Print ISSN : 2395-1990
Online ISSN : 2394-4099
www.ijsrset.com

Techno Science Academy
The International Open Access Publisher

IJSRSET

National Conference on Advance Research Trends in Information and Computing Technology NCARTICT-2018

Organised by
Information Technology Department,
L.D. College of Engineering (LDCE), Ahmedabad, Gujarat, India

VOLUME 4, ISSUE 2, JANUARY-FEBRUARY-2018

INTERNATIONAL JOURNAL OF SCIENTIFIC RESEARCH IN SCIENCE, ENGINEERING AND TECHNOLOGY

UGC Approved Journal [Journal No : 47147]

Email : editor@ijsrset.com Website : http://ijsrset.com

National Conference on Advance Research Trends in Information and Computing Technology

NCARTICT-2018

20th January, 2018

Organised by

Information Technology Department,

L.D. College of Engineering, Ahmedabad, Gujarat, India

Published By

International Journal of Scientific Research in Science, Engineering and Technology

Print ISSN : 2395-1990 Online ISSN : 2394-4099

UGC Approved Journal [Journal No : 47147]

Volume 4, Issue 2, January-February-2018

Technoscience Academy

(The International Open Access Publisher)

[www.technoscienceacademy.com]

Sponsored By

Gujarat Council on Science and Technology

Department of Science & Technology

Gandhinagar, Gujarat, India

Chife Patron

Dr. G.P. Vadodaria
Principal, LDCE, Ahmedabad

Coordinator

Prof. Hiteishi M. Diwanji
HOD, Information Technology Department,
LDCE, Ahmedabad

Co-Cordinator

Prof. (Dr.) Mehul C. Parikh
Associate Professor (IT)
Information Technology Department,
LDCE, Ahmedabad

National Advisory Committee

Prof.(Dr.) Navin Chaudhary (MPUAT, Udiapur)	Prof.(Dr.) Manoj Gor (MNIT, Jaipur)
Prof. B.V. Buddhdev	Prof. M.B. Chaudhary
Prof. (Dr.) S.M. Shah	Prof. (Dr.) M.V. Shah
Prof. Usha Neelkantan	Prof. (Dr.) K.G. Maradia
Prof.(Dr.) Amit Gantra	Prof.(Dr.) H.B.Patel

Technical Review Committee

Prof.(Dr) D. A. Parikh	Prof. (Dr.) Vibha Patel,
Prof. (Dr.) C.S. Thaker	Prof (Dr.) Vimal Vaghela,
Prof. T. J. Raval	Prof. B.B. Panchal,
Prof. M.T. Savalia	Prof. (Dr.) Mahesh Goyani,
Prof.(Dr.) Amit Thakkar	Prof. Tushar Chapaneria
Prof. J.S. Dhobi	Prof. (Dr.) Keyur Brahmbhatt,
Prof.(Dr.) Hemal Shah	Prof. Pradip Patel
Prof. C.A. Patel	Prof. Ankit Patel
Prof. (Dr.) S.D. Panchal	Prof. Swati Patel

About the Institute

Lalbhai Dalpatbhai College of Engineering, Ahmedabad is a premier engineering institute in Gujarat State set with the objectives of imparting higher education, research and training in various fields of engineering & technology. The institute is affiliated to Gujarat Technological University, Ahmedabad and administrated by Department of Technical education, Gujarat State, Gandhinagar. The institute was established in June 1948 with a generous donation of Rs. 25 lacs and 31.2 Hectres of land by the textile magnate Sheth Shri Kasturbhai Lalbhai. Hence College is named as Lalbhai Dalpatbhai College of Engineering. It is situated adjacent to Gujarat University buiding and is located at the nucleus of various national level institutes such as PRL, ATIRA, IIM etc. The campus is having buildings for various departments, offices, hostels, residences for Principal, rector and wardens. The plinth area of college and hostel buildings admeasures 45,220 sq.m. and 12,556 sq.m. respectively.

About the Conference

L.D. College of Engineering and I.T. Department always believes in providing a platform for sharing and nurturing research. This conference is aimed to Provide platform to Researchers, research scholars, Academicians, industry persons to present their Innovative idea, work and exchange views with Experts. This conference will provide

directions to all participants how to go ahead with quality work. It will also provide common platform to learn from each other's experience and knowledge. In current scenario, sound research work requires expertise and knowledge of core engineering as well as fundamental The wide spectrum of conference will make dais available to researchers working in Information technology and computer engineering area to look for experts of various domains of their interest to discuss their work and shape their future directions for smart and sustainable growth of our country.

TOPIC of Interest

- ✓ Cloud Computing, Grid Computing, IOT
- ✓ Digital Image Processing
- ✓ Data Ware Housing, Data Mining, Big Data
- ✓ Software Engineering
- ✓ Mobile Computing
- ✓ Neural Network, Machine Learning and Other Computer IT related topics

CONTENTS

© 2018 IJSRSET | Volume 4 | Issue 2 | Print ISSN: 2395-1990 | Online ISSN : 2394-4099
National Conference on Advanced Research Trends in Information and
g Technologies (NCARTICT-2018), Department of IT, L. D. College of Engineering,
Ahmedabad, Gujarat, India In association with
International Journal of Scientific Research in Science, Engineering and Technology

Critical Analysis of Cryptography and Steganography

Alpa Agath*, Chintan Sidpara, Darshan Upadhyay
*I.T. Department, V.V.P. Engineering College, Rajkot, Gujarat, India

ABSTRACT

Digital communication has become an essential part of infrastructure in the present world and it has witnessed a noticeable and continuous development in a lot of applications during last few decades. Nowadays, almost all applications are Internet-based and it is important that communication be made confidential and secure. Cryptography and steganography are the two important and widely used techniques that are used to provide information security over the open and insecure networks such as Internet. Cryptography distorts the original message itself whereas steganography hides the existence of the message. This paper gives an overview about the concepts of cryptography and steganography. Moreover, it presents a fair comparative analysis between various selected encryption algorithms on various parameters such as key size, block size, speed of encryption, level of security provided by algorithm, and memory usage. It also carries out comparison between traditional steganography methods and Hex Symbol Steganography method on basis of basic parameters such as capacity of carrier file to hide information, robustness and amount of security provided. The comprehensive analysis shows that AES and Hex Symbol Steganography provides more level of security and are robust in nature as compared to other competitors, hence providing more confidentiality.

Keywords: Information Hiding, Cryptography, Steganography, Advanced Encryption Standard (AES), Hex Symbol Steganography.

I. INTRODUCTION

In today's era, due to tremendous growth of networking technologies an enormous amount of data is being exchanged over the Internet as a result of which security of information being conveyed over the Internet is becoming more significant as sensitive data needs to be transferred securely over the internet while maintaining its confidentiality, integrity and availability [1].

To maintain the privacy and security of confidential and sensitive information there is a need of approaches which enhances the level of information security. Information hiding is one of the many available approaches which increase the level of information security. The most powerful and widely used approaches of information hiding used to contravene the threats to information security are Cryptography and Steganography [1]. Cryptography provides security by manipulating the original confidential information so that it becomes unintelligible for the intruders.

Steganography conceals the existence of communication by embedding the confidential information in some other cover medium (e.g. image, audio, video, etc.).
Cryptography is used nowadays in almost all applications that use Internet as means of communication. Real time applications of cryptography include ATM machines; password protection of email passwords, social account (Facebook, twitter, etc.) passwords; E-commerce; Defence forces; intelligent agencies.

Steganography is used to overcome the shortcomings of cryptography and support the cryptography techniques to provide better and more efficient information security. Areas where steganography is used include bank and commercial organizations, digital watermarking, E-commerce, military, and the areas where cryptography is used.

The remainder of the paper is organized as follows: Section II introduces a brief note on cryptography and steganography. Classification of cryptography and a de-

tailed comparative analysis of selected encryption algorithms are given in section III. Classification of steganography and a comprehensive comparative analysis of traditional steganography methods and hex symbol steganography are given in section IV. Finally, a brief conclusion with future work is given in Section V.

II. STATE OF ART: CRYPTOGRAPHY AND STEGANOGRAPHY

Cryptography and steganography are the two most widely and commonly used approaches to secure information being transmitted over the Internet either by encoding the information or by hiding the information [4].

A. Cryptography

Cryptography is the art and science of fabricating methods or algorithms that allow transmission of data in a secure manner by transforming the readable and understand-able data into irrational and unfathomable data in such a way that only the intended person is able to retrieve the exact original data from the data being transmitted [2][3].

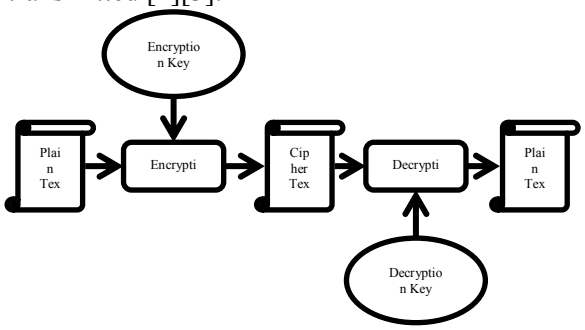


Figure 1. Process of Cryptography [2]

B. Steganography

Steganography is the art and science of camouflaging information into covert channels, hence preventing the detection of the camouflaged information from the eaves-droppers [5].

