrical Conductivity	Digital Conductivity meter
3	Bulk Density	Core cutter method
4	Dry Density	Core cutter method
5	Permeability	Constant head method
6	Moisture Content	Oven dry method
7	Specific Gravity	Pycnometer method

III. RESULTS AND DISCUSSIONS

The present paper mainly focused on identification of selected pollutants in the soil and ground water due to lechate generated from municipal solid waste landfill site.

3.1 Assessment of Ground water bodies

i. Colour : From Table.3.1. it was observed that the colour of the bore well samples are all less than 2, which falls under the desirable limit set by IS 10500:1991.

ii. Turbidity: From Table.3.1. it was observed that the amount of turbidity in the bore well samples varied from 0.5 NTU to 0.7 NTU, which is less than the desirable limit set by IS 10500:1991. Results depicts the variation of turbidity in the ground water samples

iii. pH : From Table.3.1. it was observed that the pH of the bore well samples varies from 7.72 to 8.19, which falls under the desirable limit set by IS 10500:1991. It was observed that the variation of pH in the ground water samples.

iv. Conductivity From Table.3.1. it was observed that the conductivity of the bore well samples varies from 589 µS/cm to 1451 µS/cm. Conductivity so high implies that the water sample is in fact contaminated.

v. Total Dissolved Solids, Suspended Solids and Total Solids From Table.3.1. it was observed that the TDS in the bore well samples varied from 390 mg/l to 930 mg/l, which is lower than the desirable limit set by IS 10500:1991. The amount of SS present is nil. Hence the TS also varies from 390 mg/l to 930 mg/l.

vi. Total Hardness, Calcium Hardness and Magnesium Hardness From Table.3.1. it was observed that the Total Hardness in the bore well samples varied from 380.11 mg/l to 171.23 mg/l, which is mostly under the desirable limit but under the permissible limit set by IS 10500:1991. BW-3 has total hardness more than the desirable limit. The Calcium Hardness varies from 252.50 mg/l to 95 mg/l while the Magnesium Hardness varies from 128 mg/l to 69.87 mg/l.

vii. Alkalinity and acidity From Table.3.1. it was observed that the alkalinity in the bore well samples varies from 308.80 mg/l to 183.45 mg/l. The alkalinity is greater than the desirable limit set by IS 10500:1991; for BW2(small amount), BW3 and BW4, whereas it falls under the desirable limit for the other samples From these results, it was observed that the acidity in the bore well samples varies from 2.36 mg/l to 1.07 mg/l..

viii. Chlorides From Table.3.1. it was observed that the amount of chlorides present in the bore well samples varied from 292.11 mg/l to 89.23 mg/l. BW3 has chlorides content more than desirable limit set by IS 10500:1991). The rest of the samples are found to have values within the desirable limit.

ix. Sulphates From Table.3.1. it was observed that the amount of sulphates present in the bore well samples varied from 99.11 mg/l to 31 mg/l, which falls under the desirable limit set by IS 10500:1991. Data depicts the variation of sulphates in the ground water samples.

x. Nitrates From Table.3.1. it was observed that the amount of nitrates present in the bore well samples varied from 15.24 mg/l to 7.11 mg/l, which falls under the desirable limit set by IS 10500:1991.

xi. Fluorides From Table.3.1. it was observed that the amount of fluorides present in the bore well samples varied from 0.42 mg/l to 0.24 mg/l, which falls under the desirable limit set by IS 10500:1991.

xii. Sodium From Table.3.1. it was observed that the amount of sodium present in the bore well samples varied from 144 mg/l to 56 mg/l.

xiii. Potassium From Table.3.1. it was observed that the amount of potassium present in the bore well samples varied from 8 mg/l to 4 mg/l.

xiv. Ammonia From Table.3.1. it was observed that the amount of ammonia present in BW 3 and BW5 samples was 0.24 mg/l and 0.12 mg/l respectively. The remaining samples had amount of ammonia below detection level (BDL).

xv. Iron From Table.3.1. it was observed that the iron content in the bore well samples varied from 0.15 mg/l to 0.07 mg/l, which is less than the desirable limit set by IS 10500:1991.

xvi. DO, BOD and COD From Table.3.1. it was observed that the amount of DO present in the bore well sample varied from 5.4 mg/l to 4.5 mg/l. Also, the amount of BOD present was found to be below detection level (BDL).

From Table.3.1. it was observed that the amount of COD present in the bore well sample varied from 4.89mg/l to 2.77 mg/l. Fig.4.16. depicts the variation of COD in the ground water samples.

Table 3.1.Ground water Assessment

Sl.No.	Test Parameters	Unit	BW1	BW2	BW3	BW4	BW5	BW6	BW7	IS 10500:1991	
										Desirable limit	Permissible limit
i	Colour	Hazen	< 2	< 2	< 2	< 2	< 2	< 2	< 2	5.00	25.00
ii	Turbidity	NTU	0.50	0.60	0.60	0.50	0.70	0.50	0.50	5.00	10.00
iii	pH	-	7.81	7.82	8.19	8.12	7.72	7.94	7.90	6.50-8.50	-
iv	Conductivity	µS/cm	664.00	760.00	1451.00	1327.00	612.00	589.00	1106.00	-	-
v	Total dissolved Solids	mg/l	420.00	510.00	930.00	850.00	390.00	390.00	690.00	500.00	2000.00
vi	Suspended solids	mg/l	< 1	< 1	< 1	< 1	< 1	< 1	< 1	-	-
vii	Total solids	mg/l	420.00	510.00	930.00	850.00	390.00	390.00	690.00	-	-
viii	Total Hardness as $CaCO_3$	mg/l	204.14	190.10	380.11	328.05	175.10	171.23	284.41	300.00	600.00
ix	Calcium Hardness as $CaCO_3$	mg/l	197.50	105.25	252.50	227.50	95.00	101.35	197.50	-	-
x	Magnesium Hardness as $MgCO_3$	mg/l	78.99	84.35	128.00	100.50	79.75	69.87	87.91	-	-
xi	Total Alkalinity as $CaCO_3$	mg/l	196.21	201.00	298.40	308.80	195.20	183.45	249.10	200.00	600.00
xii	Acidity	mg/l	1.07	1.75	2.36	1.79	1.20	1.12	1.42	-	-
xiii	Chlorides as Cl^-	mg/l	98.12	129.10	292.11	241.20	101.00	89.23	186.00	250.00	1000.00
xiv	Sulphates as SO_4^{2-}	mg/l	58.10	53.14	99.11	87.11	31.00	39.12	74.56	200.00	400.00
xv	Nitrates as NO_3^-	mg/l	7.11	12.10	15.24	12.14	8.23	9.01	10.24	45.00	-
xvi	Fluorides as F	mg/l	0.27	0.38	0.42	0.39	0.34	0.24	0.35	1.00	-
xvii	Sodium	mg/l	60.00	86.00	144.00	133.00	57.00	56.00	104.00	-	-
xviii	Potassium	mg/l	7.00	6.00	8.00	6.00	5.00	4.00	5.00	-	-
xix	Ammonia	mg/l	BDL	BDL	0.24	BDL	0.12	BDL	BDL	-	-
xx	Iron as Fe	mg/l	0.07	0.12	0.15	0.10	0.09	0.11	0.09	0.30	1.00
xxi	DO	mg/l	5.20	5.10	4.50	4.80	4.90	5.40	4.90	-	-
xxii	BOD	mg/l	BDL	BDL	BDL	BDL	BDL	BDL	BDL	-	-
xxiii	COD	mg/l	2.87	2.77	4.89	3.54	3.49	2.78	2.78	-	-

- BW1, BW2, BW3, BW4, BW5, BW6 and BW7 are Ground Water Samples
- BDL- Below Detective Level

3.2. Assessment of Surface water bodies

i. Colour From Table.3.2. it was observed that the colour of the lake sample was 4 Hazens, which falls under the desirable limit set by IS 10500:1991.

ii. Turbidity From Table.3.2. it was observed that the amount of turbidity in the lake sample was 21 NTU, which is more than the desirable limit set by IS 10500:1991. This may be due to the blown away leaves, sand, and also due to the villagers washing their clothes and cattle over the banks of the lake. The lake is also used by the commoners for bathing.

iii. pH From Table.3.2. it was observed that the pH of the lake sample was 7.96, which falls under the desirable limit set by IS 10500:1991.

iv. Conductivity From Table.3.2. it was observed that the conductivity of the lake sample was 3526 µS/cm. Conductivity so high only implies that the water sample is infact contaminated. This could be due to the possibility of the leachate to infiltrate through the soil and reach the lake and also other sources of pollution.

v. Total Dissolved Solids, Suspended Solids and Total Solids From Table.3.2. it was observed that the TDS, SS and TS present in the lake sample was 2270 mg/l, 50 mg/l and 2320 mg/l respectively, which is more than the desirable limit, set by IS 10500:1991.

vi. Total Hardness, Calcium Hardness and Magnesium Hardness From Table.3.2. it was observed that the Total Hardness, Calcium Hardness and Magnesium Hardness present in the lake sample was 320m mg/l, 190 mg/l and 130 mg/l respectively, which is greater than the desirable limit but under the permissible limit set by IS 10500:1991.

vii. Alkalinity and acidity From Table.3.2. it was observed that the alkalinity and acidity of the lake sample was 718.5 mg/l and zero respectively. The alkalinity is greater than both the desirable limit and the permissible limit set by IS 10500:1991.

viii. Chlorides From Table.3.2. it was observed that the amount of chlorides present in the lake sample was 723 mg/l, which is greater than the desirable limit but it falls under the permissible limit set by IS 10500:1991.

ix. Sulphates From Table.3.2. it was observed that the amount of sulphates present in the lake sample was 189 mg/l, which falls under the desirable limit set by IS 10500:1991.

x. Nitrates From Table.3.2. it was observed that the amount of nitrates present in the lake sample was 29 mg/l, which falls under the desirable limit set by IS 10500:1991.

xi. Fluorides From Table.3.2. it was observed that the amount of fluorides present in the lake sample was 0.78 mg/l, which falls under the desirable limit set by IS 10500:1991.

xii. Sodium From Table.3.2. it was observed that the amount of sodium present in the lake sample was 600 mg/l.

xiii. Potassium From Table.3.2. it was observed that the amount of potassium present in the lake sample was 41 mg/l.

xiv. Ammonia From Table.3.2. it was observed that the amount of ammonia present in the lake sample was 21.9 mg/l.

xv. Iron From Table.3.2. it was observed that the iron content in the lake sample was 0.46 mg/l, which is above the desirable limit but falls under the permissible limit set by IS 10500:1991.

xvi. DO, BOD and COD From Table.3.2. it was observed that the DO, BOD and COD of the lake sample was 0.8 mg/l, 34 mg/l and 187.14mg/l respectively, which is below the desirable limit set by IS 10500:1991.

Table.3.2. Bingipura Lake Assessment

Sl.No.	Test Parameters	Unit	L1	IS 10500:1991	
				Desirable limit	Permissible limit
i	Colour	Hazen	4.00	5.00	25.00
ii	Turbidity	NTU	21.00	5.00	10.00
iii	pH	-	7.96	6.50-8.50	-
iv	Conductivity	µS/cm	3526.00	-	-
v	Total dissolved Solids	mg/l	2270.00	500.00	2000.00
vi	Suspended solids	mg/l	50.00	-	-
vii	Total solids	mg/l	2320.00	-	-
viii	Total Hardness as $CaCO_3$	mg/l	320.00	300.00	600.00
ix	Calcium Hardness as $CaCO_3$	mg/l	190.00	-	-
x	Magnesium Hardness as $MgCO_3$	mg/l	130.00	-	-
xi	Total Alkalinity as $CaCO_3$	mg/l	718.50	200.00	600.00
xii	Acidity	mg/l	BDL	-	-
xiii	Chlorides as Cl^-	mg/l	723.00	250.00	1000.00
xiv	Sulphates as SO_4^{2-}	mg/l	189.00	200.00	400.00
xv	Nitrates as NO_3^-	mg/l	29.00	45.00	-
xvi	Fluorides as F	mg/l	0.78	1.00	-
xvii	Sodium	mg/l	600.00	-	-
xviii	Potassium	mg/l	41.00	-	-
xix	Ammonia	mg/l	2.19	-	-
xx	Iron as Fe	mg/l	0.46	0.30	1.00
xxi	DO	mg/l	0.8	-	-
xxii	BOD	mg/l	34.00	-	-
xxiii	COD	mg/l	187.14	-	-

3.2. Assessment of soil parameters

The soils in Bangalore city are mainly lateritic soil and red fine loamy to clayey soils. Red loamy soils generally occur on hilly and undulating land slope on granite and gneissic terrain. It is mainly seen in eastern and southern parts of Bangalore. Laterite soil is usually found in Anekal taluk and western parts of Bangalore North and South taluks. The results of soil analysis are tabulated in Table.3.3. It was observed from the Table 3. that the color parameter of soil near the dumping location SS1,SS2 and SS3 is dark brown to dark black. Hence an attempt has made to collect two more sample viz one at the contaminated soil at the dumping site itself and other one is 2 km away from the land fill site. The analysed results are depicted in Table 3.4.

Table.3.3. Tests on Soil

Sl. No.	Parameter	Study Site				
		SS 1	SS 2	SS 3	SS 4	SS 5
1	Colour	Dark Brown	Dark Brown	Black	Light Brown	Light Brown
2	pH	7.18	7.25	8.46	7.15	7.20
3	Electrical Conductivity, /Ω/cm	1.00	0.80	3.20	0.10	0.30
4	Moisture Content, %	10	11	13	12	13
5	Specific Gravity	2.44	2.46	2.33	2.35	2.37
6	Field Density, g/cm^3	1.38	1.44	1.56	1.81	1.74
7	Dry Density, g/cm^3	1.25	1.29	1.38	1.61	1.53
8	Dry Unit Weight, kN/m^3	12.26	12.65	13.53	15.79	15.01
9	Permeability, cm/hr	2.5 to 5 Moderate	2.5 to 5 Moderate	2.5 to 5 Moderate	2.5 to 5 Moderate	2.5 to 5 Moderate

Experimental results obtained on effect of municipal solid waste lechate on the characteristics of soil on both contaminated and uncontaminated soil presented in Table 3.4. The present paper also focused on identification of selected pollutants in the soil due to lechate generated from municipal solid waste landfill site and uncontaminated soil to serve as control. Finally comparison of both contaminated and uncontaminated soil characteristics was made

Table.3.4. Quality of soil Parameters estimated in contaminated and uncontaminated soils

Sl No	Parameters	Contaminated Soil(SS1)	Un contaminated Soil (SS5)
1	Moisture Content	14%	11%
2	Specific Gravity	2.437	2.37
3	Particle Size Distribution		

Sl No	Parameters	Contaminated Soil(SS1)	Un contaminated Soil (SS5)
	Uniformity coefficient	Cu= 5.5	Cu= 8.57
	Curvature coefficient	Cc= 2.36	Cc= 3.07
4	Permeability	0.62 Cm/S	0.069 Cm/S
5	Shear Strength	13.5 Kn/Sq M	13 Kn/ Sq M
6	Compressibility	0.82 Sqm/ Kn	1 Sqm/ Kn
7	pH	7.20	8.00
8	Chloride	108.46 Mg/L	40mg/L
9	Alkalinity	83 Mg/L As Caco3	236 Mg/L As Caco3

3.2.1. Natural Moisture Content of contaminated and uncontaminated soil The results show that the values of the Natural Moisture Content of the Uncontaminated soil is lower compared to those of the contaminated soil samples. This trend could attribute reason that the contaminated soil is expected to be damper, since the natural ground level is covered by the MSW, thereby preventing direct evaporation of moisture from the soil below.

3.2.2. Specific gravity of contaminated and uncontaminated soil The results show that, the values of the specific gravity of the contaminated soil was higher than the uncontaminated soil. It could be attributed that the specific gravity of contaminated soil is higher because of the higher moisture content of the contaminated soil as compared to uncontaminated soil.

3.2.3. Particle Size Distribution of contaminated and uncontaminated soil From the Table 3.1 the uncontaminated soil is relatively homogeneous and contaminated soil has more fines than the uncontaminated soil. The higher percentage of fine content recorded for the contaminated soil can be attributed to the fines emanating from the decomposed MSW above the soil. Also during bacterial degradation or decomposition of MSW large amount of fines are produced.

3.2.4. Permeability Test of contaminated and uncontaminated soil Laboratory falling head method was used in the determination of the coefficient of permeability of the soils. From the results, the contaminated soil has higher values of coefficient of permeability than the uncontaminated soils. These results somehow contradict the fact that the contaminated soil particles are loosely arranged which would have ordinarily increased the pore space in the soil. This anomaly may be due to particles flocculation as a result of contamination with MSW. The flocculation process may have altered the behaviours of the fine particles from clay-like to silt-like and consequently, making the soil more permeable.

3.2.5. Shear Strength Test Contaminated and Uncontaminated Soil The shear strength parameters were determined by undrained triaxial test using undisturbed soil samples. From the results, the shear strength value is higher in case of contaminated soil than those recorded for the uncontaminated soil. The relatively high value recorded for contaminated soil samples a result of pseudocohesion, brought about by leachate from the decomposing MSW. This may be due to particle flocculation as a result of contamination with MSW.

3.2.6. Compressibility Test of contaminated and uncontaminated soil Consolidation test on the undisturbed samples was use to investigate the effect of the MSW on the compressibility characteristics of the soils. The results show that the contaminated soil has relatively lower values than uncontaminated soil. The lower values obtained for contaminated soil in comparison with the values obtained for

uncontaminated soil, can be attributed to the soil immediately beneath the MSW don't undergoing any compression as a result of the weight of the MSW above.

3.2.7. pH of contaminated and uncontaminated soil We can conclude that, pH value of uncontaminated soils is higher than the contaminated soils. The pH of the contaminated soils is 7.20, it signifies that it is slightly acidic in nature compared to uncontaminated soil could be reason behind that the nature of the solid waste contribute acidity of the soil. Due to this reason the pH of contaminated soil is slightly acidic than uncontaminated soils. Alkalinity value of uncontaminated soils is higher than the contaminated soils. This could be the reason that the pH of the contaminated soil is slightly acidic than uncontaminated soil.

3.2.8. Chloride of contaminated and uncontaminated soil The chloride concentration in contaminated soil is 108.46 mg/l where as uncontaminated is 40 mg/l, it indicates that it is higher than uncontaminated soil. This contribute due to disposal of solid waste, the quality of the soil is reduced and it clearly indicated by the chloride values of contaminated soils.

3.2.9. Alkalinity of contaminated and uncontaminated soil The alkalinity concentration in contaminated soil is 83 mg/l as caco3 where as uncontaminated is 236 Mg/L As Caco3, it indicates that it is lower than uncontaminated soil. This clearly indicated by the lesser alkalinity values of contaminated soils due to acidic properities due to the concentration of leachate.

The results of contaminated and uncontaminated soils are represented in Table 3.4. The result in the table indicates that except for pH and alkalinity, all other parameters are higher in contaminated soil compared to uncontaminated soils. The study concludes based on the results obtained, the disposal site soil quality is reduced compared to uncontaminated soil. In other words, due to the disposal of solid waste on land the soil quality gets reduced.

IV. CONCLUSION

The following conclusions has been drawn based on the results obtained in the present study.

- The surface water sample is found to have significantly high salinity and alkalinity as reflected in their values for conductivity, TDS, alkalinity and pH. Hence it indicating that the surface water body is polluted.
- Test result on ground water concluded that certain bore well on the down stream side were polluted.
- The Analysis of the soil samples around the site shows that the soil has moderate permeability.
- Also Based on the experiment results obtained from the soil sample analysed in both contaminated and uncontaminated soils following major conclusions have been drawn.
    - The coefficient of permeability of the contaminated soil has higher than the uncontaminated soils. This indicates that due to disposal of solid waste the quality of the soil is reduced and it clearly indicated by the chloride values of contaminated soils.
    - Study conclude based on the results obtained, the disposal site soil quality is reduced compared to uncontaminated soil. In other words, due to the disposal of solid waste on land the soil quality gets reduced.

ACKNOWLEDGMENT

The authors are great full to the authority of Bruhat Bangalore Mahanagar Palike (BBMP) for their kind co-operation and giving permission to carry out this study at landfill site.

REFERENCES

[1] A Asadi Huat BBK, H Moayedi, Shariatmadari N, Parsaie A (2011). Changes of hydraulic conductivity of silty clayey sand soil under the effects of municipal solid waste leachate. International Journal of the Physical Sciences, 6(12): 2869-2874.

[2] APHA, Standard Methods for the Examination of Water and Wastewater , 21 ed. Washington D.C: American Public Health Association, 2005.

[3] Barjinder Bhalla, M.S. Saini, M.K. Jha (2012) "Characterization of Leachate from Municipal Solid Waste (MSW) Landfilling Sites of Ludhiana", India: A Comparative Study. International Journal of Engineering Research and Applications (IJERA) ISSN: 2248-9622 Vol. 2, Issue 6, November-December 2012, pp.732-745.

[4] Basavaraj Paruti, and Santhaveeranagoud B, (2019)." CHARACTERISATION OF LEACHATE IN SOLID WASTE LANDFILL SITE- A Case Study of S Bingipur in Bangalore", Journal of Engineering Annals, of the Faculty of Engineering Hunedoara. Tome XVII [2019] | Fascicule 1 [February] pp. 205-208

[5] Bharat Jhamnani and SK Singh (2009) "Groundwater Contamination due to Bhalaswa Landfill Site in New Delhi", International Journal of Civil and Environmental Engineering, vol 1(3), pp. 121-125 .

[6] E.Panahpour ,A. Gholami, A.H Davami (2011). Influence of Garbage Leachate on Soil Reaction, Salinity and Soil Organic Matter in East of Isfahan. World Academy of Science, Engineering and Technology, 171-176

[7] Evaluation of Technology for processing existing waste at Seven Landfill sites of BBMP, Bangalore Technical committee Recommendations on EOI application REPORT; Date: 03-01-2013.

[8] Iwekumo Ebibofe, Michael (2015) "Leachate characterization of active and closed dump sites in Port Harcourt metropolis", Nigeria. Int. J. Biol. Chem. Sci. 9(2): 1107-1119.

[9] Lovleen Gupta, Swati Rani (2014) "Leachate characterization and evaluating its impact on groundwater quality in vicinity of landfill site area".

[10] O.Schwarz-Herion, A. Omran, and H.PRapp, (2008)."A Case Study on Successful Municipal Solid Waste Management in Industrialized Countries by the Example of Karlsruhe City, Germany", Journal of Engineering Annals, of the Faculty of Engineering Hunedoara. 6 (3) pp. 266-273

[11] Ohwoghere–Asuma, Aweto (2013) "Leachate Characterization and Assessment of Groundwater and Surface Water Qualities near Municipal Solid Waste Dump Site", Effurun, Delta State, Nigeria.

[12] The Municipal Solid Wastes (management and handling) rules, 2000 (Ministry of Environment & Forests) notification New Delhi, the 25th september, 2000

[13] Naveen.B.P, Sivapullaiah.P.V, Sitharam.T.G, Ramachandra.T.V (2014) "Characterization of leachate from municipal landfill and its effect on surrounding water bodies", Mavallipura.

A survey on Machine Learning in Compiler

Sriyasree Swain*, Bhavani Supriya*, Nivedita Manohar, Vijaya Lakshmi R****

*Dept.of Information Technology
ACED, Alliance University, Bengaluru

**Assistant Professor, Dept.of CS/IT,
ACED, Alliance University,Bengaluru

Abstract: Optimization of a tuning hardwired compiler is the rapidly budding hardware which makes porting a compiler which will be optimizing in its feature and challenging. The approach for the such challenging compiler is with development of modular, self-optimizing and extensible compilers that adopt the best optimization heuristics based on the behavior of the platform. The contribution of the machine learning towards the development of compiler which is capable to adjust automatically with improved execution time, code size on different architectures is Machine Learning for Embedded PrOgramS optimization (MILE POST). Recursive queries technique can be utilized for effective execution plans and the resulting runtime plans can be executed on a single unified data parallel query processing engine.

High performance software development is also difficult task that requires the use of low-level, architecture specific programming models such as MPI for clusters, CUDA for GPUs, and OpenMP for CMPs. Probabilistic search of the optimization space can support to a significance speedup over the baseline compilers with the higher optimization settings, on several different processor architecture. Classifier systems are massively parallel, message-passing, rule-based systems that learn through credit assignment (the bucket brigade algorithm) and rule discovery (the genetic algorithm). They typically operate in environments that exhibit some characteristics such as -perpetually novel events accompanied by large amounts of noisy or irrelevant data, continual, often real-time, requirements for action, implicitly or inexactly defined goals, sparse payoff or reinforcement obtainable only through long action sequences. Classifier systems are designed to absorb new information continuously from such environments, devising sets of competing hypotheses (expressed as rules) without disturbing significantly capabilities already acquired.

The survey highlights the approaches taken so far, the obtained results, the fine-grain classification among different approaches and finally, the influential research in the field. The challenges are listed at end.

Index Terms- Optimization, Classifier systems, Machine Learning

I INTRODUCTION

In simple design of game software, any good player can complicate the system by throwing some twists. But in real world such twist will be more, and design of system should support to face such twists. For a machine learning system, the problem is one of constructing relevant categories from the system's primitives (pixels, features, or whatever else is taken as given). Discovery of relevant categories is only half the job; the system must also discover what kinds of action are appropriate to each category. The overall process bears a close relation to the Newell-Simon [3] problem solving paradigm, though there are differences arising from problems created by perpetual novelty, imperfect information, implicit definition

of the goals, and the typically long, coordinated action sequences required to attain goals. Environments wherein timely outside intervention is difficult or impossible. The only option then is learning or, using the more inclusive word, adaptation. In broadest terms, the object of a learning system, natural or artificial, is the expansion of its knowledge in the face of uncertainty. More directly, a learning system improves its performance by generalizing upon experience. Clearly, in the face of perpetual novelty, experience can guide future action only if there are relevant regularities in the system's environment. Human experience indicates that the real world abounds in regularities, but this does not mean that it is easy to extract and exploit them. The overall process bears a close relation to the Newell-Simon [40] problem solving paradigm, though there are differences arising from problems created by perpetual novelty, imperfect information, implicit definition of the goals, and the typically long, coordinated action sequences required to attain goals[1][2][3].

New compiler techniques must arise to support complex image processing applications without sacrificing programmability. This paper focuses on two image processing interfaces considered as DSLs, Simple Morphological Image Library(SMIL) and Framework for Embedded Image Applications (FREIA), supporting each a different set of hardware targets and providing different levels of programmability. We built a compiler to automatically generate lower-level but more portable FREIA DSL code from high-level SMIL DSL applications. We evaluate this compiler on a set of seven image processing applications. Some of the advanced compilers to support image processing are Simply Morphological Image Library(SMIL), Framework for Embedded Image Applications (FREIA), Cython which is a Python to C – Compiler[4].

To locate the bugs of program it is necessary to design new types of compilers which will support to debug the errors in server level as well as in kernel levels[5]. These all compilers are the challenges of the competitive industry these are discussed. The reference reveals that in the domain of Big data domain specific languages are necessary to extend with machine learning concept in bigdata. In such situation specific languages should support for all aspect with optimized way[6].