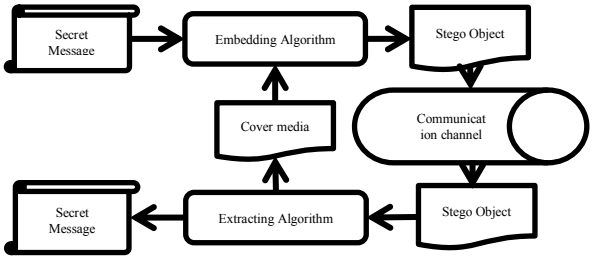


Figure 2. Process of Steganography [5]

Steganography is also known as "Disappearing Cryptography" [6].

III. CLASSIFICATION OF CRYPTOGRAPHIC ALGORITHMS

Encryption algorithms in cryptographic systems can be classified into different categories depending on the number of keys used to encrypt the plain text as shown in Fig. 3.

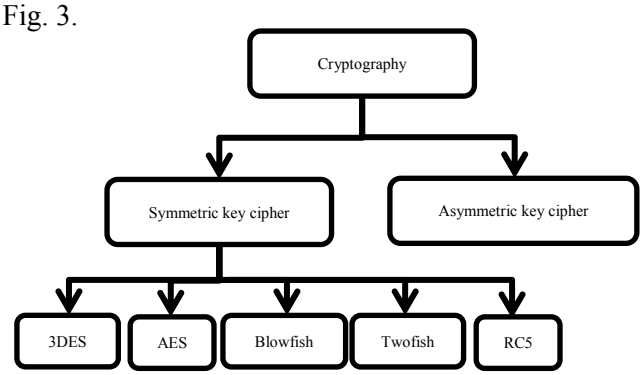


Figure 3. Classification of cryptography [11]

A. Comparison of Symmetric Encryption Algorithms

Comparison of various symmetric encryption algorithms on the basis of some selected parameters is shown in table 1.

Table 1. Comparison of Encryption Algorithms [7] [9] [10] [11]

Encryption Algorithm	Block Size (in bits)	Key Size (in bits)	Number of Rounds	Encryption Speed	Memory Usage	Flexibility	Level of Security
3DES	64	112,168	48	Very Slow	Moderate	Yes	Adequate
AES	128	128,192,256	10, 12, 14	Very Fast	Low	Yes	Excellent
Blowfish	64	128-448	16	Fast	High	Yes	Excellent
Twofish	128	128, 192, 256	16	Fast	Low	Yes	Good
RC5	34, 64, 128	128	1-255	Slow	Low	No	Good

The above comparison shows that AES provides higher level of confidentiality to sensitive information.

B. Advanced Encryption Standard (AES)

AES encryption algorithm is a symmetric block cipher that uses a secret encryption key and several numbers of rounds to encrypt the sensitive information (plain text) published by National Institute of Standards and Technology (NIST) in 2001.

1) Encryption Process of AES:

- AES deals with fixed size block of 128 bits or 16 bytes in length which is represented in 4x4 matrixes of bytes known as state array, which is modified at each round of encryption and decryption.
- The key provided as input is depicted as a square matrix of bytes and is then expanded into an array of forty-four 32-bit words, w[i]. Four distinct words (32bits * 4 = 128 bits) serve as round key in each round of the encryption and decryption.
- Based on the length of secret key used (128, 192, 256) for the encryption, the number of rounds i.e., N (10, 12, 14) in the cipher will differ accordingly. The single round of AES encryption process is shown in Fig. 4.
- First N-1 round in the AES cipher structure consists of the four basic transformations that encrypt the plain text of 128 bits in length of which one performs permutation and three perform substitution. The transformations are as follows:
  - AddRoundKey: It is a simple bitwise XOR operation of the current state array with a portion of the expanded key.
  - Substitution bytes: It uses an S-box to perform a byte-by-byte substitution of the state array.
  - ShiftRows: It is a simple permutation.
  - MixColumns: It consists of substitution operation that makes use of arithmetic over GF (28).
- The final Nth round contains only three transformations, and there is an initial single transformation i.e., AddRoundKey before the first round, which can be considered as round 0.
- The cipher begins and ends with an AddRoundKey transformation because only the AddRoundKey makes use of the key. The other three transformations together provide confusion, diffusion, and nonlinearity, but no security since they don't use the key.
- Each transformation in a round takes one or more 4x4 matrices or state array as an input and produces a 4x4 matrix as output. The output of the final N^{th} round will produce the cipher text.
- Each transformation is easily reversible. For the Substitute Byte, ShiftRows, and MixColumns stages, an inverse function is used in the decryption algorithm. For the AddRoundKey stage, the inverse is achieved by performing XOR with the same round key to the block (A XOR A XOR B = B).
- The decryption process in AES will use the expanded keys in the reverse order as used in the encryption. The decryption algorithm is not identical to the encryption algorithm. The order of transformations used in each round of decryption is different from that used in encryption.

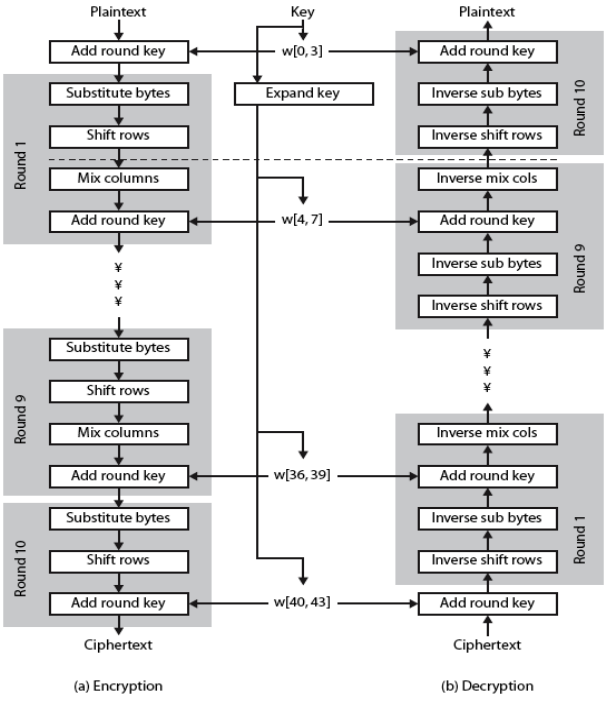


Figure 4. AES Encryption and Decryption Process [8]

IV.CLASSIFICATION OF STEGANOGRAPHY

Steganography can be classified according to type of cover medium used to hide the secret message as shown in Figure 5.

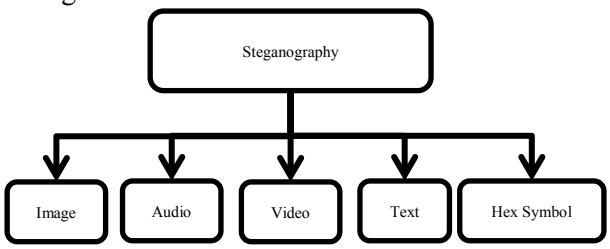


Figure 5. Classification of Steganography [5] [14]

A. Comparison of Traditional Steganography methods and Hex Symbol Steganography

Comparison between traditional steganography methods and hex symbol steganography is shown in table 2 on the basis of primary steganography measures.

Table 2. Traditional Steganography methods vs. Hex Symbol Steganography [14]

Steganography Measures	Traditional Steganography Methods	Hex Symbol Steganography
Imperceptibility	Concealed data can be recognized in the form of disruption in sound and video files, changes in image frames and colors.	The secret data will be embedded into hex symbols, so it is impossible to recognize by human eyes.
Capacity	Limited capacity to embed secret data.	More capacity to embed secret data as compared to other methods due to the nature of hex symbols.
Robustness	Less robust in nature.	More robust.
Security	Adequate	Excellent
Codes used to conceal secret data	Binary codes	Hex symbol codes

The above comparison shows that Hex Symbol Steganography method is more robust in nature and provides higher level of security (confidentiality) than traditional methods such as image, audio, video steganography.

B. Hex Symbol Steganography (HSS):

Hex Symbol steganography uses hex symbol carrier files to conceal the confidential information instead of digital multimedia carrier files such as image, audio, video and text files [14]. It uses the hex symbol codes to embed the secret information unlike the traditional steganography methods Which use binary codes to conceal the secret information [14].The stego object obtained after embedding the secret message in the hex symbol carrier file will be unfathomable (difficult to understand) by the intruders who are trying to gain access of the confidential information stored in the secret message.

1) Hex Symbol Algorithm:

- Prior to communication the authorized parties i.e., both the sender and receiver will decide certain patterns which will act as key to embed the secret information in the carrier file.
- The patterns used to conceal the secret information will be created by converting a chosen carrier file into hexadecimal symbols, segmenting the resulting hexadecimal carrier file with size of each segment being a 16x16 matrix, and then numbering the matrices (segments) sequentially [18] as shown in Figure 6.