ScalOps is a DSL with the goal of enabling Machine Learning algorithms to run on a cloud computing environment and overcoming a limitation of the traditional MapReduce programming model: the lack of iteration. ScalOps is a textual programming DSL developed in jointly by the University of California, Irvine and Santa Cruz, and the division Yahoo! Research. The DSL also has Scala language serving as a host language, which means that ScalOps is an internal DSL. Its high-level syntax makes it a declarative language, and as the type checking happens in compilation time, ScalOps is considered a statically typed. ScalOps needs to be compiled to generate lower level code, which makes it be classified as a translated language, according to the analysis of this survey. Additionally, the language supports vector, matrix, and graph operations in both parallel and cloud computing environment. To support iterations in MapReduce, ScalOps designers introduced an enhanced version of the programming model called Map-Reduce-Update. This new version consists of three user-defined functions called map, reduce, and update. The map function receives read-only global state values and is applied to training data points in parallel. The reduce function aggregates the output of the map function. Finally, the update function receives the aggregated value and produces a new global state value for the next iteration. Alternatively, when appropriate, the update function indicates that no additional iteration is necessary [9][10][11].

Compilation can be speed up the compile processes by at least a factor of two with almost the same generated code quality on the SPEC2000 benchmark suite, and that our logistic classifier achieves the same prediction quality for non-SPEC benchmarks.

II. METHODOLOGY

The compilation algorithm can be written as shown in the Table 1. The algorithm accept the Majority Inverter Graphs (MIG) M as an input to give a Programmable Logic-in-Memory (PliM) program as an output. PliM enables logic operations on a regular Resistive Random-Access Memories (RRAM) array. It uses a single instruction RM3, which computes the three major operations in which one input is inverted [10]

Table 1: Compilation Algorithm

Input MIG M

1. Foreach leaf in M
2. Do set COMP[v]← T
3. End
4. Foreach MIG node in M
5. Do if all children of v are computed then
6. Q.enqueue(v)
7. End
8. End
9. While Q is not empty
10. Do set c ← Q.pop();
11. Set P ← P U `translate( c );`
12. Set COMP[c] ←T
13. Foreach parent of c do
14. If all children of v are computed then
15. Q.enqueue(v);
16. End
17. End

SMIL Python Code:

```
import smilPython as smil
imin = smil.Image("input.png")
imout = smil.Image(imin)
smil.dilate(imin, imout)
imout.save("output.png")
```

FREIA C Output Code:

```
#include "freia.h"
#include "smil-freia.h"
int main(int argc, char *argv[]) {
/* initializations... */
freia_data2d *imin;
imin = freia_common_create_data(/* */ );
freia_data2d *imout;
imout = freia_common_create_data(/* */ );
#define e0 SMILTOFREIA_SQUSE
#define e0_s 1
freia_cipo_dilate_generic_8c(imout, imin, e0, e0_s);
freia_common_tx_image(imout, &fdout);
freia_common_destruct_data(imout);
freia_common_destruct_data(imin);
/* shutdown... */
}
```

Figure 1 SMIL Input Code and FREIA C Code

For SMIL applications, the compiler such as smiltofreia generates directly FREIA C code from SMIL Python programs. The sample of such code is as shown figure 1.[13]

III CONCLUSION

The study throws many challenges such as requirement of optimized compiler for processing image processing languages, conversion from one languages to other many languages. The translation of languages is essential in case of multi programming development. This makes user to adopt any one language. Hardware specific optimization compilation is also required which is one of the challenging.

REFERENCES

[1] Grigori Fursin,et al, "MILEPOST GCC; machine learning based research compiler".

[2] L.B.Booker,D.E. Goldberg and J.H Holland," Classifier Systems and Genetic Algorithms", Artificial Intelligence, Elsevier Science Publisher, 1989.

[3] Newell A and Simon H A, Human Problem Solving (Prentice-Hall, Englewood Cliffs,NJ,1972).

[4] Pierre Guillou, Bnoit Pin, Fabien Coelho, Francois Irigoin, " A Dynamic to Static DSL Compiler for Image Processing Application", Processing Applications. Compilers for Parallel Computing 2016. . July 8, 2016, Valladolid, Spain.

[5] Chengnian Sun, VuLe, Qirun Zhang, Zhendong Su, " Toward Compiler Bugs in GCC and LLVM", ACM, ISSTA '16, July 18-22

[6] Ivens Portugal David R. Paulo Alencar David R. Donald Cowan David R. ," A Survey on Domain-Specific Languages for Machine Learning in Big Data",

[7] Gennady Pekhimenko , Angela Demke Brown, "Efficient Program Compilation through Machine Learning Techniques", http://sysweb.cs.toronto.edu

[8] Borkar, V., Bu, Y., Carey, M. J., Rosen, J., Polyzotis, N., Condie, T., Weimer, M., & Ramakrishnan, R. (2012). Declarative Systems for Large-Scale Machine Learning. IEEE Bulletin of the Technical Committee on Data Engineering, 35(2), 24-32.

[9] Mathias Soeken et al, "An MIG based Compiler for programmable Logic-in-Memory Archtectures

[10] Weimer, M., Condie, T., & Ramakrishnan, R. (2011). Machine learning in ScalOps, a higher order cloud computing language. In NIPS 2011 Workshop on Parallel and Large-scale Machine Learning (BigLearn) (Vol. 9, pp. 389-396).

[11] Odersky, M., Spoon, L., & Venners, B. (2005). Programming in Scala. École Polytechnique Fédérale de Lausanne.

[12] Pierrc Guilou et al " A Dynamic to Static DSL Compiler for Image Processing Application",

Image Scrambling through Two Level Arnold Transform

Satish A[*], Erapu Vara Prasad[*], Tejasvi R[*], Swapna P[*], Vijayarajan R[*]

[*] Department of ECE,
[*]RGM College of Engineering and Technology, Nandyal, AP
anchulasathish@gmail.com, varaprasad.erapu@gmail.com, reddytejasvi1998@gmail.com
swapna1cube@gmail.com, viraj2k@gmail.com

Abstract- In today's world of multimedia applications, digital image security is of utmost importance. This paper proposes a two level image scrambling algorithm using Arnold transform with different number of iterations. Pixels are scrambled for each iteration in the conventional Arnold scrambling. Due to its periodic property, the original image can be retrieved by experimenting the inverse Arnold transform on the scrambled image. The main objective of this paper is to improve the complexity of malicious decryption by considering Arnold scrambling for pixels and blocks. In the first level, Arnold pixel scrambling is performed on the original image of size N × N, where the pixels are scrambled based on number of iterations. In the second level the scrambled image is divided in to 8×8 blocks for which Arnold scrambling is conducted block-wise. In the decryption process, inverse Arnold block scrambling is applied first for the scrambled image and Arnold pixel scrambling is conducted second. Decryption process needs to be carried out with proper order of scrambling and iterations. This order of scrambling and number of iterations increase the complexity of malicious decryption. Correlation coefficient, number of pixel change rate and unified average changing intensity are computed between the scrambled and original images for performance analysis. The proposed method is also analysed in each level for its robustness.

Index Terms-Arnold scrambling, encryption, decryption, structural similarity index

I. INTRODUCTION

Due to recent advancement in imaging technologies and multimedia applications, a huge amount of multimedia content in the form of text, audio, image, video and computer generated contents are exchanged in various networks. These multimedia contents often exposed to malicious access, tampering, copyright and ownership issues. All these issues are dealt with the solutions in the form of watermarking [1, 2, 3], authentication and cryptography algorithms [4, 5]. High security should be provided for an image that contains confidential information. Subsequently, the confidential information will be hidden in the image in the form of secret authentication codes by using different algorithms [6]. In order to differ from conventional methods to hide data in the image, an algorithm can be developed using scrambling. Digital image scrambling transforms an image in such a way that an intruder can't recognise the image as it is a disguised one. Image scrambling technology provides an algorithm which doesn't involves any password or any authentication codes [7] but provides high security for the image that is being transmitted. The main purpose is to transmit the image through the public networks. After scrambling an image, the scrambled image will be disorganised, so an intruder can't decipher it [7].

Image scrambling can be performed by using Arnold scrambling algorithm [8]. The special feature of Arnold scrambling is that it uses periodicity concept [7]. According to Arnold scrambling, the original image can be recovered after a certain number of iterations based on the size of the image. But the number of iterations will be different for different size of the images and the number of cycles does not follow any order. Currently, Arnold scrambling is applied to pixels only but it can be extended to blocks

of the image also. If the scrambling is performed on both pixels and blocks the robustness and security of the image can be improved. Arnold scrambling for pixels can be applied to any image of any size. But to apply Arnold scrambling to an image which is divided into blocks, the image size should be of order M×M. If the size of the image is not M×M, it can be made M×M by adding zeros to the image which is called as padding. Arnold transform is widely used in image stenography, authentication, tamper detection, self-recovery and image cryptography algorithms. In all these cases, Arnold transform is used as a scrambling step in which the number of iterations is used as a key. Arnold transform for pixel scrambling is used in most of the applications and hence provides one key for the security. This paper proposes Arnold transform for pixel scrambling and block scrambling. The coordinates of the pixels are scrambled first which is followed by the coordinates of the blocks and thus providing two levels of security for scrambling. If the first level descrambling is successful, then only the second level descrambling can be carried out. This increases the complexity of malicious and unauthorized descrambling of images.

This paper proposes a two level image scrambling to increase the robustness of Arnold transform. First, the plain image is divided into blocks and each block is assigned a coordinate. The block coordinates can be transformed through Arnold scrambling. Hence, each block of the image will get a new coordinate and gets scrambled. Once the blocks are scrambled and arranged as an image, pixel scrambling can be carried out to scramble all the pixels in the image. This two level can also be implemented by doing pixel scrambling first which is followed by block scrambling. This paper analyses the two methods and the performance is analysed by Correlation coefficient (CC), number of pixel change rate (NPCR) and unified average changing intensity (UACI).

This paper is organized as follows. Section 2 deals with conventional and the proposed two level Arnold pixel scrambling which is followed by experimental results and analysis in section 3. Conclusion is given in section 4.

II. ARNOLD TRANSFORM

Arnold scrambling transforms the position of a pixel from (x, y) to a new position (s, t) [7]. The position of pixels changes from one point to another [7,9,10 & 11] based on the equation (1)

$$\begin{bmatrix} s \\ t \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} (\text{mod } 1) \tag{1}$$

In order to apply the transformation to a digital image the term mod 1 can be replaced by mod N where N is the size of the digital image [12].

$$\begin{bmatrix} s \\ t \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} (\text{mod N}) \tag{2}$$

The transformation matrix of Arnold scrambling is a mod2 transformation matrix which is given by A. (x, y) T is the input and (s, t) T is the output. Considering the feedback and iterative process, the equation (2) can be written as follows

$$P_{xy}^{n+1} = A\, P_{xy}^{n} \text{ (mod N)} \tag{3}$$
$$P_{xy}^{n} = (x, y)^T$$

Where n specifies the number of iterations i.e., n=0, 1, 2, … After applying scrambling to an image, the scrambled image will be different from the original image. Apart from being simple and easy, Arnold scrambling has the advantage of periodicity.

Inverse Arnold transform [10] can be given by

$$\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 2 & -1 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} s \\ t \end{bmatrix} (mod\ N) \tag{4}$$

III. TWO LEVEL ARNOLD SCRAMBLING

This paper proposes a two level image scrambling to increase the robustness of Arnold transform. First, the plain image is divided into blocks and each block is assigned a coordinate. The block coordinates can be transformed through Arnold scrambling. Hence, each block of the image will get a new coordinate and gets scrambled. Once the blocks are scrambled and arranged as an image, pixel scrambling can be carried out to scramble all the pixels in the image. This two level algorithm can also be implemented by doing pixel scrambling first which is followed by block scrambling.

A. Block Scrambling

In block scrambling, image is divided into M×M blocks and each block is assigned a coordinate {m,n} according to their spatial orientation. For an image of size 512×512, image can be divided into 64×64 blocks. Hence, there are spatial coordinates in the set of {(1,1), (1,2),…,(1,8),…,(8,1),(8,2),…,(8,8)}. Arnold scrambling is applied to the coordinates of blocks and hence each coordinate is assigned a new coordinate as given by equation (2). This paper proposes a block scrambling method for the blocks with same spatial resolution. This is possible only when the input image I(x,y) is of spatial resolution $2^n \times 2^n$; where n=1,2,3,…,N. Scrambling of blocks through Arnold transform is given as follows.

$$\begin{bmatrix}\{B(x_i)\}\\ \{B(y_i)\}\end{bmatrix} = \begin{bmatrix}1 & 1\\ 1 & 2\end{bmatrix}\begin{bmatrix}\{B(x_{i-1})\}\\ \{B(y_{i-1})\}\end{bmatrix} \pmod{M} \tag{5}$$

Where $B(x_i,y_i)$ is the coordinate of the block of an image. M is number of rows or number of columns of all the blocks.

B. Illustration of Arnold block scrambling

Fig. 1 Illustration of Arnold block scrambling

First, the image is divided into blocks. For example, a 512 x 512 image is converted into blocks of size 128 x 128. Then the image will be divided into 16 blocks as shown in Fig. When Arnold scrambling is applied to blocks then the positions of blocks will get shifted. Let us consider block6 whose original position is at (2, 2) but after scrambling the position of block6 is shifted to new position i.e., (4, 3) for one iteration. For further iterations the input will be the output of previous iteration.

C. Scrambling Process

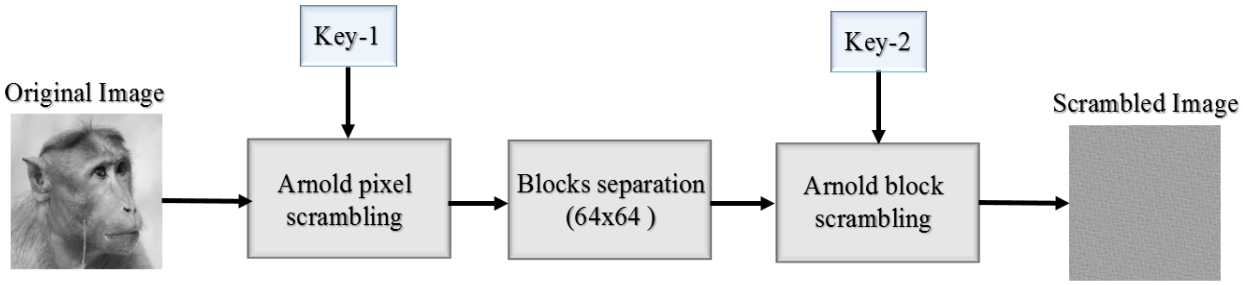


Fig. 2 Scrambling process

In the scrambling process, the original image is scrambled in two levels as shown in Fig. 2. In the first level the image is subjected to pixel Arnold scrambling with a specific number of iterations. The information related to number of iterations of level1 will be in key1 (K1). The scrambled image is divided into 8 x 8 blocks. In the second level the divided image is subjected to block Arnold scrambling with other specified number of iterations. The information of number of iterations of second level scrambling will be in key2 (K2). The image obtained after the second level scrambling is the image that will be transmitted.

D. Descrambling Process

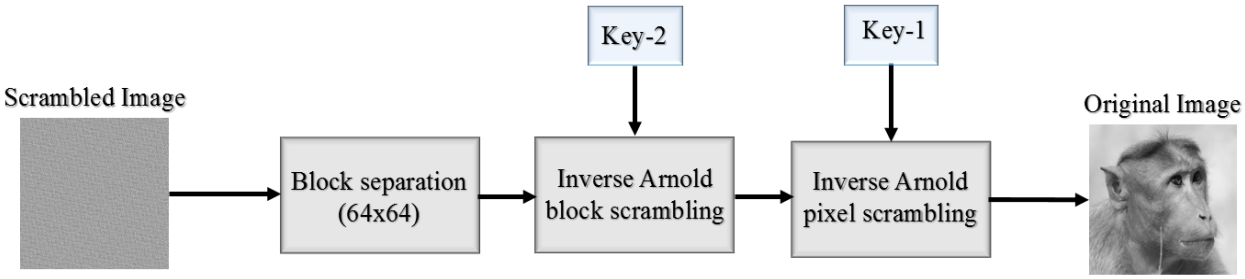


Fig. 3 Descrambling process

Same as scrambling, the original image will be extracted in two levels as shown in Fig 3. The scrambled image is divided into blocks and inverse block Arnold scrambling is applied based on key2. In the second level, the obtained image after the inverse block Arnold scrambling will be subjected to pixel level inverse Arnold scrambling based on key1. The obtained image after the second level is the original image.

E. Algorithm

The scrambling and descrambling steps for pixel & block two level Arnold scrambling are stated below.

Scrambling steps

1. Arnold pixel scrambling with certain number of iterations (key1) as given in equation (1) is applied to the input image I(x,y) and denoted as Ip(x,y).
2. Ip(x,y) is divided into blocks of size 64×64.
3. Each block is arranged in a matrix format and assigned a spatial coordinate (m, n).
4. Arnold scrambling with certain number of iterations (key2) is applied to the coordinates of the blocks and the blocks are rearranged with the new coordinate as an image.
5. The final two level scrambled image I'(x,y) is obtained.

Descrambling steps

1. I'(x,y) is divided into blocks as done in the scrambling
2. Blocks with the coordinates are descrambled and all the blocks are arranged in an image format.
3. Arnold descrambling is applied to the output image of block descrambling image to get the original image I(x,y).

IV. Experimental Results and Analysis

The proposed two level pixel scrambling is carried out in two ways. In the first method, block level scrambling applied first which is followed by pixel level scrambling. In the second method, pixel scrambling is followed by block scrambling. Images with spatial resolution of 512×512 or 2n×2n; where n=4,5,6,7,8,… are considered for experiments and block separation is also carried out with the block size of 2m×2m; where m=4,5,6,7,8,… and m < n. Images with other spatial resolution can also be considered with proper padding to carry out block processing.

A. Performance metrics

Performance of the two level Arnold scrambling is analysed by the metrics such as CC, NPCR and UACI.

Correlation Coefficient (CC):

The correlation coefficient analyzes the relationship between two normally distributed random variables. It is evaluated through covariance and variance of the plain-scrambled images [13, 14]. The covariance between a plain image (Ip), and scrambled image (Is) is given by cov (Ip,I_s) and variance is represented as σ_{I_p} and σ_{I_s}. The correlation coefficient is given by

$$r_{xy} = \frac{|\mathrm{cov}(I_p, I_e)|}{\sqrt{\sigma_{I_p}}\sqrt{\sigma_{I_e}}} \tag{5}$$

NPCR

NPCR [15] measures the different pixel numbers between the scrambled image and the plain image.

$$\mathrm{NPCR}\ \% = \sum_{i=1}^{M}\sum_{j=1}^{N}\frac{P(i,j)}{MN} \times 100 \tag{6}$$

Where P(i,j)=0 if $I_p(i,j)=I_s(i,j)$ and P(i,j)=1 if $I_p(i,j)\neq I_s(i,j)$

For this experiment, original image is scrambled by the proposed two level Arnold scrambling algorithm. The scrambled images are compared with the original image and NPCR value is evaluated. If NPCR is high that denotes good performance of the scrambling algorithm.

UACI

UACI [15] evaluates the average intensity difference between the plain and scrambled images. High UACI represents good performance of the scrambling algorithm.

$$\mathrm{UACI}\ \% = \frac{1}{MN}\left[\sum_{i=1}^{M}\sum_{j=1}^{N}\frac{I_P(i,j)-I_s(i,j)}{255}\right] \times 100 \tag{7}$$

B. Selection of Keys for two level scrambling

Scrambling and descrambling of blocks and pixels of images are effectively carried out with key1 and key2. The keys are number of iterations for Arnold scrambling that is based on spatial resolution of the image and coordinates assigned to the blocks. For different spatial resolution, number of iterations are listed out in Table 1. Keys for two level scrambling can be selected according to the number of iterations. For descrambling, number of iterations is given by

$$\begin{aligned} &Number\ of\ iterations\ for\ descrambling \\ &\quad = \{maximum\ number\ of\ iterations\ for\ the\ spatial\ resolution\ (2^n \times 2^n)\} \\ &\quad - Key\ selected\ for\ scrambling \end{aligned}$$

Table 1. Number of iterations for different sizes of an image

Size of Image	Iterations
4×4	3
8×8	6
16×16	12
32×32	24
64×64	48
128×128	96
256×256	192
512×512	384

C. Block-pixel scrambling

This method is accomplished by block scrambling followed by pixel scrambling. The reverse is done for descrambling. Block separation is a trivial task in which the spatial resolution of the blocks can be $2^n \times 2^n$. The spatial resolution of the blocks can be decided by the user and hence helps to improve the robustness of the method. Once the block separation is carried out with certain number of iterations (key1), all the blocks are arranged to get the block scrambled image. Arnold pixel scrambling is applied to this scrambled image with certain number of iteration (key2) to get the final scrambled image.

Table 2. Performance analysis of Block-pixel scrambling

	Boat	Barbara	Baboon	Airplane	Zelda
CC					
Block scrambling	0.957443	0.944454	0.743324	0.948021	0.976404
Pixel Scrambling	0.21381	0.14124	0.09179	0.07031	0.11767
NPCR					
Block scrambling	97.5399	97.826	97.81647	96.96465	97.74857
Pixel Scrambling	99.02649	99.4606	99.28589	98.54507	99.29657
UACI					
Block scrambling	19.6261	23.395	19.03114	18.91619	17.14758
Pixel Scrambling	19.52677	24.6594	18.94954	18.31416	18.26919

Descrambling is done with the reverse of the above mentioned steps and keys. Pixel descrambling with the certain number of iterations is applied to the scrambled image that delivers another scrambled image. From this scrambled image, one malicious user can guess some details about block scrambling that obstructs the robustness objective of this method. This is illustrated in Fig. 4. From Table 2, it is observed that block scrambling does not scramble the image completely. With the pixel scrambling step, robustness is improved.

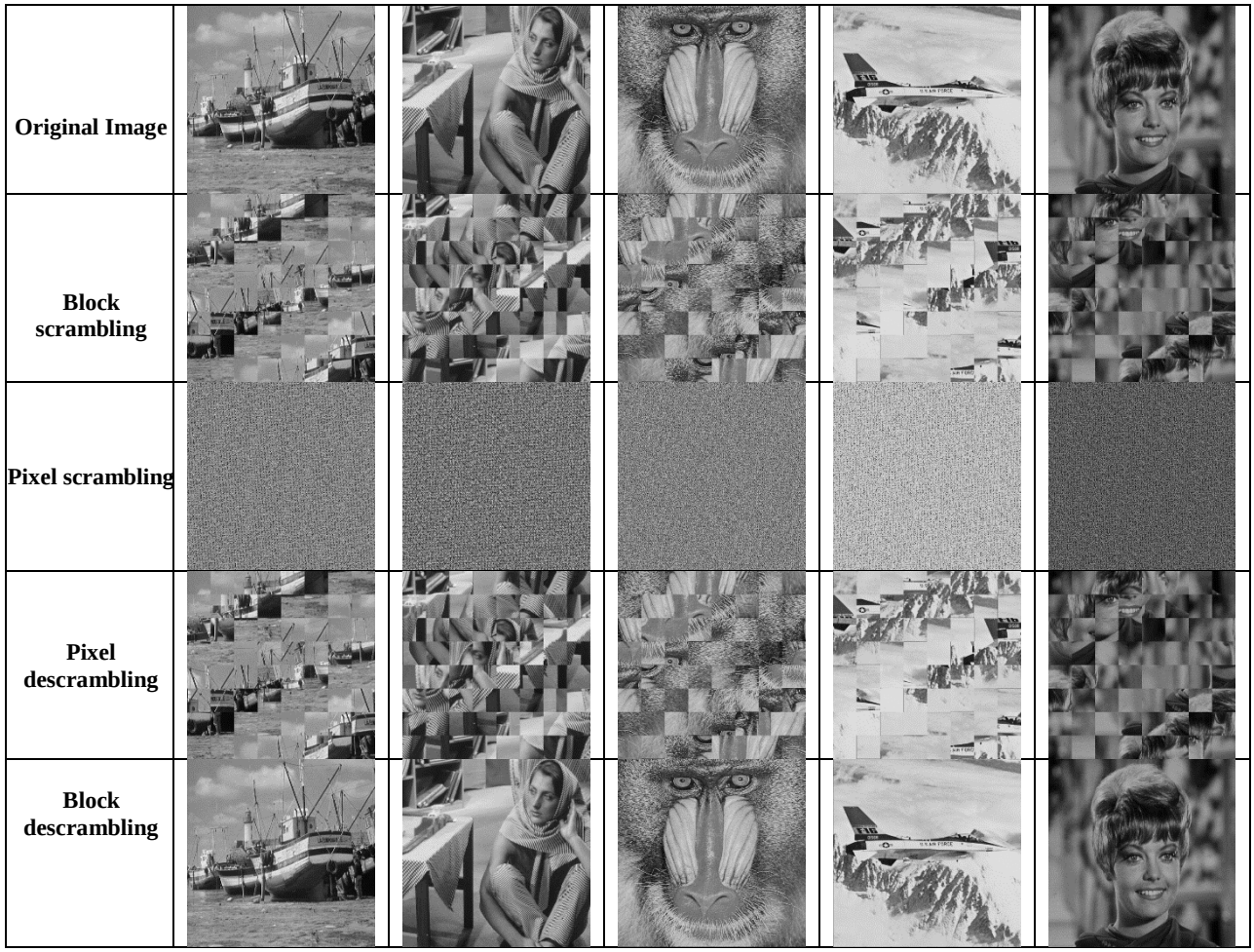


Fig. 4 Block-pixel scrambling plain and scrambled images

D. Pixel-block scrambling

This method deals with pixel scrambling followed by block scrambling with key1 and key2. Descrambling is done vice versa. In this method, block level descrambling should be carried out first with two factors, size of the blocks and number of iterations (key2). If both factors are correctly estimated then only the malicious user can do successful block level descrambling. Even after successful descrambling, the malicious user will get another scrambled image. Pixel level descrambling with correct key (key1) should be carried out to get back the original image. Compared to method1, the first descrambling step in this method does not reveal any information about the second descrambling step. This increases the robustness of this two level scrambling method. This is illustrated in Fig. 5. By analysing the descrambled images in Fig. 4 and Fig. 5, it can be observed that pixel scrambling followed by block scrambling provides more security to the plain images. From Table 3. It is revealed that the first pixel scrambling step itself, the plain image is scrambled. The scrambled image is again scrambled block wise. For malicious users, estimating the size of the blocks for descrambling is complex task. If this step is completed correctly, then only the pixel descrambling can be carried out to get the plain image.

Table 3. Performance analysis of pixel - Block scrambling

	Boat	Barbara	Baboon	Airplane	Zelda
CC					
Pixel Scrambling	0.13751	0.11479	0.047531	0.13927	0.04106
Block scrambling	0.21431	0.1397	0.091503	0.07029	0.11699

NPCR					
Pixel Scrambling	99.05396	99.46938	99.3309	98.57788	99.31679
Block scrambling	99.01886	99.43352	99.30725	98.53745	99.27635
UACI					
Pixel Scrambling	19.56608	24.66429	18.9602	18.35283	18.30821
Block scrambling	19.53834	24.67745	18.95002	18.32674	18.27697

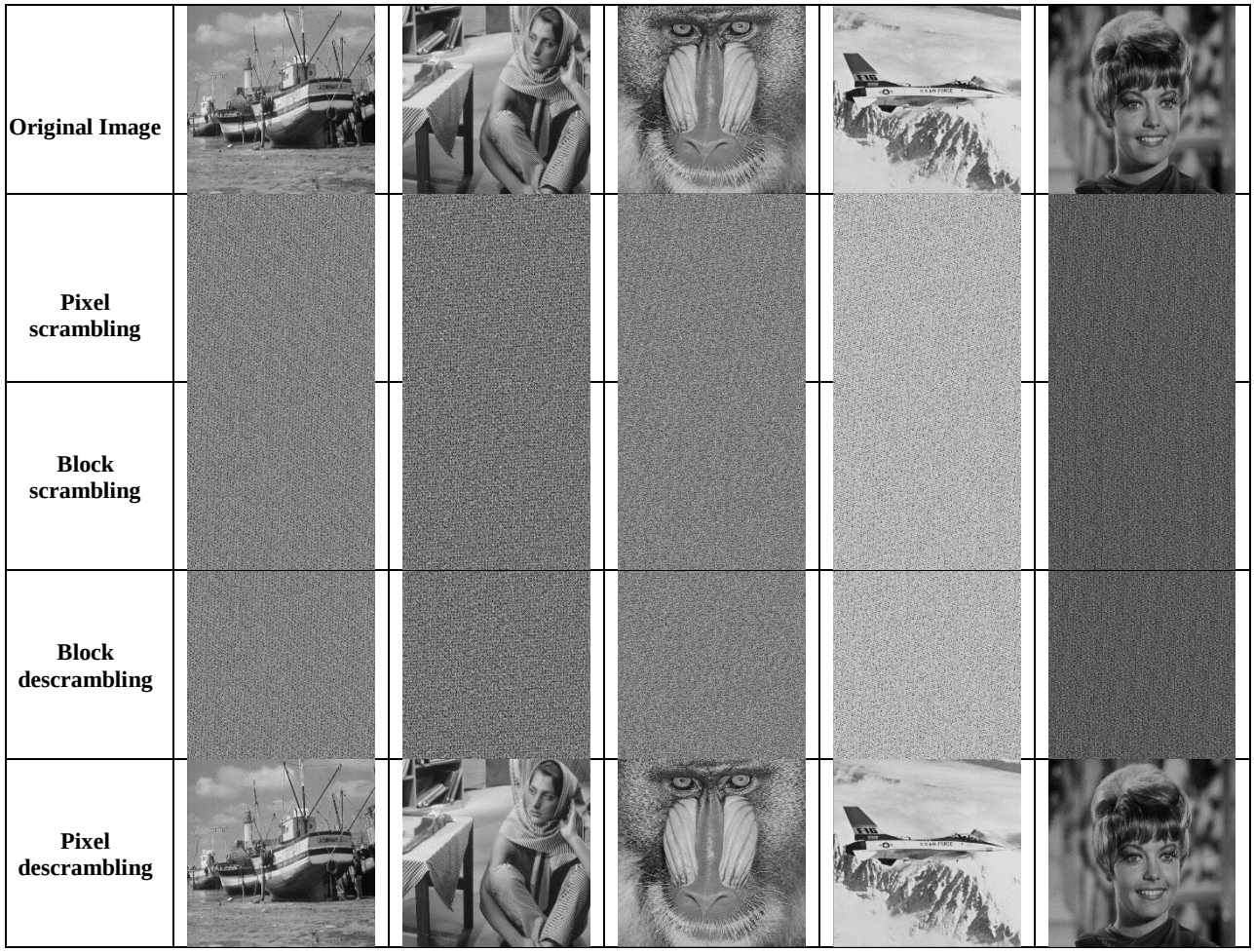


Fig. 5 Pixel-block scrambling plain and scrambled images

V. CONCLUSION

This work proposes a two level Arnold scrambling method which increases the security of scrambling technique. Conventional Arnold scrambling has one level of scrambling in which the original image can be obtained by applying descrambling iteratively. This paper introduces Arnold scrambling for pixels and blocks of the plain image. This pixel and block scrambling can be applied alternatively, but pixel scrambling followed by block scrambling delivers more robustness compared to block scrambling followed by pixel scrambling. This is substantiated by both subjective and objective analysis. Performance of this two level Arnold scrambling is also analysed by Correlation coefficient, number of pixel change rate and unified average changing intensity. This two level Arnold scrambling can be used an image scrambling step in image watermarking, image tamper detection & recovery and image authentication algorithms.

REFERENCES

[1]. Matsuo T and Kurosawa K (2004) On parallel hash functions based on block-ciphers, IEICE Trans. Fundam. Electron. Commun. Comput. Sci.,vol. E87-A, no. 1, pp. 67_74.

[2]. Li N, Du W, and Boneh A D (2005) Oblivious signature-based envelope, Distrib. Comput., vol. 17, no. 4, pp. 293_302, 2005.

[3]. Yeung M M and Mintzer F (1997) An invisible watermarking technique for image verification, in Proc. Int. Conf. Image Process., vol. 2. Oct. 1997,pp. 680_683

[4]. Jan Z et al (2017) A review on automated diagnosis of malaria parasite in microscopic blood smears images, Multimedia Tools Appl., pp. 1_26. Available: https://doi.org/10.1007/s11042-017-4495-2

[5]. Muhammad K, Sajjad M, Mehmood I, Rho S and Baik S W (2018) Image steganography using uncorrelated color space and its application for security of visual contents in online social networks, Future Generation Computer Systems, Volume 86. Pages 951-960

[6]. Abdulaziz shehab et al. (2018) Secure and robust watermarking scheme for medical images, IEEE Access, Vol. 6, pp. 10269 – 10278.

[7]. Min Li et al. (2013) Arnold Transform Based Image Scrambling Method, 3rd International Conference on Multimedia Technology（ICMT 2013), pp. 1309-1316.

[8]. Wu L, Zhang J, Deng W and He D (2009) Arnold transformation algorithm and anti-Arnold transformation algorithm. in Proc. 1st IEEE Int. Conf. Inf. Sci. Eng., Dec. 2009, pp. 1164_1167.

[9]. Vaiopoulos, A. D Developing Matlab scripts for image analysis and quality assessment. Proceedings of SPIE - The International Society for Optical Engineering 8181:6, DOI: 10.1117/12.897806

[10]. Siva Shankar S (2016) Data Hiding In Encrypted Images Using Arnold Transform. ICTACT Journal On Image And Video Processing, Volume: 07, Issue: 01.

[11]. Sukhwinder Kaur (2017) Arnold transform based Security Enhancement using Digital Image Watermarking with Complex Wavelet Transform. International Journal of Electronics Engineering Research. ISSN 0975-6450 Volume 9, Number 5 pp. 677-693.

[12]. Sujatha S S (2010) Feature Based Watermarking Algorithm by Adopting Arnold Transform. V. V Das, R. Vijaykumar et al. (Eds.): ICT 2010, CCIS 101, pp. 78–82, 2010.

[13]. Hung I. Hsiao and Junghsi Lee (2015) Fingerprint image cryptography based on multiple chaotic systems. Signal Processing, 113, pp. 169–181.

[14]. Tieyu Zhao, Qiwen Ran and Yingying Chi (2015) Image encryption based on nonlinear encryption system and public-key cryptography. Optics Communications, 338, pp. 64–72.

[15]. Saeed Bahrami and Majid Naderi (2012) Image encryption using a lightweight stream encryption algorithm. Advances in Multimedia, Article ID: 767364, pp. 1-8.

BRAINY WHEEL CHAIR FOR PHYSICALLY AND VISUALLY IMPAIRED

K Sukesh, Sujay P, S Tanmoy, Gagan Raj A

Department of CSE, Alliance College of Engineering and Design,
Chikkahagade Cross, Chandapura-Anekal Main Road, Anekal
Banglore-562106

***Abstract* -** The Brainy Wheelchair concept is devised to provide a smart electronic aid for the physically and visually impaired people. The system consists of ultrasonic sensors and GPS module, and the feedback is received through audio. Voice output is achieved through TTS (text to speech). The system is intended to provide object detection, and real time assistance via GPS, by making use of Raspberry Pi. The proposed system detects an object around the person and sends feedback in the form of speech and warning messages via earphones. It also provides navigation to specific location through GPS. The aim of the overall system is to provide a low cost, efficient navigation and obstacle detection aid for blind. This gives a sense of artificial vision by providing information about the environmental scenario of static and dynamic object around them, so that they can move independently.

***Index Terms* -** GPS Module, Obstacles, Raspberry Pi, TTS, Ultrasonic Sensor.

I. INTRODUCTION

Blindness or visual impairment is a condition that affects many people around the world. This condition leads to the loss of the valuable sense of vision. Visually impaired people suffer inconveniences in their daily and social life. Eyesight plays a major role in collecting most of the information from the real world and that information will be processed by the brain. Across the world, there are millions of people who are visually impaired, out of which many are blind. The need for assistive devices was and will be continuous. There is a wide range of navigation systems and tools existing for visually impaired individuals. India's population is currently at a whopping 133 crores, out of which, about 1.5 crore people are visually impaired, and 2.7 crore people are physically disabled.

II. OBJECTIVE

The main objective of this project is to provide a voice-based assistance to blind people. Here we have developed an intelligent system that helps a blind person to travel without the help of anyone and that works efficiently. The current navigation system for the visually impaired focus on travelling from one location to another. The device is used to help blind people to move with the same ease and confidence as a sighted people.

III. LITERATURE SURVEY

For the thorough development of the brainy wheel chair device for Blind Using Raspberry Pi, we need to go through each and every technical aspect related to it. This chapter provides an introduction to the area of research. A Brief Study and Survey has been carried out to understand various issues related to the project, which involves providing a smart electronic aid for blind people to provide artificial vision and object detection, real time assistance via GPS module by using Raspberry Pi .A survey is made among Blind people, who find it difficult to detect obstacles while moving in the street .The focus is on the visually impaired, who cannot walk independently in unfamiliar environment .The main aim of our project is to develop a system that helps the blind people to move independently. The Brainy wheel Chair for the Blind consist of three parts to help people

travel with a greater degree of psychological comfort and independence: sensing the immediate environment for obstacles and hazards, providing information to move left or right and orientation during travel.

- "Navigation Tool for Visually Challenged using Microcontroller", Sabarish.S.
- "Smart walking stick - an electronic approach to assist visually disabled persons", Mohammad Hazzaz Mahmud, Rana Saha, Sayemul Islam
- "Ultrasonic smart cane indicating a safe free path to blind people", Arun G. Gaikwad 1, H. K. Waghmare2 1ME Embedded system Design, MIT Aurangabad 2 Assistant Professor Department of E&TC, MIT Aurangabad
- "A Multidimensional Walking Aid for Visually Impaired Using Ultrasonic Sensors Network with Voice Guidance", Olakanmi O. Oladayo

IV.EXISTING SYSTEM

Blind people generally use either the typical white cane or a guide dog to travel. The white cane is a widely used mobility aid that helps blind people to navigate in their surroundings. The idea of designing and manufacturing ultrasonic sensor combines the properties of sound monition, which benefit the blind and a vibrating alert feature, which benefit the people that experience deafness. Although the wheelchair gives a warning about few meters before the obstacle, for a normal moving speed, the time to react is very short. A sensor is used detect obstacles within the designed range for a blind person to avoid them, through the issuance of distinctive sound or vibration. This can even be issued by the deaf, by putting their finger on the button at the top of the device, which vibrates when there is a risk. This system involves more manual work and it does not provide a good enough result. The existing system doesn't provide proper navigation and therefore, is not much effective.

V. PROPOSED SYSTEM

The proposed system consists of three main units:

- Ultrasonic Sensor unit.
- GPS Module unit.
- Espeak Text to Speech unit.

"Brainy Wheelchair for blind using Raspberry Pi" system is easy to understand and maintain. This system uses Raspberry Pi, which is a small processing device that works as a computer at relatively low cost. The system consists of ultrasonic sensors, GPS module and the feedback is received through audio. Voice output works through TTS (text to speech).

The proposed system detects an object around them and sends feedback in the form of speech, i.e. warning messages via earphones. It also provides navigation to specific location through GPS. The aim of the overall system is to provide a low cost, efficient navigation and obstacle detection aid for the blind which gives a sense of artificial vision by providing information about the environmental scenario of static and dynamic objects around them, so that they can move independently.

A. Ultrasonic Sensor

High frequency sound waves is generated by ultrasonic sensor. It evaluates the echo which is received back by the sensors. The time interval between sending the signal and receiving the echo is calculated by a sensor to determine the distance to an object. Ultrasonic is like an infrared where it will reflect on a surface in any shape, but ultrasonic has a better range detection compared to infrared. In robotic and automation industry, ultrasonic has been highly accepted because of its usage. In our Project the Ultrasonic sensor distance measurement Module deals with the distance measurement between the obstacle and the blind person. This module starts the process when the user turns on the device using power supply. Firstly, when the device turns on, the ultrasonic sensor will automatically give the distance measurement of the obstacle in front of the blind, and then the distance measured is stored in the SD card.

FIG 5.1 ULTRASONIC SENSOR

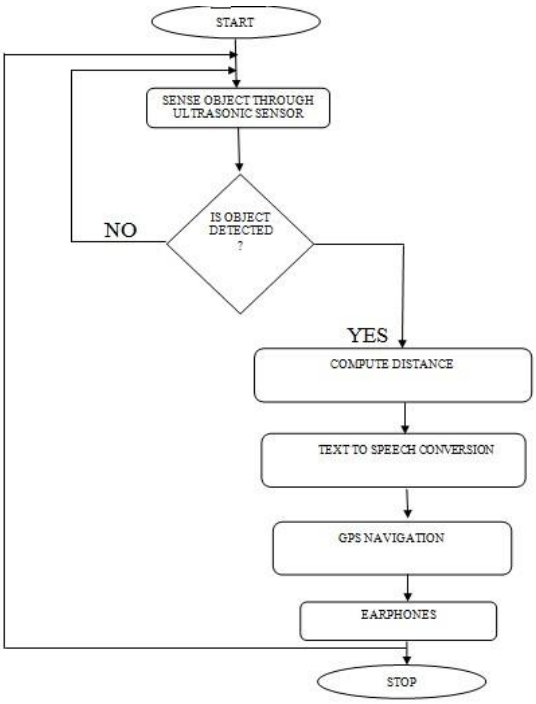


FIG 5.2 PROPOSED FLOW CHART

B. GPS Module

This module deals with the navigation of blind person from particular source to destination. This phase starts by Obstacle Detection. First the ultrasonic sensor gives voice command about the distance measurement between the obstacle and the blind person, based on that the navigation route instruction will be provided to blind by GPS Module via voice command. The navigation route is provided based on the latitude and longitude values. The latitude and longitude values will be stored so that when that value is matched the blind person gets the voice command to move left or right.

FIG 5.3 GPS MODULE

C. Voice Command Module

This module deals with giving the instructions to the blind user about the obstacles via Earphone. After detecting the Obstacles it gives the instructions about the obstacle and based on that GPS Module Provides route to the Blind. The function of a TTS (Text to Speech) is to convert the given text into spoken waveform. In order for us to give verbal instructions to the user, we need to convert our text instructions to audible speech.

D. Vibrating motor

A vibrating motor is essentially a motor that is improperly balanced. Our program triggers the vibrating motor when there is an obstacle in the way. As the obstacle gets closer, the intensity of the vibrating motor increases, thereby alerting the user.

HARDWARE REQUIREMENTS

The working of the system begins when the power supply is given. The ultrasonic sensor is then used to detect obstacle and provides distance between obstacle and the device. GPS Module provides navigation. When obstacle is detected, the distance and the navigation will be processed using Raspberry Pi device. The processing happens in such a way that if the obstacle is on to the right side, a voice command will be given to take left and vice versa.

VI. RESULTS

To evaluate the performance of the proposed method the experiments were conducted. The results in this paper shows the beginning of our efforts to build a compact travelling aid that allows the visually impaired to negotiate everyday environment. As previously mentioned, the sensor circuits give information about the environment. The circuit that has been designed for the object detection has provided an accuracy of 1 meter. For providing navigation GPS module has been used.

VII. CONCLUSION

The project “Brainy Wheelchair for the Physically and Visually Impaired” is designed to create a system using ultrasonic sensors, GPS module and providing voice command through headphone to the blind people. It would help a visually impaired person navigate through a public place independently. The proposed system tries to eliminate the faults in the previous system. The system takes measures to ensure their safety. It also aims to solve the problems faced by the blind people in their daily life. The design brainy wheel chair for Blind using ultrasonic sensors and GPS with voice output is of great benefit to blind people when it comes to independent mobility. The advantage of the system lies in the fact that it can prove to be a very low-cost solution to millions of blind people worldwide. The proposed combination of Ultrasonic Sensor and GPS makes a real-time system that monitors position of the user and provides feedback making navigation more safe and secure. We are using eSpeak text to speech conversion to provide voice command as output. Blind people can easily navigate from one place to another as our product provides voice messages. It is therefore capable of guiding a visually impaired person reach his/her destination.

ACKNOWLEDGMENT

We would like to thank H.O.D i/c of Computer science and Engineering Department Dr. Shekhar R**,** Alliance College of Engineering and Design for providing us an opportunity and his valuable support for her keen interest and boundless encouragement in preparing this work.

REFERENCES

[1] “Smart walking stick - an electronic approach to assist visually disabled persons”, Mohammad Hazzaz Mahmud, Rana Saha, Sayemul Islam.

[2] “A Multidimensional Walking Aid for Visually Impaired Using Ultrasonic Sensors Network with Voice Guidance”, Olakanmi O.

Oladayo Electrical and Electronic Engineering, Technology Drive, Office 6, New Faculty of Engineering Building, University of Ibadan, Ibadan, Nigeria.

[3] "Ultrasonic smart cane indicating a safe free path to blind people", Arun G. Gaikwad 1, H. K. Waghmare2 1ME Embedded system Design, MIT Aurangabad ,2 Assistant Professor Department of E&TC, MIT Aurangabad.

[4] "Smart chair for Blind: Obstac le Detection, Artificial vision and Real time assistance via GPS ", Shruti Dambhare M.E 3rd SEM (ESC) G.H.R.C.E. Nagpur, Prof. A. Sakhare M.Tech (ESC) G.H.R.C.E. Nagpur.

"Sensor assisted stick for the blind people", G. Prasanthi1 P. Tejaswitha 2 Professor, Dept. of MechanicalEngineering, JNTUA

College of Engineering, Ananthapuramu, A. P.PG Research Scholar,

Product Design, Dept. of Mechanical Engineering, JNTUA College of Engineering, Ananthapuramu, A. P.

[5] "Navigation Tool for Visually Challenged using Microcontroller

", Sabarish.S.

[6] "Smart walking stick for visually impaired", g. gayathri#1, m. vishnupriya r.nandhini3,ms.m.banupriya#4 1, Department Of ECE,

SNS College of Engineering, Coimbatore - 641107 4 Assistant Professor, Dept of ECE, SNS College of Engineering, Coimbatore – 107.

[7] "A Smart Pre-Warning, Guide, Alarm, Recovery a Detection (GUARD) Network System Electrical Engineering", Tamkang University, Tamsui, Taiwan 251, R.O.C.Department of Communication Engineering, National Central University, Jungli, Taiwan 320, R.O.C.Department of Computer Science and Information

Engineering, Tamkang University,Tamsui, Taiwan 251, R.O.C. for the Blind Ching-Chang Wong1, Yih-Guang Jan1, Yang-Han Lee1*, PoJen Chuang1.

[8] " An Intelligent Walking Stick for the Blind", Kher Chaitrali S., Dabhade Yogita A., Kadam Snehal K., Dhamdhere Swati D., Deshpande Aarti V. JSPM's Jayawantrao Sawant College of Engineering.

[9] "Smart Stick for Blind Man", Nitish Sukhija1, Shruti Taksali2, Mohit Jain3 and Rahul Kumawat4 1, 3 Student JECRC UDML COLLEGE of Engineering.

Quality of Service in Wireless Sensor Networks: A Review and Challenges

Seema.J.K[*], Dr. D.Ramesh [**]

[*] Research scholar, Sri Siddhartha Academy of Higher Education [SSAHE], Tumkur.
[**] Professor, Department of CSE, Sri Siddhartha Institute of Technology, Tumkur.

Abstract- Reliability plays an important role in Wireless Sensor Network (WSN). Nodes to sink data transmission reliability is important for monitoring the sensor field, where the sensors are deployed and sink to nodes data transmission reliability for acknowledgement and software configuration. Only one transmission protocol for reliable transmission cannot be adopted for all applications and scenarios of WSNs. In this paper we discuss existing works which implement data transmission protocols that ensure reliable transmission of information over the channel to improve Quality Of Service (QOS) of WSN. Existing works are compared to help the research scholar to choose proper technique based on WSN application type and its constraints.

Index Terms- Wireless sensor networks, Reliability, Quality of service, Protocol

I. INTRODUCTION

Wireless sensor network is internetwork of sensor nodes which communicate wirelessly to send the sensed information from the sensing field to base station through sink node, in turn the base station is connected to external network to reach end user[9]. Sensor nodes communicate in multihop manner [16].Various protocols are used for reliability of data transmission in WSN. Two types of data transport protocols used during erroneous channel condition for data transmission are

A. Automatic Repeat Request (ARQ) protocol
B. Forward Error Control (FEC) protocol

A. Automatic Repeat Request (ARQ) protocol

In this scheme when receiver finds any error in received data it sends a negative acknowledgement to transmitter so that it can retransmits the message and this continues until positive acknowledgement is received from receiver. ARQ-technique depends on the retransmission strategy to recover the original data packets [11].

B. Forward Error Control (FEC) protocol

FEC protocol depends on check bits added to the transmitted packet so that it can be corrected for errors at the receiver to recover packets without any error [13]. Two types of classification in FEC are block codes and convolution codes [12].

II section discusses different optimization metrics used to compare performance of different data transmission protocols.III section provides various energy models for power consumption metrics analysis in WSN. IV section provides the survey on existing work[19][20]. V section concludes the discussion.

II. OPTIMIZATION METRICS

Optimization metrics used to evaluate performance of various data transmission protocols are

1) Coding gain.
2) Throughput.
3) Bit error rate (BER).
4) Symbol error rate (SER).
5) Packet error rate (PER).
6) Latency.
7) Power consumption.
8) Decoding energy per bit.
9) Packet delivery rate.

Coding gain: For given BER, difference between the Signal to Noise Ratio (SNR) values of coded and uncoded system is coding gain [6].

Throughput: Throughput is the successful delivery of a message over a communication channel[4].

Bit error rate: The bit error rate is the number of bit errors divided by the total number of transferred bits during a studied time interval. BER is a performance measure of reliability and does not have any unit. It is expressed as a percentage [21]

Symbol error rate: The symbol error rate is the ratio of number of incorrect received symbols to the total number of symbols received. Symbol is declared incorrect if minimum one bit in symbol is erroneous [18].

Packet error rate: The packet error ratio is the ratio of number of incorrectly received data packets to the total number of received packets. A packet is considered incorrect if minimum one bit is erroneous [3].

Latency: Delay between transmission and reception of data [4].

Power consumption: Power consumption of the network [1].

Decoding energy per bit: Energy required for decoding a bit [5].

Packet delivery rate: Successful delivery of packets per unit time [2].

III. ENERY MODELS

Energy models are used in WSN for analysis of power consumption based on transmission strategies. Two types of transmission strategies are end to end strategy and node to node strategy.

A. *Energy model for end to end strategy*

Total energy consumed in network is given by

$$Etotal=Etx+Erx \tag{1}$$

Where *Etx* is energy consumed while transmitting the data, *Erx* is energy consumed while receiving data. Further

$$E_{total}=\sum_{i=1}^{m} NEtx(b) + \sum_{i=1}^{m} NErx(b) \tag{2}$$

Where N is number of bits. $Etx(b)$ and $Erx(b)$ are the power required to transmit and receive single bit respectively.

B. Energy model for node to node strategy

The energy required to transmit and receive single bit information is given by

$$E_b=E_{tx} + E_{rx} + E_{Dec} \quad (3)$$

Where E_{Dec} is energy spent on decoding single bit.Energy consumed while transmitting and receiving can be written as

$$E_{tx}=((P_{tx}+P_o)\ n/R+P_{st}T_{st})/k \quad (4)$$

$$E_{rx}=(P_{rx}(n/R)+P_{sr}T_{sr})/k \quad (5)$$

Where Po is transmit power, Ptx, Prx are power consumption in the transmit and receive circuitry. Pst, Psr are startup power consumption at transmitter and receiver respectively. Tst, Tsr are startup time in transmitter and receiver respectively, R is data rate, n is packet length and k is number of information bits.

IV. EXISTING WORK

A. Error correction codes in WSN: an energy aware approach

[15] Mohammad Rakibul Islam develops a framework for finding the suitable error control code for WSN. First bit error characteristics in sensor network is analyzed to determine whether their BER varies smoothly enough to be traced down. State machine is used to model the wireless channel, where each state corresponds to BER. Two parameters are analyzed; one is duration for which it exists in any of the states and BER variation between two states. Above two parameters decide usefulness of adoptability technique in FEC. Adoptability technique fails in two cases one is when BER rate is constant and the other case is when there is fast variation of BER over a time, then it is not possible to find suitable FEC because of the processing time required, therefore it hardly accomplishes any improvement. Graph of number of erroneous bytes per packet (NCBPP) distribution standard deviation distribution for 10 traces at Transmitter Receiver (TR) distance (6meters to 13meters) is analyzed. In this work set up made are 1) 4 hour traffic flow, 2) 3.2kbs speed, 3) transmission power is 90mW. NCBPP increases gradually with TR distance. Multipath interferences get strong when signal power becomes weak. Allan deviation (in bytes) is plotted versus time interval (in seconds) to show how fast the bit error rate changes. Frequency versus burst error length analysis graph for 1000 packets is plotted and it is observed that most of the bit errors are either single or double bit errors; presence of burst error is rare. Reed solomon codes(RS) with different error correcting capabilities like RS(15,11), RS(31,26), RS(31,21), RS(31,16) and RS(31,11) are simulated. Power consumption and BER analysis is made for all RS codes and RS(31,21) turns out to be a optimal choice.

B. Forward Error Correction in Sensor Networks

[7] Jaein Jeong, Cheng-TienEe, implemented and tested few versions of FEC protocols.WSN channel error characteristics are analyzed in this work using Mica Mote transmitter and receiver.analysis shows that most errors are single bit or double bit errors and burst errors are rarely present. This work implements three types of simple encoding scheme in indoor and outdoor environment which are less complex. Simple encoding schemes chosen for low power and small memory WSNs are Odd weight column code with 13 bit code word and 8 bit data (SECDED(13,8)),Odd weight column code wit 30 bit code word and 24 bit data (SECDED(30,24)) and quasi cycle code with 16 bit code word and 8 bit data(DECTED(16,8)).All codes reduce the packet error rate close to zero in outdoor environment where most errors are single bit or double bit. When most of the errors are burst errors in indoor environment,

the codes are not efficient in reducing packet losses, but packet loss is still lower than that of not using error control codes.SECDED (13,8) produces smallest packet drop rate among three. Flat form used is chipcon CC 1000 radio, in indoor and outdoor environment.