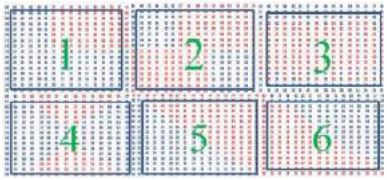


Figure 6. Segmented Hex Symbol Carrier File [14]

- Some of the segments of the hexadecimal carrier files are used to conceal the secret information on the basis of the pattern chosen by the sender and receiver.
- A codebook is prepared which consists of a letter representing the pattern with the sequence of segment numbers that are used to embed the secret message [14]. The letter representing the pattern will act as the stego key in the process of steganography. The example of code book is shown in Table 3.

Table 3. Example of Shared Secret Codebook [14]

Key Symbol	Selected Random Sequence

S	524316
O	643125
M	365124
Y	513642

- Once the codebook is prepared, it is shared between the sender and receiver in a secure way and then the secret message containing the confidential information can be embedded in the carrier file.
- The secret message is first converted into its equivalent hexadecimal representation.
- The resulting hexadecimal values for each character are then inverted in order to increase the security. For example if letter 'n' is represented as 68 in hexadecimal, then it is inverted as 86. Now 86 will represent the letter 'n' in the secret message [14].
- The content of the resulting secret message is concealed into the hex symbol carrier file according to the chosen key symbol and pattern from the codebook shared between the sender and the receiver.
- The contents of the matrix segments used for embedding the secret message are relocated by exchanging the rows with the columns of the matrix in order to increase the complexity of the embedding process [14].
- Once the entire secret message is concealed in the hex symbol carrier file according to chosen pattern, the resultant stego file is concatenated with the key symbol (which represents the pattern of segments used in embedding process) and sent to the authorized receiver.
- The hex symbol steganography with its embedding process and extracting process is shown in Fig. 3.5.1 (ii).
- The receiver will receive the stego file containing the confidential information along with the key symbol which indicates the chosen pattern used for embedding the secret message.
- On receiving the key symbol, the receiver will now be able to recognize the arrangement of the matrix segments by referring to the secret codebook.

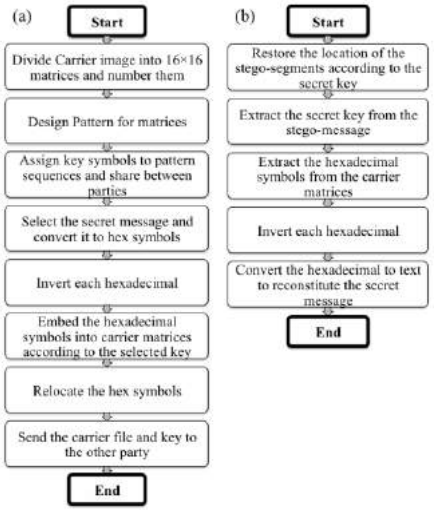


Figure 7. (a) Embedding of the secret message in carrier file by the sender (b) Extracting of the secret message from stego file by the receiver [14]

- Once the receiver is able to recognize the pattern, steps of the embedding process can be executing in the reverse order so as to extract the secret message from the stego file.

V. CONCLUSION AND FUTURE WORK

The comprehensive analysis shows that among the symmetric encryption algorithms, Advanced Encryption Standard (AES) and among the steganography methods, Hex Symbol Steganography is more efficient and boosts the level of confidentiality when compared with their respective competitor.

Future contribution will focus on proposition of an approach which consists of merging cryptography and steganography into a single security system and hence enhancing the level of information security by providing high level of confidentiality while maintaining the privacy and secrecy of information stored and transmitted through insecure communication channels.

VI. REFERENCES

[1] S. Almuhammadi and A. Al-Shaaby, "A Survey on Recent Approaches Combining Cryptography and Steganography", Computer Science & Information Technology (CS & IT), 2017.

[2] S. Mishra, P. Pandey, "A Review on Steganography Techniques Using Cryptography", International Journal of Advance Research In

Science And Engineering, Volume 4, Special Issue (01), 2015.
[3] M. Pandey and D. Dubey, "Survey Paper: Cryptography The art of hiding Information", International Journal of Advanced Research in Computer Engineering & Technology (IJARCET), Volume 2, Issue 12, 2013.
[4] K. Rahmani, K. Arora and N. Pal, "A Crypto-Steganography: A Survey", International Journal of Advanced Computer Science and Applications (IJACSA), Volume 5, No. 7, 2014.
[5] P. Joseph and S. Vishnukumar, "A Study on Steganographic Techniques", Proceedings of Global Conference on Communication Technologies (GCCT), IEEE, 2015.
[6] http://io.acad.athabascau.ca/~grizzlie/Comp607/menu.htm
[7] M. Kumar, V. Kumar and A. Sharma," A Survey on Various Cryptography Techniques", International Journal of Emerging Trends & Technology in Computer Science (IJETTCS), Volume 3, Issue 4, 2014.
[8] A. Abdullah, "Advanced Encryption Standard (AES) Algorithm to Encrypt and Decrypt Data", Cryptography and Network Security, 2017.
[9] G. Yadav and A. Majare," A Comparative Study of Performance Analysis of Various Encryption Algorithms", International Conference on Emanations in Modern Technology and Engineering (ICEMTE), Volume: 5 Issue: 3, 2017.
[10] D. Talukdar and L. Saikia," A Review On Different Encryption Techniques: A Comparative Study", International Journal of Engineering Research and General Science, Volume 3, Issue 3, 2015.
[11] S. Swathi, P. Lahari and B. Thomas," Encryption Algorithms: A Survey", International Journal of Advanced Research in Computer Science & Technology (IJARCST), Volume 4, Issue 2, 2016.
[12] N. Singh, "Survey Paper on Steganography", International Refereed Journal of Engineering and Science (IRJES), Volume 6, Issue 1, 2017.
[13] A. Rashid and M. Rahim, "Critical Analysis of Steganography "An Art of Hidden Writing"", International Journal of Security and Its Applications, Volume 10, No. 3, 2016.
[14] S. Asbeh, H. Al-Sewadi, S. Hammoudeh and A. Hammoudeh, "Hex Symbols Algorithm for Anti-Forensic Artifacts on Android Devices", International Journal of Advanced Computer Science and Applications (IJACSA), Volume 7, No. 4, 2016.

© 2018 IJSRSET | Volume 4 | Issue 2 | Print ISSN: 2395-1990 | Online ISSN : 2394-4099
National Conference on Advanced Research Trends in Information and
g Technologies (NCARTICT-2018), Department of IT, L. D. College of Engineering,
Ahmedabad, Gujarat, India In association with
International Journal of Scientific Research in Science, Engineering and Technology

A Survey on Social Media Security

Khevana Shah[1], Prof. Mital Panchal[2]
Information Technology, L.D. College of Engineering, Gujarat Technological University, Ahmedabad, Gujarat, India

ABSTRACT

A user authentication scheme that uses any form of social knowledge, utilizes users' trust relationships, monitors users' social contexts, or records users' friend associations for granting or denying access to any resource is considered a social authentication scheme. "The direct or indirect utilization of social knowledge or trust relationships in human-computer authentication systems deployed in online or offline contexts." In this study we analyze the security of social media in basic prospective of type of authentication, possible threats and terms to protect authentication from that threat.

Keywords: Social Media, Security, Session Hijacking, Cookie

I. INTRODUCTION

Social media security is the process of analysing dynamic social media data in order to protect against security and business threats.Influencing factor in use of social media
(I) Ease of use, (II) Globally acceptable and (III) Used for faster, easier, better communication.

Type of data shared on social media can be of Profile data,
Pictorial data and Activity data. The main reason for the security risk can be of multiple degree depending on the type of data and priority for the person holding the same. The main reasons for this are (I) Third party application access, (II) Poor implementation and (III) Lack awareness

A user authentication scheme on social media, utilizes users' trust relationships, monitors users' social contexts, or records users' friend associations for granting or denying access to any resource is considered a social authentication scheme. "The direct or indirect utilization of social knowledge or trust relationships in human-computer authentication systems deployed in online or offline contexts."

In this study we analyse the security of social media in basic prospective of type of authentication, possible threats and terms to protect authentication from that threat.

The rest of this paper proceeds as follows. In Section 2, we present our secure cookie scheme in detail. In Section 4, we discuss the implementation of our secure cookie scheme and its performance. In Section 5, we review and examine existing cookie schemes. We give concluding remarks in Section 6.

II. METHODS AND MATERIAL

Many collaborative websites and social media networks utilize session cookies as a cheaper alternative to the wide utilization of the secure HTTPS protocol. The unprotected nature of cookies can compromise the collaborative environment. Evidently, the availability of social networks and collaboration websites where access to the website is extended to long durations has made this issue even more pressing. Although using a secure protocol (e.g. HTTPS) to connect to the web provides higher levels of security, it is not always applied by many web servers and is replaced by cookie protection. The nature of cookiesThe issue of session hijacking or 'side jacking' due to sniffing out of Internet cookies is one of the important Internet security concerns. Session

hijacking results from unlawful control over cookies during an ongoing internet session in an unprotected network where plaintext traffic is unencrypted. Cookies are vulnerable to attacks, which makes their current deployment questionable and warrants a search for more reliable and secure techniques. Several researchers have tried to solve the vulnerability of cookies. For example, the use of an external proxy where authentication and sensitive information management is carried out completely at the proxy or some other external device. However, this solution's implementation can pose difficulties as it might not be optimal in all situations. Specifically, if a user does not have access to the proxy for any reason or in case the external device is not available at the time when the service is desired (e.g. cell phone battery dead, no coverage…etc.), he will not be able to use the service.

A. Social Authentication: A Definition

As we focus on reviewing the user authentication schemes that leverage information extracted from users' social contexts or intermediate humans in their identity verification processes, all human-computer authentication techniques that rely on eliciting unique characteristics from individuals' social interactions with others are considered social authentication schemes.

B. Social Authentication Techniques:

A Review Previously developed socially aware authentication systems have either leveraged social knowledge to authenticate users

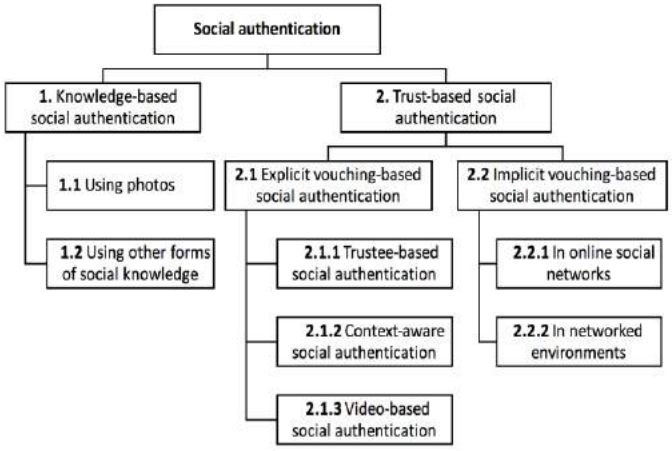


Figure 1. Taxonomy of social authentication schemes in the literature [1]

1) Knowledge-Based Techniques: Knowledge-based social authentication mechanisms rely on the design of the security questions that ask the user about his/her social context, such as social relationships, conversations, or shared knowledge. These mechanisms may require the user either to recognize or recall some information about people he/she knows. For challenge questions to be effective, their answers should be easily memorized, difficult to guess, and unreachable by unauthorized users. Techniques that can be built on top of Facebook's photo-based two factor social authentication system few research attempts have focused on investigating other types of social knowledge that can be uniquely used for authenticating people. This paucity could be linked to the difficulty in analysing individuals' social contexts, identifying the private social information that users can easily remember, and measuring the accuracy and uniqueness of social data. Since currently used knowledge-based social authentication schemes, which were developed as variants of some existing two-factor authentication systems (e.g. Facebook's Login Approval and Google 2-step authentication techniques)

As a sum of all techniques we can accumulate the list of all as (I) Photos Questionaries' (Node, Edge, Pseudo edge), (II) Vouch ID Choose trustees Trusted Contact tracking gestures and motions using wearable sensors Video Notarization Process and (III) User's digital certificate, Biometrics, PINs, sensor data, and e-mails.[1]

A novel protocol for protecting transmitted cookies using two dimensional one-way hash chains. In the first dimension, there is a hash chain that computes secret values used in the second dimension hash function. The optimal lengths of the chains are derived when the number of transactions in the session is known by adopting the position-indexed hashing protocol, energy consumption is reduced significantly especially with longer sessions making our protocol ideal for battery operated devices. Once the authentication credentials are used, they are recycled and never used again.

C. One-way Hash Cookie (OHC) Protection

Since we are using the one-way hash cookie protection Scheme as the backbone for our solution, it is worth illuminating its main aspects and how its hashing operation is carried out to protect cookies. In the OHC scheme, a one-way hash chain of length N is used to protect a stream of N transactions of a web session. During the initial HTTPS login step, the server and the client exchange a shared secret value S0, and a value N

which refers to the chain length or number of transactions expected to be handled during a session. The OHC protects the jth transaction by computing an authentication token VJ=HN-j+1(S0), where the notation Hm(x) implies applying the hash function m times, for example, H2(x) = H (H(x)). For instance, if N=100, then
The authentication tokens for the 1st, 2nd, and 3rd transactions are V1=H100 (S0), V2=H99 (S0), V3=H98 (S0), respectively.

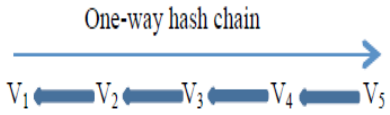


Figure 2. One-way hash chain [2]

Figure 2 illustrates how the one-way hash chains are configured. The straight arrow going from the left to the right corresponds to the length of the chain. In this specific figure, the length is 5 transactions. The small arrows going from the right to the left represent the points where authentication tokens are generated and checked. At each point in the hash chain, the server and client must be able to derive the same value of the authentication token. [2]

This solution achieves its goal by utilizing two one-way hash chains; one is responsible for updating the secret key and the other for creating the authentication tokens attached to the cookies using the secrets produced by the first chain. Use of SHA-1 for hash function to produce authentication tokens. Energy consumption is largely influenced by the cryptographic hash function used in the authentication scheme.

Table 1. Methodology with proposed one

Method	Description
HMAC (m ; k) Keyed-Hash Message Authentication Code	message m using key k
	Sk Server Key
	(m)k Encryption of message m using key k
High level confidentiality	Replay attack
	To provide authentication, integrity, and anti-replay.
HMAC (user name, expiration time, data, session ID, sk)	The cookie becomes session specific For example, in javex.net.ssl package, the function getID ()
	SSL session IDs are easier to obtain than SSL session keys.

The state-of-the-art secure cookie schemes was described by Fu et al. in their seminal paper [5]. In this section, we first examine this scheme, which we refer as Fu's cookie scheme. We show that this scheme has three major limitations, and we give a solution to each of them. Finally, we present our secure cookie scheme as shown in table 1.

III. RESULTS AND DISCUSSION

We prove that our cookie scheme is secure. First, our cookie scheme achieves authentication. A cookie created using our scheme can be used as an authentication token because no one can forge a cookie without knowing the server key sk, which is only known to the server. Note that HMAC is a one-way collision resistant hash function.

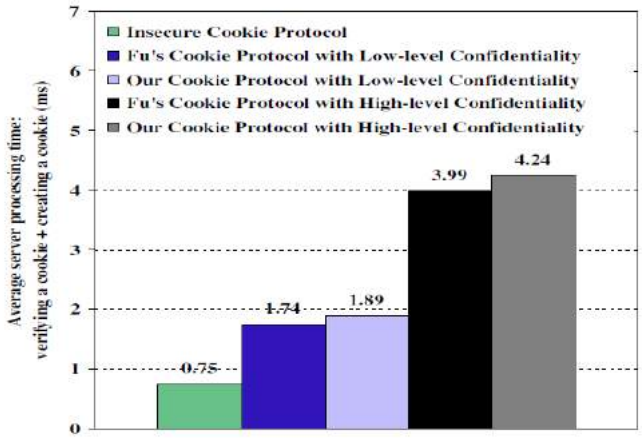


Figure 3. Server side performance comparison [2]

Second, our cookie scheme achieves high-level confidentiality. No one can obtain the key k for decrypting data without knowing the server key sk.

Third, our cookie scheme is secure against replaying attacks. Each SSL session ID is uniquely generated by a server; thus, the stolen cookie in one SSL session is invalid in another SSL session as the session IDs are different.

Fourth, for the user who receives a cookie from a server, from the hash HMAC (username, expiration time, data, session ID, k), they cannot infer any information about the data and the server key k because HMAC is a one-way collision resistant hash function. In other words, the user will not be able to choose a user name and even data that will allow them to infer the server key k. Note

that SSL session ID is sent in clear and scheme is secure against volume attacks because the data encryption key is used only in one SSL session.

IV.CONCLUSION

Cookies are the essential part for social media site to be more user friendly and user personalized. But the security of information residing in it is the questionable thing which should be analyse with appropriate solutions available. The time complexity and overhead are two concerns in which more evolution scope is exist which can be used as future work of research and analysis.