C. Experimental investigation of Reed Solomon error correction technique for wireless sensor networks:

[10] Cheng-Lai Cheah, Poh-Ling Tan, and Chee-Kit Ho proposed forward error correction code which reduces the PER for distances less than 40 meters. Experimental setup is done for investigation of error pattern for WSN (At transmitting end, the CC2520 IEEE 802.15.42.4Ghz RF transceiver is used to send 10000 random data packets, with 114 bytes length which is approximately equal to maximum packet length of IEEE 802.15.4 Wireless sensor network of 122 bytes. Receiver used is CC2520 2.4 Ghz RF transceiver) .Experimental investigation shows that 82% of error packets occur because of burst error and 18% are caused by random errors. RS code is used for burst error correction. PER is investigated before and after correcting errors. PER v/s distance graph before and after error correction is plotted. For the target of 10-3 packet error rate (acceptable by most of internet protocols (IP) applications, proposed work improves the distance by about 10m compared to the WSN without error correction. Packet error rate v/s received signal strength indicator (RSSI in –dBm) before and after error correction is plotted. For the target of 10-3 Packet error rate, proposed technique improves the RSSI by about 8DB compared to the wireless sensor network which does not uses error correction technique.

D. A hybrid adaptive coding and decoding scheme for multihop wireless sensor networks

[8] Imad EZ-zari,Mounir Ariona, Ahamed El Oualkadi, PascalLorenz, proposes an approach for reducing decoding power consumption and to increase the lifetime of the network and also to improve the reliability of the transmission in multihop sensor networks. Hybrid adaptive coding is implemented based on inter node distance. Strong low density parity check(LDPC) codes and RS codes are considered adaptively on the basis of channel conditions and inter node distance. If d<dcrossover Friss free space model is used. If d>dcrossover two ray ground reflection model is used. This work uses MTE routing algorithm. If d<dcrossover data is encoded by RS code else data is encoded by LDPC code. Energy consumption is less and network life time is improved. It is observed that performance for clustered wireless networks is not analyzed in this work.

E. Adaptive forward error correction(AFEC) for best effort wireless sensor networks

[22] KanyU, FilipBarac, mikaelGidlund and johan Akerberg proposes an adoptive FEC protocol scheme on the medium access control. RS codes with different error correcting capacity are employed..Markove model with M states is used with S(i) states. Transmission starts from state (0) that is the FEC code with low error correcting capability, S(M) is the state with high error correcting capability. In adoptive switching concept transmitter evaluates the channel conditions and makes the changes to coding according. If channel error conditions improve the receiver will switch to a low power FEC .If channel is distorted, the transmission will transit to a more powerful code. Switching technique is based on the number of acknowledgements (ACKS) received inside a window of L previously transmitted packets. Packet error rate window$(PER)_{WIN}$ within the window is found and it is compared with a defined switching threshold PER_T for determining whether to switch to a error correcting code with higher capacity or a error correcting code with lower error correcting capability. Adoptive algorithm does not require a dedicated feedback channel.

In AFEC code ranking is given for different codes, the highest code rate will give the lower overhead in to encoded packets, resulting in high capacity for information payload. Error correction capability of RS (n, k) code is given by

$$t= [(n-k)/2 \quad (6)$$

Where *n* is total number of symbols, *k* is total number of information symbols in a code word. AFEC can provide higher throughput and reduced power consumption compared to static FEC. Since AFEC cannot provide the Packet delivery rate as high as strongest static FEC scheme, it is suitable for packet loss tolerant WSN applications. It is observed that reliability of acknowledgement is not considered in this work.

In [15] step by step approach for finding suitable data transmission protocol is provided. In[7][10]emphasis is given for error pattern rather than power consumption, computational requirement and storage space. Hybrid adoptive coding[8] is implemented by considering inter node distance. Channel conditions are estimated for adaptive FEC [22] based on acknowledgement received.

Table 1.Provides the comparison of reviewed works and their suitability for particular WSN application.

Table 1: Comparison Table

Author	Network Topology	Type of protocol	Application suitability
Mohammad Rakibul Islam	Multihop	FEC,RS CODE RS(15,11), RS(31,26), RS(31,21), RS(31,11), RS(31,16)	WSNs which depend on BER of channel
Jaein Jeong	Point to point	FEC	WSNs with single and double bit error pattern
Cheng-Lai Cheah	Point to point	FEC,RS code	WSNs with bursty channel and which are not tolerant to PER over short distance.
Imad EZ-zari	Multihop	FEC,RS,LDPC	WSNs with more life time and reliability in transmission
KanyU	Point to point	FEC,RS CODE,RS(15,5),Rs(15,7) Rs(15,9),Rs(15,11),Rs(15,13)	Packet loss tolerant WSNs

V. CONCLUSION

This paper provides review on different protocols for reliable transmission of data in WSN during bad channel conditions. Different existing works are discussed to improve reliability of transmission.WSN and their suitability for different applications and WSN constraints are analyzed. In existing works QOS is improved by means of reliable data transmission protocols. Future challenges involved to improve QOS of WSN are 1) resource limitations like power required for transmission, Memory size.2) Data redundancy that is, similar event is sensed by the nodes.3) Dynamically changing network like node failure, link failure and topology.

ACKNOWLEDGMENT

The authors wish to thank the anonymous reviewers for their very helpful suggestions and comments.

REFERENCES

[1] Mohamed Elshrkawey, Samiha M. Elsherif , M. Elsayed Wahed Suez," An enhancement approach for reducing the energy Consumption in wireless sensor networks", Egypt. Journal of King Saud University –Computer and Information Sciences 30 (2018), 259–267

[2] Dr. C. Suresh Gnana Dhas2 , Dr.V.Ganesa R.Mahaveerakannan1,"Exploration on increasing packet delivery rate in wsn using cluster approach" EAI Endorsed Transactions on Energy Web and Information Technologies 07 2018 - 09 2018 ,Volume 5 , Issue 20 , e1

[3] Chunliang Zhou,1 Ming Wang,2 Weiqing Qu,1 and ZhengqiuLu11Ningbo Dahongying University, Ningbo, "Research articlea wireless sensor network model considering energy consumption balance", Published 3 June 2018Academic Editor: Emilio Insfran Pelozo,Copyright © 2018 Chunliang Zhou et al.

[4] Afandi Ahmad, Muhammad Faris Roslan, and Abbes Amira ," Throughput, latency and cost comparisons of microcontroller-based implementations of wireless sensor network (WSN) in high jump sports" AIP Conference Proceedings 1883, 020010 (2017); doi: 10.1063/1.5002028 Published by the American Institute of Physics

[5] Imad Ez-zazi_, Mounir Arioua, Ahmed El Oualkadi" On the design of coding framework for energy efficient and reliable multi-hop sensor networks" The 8th International Conference on Ambient Systems, Networks and Technologies (ANT 2017) Laboratory of Information and Communication Technologies (LabTIC), National School of Applied Sciences, Abd Procedia Computer Science 109C (2017) 537–544

[6] Amit Singh andAitha Nagaraju,"Article An Artificial Bee Colony-Based COPE Framework for Wireless Sensor Network "Published: 6-May2016,Computers-2016,5,8; doi:10.3390/computers5020008 www.mdpi.com/journal/computers

[7] Jaein Jeong, Cheng-Tien Ee,"Forward Error Correction in Sensor Networks" EECS Department, University of California, Berkeley, California 94720, USA.

[8] Imad EZ-zari,Mounir Ariona,Ahamed El Oualkadi,Pascal Lorenz,"A hybrid adaptive coding and Decoding Scheme for Multihop Wireless Networks",Wireless Pers Commun,Springer 24 September 2016.

[9] Abdelhak, Nabia,Blaise Omer ”Concepts and evolution of research in the field of wireless sensor networks” International journal of computer and communication(IJCNC) Vol.7,No.1,January 2015.

[10] Cheng-Lai Cheah, Poh-Ling Tan, and Chee-Kit Ho," Experimental Investigation of Reed-Solomon Error Correction Technique for Wireless Sensor Network ", funded by Malaysia’s Ministry of Higher Education (MOHE) Exploratory Research Grant Scheme (ERGS) International Journal of Information and Electronics Engineering, Vol. 4, No. 2, March 2014.

[11] Nabil Ali Alrajeh, Umair Marwat, Bilal Shams and Syed Saddam Hussain Shah "Error Correcting Codes in Wireless Sensor Networks: An Energy Perspective"An International Journal Appl. Math. Inf,Sci. 9, No. 2, 809- 818 ,2015.

[12] Kheirandish, D.; Safari, A.; Yinan Kong, "A novel approach for improving error detection and correction in WSN," in Electrical and Computer Engineering (CCECE), 2014 IEEE 27th Canadian Conference on , vol., no., pp.1-4, 4-7 May 2014 .

[13] Solomon Error Correction Technique for Wireless Sensor Network ", funded by Malaysia’s Ministry of Higher Education (MOHE) Exploratory Research Grant Scheme (ERGS) International Journal of Information and Electronics Engineering, Vol. 4, No. 2, March 2014.

[14] Vibhav Kumar Sachan, Syed Akhtar Imam, M. T. Beg "Energy efficient Communication Methods in Wireless Sensor Networks: A Critical Review" International Journal of Computer Applications (0975 – 8887) Volume 39– No.17, February 2012.

[15] Mohammad Rakibul Islam "Error Correction Codes in Wireless Sensor Network: An Energy Aware approach" World Academy of Science, Engineering and Technology International Journal of Electrical, Computer, Energetic, Electronic and Communication Engineering Vol: 4, No:1, 2010.

[16] Chiara Buratti , Andrea Conti , Davide Dardari , Roberto Verdone "An Overview on Wireless Sensor Networks Technology and Evolution" open access, Sensors, 9, 6869-6896; doi:10.3390/s90906869,2009. [17].

[17] Mehmet C. Vuran, Member, IEEE, and Ian F. Akyildiz, Fellow, IEEE" Error Control in Wireless Sensor Networks: A Cross Layer Analysis " IEEE/ACM transactions on networking, vol. 17, no. 4, august 2009.

[18] Shaoqing Wang and Jingnan Nie "Energy Efficiency Optimization of Cooperative Communication in Wireless Sensor Networks", Hindawi Publishing Corporation EURASIP Journal on Wireless Communications and Networking, Article ID 162326, 8 pages doi:10.1155/2010/162326, Volume 2010.

[19] A.Angelin, B. Revathi, T.Gayathri, Mr.D.Balakumaran " Channel Coding in WSN for Energy Optimization " International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering(An ISO 3297: 2007 Certified Organization) Vol. 3, Issue 3, March 2014.

[20] Alireza Masoum,Nirvana Meratnia,Arta Dilo,Zahra Taghikhaki,Paul J. M. Havinga, "Cross-Layer Analyses of QoS Parameters in Wireless Sensor Networks" International Conference on Computer Science and Information Technology.CCSIT 2011: Advances in Networks and Communications pp 595-605.Springer

[21] M.R.Ebenezar, jebarani, and T.Jayanthy. "An analysis of various parameters in wireless sensor networks using adaptive FEC technique" International Journal of Ad hoc, Sensor & Ubiquitous Computing (IJASUC),Vol.1, No.3, September 2010 .DOI : 10.5121/ijasuc.2010.1303 33

[22] Kan yU, Filip Barac, mikael Gidlund and johan akerberg "Adaptive forward error correction for best effort wireless sensor network" ICC communications, IEEE international conference on communication, June 2012.J. Wang, “Fundamentals of erbium-doped fiber amplifiers arrays (Periodical style—Submitted for publication),” *IEEE J. Quantum Electron.*, submitted for publication.

An AI driven approach for Smart refrigerator to enhance family diet and sustainability

Naveen Ananda Kumar J*, Shivani Suresh, Kannika Nirai Vaani****

*Institute of management,
Christ (Deemed to be University), Bengaluru
**Institute of management,
Christ (Deemed to be University), Bengaluru

Abstract- With smart homes changing the way of life of people, a smart refrigerator is an addition to their family. The fast pace of our life has led to an alarming consumption of junk food, expired products or vegetables at homes. In this paper, we are introducing a novel concept of a smart refrigerator which is content aware and focuses on personalization for each of its users, which, in the long-term, would help to tackle the problem of effective waste management. It is an AI-driven solution, aiming at resolving the problems faced by the current generation and also has various other salient features like health data analysis, framing nutrient plans, recipes, dishes, etc. Various sensors like cameras, bar code scanners and weight gauge help to calculate the nutrients present. AI and ML continuously evolve, adapt and help to better understand the users' tastes and preferences. Analysis of the consumption pattern will lead to the generation of what all items are to be bought and in what quantity to be stored in the refrigerator. This would lead to healthier homes and a healthier planet.

Index Terms- Artificial Intelligence (AI), barcode scanners, cameras, machine learning (ML), smart refrigerator

I. INTRODUCTION

The advances in AI and ML powered by Cloud-based computing has revolutionized the home appliance industry, whether it is a good old coffee maker, or a virtual assist. There is no segment that has been unexplored. Integration of technology with homes has made smart living possible and is taking us towards the future. With the rising advancements, demands and expectations are also rising at a tremendous pace. Consumers are becoming more tech-savvy and want the next big thing in the market for their homes.

There have been many efforts in developing a smart refrigerator in the industry. A smart refrigerator that has both personalization and sustainability-driven holds its own USP. The traditional notion of a refrigerator has changed, from a storage unit, to keep food cold, to a device that integrates with the internet and the cloud, giving it capabilities to develop data and applications, focusing on a healthier lifestyle.

In this project, we aim at developing a smart refrigerator that driven by both, personalization and sustainability. Interweaving itself into this complex system of devices talking to each other, making routine tasks simpler [1]. Using these intelligent devices in the kitchen, which is the heart of every home, would help to bring in a range of products and appliances to tackle the various problems faced by homemakers. With weighing scales and image sensors, it would be easy to identify which product is going to consumed the most and at the earliest. Most of the existing smart refrigerators do not provide a solution for these problems. Smart refrigerators are the next evolutionary step in home appliances. The refrigerator would have sensors that would monitor the products kept inside, and accordingly would pan out recipes which can be viewed with a feed of the demo, by linking it to the

internet for real-time content. Also, with the help of smart wearable technology, access to health related data of the people living in the house would be possible. This way, the refrigerators can be in constant touch with their users, letting them know if any product has to be bought, or if a product was about to get over and has to be replenished.

There has been a drastic change in people's lifestyle, which is a result of the developments taking place in the technological sector. This lifestyle has reduced the level exercise and lead to the consumption of unhealthy food [2]. The eating pattern of an individual can cause diet related illnesses during later stages of our lives [3]. Changing lifestyle has led to an unhealthy dietary pattern and the lack of exercise has given rise to obesity, which is becoming a major health hazard. From this perspective, the smart refrigerator would be focusing on the nutrition and health habits of its users. Refrigerators are a part of the majority of households. As results play an important role in the user's consumption pattern, smart refrigerators will make it ideally suitable to tackle the problem of family healthcare.

II. PRESENT SCENARIO AND RESEARCH GAP

A. Present Scenario

The smart refrigerator industry has seen a lot of changes in the recent times, with some of the market leaders in the e-commerce sector venturing into the industry as well. Some of the popular ones area Kenmore Smart sold on Amazon, which is the best match for Alexa lovers, Samsung RF28JBE, LG InstaView Door-to-Door, all unique in their own ways. However, all of them have various shortfalls, mainly:

- Most of these companies don't give a guarantee on software updation which eventually leads on to issues, such as unable to connect to the server, making it vulnerable and unintelligent in due course. Considering refrigerators are part of a household for a long period of time, this is a major issue. Once the refrigerator becomes vulnerable it is prone to DDoS attacks which can leak sensitive information of the users.
- The price of these refrigerators is generally high, making buying a high end refrigerator with any of these features a sweeter deal. These refrigerators would last longer and also don't have a software updation issue.
- There isn't a standardised system to record bar code information, for instance, the expiry, or the quantity of the product.
- There aren't any functions helping with nutrient and dietary control, which would pose to be a very attractive feature in any refrigerator.

B. Research Gap

The global smart appliance market is believed to grow at a CAGR of 23.48% during 2016-2020 [4]. Businesses are starting to intensely focus of this sector, trying to come up with numerous advancements and digital technologies that can be integrated into the daily lives of their customers. The present scenario gives insight into the fact that the existing refrigerators are made with advanced technology and are more concentrated on the combination of intelligent sensor [5][6] networks and information. However, most companies overlook the fact that they can deliver more to their customers by integrating personal factors like their health, building a stronger brand loyalty. This segment has a lot of potential and further exploration would lead to more opportunities to bring the development of the same.

III. PROPOSED MODEL

A. Proposed Model

The project is focused around developing a smart refrigerator which enables their customers to enhance their diet, including more nutritious content and endorsing a healthier dietary plan. The system is built around cloud computing and AI, eventually branching out into ML. The system has a database of its users and a locally built database which has nutrient information [7]. This goes in accordance to the user's tastes and preferences and helps to achieve the maximum level of satisfaction from the customer's point of view. The database contains information about the user's age, weight, height, medical record, allergies etc [8]. The in-built sensors scan the bar code and also scan an image of the product every time it passes the door. This way, the refrigerator is able to record what food item it is and with the help of the weighing scales it will be able to have a clearer idea of how much of the product is consumed, enabling it to identify the consumption pattern. This would let it generate grocery lists and enable it to find recipes and dietary plans which fit the same [9]. Also, it would be able to let the user know when a particular product is about to expire. Notifying them regarding the same when they are near a store or when they go shopping would make their life easier. The screen on the door of the refrigerator displays all this information and gives out warning signals to know if the user still wants the product or not.

B. USP of the model

The USP of our prototype is the cameras, bar code scanners and weight gauge which helps to calculate nutrients, giving a clear picture of the contents present inside the refrigerator. By integrating ML and AI, we are able to better understand the customer's tastes and preferences. It analyses the consumption pattern and notifies the user not only what to buy, but also how much to buy. The cloud-based system tracks the nutrition intake and identifies which nutrients are lacking in the user's diet, which is based on the intake. It also develops customised recipes based on the number of members in the house to improve taste and reduce waste. It calculates the BMI of its users when the information is fed in, which leads to more efficient nutrient suggestions.

C. System Design

The smart refrigerator application requires the usage of Python for its implementation. The users interact with the interface of the smart refrigerator. The database design is vital for the system design. Below is the entity-relation diagram of the smart refrigerator. In order to explain the working of our model we have devised an architectural design of the entire process subdivided into 3 parts. The first is wearables and smart phones, next by the smart refrigerator and the last being cloud. The Fig 1.1 explains the entire architectural framework.

C.1. Smart wearables and smart phones

Smart wearables record the day to day activities such as exercise, physical fitness, sleep cycle, heart rate, calories burnt etc. This health data is gathered and sent to the users' smart phones. Smartphones nowadays have in-built health applications which are equipped with a Medical ID. This Medial ID is unique to its user, containing all the information generated from these smart wearables and also has its own input on the same, for instance, the previous medical records which include user's BMI, allergies, medical condition and other issues. When the smart phone is connected to the smart refrigerator it can transfer this data to the refrigerator.

C.2. Smart refrigerator

Each smart refrigerator is assigned with a unique ID and has in-built cameras which have a 360° view which recognise what comes in and goes out of the refrigerator. There is also a bar code scanner which scans the codes of packed goodss. The weight gauge measures the weight distribution and the

consumption of the what is kept in the refrigerator. Combining these two helps to track the consumption pattern of the users, leading to higher accuracy in predictability of dishes and generation of grocery list. The data obtained from the smart refrigerator is sent to the cloud.

C.3. Cloud

Cloud serves two important functions, the first one being, it has a database of all the recipes and the unique smart refrigerator ID. The second is that it has the AI and ML based algorithms. Cloud has database of recipes which is stored on it. It further segregates data with reference to certain keywords and hashtags. For example, if the users use certain hashtags like #spicy, #south Indian, etc., the system will match these words with recipes, which would enable the users to relive the same.

Once this data is sent, the customer is able to view customised dishes, which is made available with the help of cloud. These dishes are orientated around the user's health condition. With the help of these devices, it will be capable to monitor the customer's health more accurately. If the user is an active gym user, then it would also pick out those dishes which would complement their work-out pattern.

A separate database is maintained in cloud for each smart refrigerator based on its unique ID. The user's information such as the medical records and the inventory present in the refrigerator are sent to the cloud and based on this, coupled with his/her preferences, it will select the best recipe.

D. Working Process

A separate database is maintained in cloud for each smart refrigerator based on its unique ID. In Fig 1.2 the smart refrigerator with the help of the in-built camera, would scan what comes inside the refrigerator. For instance, if it is a fruit, vegetable or dairy product, it would classify this on the bases of certain predetermined criteria in case it is a packed product and has a bar code. Then, the bar code scanner present would scan the code, getting all information regarding what item it is, when it would expire and lastly, the weight gauge would let us know what all is being consumed and how much of it is being consumed. All this information is derived from the smart refrigerator, and combined with the information with the data obtained from the smart phone which tells us regarding the medical records, sleeping patterns etc. In the end, the final information that is sent to the processing unit is the user's medical records. The inventory present in the refrigerator are sent to the cloud. Based on this and his preferences, it will select the best recipe.

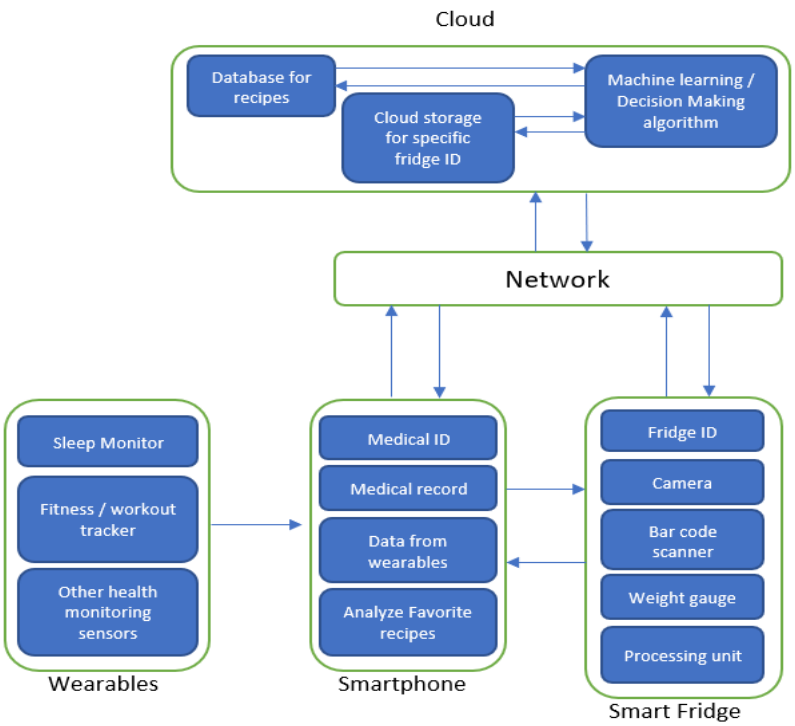


Fig 1.1 System design and Components of smart refrigerator

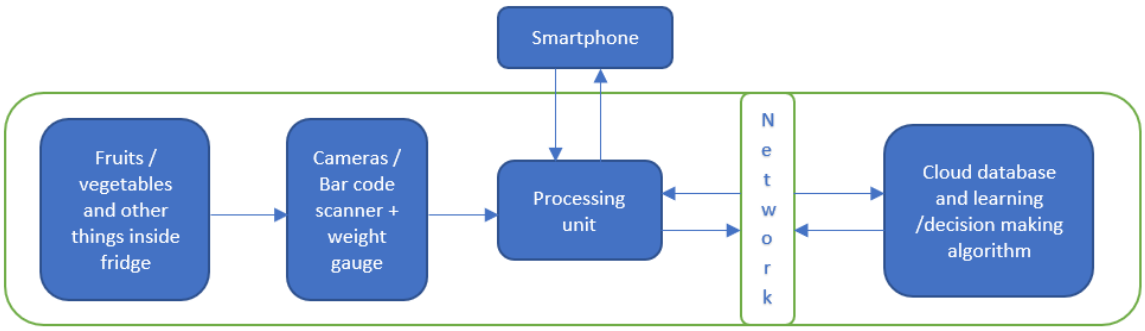


Fig1.2 Working process of Smart Refrigerator

The processing unit, where with the help of the cloud database, which contains millions of recipes, an algorithm is created after careful assessment of keywords. A proper match is made with the database, helping to achieve a personalised experience for the user. It gives recipes based on what is available inside the refrigerator which it is able to perceive with the help of the consumption pattern. This also enables it to give out suggestions regarding what is to be bought. The ML mechanism present would help the machine to predict on its own in the future, without requiring access to all this information and after rectifying past errors. If the customers refrigerator has meat stored in it previously, then it would be able to detect that the user is a non-vegetarian and would recommend in accordance to that. Based on the previous consumption pattern of the user, the machine would be able to detect more accurately what to recommend its customer. The consumption pattern of the user would also enable the system to tell the users which all products are to be bought and in what quantity. This would help in waste management as the system would know which all products are not being consumed and is going to be wasted. So, by avoiding purchase of those products, or reducing its quantity, it will be help to minimize the waste. The final recipe or ingredients required would be displayed on the screen in the refrigerator, or on the user's phone, which is even more convenient to the user. The user would

be notified if a product has expired, when it should be removed from the smart refrigerator and when it should be bought to replace it as well.

E. Sustainability

The smart refrigerator would continuously analyse and monitor consumption, due to which it becomes easy to predict the future consumption. This estimation would enable in reducing the waste emission from households as people would be aware what quantity is apt for the user and buy goods in accordance to that. This would lead to waste minimisation and installing these refrigerators in an apartment complex, proceeding into a locality, which would help in waste minimisation at a greater level. Its further expansion can help to reach this at a global scale.

F. Example

In order to comprehend the application of this smart refrigerator, we were given a problem along with how we would tackle this issue.

Problem **–**Jenny who is a 40-year old female who has early onset of Type 2 diabetes. With her hectic lifestyle she was barely able to manage her health and the lack of unawareness had made Google her virtual dietician. As most of the information was not personalised it left her dissatisfied. For instance, being South Indian, she preferred South Indian cuisine, however the options made available to her are mostly bland and tasteless. And more often, she found that she did not have the ingredients in the refrigerator for the dishes she liked. The unreliability of the information of these websites was extremely high. She was persistently looking for a way out of this.

Solution **-** With the help of our concept we believe we would be able to address her problem. The first step would be accessing her medical data from her smartphone. With most of smartphones having Medical ID and day to day activity through her fitness tracker, it would be easier to get information that is required to assess her medical condition and recommend dishes to her in accordance to that. Based on the contents inside the smart refrigerator we would be able to analyse and recommend dishes which would be customised to her taste. Giving preference mainly to the dishes which she would be able to cook based on what is available in the refrigerator.

IV. FUTURE WORK

The future is smart living and it begins at home. The smart refrigerator sector is growing rapidly and we believe our product has tremendous potential and, in the future, it can be integrated with super markets to supply the products at the user's door step. It can predict the demand for products and automatically notify the user enabling him to place the order. The proposed model is still in an infant stage and therefore, still has scope of improvement. With access to the user's medical record and their consumption pattern it would be possible to develop nutrition chart for their users. Upon following this chart by the user, the system would recommend the dishes and workout plans they can use to make it more effective as well. The ultimate aim is to achieve a smart home which is driven towards making our lives easier, and at in the long term not harm our planet.

Energy conservation is becoming a vital part of sustainability, so in future scope we can include an energy conservation system. For this, we would initially connect the smart refrigerator with a smart energy meter which receives load scheduling data from the smart grid. This would help us to monitor and analyse the peak hours enabling us to conserve the energy by automatically managing the temperature inside the smart refrigerator. This can reduce power consumption of household appliances, helping to cut down on expenses and also reducing the carbon footprints as well.

V. CONCLUSION

This novel concept which is an integration of AI, ML and cloud-based computing. It has a bright future and is bound to branch out into various other dimensions of kitchen appliances, especially with smart homes becoming more common. The refrigerator would enable the users to have a better nutrient consumption, in turn improving their health. The design is in a manner wherein the items stored are used for advising the users dishes that can be made by guiding them step by step. It can also track the consumption pattern of the customers and do a health data analysis based on this.

ACKNOWLEDGMENT

We thank our colleagues from Institute of management, Christ (Deemed to be) University and Alliance University who provided us with insights which was a great assistance in shaping our research. We would like to express our gratitude to Prof. Dr. Lakshmi Shankar Iyer, Head of Specialization - Business Analytics, Institute of Management and Dr. Manohar Kapse, Assistant Professor, Faculty- Business Analytics, Institute of Management, for sharing their pearls of wisdom, reviewing our research and giving us key inputs.

REFERENCES

[1] Desai Professor SAL, S. (2016). Understanding IoT Management for Smart Refrigerator, (January), 2–5.

[2] C. E. Woteki and P. R. Thomas, Eat For Life: The Food and Nutrition Board's Guide to Reducing Your Risk of Chronic Disease Washington, D.C. National Academies Press, 1992.

[3] N. King, Smart Home-What do consumers want?, Intertek Research&Test Centre, 2003.

[4] N. M. White, et al., Sensors in Adaptronics, Springer Berlin Heidelberg, ISBN, 978-3-540-71965-6, 2007.

[5] J. Li, , Y. Guo, , and G. Poulton, , Critical Damage Reporting in Intelligent Sensor Networks, Proceedings, 17th Australian Joint Conference on Artificial Intelligence – AI 2004: Advances in Artificial Intelligence, 6th - 10th December 2004, Cairns, Australia, vol. 3339, pp. 26-38, 2004.

[6] G, M. N., M, E., & S, A. (2017). IoT Based Interactive Smart Refrigerator. Conference: 3rd International Conference on Computers and Management (ICCM 2017) At: Jaipur, India, (December 2017).

[7] Mahajan, M. P., Nikam, R. R., Patil, V. P., & Dond, R. D. (2017). Smart Refrigerator Using IOT. International Journal of Latest Engineering Research and Applications (IJLERA), 2(3), 86–91.

[8] Singh, D., & Jain, P. (2016). IoT BASED SMART REFRIGERATOR SYSTEM. International Journal of Advanced Research in Electronics and Communication Engineering (IJARECE), 5(7), 2080–2084.

[9] Ghayvat, H., Mukhopadhyay, S., Gui, X., & Suryadevara, N. (2015). WSN- and IOT-based smart homes and their extension to smart buildings. Sensors (Switzerland), 15(5), 10350–10379. http://doi.org/10.3390/s150510350

Drone based Solid Waste Detection using Deep Learning & Image Processing

Anjali Pradipbhai Anadkat , B V Monisha, Manasa Puthineedi, Ankit Kumar Patnaik, Dr. Shekhar R, Riyaz Syed*

Computer Science and Engineering , Aerospace Engineering*,Alliance University, Bangalore, India

***Abstract*—** A significant cause of waste generation is littering. Whether intentional or unintentional, littering has an adverse impact on the environment and it also costs municipalities millions of dollars annually to clean up the waste. Traditional systems for waste management include regular waste collection by assigned groups such as the municipal corporation trucks for cities, janitors for an institution etc. Due to the shortcomings of existing waste monitoring and management systems, this paper proposes an alternative and efficient approach for waste monitoring in large. This paper presents a smart solid waste detection system which uses the motion of an Unmanned Aerial Vehicle (UAV) such as drone for waste detection based on image processing techniques which help identify places contaminated with waste and sends the location coordinates as a notification. This system uses the concepts of deep learning, image stitching and image processing through feature extraction. In the last section, we will discuss the future applications of this smart system which will tackle both garbage detection and collection. Finally, the future scope and relevance of this system will be discussed.

Keywords—*solid waste detection, waste management, UAV, drone, image processing, litter, deep learning*

I. INTRODUCTION

Solid waste management has been a major concern for a long time. India produces 42.0 million tons of Municipal Solid Waste (MSW) annually at present. According to a recent consensus, the per capita waste generation in India is increasing at a rate of 1.3% per annum. Garbage generated annually across the globe amounts to billions of tons today, with almost one lakh metric tons of garbage generated in India per day. This can cause a major impact on the environment. With urbanization in India, the scale of cities and the number of residents is rising and with it is increasing the amount of urban waste generated [1][2][3].

Often the presence of litter in a given area results in the intentional throwing of litter at that particular spot as it gives the impression that it is a right place for discarding the waste. Throwing kitchen generated waste out of the window or chucking chewing gum wrappers along the roadside for instance, without realizing the harmful effects it might have on the environment have become a commonplace. It costs residents, local, state and federal governments millions of dollars to pick up litter, reverse the effects of littering and prevent it. Municipal solid waste (MSW) management includes collection, storage, transportation and disposal of solid waste. Poor collection and inadequate transportation leads to heap of MSW at many places, which causes health and environmental problems. Governments across the world are making efforts to improve solid waste management in their respective countries. [4] The accumulated waste causes environmental problems such as the spread of pest species and diseases, marine plastic pollution, damage to wildlife ecosystems and toxicity in the land/waterways.

While most urban waste management systems in place are designed for regular collection of waste from fixed locations at regular intervals, there are certain places (remotely located) which go unnoticed and are therefore, not cleaned. Many solutions have been designed for effective garbage management. There are applications available on the Google Play Store which allow the user to send an alert message to the municipal corporation about the status of the garbage container. However, this process is time-consuming as it relies on human discretion and action and communication via an intermediary. A better solution designed for this problem is to use sensors for obtaining information

about the level of garbage in containers and provide alert messages to the Garbage Collector Truck about the level of waste in a garbage container. Nonetheless, the problem of waste accumulation in remote places persists. Thus, it is crucial to detect and manage litter accumulation at remote and undesired locations with little or no human interference.

Through extensive background study and data collection, it can be concluded that most of the current automated waste management systems rely on IoT and hardware devices such as infrared sensor, ultrasonic sensor, metal sensor, etc. for detection of garbage and subsequent communication of the bin status. Other than this, there are waste sorting systems based on image processing which identify metal materials or other special types of waste. However, these become dysfunctional in case of paper and plastic products. Moreover, there are very few existing artificially intelligent systems which are self-sufficient for waste classification. With a combined approach consisting of the hardware devices such as a drone, Arduino, GPS and GSM module and image processing software algorithms, this paper offers a better and innovative solution to the problem of effective waste detection and management in wide and remote areas.
This system aims to achieve the following:

- To capture Ariel view images using Drone and further process the same to identify solid waste disposed at inappropriate places
- To notify the respective person-in-charge for taking the necessary action to clean the litter/solid waste
- Help maintain cleanliness in places such as beaches, institutions, cities etc.

II. BACKGROUND

A. *UAV*

Drones are Unmanned Aerial Vehicles (UAV) which operate without a human pilot aboard. They can either be flown using a ground-based remote controller operated by a human or autonomously by onboard computers. They are mostly used in military, commercial, scientific, agricultural, peacekeeping and other applications. They have the ability to navigate through large and remote places which security cameras otherwise fail to cover. The drone that will be used in this project is an assembled flamewheel 450 mm model equipped with first-person-view camera. The utilization of the drone will allow for real-time transmission of video image set to be analyzed by the litter detection deep learning framework.

B. *Image Stitching*

Image Stitching is a technique used for attaining high-resolution panoramic image from multiple images combined together. Image stitching techniques can either be direct intensity-based or feature based. This project focuses on using feature-based image stitching algorithms such as Scale Invariant Feature Transformation (SIFT) to determine a relationship between the images through distinct features extracted from the processed images. Snapshots from the live video feed captured with the help of the drone will be taken at regular intervals and stitched together for further analysis.

C. *Deep Learning and Computer Vision*

Concepts of Deep Learning and Computer Vision have been used in this project to achieve the automated detection of litter. Deep learning is a subpart of Artificial Intelligence (AI) which is concerned with emulating the learning approach that humans use to gain knowledge from data patterns. In contrast to machine learning algorithms, deep which are supervised and linear, deep learning algorithms are stacked in a hierarchy of increasing complexity and abstraction. The advantage of deep learning is that the program builds the feature set by itself without any sort of

human supervision and hence are generally faster and more accurate. A Deep Neural Network (DNN) is a neural network with a certain level of complexity, they have multiple layers between the input and output layers. The DNN finds the correct mathematical manipulation to turn the input into the output. Deep learning within the field of Computer Vision is concerned with the automatic derivation of useful information from visual data. Image Processing can be done using 'TensorFlow', a dataflow framework developed by Google for high performance computation. Tensor Flow object recognition algorithms are used to classify and identify arbitrary objects within larger panoramic images obtained from multiple datasets. TensorFlow is well suited for deploying deep learning models.

D. Python

Python is a high-level object-oriented programming language with emphasis on code readability and built-in libraries to provide various functionalities. Being a general purpose programming language, it provides powerful implementations to facilitate large data and advanced calculations with libraries such as TensorFlow, Scipy, Scikit-Learn, and NumPy. Tensorflow and Scikit-Learn are libraries that provide various machine learning frameworks and algorithms, such as neural networks and support vector machines. TensorBoard is a suite of visualization tools that make it easier to understand, debug, and optimize TensorFlow programs. NumPy is a library for applying various advanced mathematical functions with arrays and matrices, both key components of machine learning algorithms. Scipy is the optimized core package for scientific routines in Python; it is meant to operate efficiently on NumPy arrays, so that Numpy and Scipy work hand in hand. To use computer vision, Python was also implemented with Scikit-Image and OpenCV. These libraries provide important image manipulation tools computer vision techniques, such as feature extraction and image classification.

E. Hardware Components

Arduino is a programmable controller which can be used for handling file transfer operations. It can "talk", (transmit or receive data) via a serial channel, so any other device with serial capabilities can communicate with an Arduino. The drone will be fixed with an arduino to facilitate the interfacing of other hardware components such as a Global System for Mobile Communication (GSM) and Global Positioning System (GPS) modules. The GSM can be integrated with the Arduino to connect with the computer system over a network and communicate. A GPS module integrated with the drone will fetch coordinates (latitude and longitude) of the places that the drone will fly through.

F. Cloud Platform

A cloud platform allows for data storage, remote access and encapsulates various tools for data analytics and machine learning. A cloud platform such as Amazon Web Services (AWS) Elastic Cloud 2 (EC2) instance will be used to store the serially transmitted location coordinates (recorded by the GPS module) sent via the GSM module integrated with the drone. The use of cloud is to allow remote file transfer and further processing of the data recorded for comparison purposes.

III. EXISTING SYSTEMS

A. Waste Monitoring and Management Systems

In the past decade, many smart solutions for waste detection and management involving the use of IoT have emerged. These systems use mechanisms to check for the status of garbage containers and eventually indicate the same to the respective in-charge. RFID technology is used for collection of

data regarding garbage container in [5]. RFID tag works to detect within the frequency range and when any tag comes to the range of RFID reader, it automatically reads data from RFID reader, then filters collected data and arranges it into specific formatted SMS. Subsequently, the data is sent to central server sends the information to the web server as well as authorized person's mobile phone. A similar system design is described in [6] is which uses RFID technology to avoid overflow of the bin by sending an alert message. It uses Arduino Uno R3 as a microcontroller for reading data from sensors. When RFID tag interrupts the RFID reader, the sensor will check the status of the bin and send it to the web server.

Another paper on waste management [7] uses ultrasonic sensor to detect the level of garbage in the bin and communicates to control room using GSM system. Four IR sensors are used to detect the level of the garbage bin. When the bin is full the output of the fourth IR is active low and this output is given to microcontroller to send a message to control room through GSM. A similar system [8] consists of waste bins equipped with ultrasonic sensors which are interfaced with Arduino Uno and a Wi-Fi module which collect the waste fill level status and upload the data to database. This data appears on the android application which notifies the appropriate collector client based on their location once the bin gets filled up. [9] introduces an Android app, SpotGarbage that can automatically detect and localize regions containing garbage in user-clicked unconstrained geo-tagged real-world images. The app utilizes the proposed fully convolutional network, GarbNet for coarsely segmenting image regions containing garbage and the locations of the images classified as garbage are marked and plotted on Google Maps.

In this paper [10] ZigBee, GSM and ARM7 controller is used to monitor the garbage bin level. When garbage bin is full, this message of garbage level is sent to ARM7 controller. Then ARM7 will send the SMS through GSM to authority as to which bin is overflowing and requires cleaning up. [11] uses a Raspberry Pi to capture images and performs image processing using edge detection algorithm to identify waste. The detection is based on computer vision and is carried out with the help of camera and opt couplers. Waste is then collected by a vacuum unit which takes all the garbage and cleans the area. A related framework is described in [12] which consists of two parts, one of which is the hardware platform with Raspberry Pi as the core module and the other is the software platform based on SURF-BoW and multi-class SVM algorithm. This paper demonstrates the effectiveness of Bag of Visual Words algorithm in image processing.

B. Surveillance Systems

Drones are unmanned aerial vehicles which often find their application in surveillance and monitoring of wide areas and remote places which are otherwise inaccessible. [13] describes a smart transport infrastructure system based on drones and feature extraction techniques to determine the level of traffic congestion on roads and also indicates vacant parking slots in a particular area. This system is an instance of the accuracy and feasibility of using drone based image processing to achieve useful results. A similar approach is discussed in [14] where a custom-built drone is used to capture images of large fields, crops in particular, to determine the type of soil present in a specific area. Image processing is used for classification of images with reference to a predefined train dataset. This paper suggests the possibility of using a GSM integrated drone for instant processing of captured images on a system upon retrieval. Using a GSM would be beneficial as it has no range limit of transmission which is useful for real-time processing. This paper [15] presents the design and implementation of SkyStitch, a multi-UAV-based video surveillance system that supports real-time video stitching implemented using commercial off-the-shelf hardware. It addresses two key design challenges: (i) the high computational cost of stitching and (ii) the difficulty of ensuring good stitching quality under dynamic conditions. Deriving valuable insights from the past works, we aim to build a system which takes live video feed recorded by drone as input which can be retrieved by the system in real-time for image processing and feature extraction to identify and classify images of

garbage into various categories. The presence of a GSM (communication) and GPS (georeferencing) module allows for real time communication and location tracking.

IV. PROPOSED SYSTEM

As part of the proposed solution, the focus of this project is to build a system which takes a live video feed (or multiple images) captured by a drone as its inputs and extracts information from the images in order to identify places contaminated with solid waste/litter. The drone would capture a wide area such as a college campus or a locality and provide an aerial view of the location. The drone will be made to transfer the visual data recorded by it to the local server which is a computer system performing the processing operations. Arduino will be used for handling file transfer operations which will allow for real time retrieval of captured images. GSM and GPS module placed over the drone would provide the functionality of mobile communication and real time location referencing. With the help of image stitching techniques, the live video feed will be stitched based on similar features producing panoramic images of a place. This data can then be processed using Deep-Learning image processing algorithms coded in Python. The predefined dataset would encompass images of litter such as waste metal cans, bottles, crumpled paper, plastic bags etc. By testing the input data against the classes/categories of the predefined dataset in Tensor Flow framework, the algorithm will produce an output which assigns a distinct class to the given input image. The end goal is to send a notification which contains geographic coordinates (GPS based location) of the place contaminated with garbage to the respective person-in-charge prompting him/her to take the necessary action. The place coordinates corresponding to the timestamp of the output image (where litter was detected) will be fetched and subsequent message will be sent. The key components for this system include the following:

- Remote controlled Drone for surveillance and capturing images
- Arduino for managing the captured images and handling file operations
- Python IDE to run Python based programs
- Interfacing circuitry for making the connections
- GPS and GSM module for position tracking and communication respectively

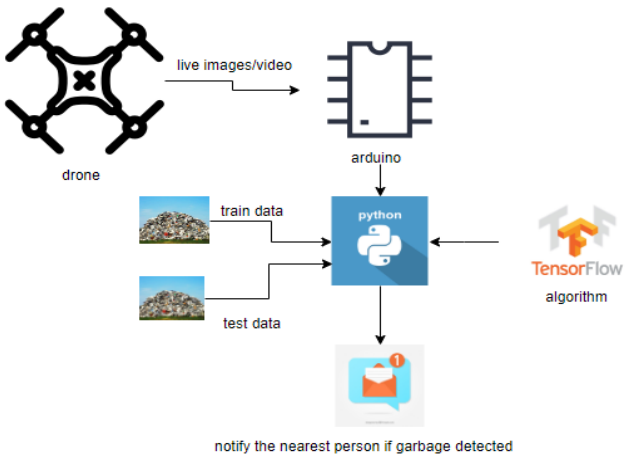


Fig. 1. Working of the proposed garbage monitoring system

The general working of the proposed system containing the hardware and software components has been shown above in fig 1.

V. EXPERIMENTAL PROCEDURE

The popular machine learning techniques such as SIFT and Bag of Visual Words (BoVW) have been used extensively for image processing through feature extraction over the past decade.

TensorFlow based algorithm on the other hand is relatively new as it was developed by Google recently. Being specialized for numerical computation and large-scale machine learning, TensorFlow provides a convenient front-end API for building applications with the framework. After performing image processing using both BoVW (which includes SIFT for feature extraction and BoVW for vocabulary formation) and TensorFlow, it was concluded that Tensor Flow is more efficient and performs better among the two for image processing and classification. The algorithms were tested using a sample dataset of images of trash taken from GitHub. The output images of TensorFlow algorithm are shown in fig 2.

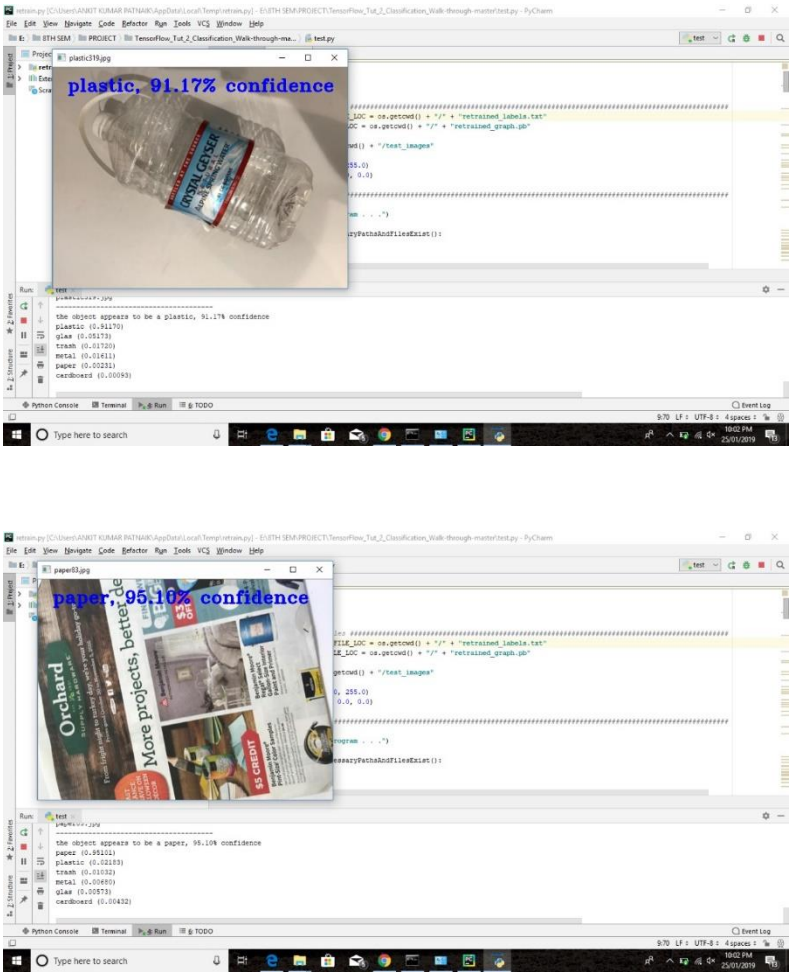


Fig 2: Output images of TensorFlow algorithm, displaying the classification of the given objects as trash with the confidence percentage

In a similar manner, Image Stitching algorithm was implemented using sample aerial images taken using first-person view (FPV) camera. OpenCV is an image and video processing library present in Python with SIFT functionalities which are used for feature extraction and image classification. Based on opencv_contrib's SIFT descriptor, the images are stitched together into a single panorama. The input images are shown in fig 3. Fig 4 shows the output of the image stitching algorithm which is a stitched image. Image stitching is needed in order to stitch the images taken by the drone in order to provide a 180 degree panoramic view of a specific location which will be used for detecting the presence of waste in the location. Real-time implementation of the algorithms will be done using images taken from the drone.

Fig 3: Input aerial images (left, centre, right) taken from FPV camera

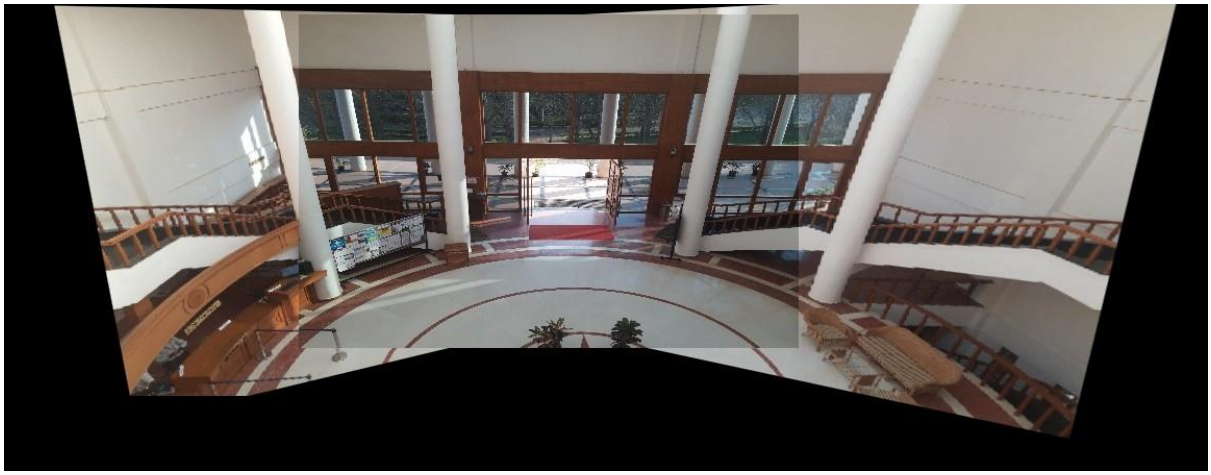


Fig 4: Stitched image output

VI. CONCLUSION

As this system is entirely automated, it will be efficient and applicable to all environments. The use of a drone adds to effective monitoring of waste as it removes the need for humans to manually travel to places for surveillance. This system offers the benefits of maintaining cleanliness in the surroundings and the environment. Waste accumulation in undesirable places can be minimized, even eliminated and waste management can be improved to a great extent. The implementation of this system will also assist organizations in effectively managing waste in remote places.

However, despite its benefits, this project specifically focuses on waste monitoring while waste disposal and cleaning remain out of its scope. Therefore, future works can be based on developing an automated robotic system to handle the waste detected. A robotic garbage collector can be integrated

with the drone. The collector can be essentially be programmed to receive inputs from the drone about the type of garbage detected upon which the collector will clean up the litter, segregate the waste into recyclable and non-recyclable materials. Disposal of this waste will also be looked after by the robotic garbage collector. Such an integrated system would complete the entire process of waste detection and collection single handedly without requiring human interference.

REFERENCES

[1] Sharma, S., Shah, K.W., 2005. Generation and disposal of solid waste in Hoshangabad. In: Book of Proceedings of the Second International Congress of Chemistry and Environment, Indore, India, pp. 749–751.

[2] Central Pollution Control Board (CPCB), 2004. Management of Municipal Solid Waste. Ministry of Environment and Forests, New Delhi, India

[3] Shekdar, A.V., Krshnawamy, K.N., Tikekar, V.G., Bhide, A.D., 1992. Indian urban solid waste management systems – jaded systems in need of resource augmentation. Journal of Waste Management 12 (4), 379–387.

[4] Arti Pamnani, Meka Srinivasarao, "Municipal Solid Waste Management in India: a review and some new results", vol 5, Issue 2, February (2014), pp. 01-08

[5] Dr.N.Satish Kumar, B.Vijayalakshmi, R. Jenifer Prathana, A.Shankar "IOT Based Smart Garbage alert system using Arduino UNO", published in Region 10 Conference (TENCON) on 22-25 Nov 2016.ISBN 978-1-5090-0751-6/16/2016 IEEE.

[6] Trushali S. Vasagade, Shabanam S. Tamboli, Archana D. Shinde "Dynamic Solid Waste Collection and Management System Based On Sensors, Elevator and GSMA" in International Conference on Inventive Communication and Computa-tional Technologies on 17 April 2017

[7] Ruhin Mary Saji, Drishya Gopakumar, Harish kumar, Lakshmi "A Survey on Smart Garbage management in cities using IOT" in international journal of enginnering and computer science ISSN: 2319-7242 on 11 Nov 2016.

[8] Vaibhav E. Pawar, Mustafa Bhatkar , Omkar Jadhav , Amey Mhatre "Smart Garbage Monitoring System Using IoT", Vol 8 Issue 4, IJESC journal

[9] Gaurav Mittal, Kaushal B. Yagnik, Mohit Garg, and Narayanan C. Krishnan, "SpotGarbage: Smartphone App to Detect Garbage Using Deep Learning", UBICOMP '16, SEPTEMBER 12–16, 2016, HEIDELBERG, GERMANY

[10] Pavithra, "Smart Trash system: An Application using ZigBee", IJISET - International Journal of Innovative Science, Engineering & Technology, Vol. 1 Issue 8, October 2014

[11] Samruddhi There, Chetan Shinde, Ashish Kumar Nath, Shubhangi A Joshi "Garbage Detection And Collection Of Garbage Using Computer Vision"

[12] Yijian Liu, King-Chi Fung, Wenqian Ding, Hongfei Guo, Ting Qu & Cong Xiao "Novel Smart Waste Sorting System based on Image Processing Algorithms: SURF-BoW and Multi-class SVM"

[13] G.Maria, E.Baccaglini, D.Brevi, M.Gavelli, R.Scopigno "A drone-based image processing system for car detection in a smart transport infrastructure" Proceedings of the 18th Mediterranean Electrotechnical Conference MELECON 2016, Limassol, Cyprus, 18-20 April 2016

[14]"Drone based image processing through feature extraction", 2017 2nd IEEE International Conference on Recent Trends in Electronics Information & Communication Technology (RTEICT), May 19-20, 2017, india

[15] Xiangyun Meng, Wei Wang and Ben Leong, "SkyStitch: a Cooperative Multi-UAV-based Real-time Video Surveillance System with Stitching", 2016

Revolutionary Approach for Smart Washing Machine using Machine Learning & Artificial Intelligence

Naveen Ananda Kumar J[*], Srikaanth GV[**], Kannika Nirai Vaani[**]

[*] Institute of management,
Christ (Deemed to be University) Bangalore
[**] Institute of management,
Christ (Deemed to be University) Bangalore

Abstract- Smart homes have started reshaping the place where people live and a smart washing machine is an addition to it. The fast pace of our life has led to automation of many of the household items. The next step in evolution after automation is smart and sustainable machines. In this paper we are introducing a ground-breaking concept of smart washing machine which analyzes the clothes and based on their specifications chooses the most appropriate type of wash for the clothes. This will reduce the fading of the clothes and can reduce the usage of soap than the conventional washing machine. It uses Artificial Intelligence and Machine Learning for enhancing the use of washing clothes more efficiently and effectively. By using less water than the conventional washing machine, this technology is paving a way to sustainability. Cameras, RFID patches, RFID reader are some of the technologies used for analyzing the material of the cloth. Using AI and ML we can choose the suitable type of wash for the clothes. This will increase the life span of clothes by mixing the right amount of detergent which in turn will reduce the water used. This will lead to a smart and sustainable lifestyle.