V. REFERENCES

[1] Noura Alomar, Mansour Alsaleh, Abdulrahman Alarifi, "Social Authentication Applications, Attacks, Defense Strategies and Future Research Direction : A systematic review": IEEE , 2016 DOI 10.1109/COMST.2017.2651741

[2] Amerah Alabrah, Mostafa Bassiouni "Robust and Fast Authentication of Session Cookies in Collaborative and Social Media Using Position-Indexed Hashing" IEEE, 2013

[3] Alex X. Liu, Jason M. Kovacs , Mohamed G. Gouda "A secure cookie scheme" elsevier, 2012 Department of Computer Science and Engineering, Michigan State University, East Lansing, MI 48824-1266, USA Exis Web Solutions Department of Computer Sciences, The University of Texas at Austin, Austin, TX 78712-0233, USA DOI: 10.1016/j.comment.2012.01.013

[4] Joon S. Park and Ravi Sandhu, George Mason University "Secure cookie on the web" IEEE, 2002

[5] Chuan Yue, Mengjun Xie, Haining Wang "An automatic HTTP cookie management system", ELSEVIER, 2010 Department of Computer Science, The College of William and Mary, Williamsburg, VA23187, United States DOI: 10.1016/j.comment.2010.03.006

[6] Paul Rabinovich "Secure cross-domain cookies for HTTP"
Springer,2016 Security Software Development, Exostar, Herndon,USA

© 2018 IJSRSET | Volume 4 | Issue 2 | Print ISSN: 2395-1990 | Online ISSN : 2394-4099
National Conference on Advanced Research Trends in Information and
g Technologies (NCARTICT-2018), Department of IT, L. D. College of Engineering,
Ahmedabad, Gujarat, India In association with
International Journal of Scientific Research in Science, Engineering and Technology

A Survey of Sign Recognition Approaches for Indian Sign Language

Twinkal H. Panchal[1], Pradip R. Patel[2]

[1]M.E. Student (Information Technology Department), L.D. College of Engineering, Gujarat Technological University, Ahmedabad, Gujarat, India

[2]Assistant Professor (Information Technology Department), L.D. College of Engineering, Gujarat Technological University, Ahmedabad, Gujarat, India

ABSTRACT

Sign Language (SL) is a communication tool for deaf & dumb people. It is a subset of gestures made with fingers, hands, and face etc. Each gesture in SL has a particular meaning & that is assigned to it. Deaf person directly not communicate with the normal person because normal person never try to learn the sign language. To solve this problem, there exists a need of system that can recognize gesture. Different country has different sign languages. For India, this is called as "Indian Sign Language (ISL)". Only little research work has been carried out in this area for ISL. Several methods have been used to recognize of ISL alphabets and numerals. Many of them method is used to recognize static gesture. Only few works have been reported for dynamic gesture recognition of ISL. The mainly four steps are involved to recognize the sign: gesture acquisition, tracking and segmentation, feature extraction and gesture recognition. This paper presents a survey on various sign recognition approaches for ISL.

Keywords: Indian Sign Language, Gesture acquisition, Hand Gesture Recognition, Segmentation, Feature Extraction

I. INTRODUCTION

Sign language is the only and cheapest approach of communication for the people who is suffering from disability like hearing and speaking. According to the survey taken by government of India, it is reported that in year 2011 census data over 2.68 crore of people in India suffer from some form of the disability. Out of this 18.9% people have speech and 7.5% people have hearing disability [1]. Sign language is composed of visual gestures and signs where every sign has a specific meaning allotted to it. There are 143 existing different sign languages all over the world, mainly American Sign Language, British Sign Language, French Sign Language, Japanese Sign Language, and Indian Sign Language [2]. The gestures are mainly divided into two types: Static gestures and Dynamic Gestures. Static gestures include only poses and configurations whereas dynamic gestures include strokes, pre-strokes, postures and phases. The dynamic gestures include movement of body parts. It may also include emotions depending on the meaning that gesture conveys [6].

ISL alphabets and numeric signs [8] are represented in Figure 1 and Figure 2 respectively.

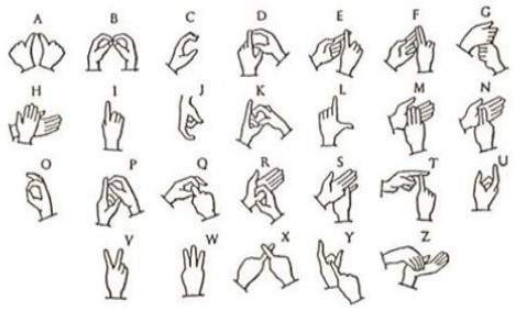


Figure 1. Representation of ISL Alphabets.

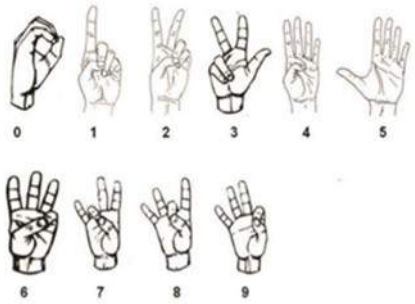


Figure 2. Representation of ISL numerals.

To make communication easy between deaf-dumb and normal people, sign recognition system that first recognize the gestures and convert it into text or voice. Developing such sign recognition system for ISL is more challenging than other sign languages due to the following reasons:

- ✓ In ISL both static & dynamic gestures are used.
- ✓ One hand move faster than other in Dynamic hand Gesture.
- ✓ ISL include facial expression.
- ✓ Complex hand shapes.

A very few work has been carried out to develop a system which converts ISL into text and voice because of these challenges. In this paper, various method used for hand gesture recognition system in ISL.

II. LITERATURE SURVEY

Surbhi Rathi and Ujwalla Gawande proposed a full duplex communication system for deaf & dumb people [5].The proposed system for recognizing a dynamic hand words gesture of Indian sign language and conversion of recognized gesture into text and speech and vice versa i.e. dual way communication. Pre-processing is divided into two steps: Segmentation & Filtering. In this system skin color filtering technique has used for segmentation. Eigen vectors and Eigen values technique has used for feature extraction. For classification, Eigen value weighted Euclidean Distance based classifier has used. The proposed system used minimum Eigen Value distance for recognition of sign. In dual way communication system first the recognized gesture has converted into text message and voice format, so that normal person can understand it. Second, voice has converted into text and its corresponding gesture, so that physically impaired humans can understand it.

Purva C. Badhe and Vaishali Kulkarni proposed a system for dynamic recognition of ISL gestures [6]. ISL alphabets & numerals convert into text in the proposed system. For tracking hand movement they used a combinational algorithm which included canny edge detection, skin color detection with YCbCr, thresholding, etc. For feature extraction they used Fourier descriptors method. In this system template matching is used to recognize sign. Self-created database is used for implementation. The proposed system gives an accuracy of 97.5%. MATLAB is used to creating a Graphical User Interface (GUI) for developing this system.

Bhumika Gupta, Pushkar Shukla, Ankush Mittal proposed a system for recognize a static images of the signed alphabets in the Indian Sign Language[7].This system first categorize sign into single handed or double handed. They used Histograms of Oriented gradients(HOG) Descriptor and Scale Invariant Feature Transform(SIFT) Descriptor for feature extraction. They used feature Fusion using K-Nearest Correlated Neighbour Algorithm for classification. The dataset constitutes of 520 images for the training segment and 260 images for testing. Fusion of both kinds of descriptors gave an accuracy of 90% without categorization into single handed &double handed gesture. Fusion of both kinds of descriptors gave an accuracy of 97.50% for single handed gestures and 91.11% for double handed gestures.

Aditya V., Vinod P.R., Usha Gopalkrishnan proposed a system based on Artificial Neural Network[8]. They proposed system recognition the static gestures in ISL alphabets and numerals.ISL fingerspelling translate into text in this system. The gestures included English alphabets (26 letters) and numerals (0-9). YCbCr model, filtering and morphological operations for hand segmentation is used in this system. For classification they used artificial neural network. 36 signs with 15 images of each were tested in this system. The system was implemented using MATLABR2010a.Low Computational capacity in this system. This proposed System gives an accuracy of 91.11%.

Shreyashi Narayan Sawant, M.S.Kumbhar proposed a system for real time sign recognition using PCA[9]. For Segmentation they used Otsu algorithm. 10 each of the 26signs total 260 images consisting in dataset.At a resolution of 380 x 420 pixels the image were captured. In this method test and train image Euclidean distance was calculated and gestures have a minimum Euclidean distance. Gesture was transformed into text and voice format.

The proposed system in [10] is used for dynamic hand gesture recognition in which a key frame extraction technique is used. This algorithm is used for to find out the most important frames from the video, based on the change in hand shape and position using certain

parameters and dynamic threshold. For Pre-processing they used YCbCr model. The main parameter to be considered are hand shape, hand motion and hand orientation for feature extraction. Multiclass Support Vector Machine (MSVM) is used for classification. The use of key frame extraction algorithm speeds up the system by selecting essential frames and eliminating extra frames**.** The proposed system gives accuracy of 90.46%.

Joyeeta Singha, Karen Das proposed a system using Eigen value weighted Euclidean distances as a classification technique for recognition sign alphabets of ISL[4].They Proposed system for static sign alphabets of ISL.For Pre-processing they used Skin filtering & hand cropping. Eigen value &Eigen vector is used for feature extraction. They used two level of classification. First, Classification based on Euclidean Distance. Second, Classification based on Eigen value weighted Euclidean distances. The proposed system gives accuracy of 97%.