Index Terms- Artificial Intelligence, Machine Learning, RFID, Smart washing machine, Sustainability

I. INTRODUCTION

Artifical Intelligence and Machine Learning are the new emerging trend in the home appliance industry. Homes are getting integrated with technology with almost everything in the house which is taking us to the future. Consumers are becoming more tech-savvy and their expectation rises with the advancement in technology [1].

The current washing machines have different washing method for textiles, which will not allow clothes to get knotted, disarranged or entangled by intensive agitating cycles [2]. Damping the vibrations of the drum of the washing machine is one of the latest designs implemented in the current washing machine products [4]. [10] speaks about how only three wires are required for communicating inside the machine. The network can be used for communicating via home network to manage peak energy usage and add a new level of troubleshooting and field upgrades. A smart washing machine is a device which learns from the previous data and improves the quality of the next wash [6]. Nowadays the conventional washing machine has changed to a machine that can not only wash but also dry the clothes.

We are proposing a concept project where in, we are developing a smart washing machine that interacts with you and is more sustainable. The usual problem with traditional washing machines is that different materials of clothes need different types of wash. Using the same type of wash for different cloth materials damage the clothes and its durability decreases. The amount of soap that is used to wash also

plays an important role in durability of the clothes [7]. Conventional washing machines tend to use a lot of water than required while washing the clothes.

The smart washing machine analyses these factors and accordingly washes the clothes without damaging them and by using less water. It analyses the data of the clothes and accordingly chooses the appropriate type of wash. This solves most of the above-mentioned problems faced in a conventional washing machine.

II. CURRENT SCENARIO AND IT'S PROBLEMS

A. Present Scenario

In the recent times, textile industries have evolved to smart clothing. These involve smart clothes which has various sensors in the fabric which gives various inputs such inbuilt music players, health monitoring sensors, etc. [9]. Many of the textile manufacturing companies are adapting RFID technology by embedding RFID tag in their products to track them from the warehouse to the customers [8]. The RFID tag contains all the details about the product so it is easier to keep track of the products. [3] compares on how much water is being used and how much water is actually required for washing.

[5] talks on how the data can be gathered from the washing machine. It also talks about bidirectional data flow between washing machine and cloud for analytical purpose. A survey conducted by IBM found that 99.9% of the customers use only 3 of the many wash modes. This made the washing machine manufacturers to cut down the uses of other wash modes which reduced their engineering and R&D costs [5].

B. Research Gap

Washing machines have not evolved parallelly with other home appliances in the area of cloud technology. Other home appliances such as television has moved from CRT TV's to LED to Smart TV's, Refrigerators have moved to smart technology fridges and many more. There are lot of research papers on the technical aspects such as reducing the sound and damping of the washing machine. Most of the research papers are based on the automatic washing machines. This paper aims to bring washing machine into the smart home sector.

C. Problem statement

To develop a smart washing machine, which analyzes the data about the clothes, then selects the most appropriate method and uses minimum amount of water for the wash. It also uses the appropriate amount of soap, so that the color of the cloth does not fade faster.

D. Objectives

This paper is intended to integrate AI and ML into washing machine so that it identifies the type of cloth using RFID reader and based on that chooses the appropriate mode of washing. Based on the selected mode, appropriate amount of soap and water is to be added with the right temperature.

III. PROPOSED MODEL

A. Components Used

A.1. RFID tag

It has all the details about the clothes such as the color, material of the cloth, the temperature at which it should be washed and the appropriate type of wash.

A.2. RFID Scanner

The RFID scanner scans the RFID tag which is embedded into all the clothes. Then it sends the information to the microprocessor.

A.3. Microprocessor

The microprocessor collects all the data and sends it to the cloud. Based on the information it gets from the cloud it starts with the washing process accordingly. It controls all the sensors in the smart washing machine.

A.4. Sensors

A PH sensor is also fit to identify the PH level of the water entering the washing machine.
Water level sensor- To sense the amount of water to be filled.
Detergent level sensor- A sensor to measure the amount of soap present and the quantity that is dispersed into the water for washing.

Fig.1.1 shows the components and flow of data from scanning clothes till washing of clothes in the smart washing machine.

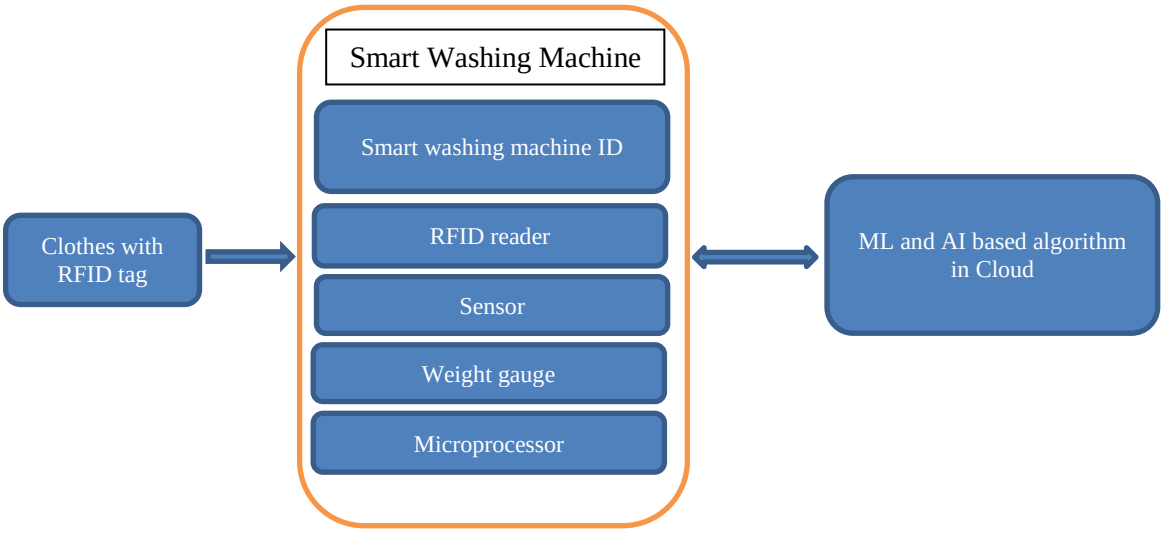


Fig.1.1: Architecture of Smart Washing Machine

B. Working process

The smart washing machine is first connected to the home WIFI network. Then the clothes are put in the smart washing machine drum where the RFID tags embedded on the clothes are scanned by the RFID scanner. This gives the details of the cloth such as the color, material type, the temperature at which it should be washed, type of wash and weight, etc. The microprocessor takes the data of clothes from the RFID scanner and send it to the cloud. The cloud which has machine learning and artificial intelligence algorithms process the data and sends the appropriate temperature, the amount of detergent soap, the amount of water required and best appropriate mode for the wash. This information is then sent to the microprocessor, based on that it chooses the amount of detergent, water and the temperature of the water and then adds it to the drum. Based on the type of wash, the microprocessor chooses starts the process of heating the water and soaking the clothes for certain period. If there are 2 or more clothes which have to be washed differently, a notification is sent to the mobile saying that there is a contradiction. So, the user

can remove the cloth, or the user can ignore the message, which the washing machine automatically chooses the best type of wash for it. After taking all these into considerations, it starts the wash. Once it has finished washing, it will send a notification to the user that the wash is complete. The smart washing machine then takes a feedback from the user and then sends the results to the cloud. These results will be analyzed by the AI and ML which will use it for future reference. Fig.1.2 shows the working processes of the washing machine.

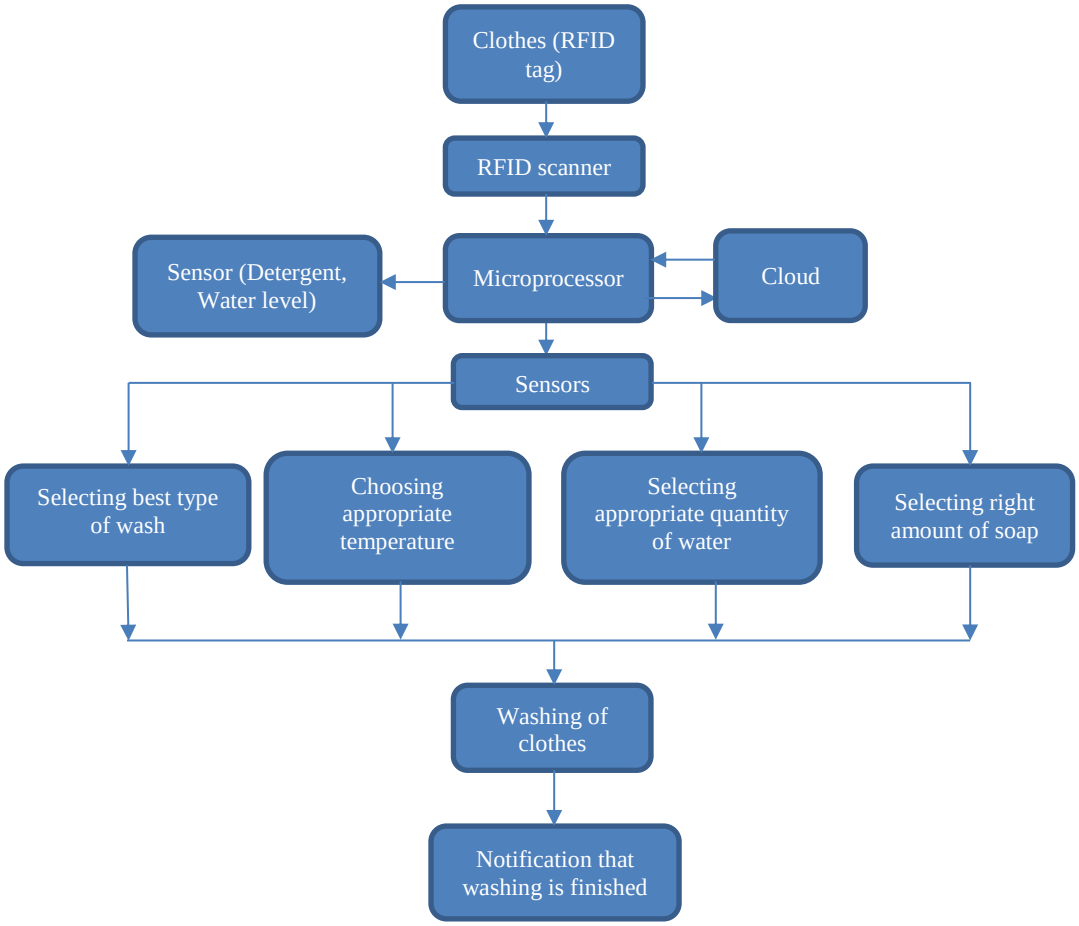


Fig.1.2: Flowchart of working procedure

B.1. Working of detergent soap sensing unit

The microprocessor checks the quantity of the detergent soap left in its storage. If it is less, it sends a notification that the detergent level is less so that the user can refill it. After this the microprocessor calculates the amount of soap and water required for the wash. Fig.1.3 shows the working of the detergent soap storage in the smart washing machine.

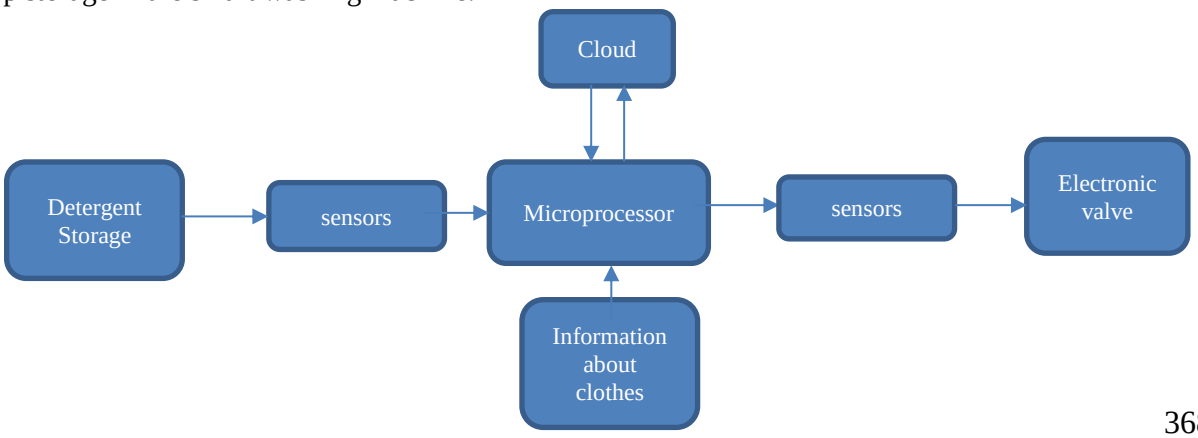


Fig.1.3: Working of the detergent soap sensing unit

IV. EXAMPLE

A. Conventional Washing Machine

Let's consider a situation where Ms. Padveeni is using an Automatic Washing Machine. She puts different types of clothes (4 cotton, 1 wool and 1 silk) for wash in the washing machine. She sets the temperature to 50 degree Celsius. She puts an approximate level of detergent soap in the machine. She set the mode in Cotton mode because majority of the cloth put for wash were cotton clothes. Then started the wash. After the wash when she removed the clothes, the cotton clothes were washed properly, but she was not happy with the condition of the wool and silk clothes. This is because, wool cloth should have been washed in a different temperature than cotton or silk and needs special care while washing. The silk cloth requires very less amount of detergent soap than the other type of clothes.

B. Proposed model (Smart Washing Machine)

Let's consider a situation where Ms. Padveeni is using the Smart Washing Machine. She puts different types of clothes (4 cotton, 1 wool and 1 silk) for wash in the washing machine. The RFID scanner reads the RFID tag of all the clothes put in the washing machine. It sends a notification that there are silk and wool clothes which cannot be washed with the same setting as cotton clothes. She ignores the warning and presses the start button. The Smart Washing Machine then finds the best wash for this set of clothes i.e. it sets the temperature between 35- 40 degree Celsius.

It uses the approximate amount of detergent soap and correct level of water required for the wash. This pattern of clothes and the appropriate type of wash gets stored in the cloud, so that the cloud can use the same data for the next wash for the same type of pattern. After the wash, the Smart Washing Machine collects the satisfaction level of the user and if there is any problem with clothes. This data gets stored in the cloud. So, when the same pattern comes for wash next time, ML and AI can analyze the satisfaction level and the type of wash used and corelate it. So, this can make the next wash more efficient and productive.

V. FUTURE WORK

This proposed model can be used a base model for finding out the actual water requirement used by the washing machine and also the amount of detergent soap which will be used. This can help the user quantify the amount of detergent soap to be purchased. This can be used by marketers for finding out the amount of detergent soaps required by every households and based on the consumption pattern they can maintain the stock accordingly.

VI. CONCLUSION

This novel concept which is an integration of AI and cloud-based computing has a bright future and is bound to branch out into various other dimensions of the home appliances especially with smart homes which are becoming more common. The smart washing machine will enable its users for a smarter and more sustainable use of water and detergent soap. As this washing machine programs and reprograms itself based on Machine Learning and Artificial Intelligence algorithms in the cloud, preprogramming of the washing machine is not required. This reduces the manufacturing and R&D costs for the manufacturers. This Smart washing machine will open new opportunities and take smart homes appliances to a new revolutionary level.

ACKNOWLEDGMENT

We thank our colleagues from Institute of management, Christ (Deemed to be) University and Alliance University who provided us with insights which was a great assistance in shaping our research. We would like to express our gratitude to Prof. Dr. Lakshmi Shankar Iyer, Head of Specialization - Business Analytics, Institute of Management and Dr. Manohar Kapse, Assistant Professor, Faculty- Business Analytics Institute of Management for sharing their pearls of wisdom and for reviewing our research and giving us key inputs.

REFERENCES

[1] Hultink, E. J., & Rijsdijk, S. A. (2009). How Today's Consumers Perceive Tomorrow's Smart Products. Journal of Product Innovation Management, 26(1), 24–42. https://doi.org/10.1111/j.1540-5885.2009.00332.x

[2] Islas, F. (2017). | mit kudo tori tutumie trolloni matematik, 1.

[3] Kiruthika, J., & Arulanantham, D. (2016). Making Washing Machines Smart through IoT, (April), 1–4.

[4] Kolhar, S. S., & Patel, D. R. (2013). Optimization of a Drum Type Washing Machine By Analytical and Computational Assessment, 4(6), 2759–2763.

[5] Locke, Dave and Lewis, K. (n.d.). Connected homes use case: Using washing machines to explain the IoT. Retrieved March 13, 2019, from https://www.ibm.com/blogs/internet-of-things/washing-iot-solution/

[6] Milasi, R. M., Jamali, M. R., & Lucas, C. (2007). Intelligent Washing Machine : A Bioinspired and Multi-objective Approach, 436–443.

[7] Nakajima, R., Asano, Y., Hida, T., Matsumoto, T., Nakajima, R., Asano, Y., … Study, A. (2016). A Study on the Effect of Dirt on an Inspection Surface on Defect Detection in Visual Inspection Utilizing Peripheral Vision To cite this version : HAL Id : hal-01417598.

[8] Nayak, R., Singh, A., Padhye, R., & Wang, L. (2015). RFID in textile and clothing manufacturing : technology and challenges. Fashion and Textiles. https://doi.org/10.1186/s40691-015-0034-9

[9] Suh, M., Carroll, K., & Cassill, N. (2010). Critical Review on Smart Clothing Product Development, 6(4), 1–18.

[10] Vanzoest, D., & Otten, D. (n.d.). Implementation of a Simulated Washing Machine with an Integrated Control Panel and Intra-Appliance Local Interconnect Network Bus.

Machine Learning Approach to Predict the Accident Risk during Foggy Weather Conditions

Dr. A.Pranayanath Reddy *, Dr. Shekhar R ** , Dr.Suresh Babu *

* Department of CSE, KL University (Hyderabad Campus), Hyderabad, Telangana, India
**Department of CSE, Alliance University, Bengaluru, Karnataka, India

Abstract- Vehicle mishaps in foggy climate conditions have expanded in the proper method of time. One of significant reason is contamination rate has expanded in the earth we are seeing a dynamic difference in atmosphere independent of seasons. Visibility level gets decreased because of foggy climate conditions is a typical factor for street mishaps on roadways. This subsequently there has been an expansion of interest to build up a smart framework which can keep away from/mishaps or smashing of vehicles from behind utilizing a visibility go estimation framework. The framework would caution the driver if there is any obstruction before the vehicle. In this paper, we give a short outline of the best in class commitment in connection to assessing visibility separate under foggy climate conditions. We at that point present a neural system approach for evaluating visibility separations utilizing a camera that can be fixed to a street side units (RSU) or mounted locally available a moving vehicle or long separation Sensors .The proposed strategy can be made as inherent item for 4 wheeler or more vehicles which make the vehicle clever.

Index Terms*:* Visibility Distance; Fog Detection; Intelligent Transportation Systems; Neural Networks; Machine Learning ; Koschmieder Law; Computer Vision ;Road Accident.

I INTRODUCTION

Foggy climate conditions speak to an up and coming risk to street wellbeing, frequently prompting lethal street mishaps on the grounds that debased street visibility can possibly (1) shock even experienced drivers, (2) adjust the drivers' driving conduct, and (3) mutilate drivers' view of profundity, separate, and speed[1].

These issues have caused gigantic financial misfortune just as human setbacks. As indicated by Global Status Report on Road Safety, distributed by World Health Organization in 2015, about 1.25 million individuals were executed and 20-50 million people groups got injured non-deadly wounds (WHO report, 2015) consistently.

As per the street auto collision information given by states, Delhi records one of the most astounding number of deadly mishaps (NHAI report, 2016). Street auto collision recorded information is acquired from National Highway Authority of India (NHAI). Not with standing, this pattern can change in future as it is difficult to anticipate the rate at which street auto collisions happen as it can happen in any circumstance.

Prior research[1,2,3] uncovered that albeit foggy climate conditions are not intermittent marvels, the quantity of related various impacting vehicles, wounds and fatalities are a lot higher than normal.

Since thruway mist reduction advances have not come to yet the ideal dimension of development and financial feasibility, a few roadway crash countermeasures and new vehicle plan innovations have been proposed to help drivers adapt to foggy climate conditions. These incorporate reflectorized paints on asphalt edge striping, beaded path delineators, squinting strobe lights, and installed hardware including haze lights, Light Detection and Ranging (LiDAR) sensors and Autonomous Emergency Braking (AEB) frameworks, among numerous others.

II RELATED WORK

1) **Identification of Traffic Accident Trigger:** Tremendous endeavors have been given to the recognizable proof of key conditions or specific traffic designs that could prompt car crash. For example, Oh proposed the supposition that problematic traffic stream is a trigger to crash [4].

2) **Real-time Traffic Accident Prediction**: With the improvement of AI, numerous specialists begin to concentrate on constant car crash forecast. Lv picked highlight factors dependent on Euclidean measurement and used k-closest neighbor technique to anticipate auto collision [5].Park gathered enormous car crash information of thruway in Seoul and fabricate an expectation work process dependent on k-implies bunch investigation and calculated relapse [6]. One impediment of these works is that, they didn't consolidate a few significance factors, for example, traffic stream, climate condition, air quality into their model.

3) **Deep Learning:** The achievement of profound learning has demonstrated its capacity in finding mind boggling structures in high dimensional information. With respect to explores on insightful transportation framework, various investigations center around traffic stream forecast dependent on profound learning [7]. In a more drawn out time scale, a few examinations attempt to anticipate the blockage development of vast scale transportation arrange [8]. Another fascinating application used profound fortification figuring out how to control the planning of traffic flag [9].

Mist definition and visibility models

Mist is a sort of cloud on the ground and is shaped by the suspension of tiny dampness dewdrops into airborne particles. As per the Meteorological Office 1969[10], mist is characterized as the condition of climatic shadowiness where visibility of article is decreased beneath 1 Km. In the event that visibility dips under 40 meters, mist is qualified as being "thick". A visibility between 40 meters and 200 meters compares to a thick haze situation.

For street security applications, the visibility scope of intrigue is somewhere in the range of 0 and 400 meters. The glowing transition exuding from obvious light (400 nm $\leq \lambda \leq$700 nm) gets dispersed every which way when it hits a water bed and assimilation is frequently immaterial for this situation. This dispersing can seriously debilitate drivers' profundity discernment and fringe vision. The lessening of obvious light is described by the termination coefficient k (m-1) which is a factor of the beads measure and concentration. The estimation of this coefficient has been the premise of numerous visibility run estimation strategies. Light proliferation through haze based on Koshmieder luminance lessening law,Duntley[11] proposed the constriction law of air differentiates under uniform illuminance which expresses that an item with inborn complexity C0 without wanting to be seen at a separation d with an evident difference C given by:

$$C = C_0 e^{-kd} \quad (1)$$

(1) The above articulation has been utilized as a reason for characterizing the "meteorological visibility remove" $d_{visibility}$ as the best flat separation at which a dark article (C_0=1) of a moderate measurement can be seen seemingly within easy reach amid daytime with a differentiation limit ε=5% , as prescribed by the International Commission of Illumination(IEC8): (2)

$$d_{visibility} = -\frac{1}{k}\ln(0.05) \cong \frac{3}{k} \quad (2)$$

The above basic articulation recommends that, from the termination coefficient k of the encompassing environment, one can determine the apparent visual scope of a dark item at daytime.

III VISIBILITY SEPARATE ESTIMATION TECHNIQUES AND ARRANGEMENTS

Estimation of range can be an overwhelming assignment given the non-uniform nature of the physical environment and the interlacing components impacting visibility which incorporate power of encompassing light, physical properties of articles, light dissipate and ingestion among numerous others. Different visibility separate estimation approaches for transportation frameworks have been proposed over the previous years. These methodologies vary in different perspectives (daytime versus evening time, mist identification or potentially visibility estimation, optical versus picture sensors, fixed cameras versus locally available cameras, video-based versus picture based; calculation utilized for picture handling and separation estimation, and so forth).

LiDARs have been utilized to assess visibility under foggy climate conditions by breaking down the flag backscattered by haze droplets[12]. As revealed by Colomb[13], this methodology, in any case, requires calibrating of the LiDAR's capacity to adjust to the elimination coefficient k. There has been a developing enthusiasm amid the previous couple of years in utilizing fixed cameras (set on the roadway) or (and to a lesser degree) locally available cameras to assess visibility remove, as these gadgets are moderately shabby and are as of now sent for traffic observing and reconnaissance on major expressways.

Camera-based methodologies can be characterized into three primary classes: The principal (type-I) approach plans to quantify the separation to the uttermost dark focus in the picture while as yet showing a differentiation more prominent than ε=5% according to the IEC recommendation[14,15]. This includes looking for focuses or districts of enthusiasm by applying thresholding and division strategies to find indicated targets, for example, path markings[16], street signs[17], street limits, or the crossing point between street surface and the sky[12]. The fundamental downside of this methodology is that it requires exact geometric adjustment of the camera and it depends on the nearness of reference objects with high differences in the scene.

The second (type II) approach depends on processing the scene differentiation and afterward playing out a direct relapse between this complexity and the visual range that is figured with the guide of extra reference sensors[13]. This methodology does not require camera adjustment or the nearness of reference objects. Nonetheless, it involves a learning stage that requires the use of extra meteorological sensors.

The third (type III) approach utilizes a worldwide descriptor vector which is registered in general picture, independent of its substance. This methodology does not require edge location or learning of the separation to different reference targets. The worldwide descriptor vector reflects data about the worldwide picture differentiation and utilizations highlights dependent on the inclination whole or the Fourier coefficients sum[15] as these are invariant to enlightenment changes. Our proposed neural system approach falls inside this class.

IV THE PROPOSED NEURAL SYSTEM APPROACH

Our methodology comprises of evaluating visibility extend through an administered preparing connected to marked models that are portrayed by worldwide instead of neighborhood highlights. Our classifier is a three-layer neural system prepared with a back-spread calculation. The principal input layer is the component vector picture descriptor. The concealed layer comprises of a lot of completely interconnected computational hubs whose number is resolved exactly.

The yield layer is a vector whose measure is equivalent to number of classes, which for our situation compares to various visibility ranges. For the worldwide element descriptor, we have settled on a Fourier Transform approach which is viewed as a standout amongst the most effective picture change techniques[18] and it catches the power range of the picture. As a result of the high dimensionality of the Fourier change extent descriptor, we have brought a dimensionality decrease through Principal Component Analysis (PCA). Notwithstanding the Fourier extent descriptor, we have likewise utilized Shannon entropy as a second descriptor that portrays a picture surface by investigating its dark dimension dissemination as per the accompanying articulation: (3) where Pi is the likelihood that the contrast between 2 nearby pixels is equivalent to I. The last total descriptor is a mix of the mean squared Fourier change vector decreased by the PCA and the picture entropy.