III. DIFFERENT TECHNIQUES FOR RECOGNITION

Gesture recognition can be divided into two types: sensor based and vision based. In sensor based method, the gestural data can be extracted using data glove or motion sensors [6]. Data glove captured Even minute details of the gesture which ultimately enhances the system performance. Wearing a hand glove with embedded sensors in this method [6]. This method also reduces the user comfort and it affects the signer's usual signing ability.

Image processing include in vision based method. This approach is comfortable to the user because no extra device wearing. With the help of camera images is captured. For identifying the gesture this method deals with the features of image such as color and texture. Complexity of the background, variation in illumination, and tracking of other skin color objects along with the hand objects are the challenges in vision based approach.

Figure 3. (A) Data-Glove based[11]

Figure 3. (B) Vision based[11]

IV. OVERVIEW OF SIGN LANGUAGE RECOGNITION SYSTEM

The mainly four steps are involved in a sign language recognition system are: gesture acquisition, tracking and segmentation, feature extraction and gesture recognition. First step of sign recognition system is to acquire gestural data. A webcam is used to capture the images. Segmentation is required to tracking the hand movement. To extract important features feature extraction is used. The process of recognition can be divided into the two stages - training and testing. In training stage the classifier is trained using the training database. The main steps in training stage are creation of database, pre-processing, extraction of features and training of the classifier. The main steps involved in the testing phase are gesture acquisition, preprocessing, feature extraction and classification. Fig.4 shows block diagram of sign language recognition system.

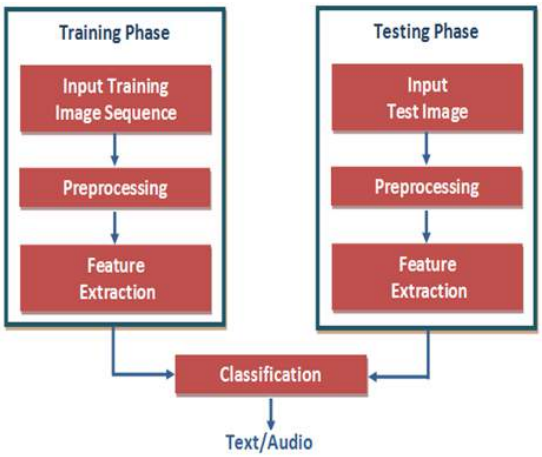


Figure 4. Sign Language Recognition System

A. Pre-processing

Preprocessing step is performed before feature extraction to produce the most useful information neglecting meaningless information. It consists of image enhancement, image filtering, image resizing, and image segmentation. To extract the region of interest the

images or videos of the test database are also preprocessed during testing phase.

In [5] skin color filtering technique is used for preprocessing. First the video is converted into frames.Then, pre-processing techniques has applied to the frames which are divided in to two parts namely filtering and segmentation. Mostly, skin color filtering technique is applied on input video frames for the detection of hand gesture from the background of the image. Hence, skin color filtering technique separates skin colored region from non-skin colored region. The entire frames are converted to HSV color space. Hue, Saturation and Value color model is used for skin detection and also it is less sensitive to illumination changes in comparison of RGB.

Hence, it has shown to be one of the most adapted techniques for skin color detection. Then, to get the rid of high frequency noise from the input image Gaussian low pass filter has applied, which has also used for the smoothning of image frames. Segmentation is used to find the only hand region from the entire image region. After removing noise from the HSV image convert this image in to gray scale image first and then binary image. So, the result obtained is a binary image with skin region is in white and non-skin region is in black color.

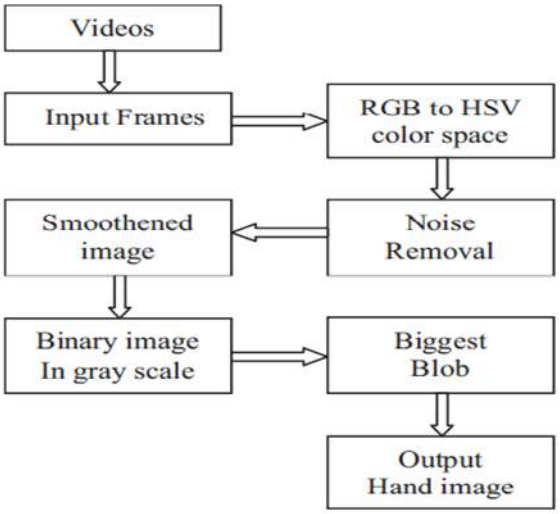


Figure 5. Block Diagram of skin color filtering

In[6] Skin color detection with YCbCr and Canny Edge detection are used for preprocessing. In this, video is first converted into frame. After cropping the frames, two subsequent frames are subtracted from each other for obtaining the difference image. Calculate the threshold & OR with the difference image. They used YCbCr for Skin color detection & Canny Edge Detection for boundary Detection. Then all are combined in one frame. Then after send it to feature extraction step.

In [7] , images are converted from a colored to a binary format by the employment of RGB thresholding . For removal of noise, the images are extracted for their HOG features. These features when fed to a standard Support Vector Machine were quite accurately able to separate single-handed gestures from the double-handed.

In [8], YCbCr color space is used for preprocessing & Hand segmentation. In order to detect the skin color in the input image it is first converted to YCbCr color space. YCbCr separates RGB into luminance and chrominance components where Y is the luminance component and Cb, Cr are the chrominance components. RGB values can be transformed to YCbCr color space using the following equations.

$$Y = 0.299R + 0.587G + 0.114B,$$
$$Cr = 128 + 0.5R - 0.418G - 0.081B,$$
$$Cb = 128 - 0.168R - 0.331G + 0.5B.$$

Thus a pixel is classified as belonging to skin if it satisfies the following relation:
$75 < Cb < 135$ and $130 < Cr < 180$ and $Y > 80$.

The result of segmentation produces a binary image with the skin pixels in white color and background in black color. The resulting binary image may contain noise and segmentation errors. Filtering and morphological operations are performed on the input image to decrease noise and segmentation errors.

In [9], the Segmentation of hands is carried out to separate object. For segmentation Otsu algorithm is used. The segmented hand image is represented certain features. These features are further used for gesture recognition To remove noises from images morphological filtering techniques are used. The preprocessing operation is done on the stored database.

B. Feature Extraction

Feature extraction stage is used to extract certain features from the hand image which is unique for each sign. Feature extraction is a method of reducing data dimensionality by removing less discriminative data. The selection of appropriate features and feature extraction methods is the most significant decision in the

design of gesture recognition system.In this step feature vector is obtained.In the next step feature vector is given. Various feature extraction techniques are Fourier descriptors, HOG features, PCA, etc.

In [5], Feature extraction stage is used to taken out certain features from the hand image which is unique for each sign . If the gesture gets match, then the last frame have chosen as a reference frame and the features like Eigen vectors and Eigen values have extracted from the reference frame. Following are the mathematical steps used for calculating Eigen values and Eigen vectors:

Step 1: Let, the last frame is assumed as 'X', 'X' has to be change the size using dimension m*m.

Step 2: Calculate the mean 'M' of the above vector X as,

$$\mathbf{M = E\{X\}} \quad (2)$$

Step 3: Calculate the covariance 'C' as given in eq. (1),

$$\mathbf{C = E\ \{(X - M)\ (X - M)\ '\}} \quad (3)$$

Step 4: From the above covariance 'C' calculate the Eigen values and Eigen vectors and arrange the Eigen vectors in descending order according to the Eigen values.

Step 5: Take only first 5 principles vectors from total m Eigen vectors that reduces the dimension of the matrix without much loss of information.

In [8],The distance transform of the image is used for feature extraction and recognition tasks. In the proposed method, the distance transform is computed by using the Euclidean distance. The equations for computing the distance between two pixels with coordinates (x, y) and (u, v) are shown below:

The city-block distance between two points P = (x, y) and Q = (u, v) is defined as

$$\mathbf{d\ (P,Q) = |\ x - u\ |\ +\ |\ y - v\ |.}$$

The chessboard distance between P and Q is defined as

$$\mathbf{d\ (\ P,\ Q\ ) = max(|\ x - u\ |\ ,\ |\ y - v\ |).}$$

The Euclidean distance between P and Q is defined as

$$\mathbf{d\ (\ P,\ Q\ ) = } \sqrt{(x-u)^2 + (y-v)^2}\mathbf{.}$$

In [7] Histogram of Oriented Gradients and Scale Invariant Feature Transform descriptors are used for feature extraction.

Histograms of Oriented Gradients (HOG) Descriptor:

A mainly used descriptor for detection of an object in an image and even its shape or color which was first given by Dalal and Triggs for Pedestrian detection in. First dividing the image into smaller units called cells and constructing single dimensional histograms for the edge orientations of pixels in each cell. In order to attain invariance from varying magnitudes of illumination, these local histograms are normalized for a group of cells which is referred to as a block. These blocks of normalized histograms are used as descriptors for an image.

Scale Invariant Feature Transform (SIFT) Descriptor:

This set of descriptors defined for an image is invariant to varying illumination as well as rotation and scaling. To define the SIFT features of an image generally constitutes of four stages. SIFT can be used both as a feature detector and a descriptor. First, the points of interest, referred to as key points, in an image are located for a scale-space. These are the locations in this scale-space that represent the local maxima or minima obtained when a set of Difference of Gaussian (DoG) filters are applied all over the image.

Table 1. Feature Extraction Techniques

Methods	Literature availability
Principle Component Analysis [9]	Occluded and overlapped gestures can be recognized ,Not scale invariant
Fourier Descriptors [6]	Simple, robust, computationally efficient and immune to rotation or Scaling of shape and noise.
HOG descriptors [7]	Invariant to illumination change and orientation ,Not scale invariant

C. Classification

The feature vector is used as the input of the classifier that recognizes the sign. Classification involves two phases: training phase and testing phase. The feature vectors are given to the classifier in training phase. Identification of classes is taken in testing phase, test image or video is given as an input and the result is produced as a text and then in voice. Classifier used are Euclidean distance [5], K Nearest Neighbor (KNN) [7], Artificial Neural Network [8], Support Vector Machines (SVM) [7]etc. The classifier's performance is measured in terms of rate of recognition.

Table 2. Recognition Techniques

Methods	Literature Availability	Acc%
Support Vector Machine [7]	Accurate results robust to noise, Computationally expensive	94.23
Euclidean distance [5]	Invariant to rotation ,Time Consuming	97.50
Artificial Neural Network [8]	Low computational complexity, Time Consuming	91.11
K Nearest Neighbor (KNN) [7]	Robust to noisy data ,to determine value of parameter K	80.00

V. CONCLUSION

This paper presents a survey on various sign recognition approaches for Indian Sign Language. The main aim of the sign recognition is remove the communication gap between deaf-dumb & normal people. More work have been carried out in static sign recognition. Only few work has been carried out for dynamic hand gesture .This paper present comparison between various recognition technique. In future there is possibility to do more work in dynamic sign, word ,sentences of ISL.

VI.REFERENCES

[1] Website: Http: // censusindia.gov.in / Census_And_You/ disabled_population.aspx.

[2] Dutta, KusumikaKrori, "Double handed Indian Sign Language to speech and text." 2015 Third International Conference on Image Information Processing (ICIIP). IEEE, 2015. pp. 374-377.

[3] AditiKalsh, N.S. Garewal," Sign Language Recognition System for Deaf and Dumb", International Journal of Advanced Research in Computer Science and Software Engineering, Vol. 3, Issue. 9, pp. 103-106.

[4] JoyeetaSingha , Karen das , "Indian Sign Language Recognition using Eigen Value Weighted Euclidean Distance based Classification Technique",2013 International Journal of Advanced Computer Science and Applications(IJACSA),2013.

[5] SurbhiRathi and UjwallaGawande, "Development of Full Duplex Intelligent Communication System for Deaf and Dumb People," 2017 IEEE7th International Conference on Cloud Computing, Data Science & Engineering – Confluence ,pp.733-738.

[6] P. C. Badhe and V. Kulkarni, "Indian sign language translator using gesture recognition algorithm," 2015 IEEE International Conference on Computer Graphics, Vision and Information Security (CGVIS), Bhubaneswar, 2015, pp. 195-200.

[7] Bhumika Gupta, PushkarShukla ,Ankush Mittal , "K-Nearest Correlated Neighbour Classification for Indian Sign Language Gesture Recognition using Feature Fusion", 2016 International Conference on Computer Communication and Informatics(ICCCI),2016.

[8] V. Adithya, P. R. Vinod and U. Gopalakrishnan, "Artificial neural network based method for Indian sign language recognition," Information & Communication Technologies (ICT), 2013 IEEE Conference on, JeJu Island, 2013, pp. 1080-1085.

[9] Shreyashi Narayan Sawant and M. S. Kumbhar ,"Real Time Sign Language Recognition using PCA", 2014 IEEE International Conference on Advanced Communication Control and Computing Technologies (lCACCCT),2014.

[10] BhumikaPathak, Subhash Chand Agrawal , Anand Singh Jalal, CharulBhatnagar , " A Framework For Dynamic Hand Gesture Recognition Using Key Frames Extraction" 2015 IEEE.

[11] Juhi Ekbote,Mahasweta joshi, "A survey on Hand gesture recognition for Indian Sign Language"International Research journal of Engineering and Technology (IRJET) 2016.

© 2018 IJSRSET | Volume 4 | Issue 2 | Print ISSN: 2395-1990 | Online ISSN : 2394-4099
National Conference on Advanced Research Trends in Information and
g Technologies (NCARTICT-2018), Department of IT, L. D. College of Engineering,
Ahmedabad, Gujarat, India In association with
International Journal of Scientific Research in Science, Engineering and Technology

A Survey on Detection of Llegally Parked Vehicle in No Parking Area

Rikita R. Nagar[1], Prof. Hiteishi M. Diwanji[2]
[1]Department of Information Technology, Government Polytechnic for Girls, Ahmedabad, Gujarat, India
[2]H.O.D., Department of Information Technology, L.D.College of Engineering, Ahmedabad, Gujarat, India

ABSTRACT

Nowadays with the increase use of vehicles traffic management becomes hectic problem. People dont find parking space easily and encouraged to do parking in no parking area. Due to vehicle parked on the busy traffic road, other vehicles may hit the car and cause the potential accidents which may create a public safety problem. The detection of illegally parked vehicles is an important part of any traffic management and video surveillance application. Camera Surveillance system becomes cheaper and due to this human activity or behavior detection and tracking also becomes more practical. So, many automated systems have been designed for number of detection tasks, but the unauthorized or illegally parked vehicle detection is still done by the human operators of such surveillance system. From the literature review different methods are identified such as background subtraction, frame differencing method, Single Shot Multibox Detector based on Deep learning, Scalable Histogram of oriented Gradient, Mixture of Gaussian modeling used to detect and track object from captured video sequence. Background subtraction method gives us better and accurate result for moving vehicles detection. Detecting an illegally parked vehicle in no parking area becomes more complex task due to illumination changes and sudden weather change condition. The goal of this paper is to analyze and review the previous approaches towards detecting illegally parked vehicle in no parking area using video sequences.

Keywords: Vehicle Detection, Real Time, Object tracking, Video Surveillance, Traffic Monitoring, Illegally parked vehicle

I. INTRODUCTION

From the study it has been found, vehicle parked on busy road or in no parking area creates heavy traffic which may cause for accidents or hitting situation. So to prevent such situation traffic management system has to detect unauthorized or illegally parked vehicle in the no parking area. Not only detecting such a vehicle in no parking area, further notification system or alarm system should be implemented to conduct immediate action by the traffic regulation authority. Recently many researchers have used different approaches to detect illegal parking of vehicle in no parking area. Generally, two types of cameras are used for any surveillance system: Static camera and moving camera. But in case of vehicle detection in no parking area static cameras are generally used. So, to detect vehicles from traffic scene most researches have used background subtraction method. After detecting vehicles in no parking area vehicle must be tracked in no parking region.

Rikita Nagar is Sr. Lecturer with the Department of Information Technol-ogy, Government Polytechnic for Girls, Ahmedabad, India.

Prof. Hiteishi M. Diwanji is H.O.D., Department of Information Technol-ogy,L.D.College of Engineering, Ahmedabad,Gujarat

If vehicle is detected static for more than some time limit, then alarm or notification should be given to traffic regulatory authority.

II. OVERVIEW

Generally, detection of illegally parked vehicle process is carried out on the basis of four generic steps: capturing image from video sequence, preprocessing, object detection, object recognition and object tracking. Basic flow diagram is shown in figure 1.

Figure 1. Basic flow Diagram of Object tracking[6]

A. PreProcessing

In preprocessing stage images captured by static cameras are generally of low quality and affected by various noise and different lighting conditions. So this images should be enhanced using different image sharpening and smoothing techniques. Morphological operators such as dilation, erosion, opening, closing may be applied on the images. As well as, different filters may also be applied to remove various types of noise from the images.

B. Object Detection

Object detection is one of the important tasks in many video surveillance systems. Frame differencing, optical flow, and background subtraction methods are used for this task, which is shown in the figure 2. In Table-I comparison between these methods are given.

C. Object Tracking

The next step is to track the object. Object tracking is used to identify location or position of moving object in consecutive video sequences. By locating a moving object in consecutive frames actually we track the object. Basically three types of object tracking approaches are used: kernel, point and silhouette based tracking. As per the literature review kernel based method is more used due to high accuracy and less computational cost compared to silhouette method. The point tracking method has less computational time but it is less accurate. The different types of object tracking techniques are shown in figure 3. Comparison of different object tracking techniques is shown in Table 2.