Fig. 1. Longest Sensors which are used in front and Rear of Vehicle.

$$E = \sum_{i=1}^{n} P_i \, log_2(P_i) \qquad (3)$$

V TRIAL ASSESSMENT AND RESULTS

In this investigation, we have considered six classes of visibility go (< 60m, 60m - 100m, 100m - 150m, 150m - 200m, 200m - 250m, and ≥ 250m). We have utilized the element vectors separated from the

pictures to gain proficiency with the required visibility classes. Tentatively, we kept up just the initial 100 eigenvectors since they speak to close 99% of the total percent of change. Figure 2 delineates the general engineering of the neural system where I1-I20 speak to the components of the information descriptor vector, H1-H12 are the shrouded layer hubs and the yields compare to the visibility ranges.

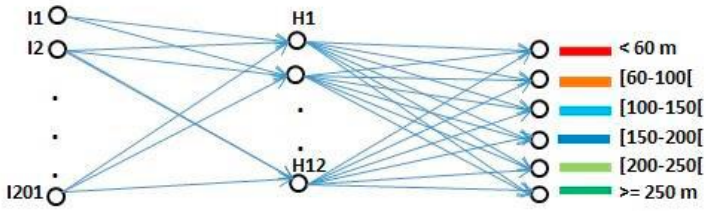


Fig.2. Neural system layers

To assess our proposed methodology, we have utilized the FROSI (Foggy Road Sign Images) database[19]. This database contains an arrangement of 400 manufactured pictures with 1000 street signs put at different extents. For each picture, a lot of 7 kinds of uniform mist thickness are accessible with visibility separations going from 50m (substantial haze) to 400m (slight haze), as appeared in the illustrative precedent in figure 3.

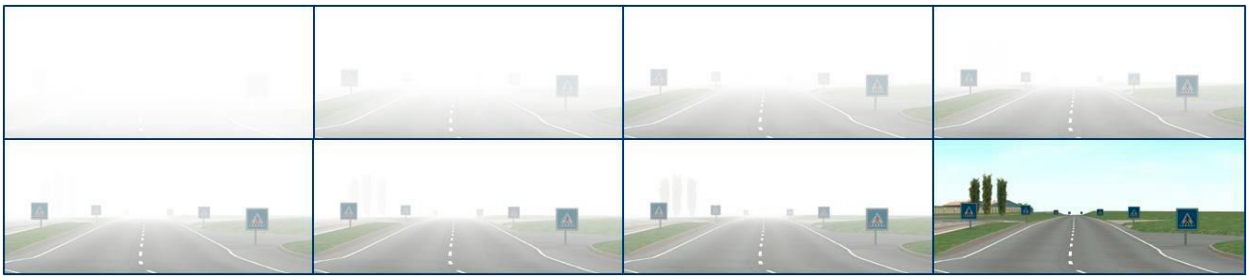


Fig. 3. Test of pictures from the FROSI database

We have likewise looked at the exactness of our proposed arrangement with that dependent on sort I approach. For our situation, no following/recognition steps were required as traffic sign positions were known from the earlier. For sort I approach, we have embraced Michelson differentiate recipe $C = (I\max - I\min)/(I\max + I\min)$, where Imax and Imin indicate the greatest and least pixel power, individually. For the neural system preparing stage, we have utilized 336 pictures under various visibility ranges. The weighted neural system was along these lines tried utilizing an alternate arrangement of 336 pictures under different haze thickness conditions. Our outcomes are abridged in the perplexity lattice portrayed in figure 4(a). In general, a 90.2% effective arrangement rate was accomplished, contrasted with a 65% achievement rate acquired through sort I approach. As appeared in figure 5(b), our methodology gives better outcomes under all visibility classes. We see that the 6th class has the base likelihood of identification (83%) because of disarray with extents somewhere in the range of 150 and 250 meters.

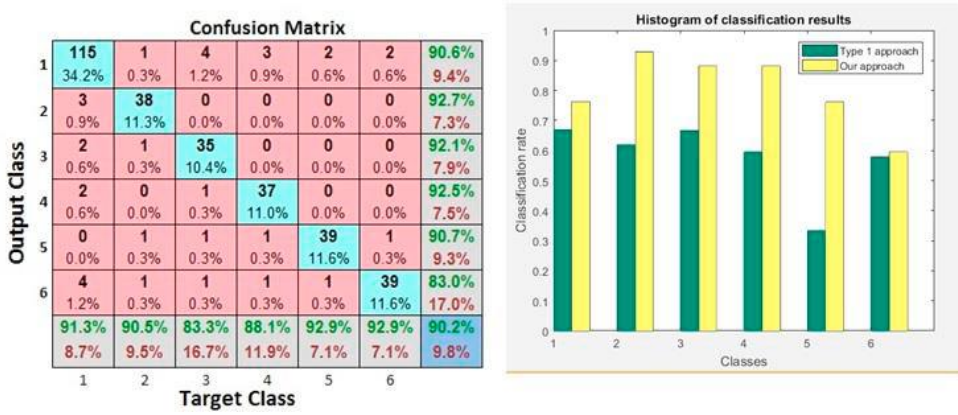


Fig. 4 (a) Confusion matrix; (b) Comparison of results

VI CONCLUSION

Road traffic accident severity keeps on changing over time and increase endlessly. The changing and increasing road traffic accident severity leads to the issues of not understanding the accident behavior, factors influencing the traffic accident severity. In this contribution, we have presented a neural network approach to estimate visibility range under foggy weather conditions. First, here we only utilized the traffic accident data itself for prediction. However, other related data, such as traffic flow, human mobility, road characteristic and special events, maybe significant to traffic accident risk prediction as well. Second, our prediction results are coarse-grained, and cannot provide road level accident risk prediction. Our solution requires a single camera which can be fixed on the roadway side or placed onboard a vehicle or high speed sensors. Our approach provided visibility range estimates that are close the expected values for a wide range of fog density scenarios. A key advantage of the proposed approach is that it is inherently generic and does not require special camera calibration or a prior knowledge of distances in the depth map. Our proposed algorithm can be implemented on existing camera-based traffic monitoring systems, which can serve as a driving aid to warn motorists and request them to adapt their speeds according to the estimated visibility distance. We are currently working on further refining our approach and comparing it with additional methods identified in our literature review.

REFERENCES

[1] Hamilon B, Tefft B, Arnold L, Grabowski J. Hidden highways: Fog and traffic crashes on America's roads, *2014 AAA Foundation for Traffic Safety Report*. July 2014. p. 1-24

[2] Creech M, Fog F. A review of the literature pertaining to highway problems and possible solutions, *Virginia Highway Research Council Report*. July 1972. p. 1-10

[3] Abdel-Aty M, Al-Ahad E, Helai H, Keechoo C. A study on crashes related to visibility obstruction due to fog and smoke, *Accident Analysis and Prevention*, 2011; 43. p. 1730-1737.

[4] C. Oh, J.-S. Oh, S. Ritchie, and M. Chang, "Real-time estimation of freeway accident likelihood," in 80th Annual Meeting of the Transportation Research Board, Washington, DC, 2001.

[5] Y. Lv, S. Tang, and H. Zhao, "Real-time highway traffic accident prediction based on the k-nearest neighbor method," in Measuring Technology and Mechatronics Automation, 2009. ICMTMA'09. International Conference on, vol. 3. IEEE, 2009, pp. 547–550.

[6] S.-h. Park, S.-m. Kim, and Y.-g. Ha, "Highway traffic accident prediction using vds big data analysis," The Journal of Supercomputing, vol. 72, no. 7, pp. 2815–2831, 2016.

[7] Wu and H. Tan, "Short-term traffic flow forecasting with spatialtemporal correlation in a hybrid deep learning framework," arXiv preprint arXiv:1612.01022, 2016

[8] X. Ma, H. Yu, Y. Wang, and Y. Wang, "Large-scale transportation network congestion evolution prediction using deep learning theory," PloS one, vol. 10, no. 3, p. e0119044, 2015.
[9] L. Li, Y. Lv, and F.-Y. Wang, "Traffic signal timing via deep reinforcement learning," IEEE/CAA Journal of Automatica Sinica, vol. 3, no. 3, pp. 247–254, 2016.
[10] Meteorological Office. *Observers' handbook*. HMSO, London, 1969.
[11] Middleton W. *Vision through the atmosphere*. University of Toronto Press, 1952.
[12] Rosenstein F. Intelligent rear light – Constant perceptibility of light signals under all weather conditions, *ISAL 2005*, Sept 2005, p. 403-411
[13] Colomb M. The main results of a European research project: Improvement of transport safety by control of fog production in a chamber ("FOG"), 3rd International Conference on Fog, Fog Collection and Dew, Cape Town , South Africa, October 2004, p. 242-251.
[14] ush C, Debes E. Wavelet transform for analyzing fog visibility, *IEEE Intelligent Systems*, 1998; 13:6. p. 66-71
[15] Hautiére N, Babari R, Dumont É, Brémond R, Paparoditis N. Estimating meteorological visibility using cameras: A probabilistic model-driven approach. In: Kimmel R, Klette R, Sugimoto A. (eds) *Computer Vision – ACCV 2010*. ACCV 2010. Lecture notes in computer Science, 6495. Springer, Berlin, Heidelberg. 2011; p. 243-254
[16] Kwon TM. Atmospheric visibility measurements using video cameras: Relative visibility, Technical Report, University of Minnesota Duluth, Minnesota Department of Transportation; July 2004, p. 1-51.
[17] Cavallo V, Colomb M, Doré J. Distance perception of vehicle rear lights in fog, *Human factors*, 2001; 43:3, p. 442-451
[18] FROSI. http://www.livic.ifsttar.fr/linstitut/cosys/laboratoires/livic-ifsttar/logiciels/bases-de-donnees/frosi/ (accessed 1 June 2017).
[19] Pavlic M, Belzner H, Rigoll G, Ilié, S. Image based fog detection in vehicles, IEEE Intelligent Vehicles Symposium (IV), Alcala de Henares, Spain, 3-7 June 2012. p. 1-6.

.

Real-Time Atonomous Military Robot Path Planning

Dr. Sandeep Bhat[1], Dr. M. Meenakshi[2]

[1]Assoc. Professor, Dept of E&CE, Srinivas Institute of Technology, Mangaluru, India
[2]Professor & Head Department of Instrumentation Technology, Dr.AIT, Bengaluru, India

Abstract: These days the robots are assuming an imperative job in the human life.In risky zones like military,medical , atomic reactor, and so on robots are utilized rather than human beings. To play out all these activities,the robot should be self-governing and furthermore fit for its way arranging by keeping away from snags. This paper clarifies about the ongoing execution of the self-governing robot way arranging and approved by utilizing an implanted framework stage.

Keywords: Autonomous, Blob Analysis,Path Planning ,Robot

I. INTRODUCTION

At the point when a service robot is asked to control an item, it ought to decide the bearings longside it can access and evacuate the article. The potential available bearings for the item are recovered from the article database. At that point, spatial prevailing upon the encompassing condition and the gripper geometry is summoned to confirm the bearings[1].An ongoing Multisensor data retrieval (MSDR), which awards offbeat access to the cloud from the robots. A market-based administration technique for proficient information recovery is approved by surveying a few quality-of-service (QoS) criteria, with accentuation on encouraging information recovery in close ongoing in regular cloud automated situations[2].

In practical cases, the human visual execution and that the robot has constraints in communication with human. The robot has constrained locally available movement and correspondence vitality and works in practical channel conditions encountering way misfortune, shadowing, and multipath[3].The issue of location and following of general objects with semi-static elements seen by a versatile robot moving in large environment. A key issue is that because of environment scale, the robot can just watch a subset of the articles at some random time. A model for the robot development in which the items normally just move locally, yet with some little likelihood they hop longer distances through what we call global movement.[4,5].

The Endoscopic Submucosal Dissection (ESD) CYCLOPS framework can accomplish powers of 46N, and demonstrated a mean error of 0.217mm amid a circular following assignment with the assistance of a robot. The workspace and instrument skill is appeared by pre-clinical preliminaries. The framework is as of now experiencing pre-clinical in vivo approval[6].

The Sum of Absolute Difference (SAD) algorithm is used for the implementation of the proposed image processing algorithm. It works on the principle of image subtraction. The developed algorithm is validated in real time by a change-based moving object detection method. The novelty of this work is the application of the developed autonomous robot for the detection of mines in the war field[7,8].

The organization of paper except this section is: Block diagram and its explanation is in the section II, Working principle is in the section III and the Results and analysis in section IV.Finally, the Conclusions are drawn in the section V.

II. BLOCK DIAGRAM AND ITS EXPLANATION

The block diagram of real-time robot's video capturing system under the servielance area is given in the figure 1.

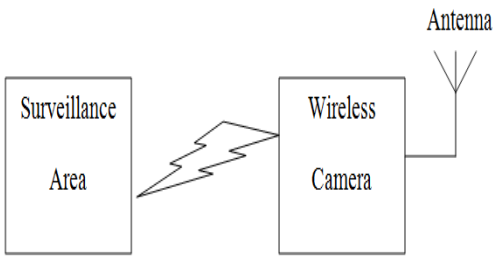


Fig. 1: The Video capturing set up.

First and foremost to check the existence of the obstacles present in the surveillance area. This is done by capturing the video from the supervision area and transmit it to the receiving system as shown in the figure 2.

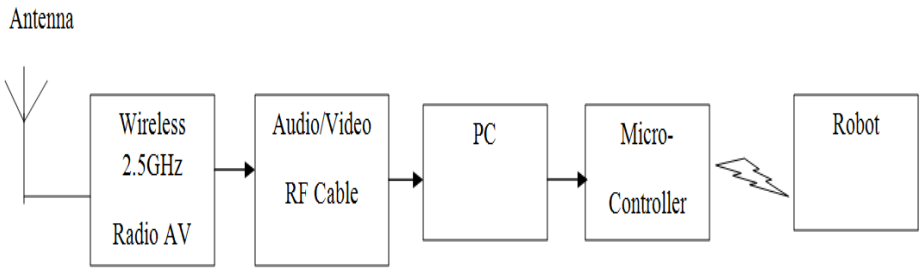


Fig. 2: Real-robot receiving system.

The captured video is transmitted to the receiver,for the further processing. The receiver system consists of a radio audio- video device to get back the video,which is already transmitted. This video is applied to the personal computer (PC),to identify the presence of obstacles using MATLAB tool with the help of the Blob analysis. Depending on the presence or absence of the obstacle,the PC sends an appropriate command to the robot through the RF communication link.

III . WORKING PRINCIPLE

The real-time robot's working principle is as described in the figure 3.

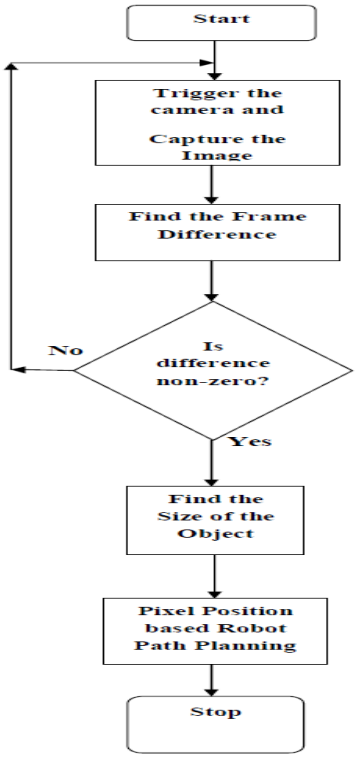


Fig. 3 : The real-time robot working principle.

First, turn ON the camera which is located in the supervision area,where actually the robot has to be worked. When the camera starts to capture the video,the PC converts the video into frames. Here, each frame is compared to its previous frame;if the difference is nonzero, then the obstacle is found. Later, the size of the obstacle is calculated using image processing based Blob algorithm. Depending on the size of the obstacle the appropriate command is transmitted through RF link in the form of ANSI code to the robot. The robot,which is developed on an embedded platform receives the command and plans its path accordingly.

IV. IMPLEMENTATION AND RESULTS

The real-time set-up of the wireless video transmission/reception system is as shown in the figure 4. The battery operated wireless camera with built-in antenna is as depicted in the figure 4a. Next, a 2.4GHz RF A/V device is shown in the figure 4b. Further, the RF A/V receiver is connected to a PC via USB dongle,as illustrated in the figure 4c.

a b c

Fig.4: The video transceiver system.

The video captured with the help of an RF system along with MultiViewer software is illustrated in the figure 5. This video is further processed for the real robot path planning applications.

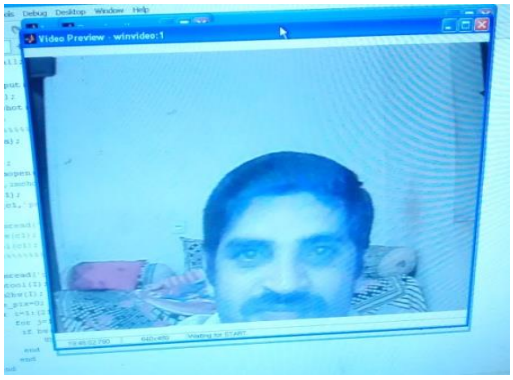


Fig. 5: Video Captured through Wireless RF Reception System.

The transmitter and receiver section of the RF communication robot are as given in the figure 6. This module is mounted on the robot and connected to the microcontroller unit to receive the video from the surveillance area.

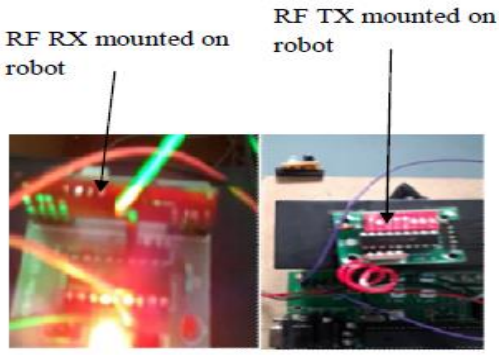


Fig. 6: The RF transceiver module.

The author's prototype of the robot with a wireless camera and RF receiver is as illustrated in figure 7.

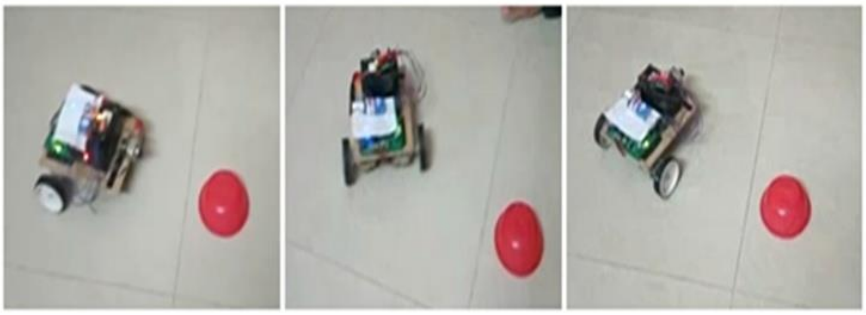


Figure 7: The final robot prototype and its movement by avoiding the obstacle(red in colour)

V.CONCLUSION

This paper offered answers for comprehend, exceptionally unsafe military applications are performed by remote robots. These applications are performed by preparing personal computer from accepting video of the observation region with the assistance of FM modulation and demodulation process. Here, in the wake of handling the video to recognize the hindrance along the way of the robot, the microcontroller interfaced with PC sends the fitting direction to autonomous robot by ASK modulation technique along with the RF modules. Thus, the upgraded way of the robot is resolved.

REFERRENCES

[1] Han-young Jang ; Hadi Moradi ; Suyeon Hong ; Sukhan Lee ; Junghyun Han, “Spatial Reasoning for Real-time Robotic Manipulation” , IEEE/RSJ International Conference on Intelligent Robots and Systems, 2006 ,pp. 2362-2367.

[2] Lujia Wang ; Ming Liu ; Max Q.-H. Meng , “Real-Time Multisensor Data Retrieval for Cloud Robotic Systems”, IEEE Transactions on Automation Science and Engineering, Volume: 12, Issue: 2, April 2015, pp. 507 – 518.

[3] Hong Cai ; Yasamin Mostofi ,Human–Robot Collaborative Site Inspection Under Resource Constraints, IEEE Transactions on Robotics , Volume: 35 , Issue: 1 , Feb. 2019 ,pp**.** 200 – 215.

[4] Jalal Heshmatpanah ; Siavash Boroumand ; Mehdi Tale Masouleh, “FPGA design and implementation for real time vision applications on NTACO mobile robot”, Second RSI/ISM International Conference on Robotics and Mechatronics (ICRoM), 2014 ,pp. 17-21

[5] Nils Bore ; Johan Ekekrantz ; Patric Jensfelt ; John Folkesson, “Detection and Tracking of General Movable Objects in Large Three-Dimensional Maps”, IEEE Transactions on Robotics, Volume: 35 , Issue: 1 , Feb. 2019 ,pp**.** 231 – 247.

[6] Timo J.C. Oude Vrielink ; Ming Zhao ; Ara Darzi ; George P. Mylonas , “ESD CYCLOPS: A New Robotic Surgical System for GI Surgery” ,Timo J.C. Oude Vrielink ; Ming Zhao ; Ara Darzi ; George P. Mylonas, IEEE International Conference on Robotics and Automation (ICRA), 2018,pp.107-110.

[7] Sandeep Bhat,M.Meenakshi, “Vision Based Robotic System for Military Applications -- Design and Real Time Validation”, Fifth International Conference on Signal and Image Processing (ICSIP), 2014 ,pp.123-125.

[8] Mustafa Engin,IEEE international conference on “Embedded and real time system design: A case study fire fighting robot”**,**5th Mediterranean Conference on Embedded Computing” "(MECO)2016, Montenegro , June 2016,pp.18-21.

Historical Kannada Handwritten Scripts Recognition System using Line Segmentation with LBP features

Parashuram Bannigidad*, Chandrashekar Gudada*

* Department of Computer Science, Rani Channamma University, Belagavi, KA, India.

Abstract- The inscriptions or Epigraphical manuscripts written on different material such as rock carving, palm leaf, cloth, metal plates and paper are the cultural heritage of our country; our aim is to recreate the cultural importance of the Kannada Language and its traditionally writing through the historical manuscripts. Most of the resources are in degraded state, the degraded manuscripts are influenced by weather condition. The offline handwritten text recognition is one of the most challenging tasks in document image analysis. In the present digital era, we need to protect and digitize the resources of our Indian culture and heritage by digitizing the manuscripts which are losing its originality and status. In this paper, we have attempted to identify and recognize the historical Kannada handwritten scripts of various dynasties; namely, Vijayanagara dynasty (1460 AD), Mysore Wadiyar dynasty (1936 AD), Vijayanagara dynasty (1400 AD) and Hoysala dynasty (1340 AD) by using the improved seam carving text line segmentation method with LBP features. For recognition and classification purpose the LDA, K-NN and SVM classifiers are used. The average classification accuracy for different dynasties are computed. The LDA classifier yields 94.2%, K-NN classifier has yielded 94.9% and SVM classifier has 96.4%. Based on the experimentation, the SVM classifier has recorded good classification performance comparatively LDA and K-NN classifiers for historical Kannada handwritten scripts. The experimental results are verified by Epigraphists and language expert, which shows the robustness of the proposed method.

Index Terms- Restoration, Seam carving, Line segmentation, Kannada, LDA, K-NN, SVM, Recognition, LBP, handwritten script, historical documents, document image analysis.

I. INTRODUCTION

India is one of the oldest and ancient civilized country in the world, its civilization started before 7000 BCE with spiritual and astrological knowledge. This knowledge of information is stored and kept preserved in the form of historical inscription and epigraphical manuscripts in the manuscript resource centres, manuscript conservation centres, manuscript partner centres, gurukula and monasteries. Due to the negligence in maintaining, these scripts are in the state of degradedness. Hence, the digitization of these degraded documents is a important task to restore the deciphering inscriptions. Particularly, in the state of Karnataka, many dynasties have ruled and contributed their knowledge to the Indian civilization. In this work, we are trying to experiment with the available historical Kannada handwritten manuscripts (inscription) written on paper from various dynasties, namely; Vijayanagara dynasty (1460 AD), Mysore Wadiyar dynasty (1936 AD), Vijayanagara dynasty (1400 AD) and Hoysala dynasty (1340 AD) collected from various institutions or universities for identification and recognition of historical Kannada handwritten manuscripts.

Very few researchers have contributed to this area in the literature; Seam carving for text line extraction on colour and grayscale historical manuscript was proposed by Nikolaos et.al.[4]. Seam carving for content-aware image resizing has been investigated by Avidan et.al.[5]. Influence of text line segmentation in handwritten text recognition was presented by Romero et.al.[6]. The identification of writer using sparse radial sampling LBP features was proposed by Nicolaou et.al.[7]. Word spotting in English handwritten historical document by using LBP features has been proposed by Dey et.al.[8]. LBP based line-wise script identification was investigated by Ferrer et.al.[9]. Ghosh et.al.[10] proposed an algorithm for text / non-text separation from handwritten document images using LBP features. The evaluation for historical document image analysis using texture feature was carried out by Mehri et.al.[11]. Laurence et.al.[12] has done a survey on text line segmentation of historical documents. Text line segmentation for gray scale historical document images was proposed by Asi et.al.[13]. Parashuram and Chandrashekar [14, 15] have proposed an image enhancement method for degraded historical Kannada handwritten document images. The Identification and classification of historical Kannada handwritten document images using LBP features was proposed by Parashuram and Chandrashekar [16].

II. PROPOSED METHOD

The objective of the proposed method is to digitize and recognize the historical Kannada handwritten manuscripts based on their age-type using improved seam carving text line segmentation approach by extracting LBP features using LDA, K-NN and SVM classifiers. The proposed method mainly consists of data collection, segmentation, pre-processing, feature extraction and classification; which are discussed as below:

A. Data collection

The availability of standard datasets of historical Kannada handwritten manuscripts are rarely found in the literature. Hence these documents are collected individually by visiting many resource centres, like; Department of P. G. Studies and Research in Kannada, Gulbarga University, Kalaburgi and Department of Hasataprati, Kannada University, Hampi. These historical Kannada handwritten manuscripts are captured through Canon 1300D, 18 megapixels DSLR Camera at 5184×3456 resolutions in the JPEG format. There are 121 manuscripts of different dynasties are captured and stored them in a separate file.

B. Segmentation

Image segmentation is a process of dividing an image into multiple regions. This is typically used to identify objects or other relevant information in digital images. For historical manuscript analysis, there are many different ways to perform image segmentation to extracts words or lines. In this work, we have considered the seam carving text line extraction method with the improved algorithm. Seam carving [5] algorithm is originally developed by Shai Avidan for content aware image resizing, it is also called as liquid rescaling. It works by building up various seams in image and consequently expels seams to diminish picture size or embeds seams to expand it. Seams cutting additionally permits manually characterizing regions in which pixels may not be adjusted and include the capacity to expel entire articles from an image. Many of the authors have used this algorithm for image retargeting, but in this work, we tried to improve the algorithm to extract text line in historical Kannada handwritten manuscripts.

C. Pre-processing

It is more important that the historical manuscript documents are badly affected by many factors [1, 2] namely; it contains smear, uneven background illumination, and spot due to age or marks resulting from the ink bleed-through. Apart from this, the style of writing varies from manuscript to manuscript [3], which leads to the confusion and complexity to recognize the historical manuscript documents. The preprocessing steps which include image enhancement and restoration, the enhancement method improves the quality of the image but also removes the unwanted objects, debris, uneven background illumination and noise, etc. In the previous papers, we have proposed a novel image enhancement and restoration technique for degraded historical Kannada handwritten manuscript document images [14, 15].

D. Feature Extraction

The feature extraction process used to obtain the feature vector sequence of individual images which describes the properties of the individual objects. The main contribution of the present work is the application of LBP (Local Binary Pattern) features for recognition of the historical Kannada handwritten scripts. The LBP characterize local image masks using binary codes that extract the relationship between a central pixel and its neighbors. LBP feature extraction usually computes the LBP descriptors at each pixel level of an image to create an image of integer code values, followed by pooling of these codes into a histogram [17].

E. Classification

The image classification mainly works with the numerical properties of an assorted image features with organized classes, the image classification contains two phases i.e., training and testing phase. In the initial training phase, characteristic properties of typical image features are isolated and based on these, unique description of each classification category, i.e. training class, is created. In the subsequent testing phase, these feature-space partitions are used to classify image features. In this experiment, we have used Linear discriminant analysis (LDA), K- nearest neighbor (KNN) and Support Vector Machine (SVM), with k-fold experimentation based on the LBP features with 59 and 19 features for classification of historical Kannada handwritten documents.

The detailed approach of the proposed method is discussed in the form of algorithm, which is described below:

Algorithm for recognition of Historical Kannada Handwritten manuscript

1. Input Camera capture historical Kannada handwritten manuscript of different age-types: namely Hoysala, Vijayanagara and Mysore dynasties.
2. Apply Improved seam carving text line segmentation method for line Extraction from historical manuscript:
   - 2.1. convert the given original colour image to grayscale image
   - 2.2. compute edge image using the Sobel edge detector
   - 2.3. medial seam computation with a projection profile matching
     - 2.3.1. Compute horizontal projection profiles of all edge image slices and find their local maxima
     - 2.3.2. Match local maxima of the projection profiles between two consecutive image slices
     - 2.3.3. Remove lines that start from some intermediate column of the image
     - 2.3.4. Extend the small lines towards the end column of the image
   - 2.4. Separating seam computation with constrained seam carving
     - 2.5.1. apply constrained seam carving for each pair of text lines
     - 2.5.2. compute minimum energy separating seam using dynamic programming

2.5.3. overlay separating seams on the original image

2.5.3.1. Compute the coordinate values of the overlay separating seam and store them in temp

2.5.3.2. Concatenate the present coordinate values with temp

2.5.4. Using coordinate value, extract the region of interest by roipoly() function

2.5.5. Apply the image enhancement technique for binarization to the extracted region of interest (text line)

2.5. Original image overlaid with both types of seams

2.6. Apply the skew correction to the segmented text line

3. Combination of Local Otsu and Global Otsu method is applied to each individual text line for binarizing the images on step 2
4. Apply size normalization to each individual text line on step 3
5. Extract LBP features of size normalized individual text line of different dynasties, namely; Hoysala, Vijayanagara and Mysore dynasties and store them as a knowledge base
6. Apply the classification techniques, namely; LDA classifier, K-nearest neighbour classifier and SVM classifier to classify and recognize the historical Kannada handwritten manuscripts, whether they belong to the Hoysala dynasty or Vijayanagara dynasty or Mysore dynasty?

The detailed approach of the proposed method is given in the Figure 1.

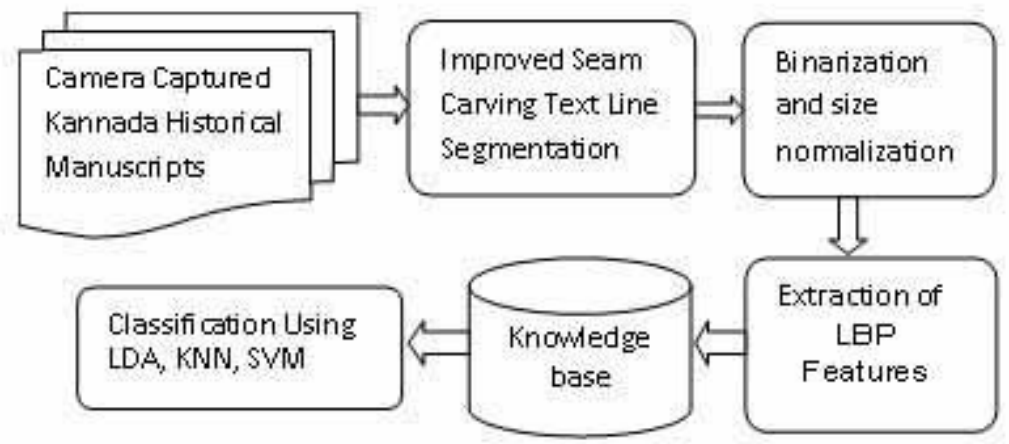


Figure 1. The detailed approach of the proposed method

III. EXPERIMENTAL RESULTS AND DISCUSSION

We have considered the datasets of different dynasties, namely; Vijayanagara(1400 AD and 1460 AD), Hoysala(1340 AD) and Mysore Wadiyar(1936 AD) (described in Sect. II) for experimentation. The experimentation is done with Intel Core i5 system using Matlab R2018b. Input the camera captured historical Kannada handwritten manuscript document images (Figure 2a) for extraction of text line segmentation. To extract text line segmentation, we applied the improved seam carving method which includes the computation of medial seam (Figure 2b), Separating seam computation with constrained seam carving based on medial seam (Figure 2c), overlaid with both type of seam (Figure 2d), region of interest i.e., text line is extracted based on the overlaid seam using roipoly() function (Figure 2e). And then apply the image enhancement method for binarization, restoration and size normalization to each

individual text line (Figure 2f). Extracted the LBP features for all the text lines and store them as a knowledge base. Finally, apply classification techniques; i.e., LDA classifier, K-NN classifier and SVM classifier for classification and recognition of the historical Kannada handwritten manuscripts based on their age-type. The other sample images of the proposed algorithm used for other dynasties namely, Mysore Wadiyar(1936 AD) dynasty, Vijayanagara(1400 AD) dynasty and Hoysala(1340 AD) dynasty, which are shown in Figure 3, Figure 4 and Figure 5, respectively. The results of the text lines extracted from the manuscripts of various dynasties based on their age-type are given in the Table 1. The average classification accuracy of the proposed method is given in the Table 4. Initially, we have extracted the LBP features with 59 features using k-fold experiment and these results are given in the Table 2. Further, to improve the results we have reduced the LBP features with 19 features and certainly results are improved. The results of these reduced features with k-fold experiments using LDA, K-NN and SVM classifiers are given in the Table 3. As per the results, it is observed that LBP features with 19 features are given better results comparatively 59 features. Hence, we propose only 19 features by reduced features from 59 and overall accuracy is calculated and represented only based on 19 LBP features.

The classification accuracy for different dynasties represents that the LDA classifier has yielded 94.2%, K-NN classifier has 94.9% and SVM classifier has 96.4%. Based on the experimentation, it seems that the SVM classifier has got a good classification performance comparatively LDA classifier and K-NN classifier for historical Kannada handwritten manuscript document images. Which indicates better recognition rates towards historical Kannada handwritten manuscript document images.

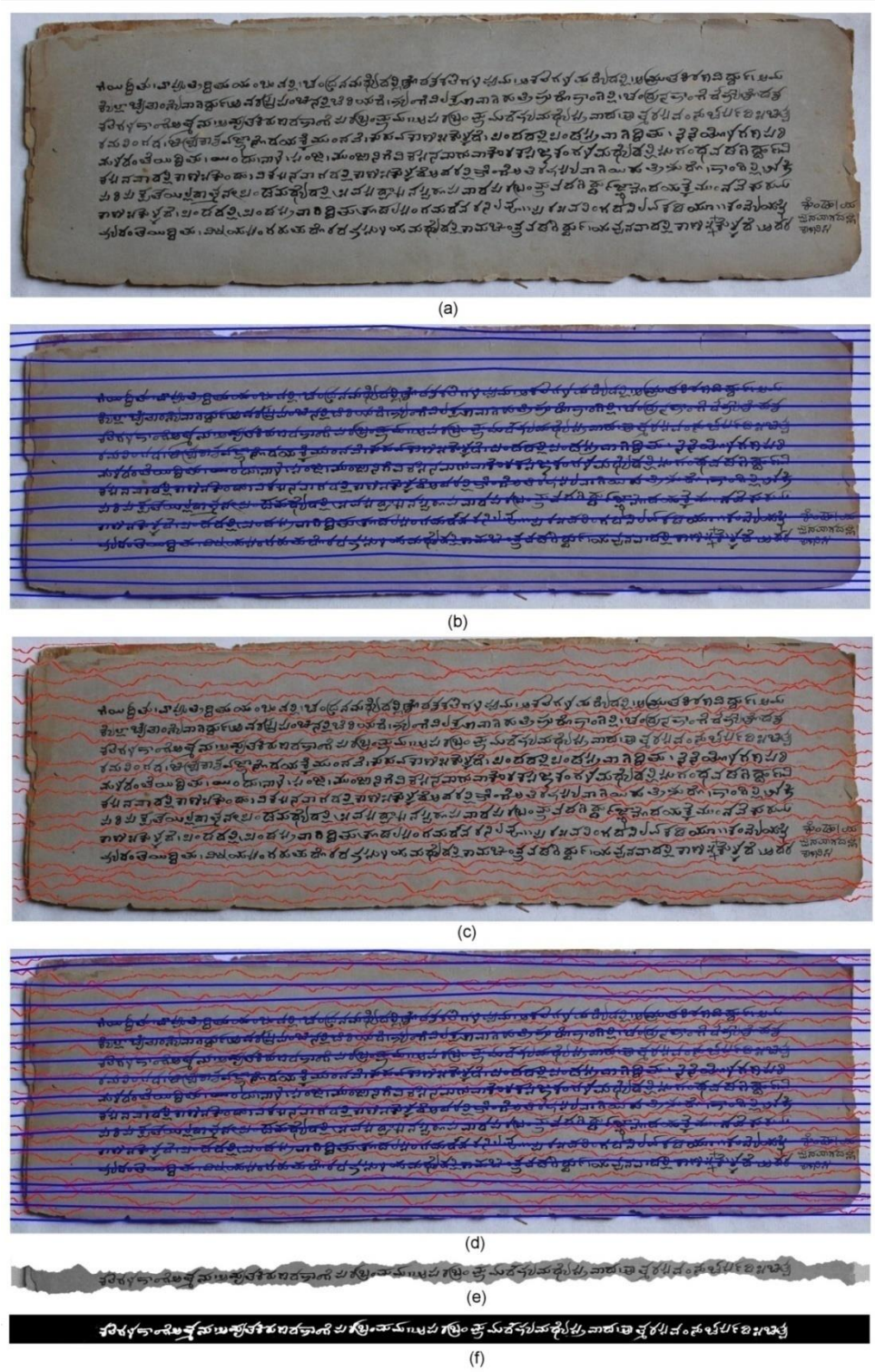


Figure 2. Sample image of the proposed algorithm (a) Original camera captured historical Kannada handwritten manuscript document image of Vijayanagara dynasty (1460AD) (b) medial seam computed image (c) Separating seam computation image with constrained seam carving based on medial seam (d) overlaid image with both type of seam (e) text line is extracted based on the overlaid seam (f) Enhanced and size normalized image

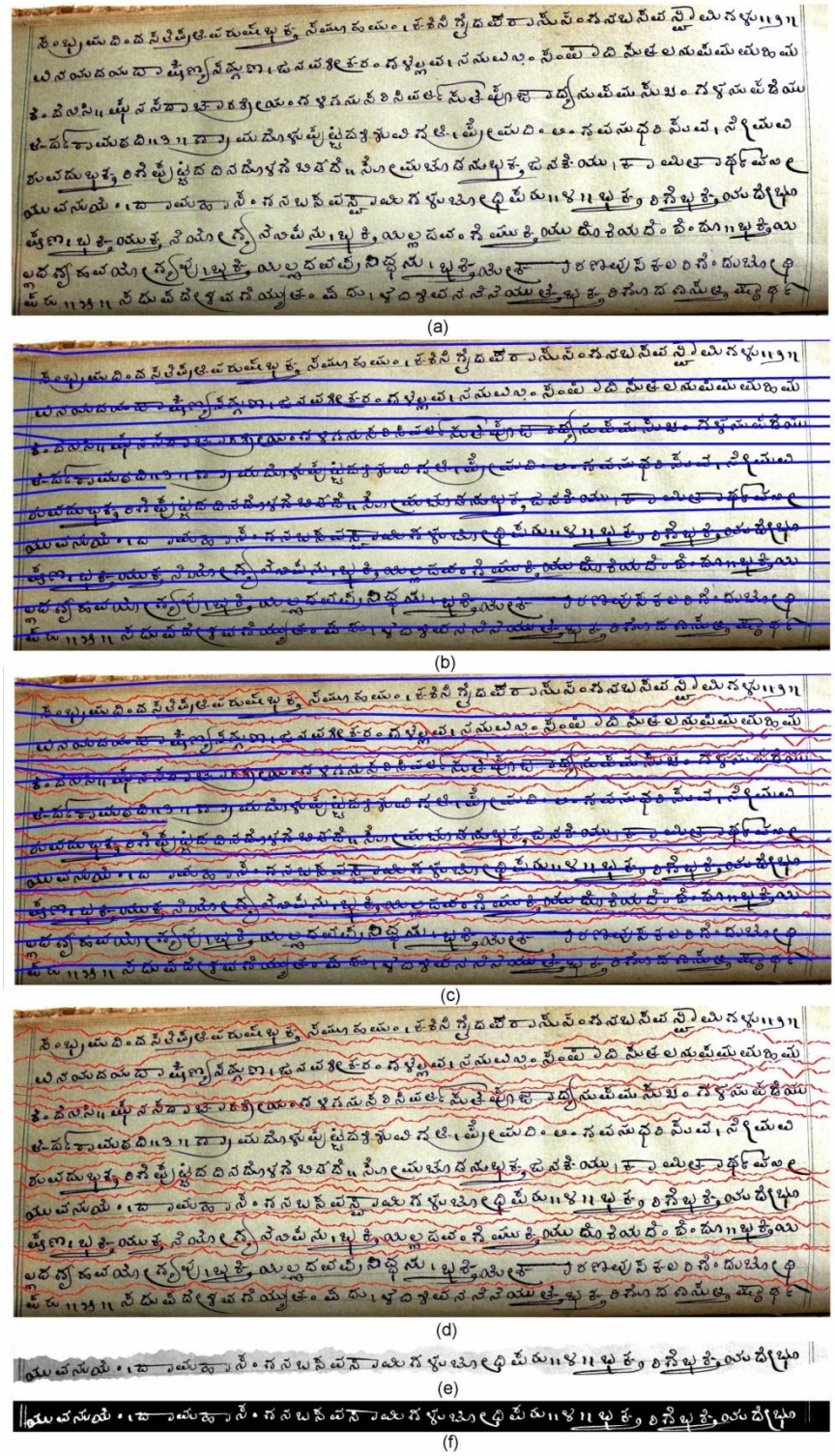


Figure 3. Sample image of the proposed algorithm (a) Original camera captured historical Kannada handwritten manuscript document image of Mysoure wodeyar dynasty (1936AD) (b) medial seam computed image (c) Separating seam computation image with constrained seam carving based on medial seam (d) overlaid image with both type of seam (e) text line is extracted based on the overlaid seam (f) Enhanced and size normalized image

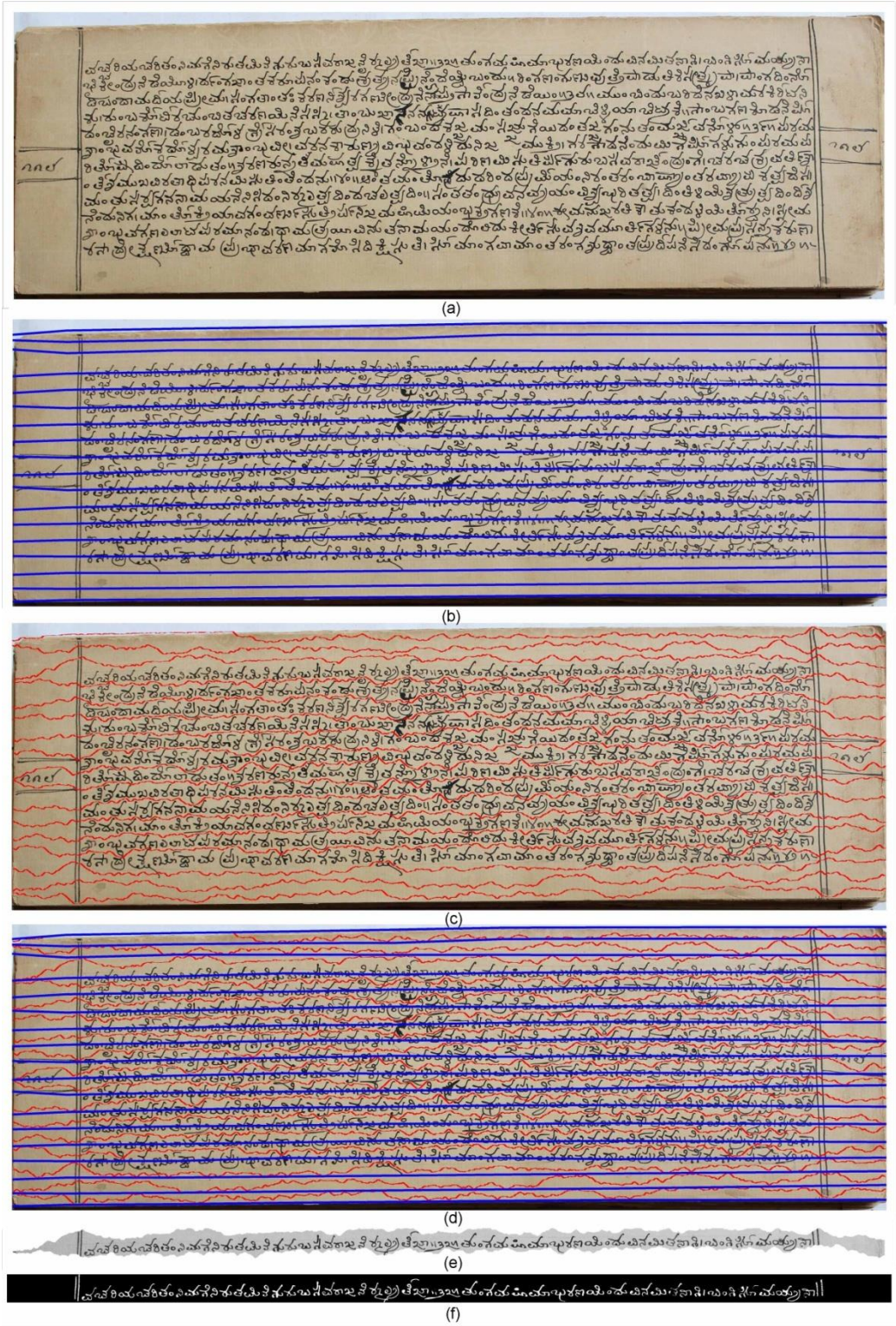


Figure 4. Sample image of the proposed algorithm (a) Original camera captured historical Kannada handwritten manuscript document image of Vijayanagara dynasty (1400AD) (b) medial seam computed image (c) Separating seam computation image with constrained seam carving based on medial seam (d) overlaid image with both type of seam (e) text line is extracted based on the overlaid seam (f) Enhancedand size normalized image

Figure 5. Sample image of the proposed algorithm (a) Original camera captured historical Kannada handwritten manuscript document image of Hoysala dynasty (1340AD) (b) medial seam computed image (c) Separating seam computation image with constrained seam carving based on medial seam (d) overlaid image with both type of seam (e) text line is extracted based on the overlaid seam (f) Enhanced and size normalized image

Table 1. The results of the text lines extracted from the manuscripts of various dynasties based on their age-type

Dynasties	No. of Manuscript Images	Total No. of Text lines in the Manuscript images	Correctly Identified and Segmented text lines	Unidentified Text lines
Vijayanagara (1460)	24	252	250	2
Mysoure Wodeyar (1936)	28	235	224	9
Vijayanagara (1400)	39	674	552	122
Hoysala (1340)	30	336	335	1

Table 2. Classification accuracy of LBP with 59 Features for different k-fold experimentation

Classifiers	5 Fold	4 Fold	3 Fold	2 Fold
LDA	NA	NA	NA	NA
KNN	94.9	94.5	94.8	93.9
SVM	96.4	96.1	95.4	95.7

Table 3. Classification accuracy of LBP with 19 Features for different k-fold experimentation

Classifiers	5 Fold	4 Fold	3 Fold	2 Fold
LDA	94.2	94.1	93.8	93.9
KNN	94.9	94.9	93.5	93.7
SVM	96.4	96.1	94.5	95.9

Table 4. The average classification accuracy of proposed method with LDA, K-NN and SVM classifiers

Dynasties	LDA		K-NN		SVM	
	Recognition Rate	Error Rate	Recognition Rate	Error Rate	Recognition Rate	Error Rate
Vijayanagara(1460)	96	4	98	2	97	3
Mysoure Wodeyar(1936)	89	11	87	13	92	8
Vijayanagara(1400)	97	3	97	3	98	2
Hoysala(1340)	93	7	95	5	97	3
Average accuracy	**94.2%**		**94.9%**		**96.4%**	

IV. CONCLUSION

In this paper, we have proposed an algorithm to identify and recognize the historical Kannada handwritten scripts of various dynasties; namely, Vijayanagara dynasty (1460 AD), Mysore Wadiyar dynasty (1936 AD), Vijayanagara dynasty (1400 AD) and Hoysala dynasty (1340 AD) by using the improved seam carving text line segmentation method with LBP features. For recognition and classification purpose the LDA, K-NN and SVM classifiers are used. The average classification accuracy for different dynasties are computed. The LDA classifier has yielded 94.2%, K-NN classifier yields 94.9% and SVM classifier has 96.4%. Based on the experimentation, the SVM classifier has proved good classification performance comparatively LDA and K-NN classifiers for historical Kannada handwritten script recognition. The experimental results are verified by Epigraphists and language expert, which shows the robustness of the proposed method. The same algorithm may be used for other dynasties with different feature sets, which will be done as future work.

ACKNOWLEDGMENT

The authors are grateful to The Chairman, Department of P. G. Studies and Research in Kannada, Gulbarga University, Kalaburgi and Dept. of Hastaprati, Kannada University, Hampi for providing the historical Kannada handwritten manuscripts and visualization of manual results.

REFERENCES

[1] Manjunath, M.G., Devarajaswamy G.K., "*Kannada Lipiya Vikasa*", Published by Jagadhguru Sri Madhvacharya Trust, Sri Raghavendra Swami Matta, Mantralaya

[2] Narasimha Murthy, A.V., "*Kannada Lipiya Ugama Mattu Vikasa*", Kannada Adhyayana Samsthe, Mysore University, Mysore (1968)

[3] Reddy D., "*Lipiya HuttuMattu Belavanige—Origin and Evolution of Script*", Kannada Pustaka Pradhikara (Kannada Book Authority), Bangalore

[4] Nikolaos Arvanitopoulos and Sabine Susstrunk, "*Seam Carving for Text Line Extraction on Color and Grayscale Historical Manuscripts*" 14th IEEE International Conference on Frontiers in Handwriting Recognition, DOI: 10.1109/ICFHR.2014.127, pp.726-731.

[5] S. Avidan, A. Shamir, "*Seam Carving for Content-Aware Image Resizing*," ACM Transactions on Graphics, vol. 26, no. 3,p. 10, 2007.

[6] Veronica Romero, Joan Andreu Sanchez, Vicente Bosch, Katrien Depuydt and Jesse de Does, "*Influence of Text Line Segmentation in Handwritten Text Recognition*" 13th IEEE International Conference on Document Analysis and Recognition (ICDAR), 2015, 978-1-4799-1805-8/15, pp.536-540

[7] Anguelos Nicolaou, Andrew D. Bagdanov, Marcus Liwickiy, and Dimosthenis Karatzas, "*Sparse Radial Sampling LBP for Writer Identification*", 2015 13th International Conference on Document Analysis and Recognition (ICDAR), doi: 10.1109/ICDAR.2015.7333855, pp.716-720.

[8] Sounak Dey, Anguelos Nicolaou, Josep Llados , and Umapada Pal, "*Local Binary Pattern for Word Spotting in Handwritten Historical Document*", Structural, Syntactic, and Statistical Pattern Recognition. S+SSPR 2016. Lecture Notes in Computer Science, vol 10029. Springer, Cham, pp.574-583.

[9] Miguel A. Ferrer, Aythami Morales and Umapada Pal, "*LBP Based Line-wise Script Identification*", 2013 12th International Conference on Document Analysis and Recognition, doi:10.1109/ICDAR.2013.81, pp.369-373.

[10] Sourav Ghosh, Dibyadwati Lahiri, Showmik Bhowmik, Ergina Kavallieratou and Ram Sarkar, "*Text/Non-Text Separation from Handwritten Document Images Using LBP Based Features: An Empirical Study*", J. Imaging 2018, 4, 57; doi:10.3390/jimaging4040057, pp.1-15.

[11] Maroua Mehri, Pierre Héroux, Petra Gomez-Krämer, Rémy Mullot, "*Texture feature benchmarking and evaluation for historical document image analysis*", International Journal on Document Analysis and Recognition, Springer Verlag, 2017, pp.1-35. doi:10.1007/s10032-016-0278-y.

[12] Laurence Likforman-Sulem, Abderrazak Zahour and Bruno Taconet, "*Text line segmentation of historical documents: a survey*" IJDAR (2007) pp:123–138, DOI 10.1007/s10032-006-0023-z

[13] Abedelkadir Asi, Raid Saabni and Jihad El-Sana, "*Text Line Segmentation for Gray Scale Historical Document Images*", HIP '11 Proceedings of the 2011 Workshop on Historical Document Imaging and Processing, DOI: 10.1145/2037342.2037362, pp. 120-126, 2011.

[14] Parashuram Bannigidad and Chandrashekar Gudada, "*Restoration of Degraded Historical Kannada Handwritten Document Images using Image Enhancement Techniques*", International Conference on Soft Computing and Pattern Recognition (SoCPaR 2016), 2016. pp. 498-508

[15] Parashuram Bannigidad and Chandrashekar Gudada, "*Restoration of Degraded Kannada Handwritten Paper Inscriptions (Hastaprati) using Image Enhancement Techniques*", IEEE International Conference on Computer Communication and Informatics (ICCCI -2017), 2017. Pp.1-6.

[16] Parashuram Bannigidad, Chandrashekar Gudada, "*Identification and classification of historical Kannada handwritten document images using LBP features*", International Journal of Intelligent Systems Design and Computing, Jan 2018, Vol. 2, Issue 2, pp. 176-188.

[17] T. Ojala, M. Pietikainen, and T. Maenpaa, "*Multiresolution gray-scale and rotation invariant texture classification with local binary patterns*," Pattern Analysis and Machine Intelligence, IEEE Transactions on, vol. 24, no. 7, pp. 971–987, 2002.

ALLIANCE INTERNATIONAL CONFERENCE ON ARTIFICIAL INTELLIGENCE AND MACHINE LEARNING (AICAAM)

ALLIANCE UNIVERSITY

Private University established in Karnataka State by Act No.34 of year 2010
Recognized by the University Grants Commission (UGC), New Delhi

BENGALURU

Chikkahagade Cross, Chandapura-Anekal Main Road, Anekal, Bengaluru - 562 106, Karnataka, India

Tel.: +91 80 3093 8100/8200 / 4619 9100 | www.alliance.edu.in