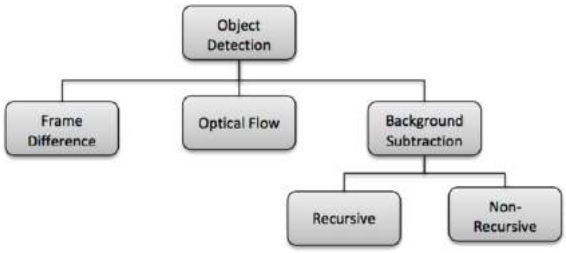


Figure 2. Types of Object Detection Method[6]

Table 1. Comparative Study Of Object Detection Technique [6]

Object Detection Method	Basic Principle	Computational Time	Accuracy
Temporal Differencing	Pixel-wise Subtraction of Current & Background frame	Low	High
Background Subtraction	Current frame is subtracted from background frame	Low to Moderate	Moderate to High
Mixture of Gaussian	Based on multimodal distribution	Moderate to high	Moderate to high
Optical Flow	Uses optical flow distribution characteristics of pixels of object	Moderateto high	High

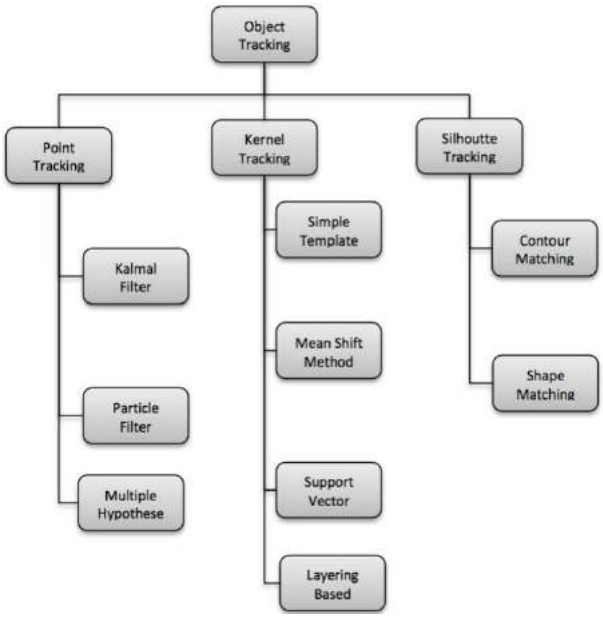


Table 2. Comparative Study Of Object Tracking Methods [6]

Object tracking method		Algorithm used	Computational time	Accuracy
Point Tracking	Kalman Filter	Kalman filtering algorithm	Low to Moderate	Moderate
	Particle Filter	Recursive Bayes filtering	Moderate to High	High
	Multiple Hypothesis tracking	MHT algorithm	Low	Low to Moder-ate
Kernel Tracking	Simple template matching	Matching region of interest in video	Low to Moderate	Low
	Mean shift method	Expression &,location object; optimal gradient decline	Low	Moderate
	Support vector Machine	Positive & negative training values	Moderate	Moderate
	Layering based tracking	Shape representation using intensity	Moderate	Moderate to High
Silhouette tracking	Contour matching	Gradient Descent Algorithm	Moderate	Moderate to High
	Shape matching	Hough Transform	High	High

Table 3. Comparative Study Of Object Classification Methods [6]

Object Classification Method	Computational time	Accuracy
Shape Based	Low	Moderate to High
Motion Based	High	Moderate
Color based	High	High
Texture Based	High	High

D. Object Recognition

After detection of object, object must be classified to the targeted object, such as vehicle, human or animal as per the requirement of the surveillance system which is known as object recognition or classification. For the object classification various approaches are used based on shape, motion, color, and texture. From the literature review comparison of these techniques are shown in Table 3.

III. LITERATURE REVIEW

1. In this paper author proposed a Methodology for detecting illegally parked vehicle in no parking area using projection of 2-D data into 1-D data. In the proposed methodology it consists four stages: Projection, Segmentation, tracking and reconstruction. In the first stage 1-D transformation is applied on source video 2-D data. By doing this they reduce complexity of segmentation and tracking. In second stage, foreground blobs are segmented which represents vehicles in 1-D transformed image. In third stage, these segmented blobs are tracked in each consecutive frames. In the fourth stage finally the transformation of illegally parked vehicles from 1-D data into 2-D original image data is performed. So, result after this stage is a 2-D image of the vehicle which is detected illegally parked vehicle which can be used for further process.

Advantages of this methodology is that it reduces complexity of segmentation and tracking process from o(n2) to o(n). But final result is fully dependent on the initial transformation results for static vehicle detection. It also results in false positives due to illumination change conditions and don't give accurate results in nighttime video frames.

2. In this paper author proposes a new methodology based on deep learning. It uses Single shot multi box detector (SSD) algorithm to locate and classify Illegally parked vehicles which are captured by camera. In this paper author used optimized SSD to improve performance of the detection algorithm. Here they adjust the aspect ratio of default box by adjusting k-means with actual dataset in efficient way. After the detection of illegally parked vehicle in region of interest tracking is performed. Thus, it is advantageous to use Optimized Single Shot Multibox Detector (SSD) to detect illegal parking , because it gives accurate results in complex weather conditions at realtime and can detect variety of vehicles like car, motor cycle, truck etc. Another advantage is that this methodology does not uses background subtraction methodology which is highly sensitive to environment changing condition, and results in higher accuracy and lower computational time.
3. In this paper author proposed two stage application framework which provides real time, illumination variation resistant and occlusion tolerant solution. It Segmentation history Images (SHI) is used to detect illegally parked vehicle in restricted parking area. SHI improves foreground segmentation accuracy for detecting stationary vehicles. In the second stage tracking of detected vehicles is performed using adaptive edge orientation based method. Advantages of this methodology are that it handles sudden illumination change condition and detects objects even when they are occluded. But often it failed to detect stationary object due to low light condition of the traffic scene.
4. In this paper author uses dual background model sub-traction to detect illegally parked vehicle. They used adaptive background model which is based on statistical information of pixel intensity. This method is highly efficient in lighting condition. To remove false region geometrical property based analysis is used. Scalable Histogram of oriented gradient (SHOG) is then applied to detect object is vehicle or not. SHOG is trained using Support Vector Machine (SVM). Then tracking is applied on the detected vehicle and time for which vehicle becomes static is counted. If vehicle stops more than some time limit than proposed system generates alert.
5. In this paper author used two background models with different learning rates for Gaussian mixture model, defined as short- and long-term models. Each model extracts foreground pixels and the stability of these pixels are then analyzed based on cumulative values and temporal positions over a certain pe-riod of time. The result shows that the proposed methodology is efficient and robust. But background subtraction method is easily affected by environments, such as illumination changing and the weather and reduces efficiency. This methodology also fail to detect vehicles which are located close to each others.

Table 4. Comparative Study Of Papers

Ref. no.	Methods Used	Advantages	Disadvantages
[1]	Background Segmentation of 1-D image	-Reduces computational complexity, -Detection rate is higher without any false positive detection	-Failed to detect closely parked vehicle, nighttime video and illumination change
[2]	Single Shot MultiBox Detector (SSD) algorithm	detection with illumination changing and complex weather conditions, robust,	Not handling occluded images
[3]	Dual Background Subtraction Scalable Histogram of Oriented Gradient (SHOG)	More robust and faster Handles Occlusion Detection problem	Fail to handle the sudden/slow changing of illumination condition
[4]	MOG,EGMM	reduce false positive detection handles occlusion detection	Affected by environment, illumination changing conditions
[5]	Foreground Segmentation	foreground segmentation adaptive edge orientation based tracking technique	sometimes failed to detect stationary object due to lighting condition of the traffic scene.

IV. RESULTS AND DISCUSSIONS

Comparative study of all papers is shown in table - IV.

1. The i-LIDS dataset is used to check the efficiency and accuracy of the proposed methodology. Results of all four daytime sequences in i-LIDS are accurate. They able to detect the illegally parked vehicles correctly as well as measure the durations of the illegal parking events with higher accuracy. The other dataset which they have taken in Austin, TX, System is able to detect two illegally parked vehicles but failed to detect illegally parked vehicle that arrive together, they could detect only one of the two illegally parked vehicles. For more accurate tracking of vehicles in a nighttime video sequence, system must be modified to accommodate the effects of headlights. One false positive was also detected due to the continuous glare of the headlights.
2. In this paper system is evaluated on their own dataset. Experiments were taken in sunny days as well as rainy days and achieved a 99 percent detection accuracy
3. In this paper the algorithm has been tested on i-LIDS dataset and their own recorded (Sussex Traffic Monitoring) dataset. i-LIDS datasets contain three progressively more demanding video sequences, taken in daylight, and one night time video sequence. From test and results shown in paper proposed system detects illegally parked vehicles more accurately than any other available system. The proposed system does not require the algorithm to shift from RGB to gray-level pixel values for the night time video sequences.
4. In this paper the proposed algorithm is evaluated using i-LIDS database as well as their own dataset. Based on exper-iment, proposed system successfully detect illegally parked vehicle for all scenarios, with zero false alarm. This method may fail to handle the sudden/slow changing of illumination condition due to transition of the day time. It is hard to design the SHOG feature, and cannot deal with complex weather conditions.
5. Proposed system is evaluated using their dataset, ISLab dataset. Proposed system successfully detect parked vehicle for all scenarios, but it produces false positives at night time due to lighting condition of scenes, the system performs average processing time

around 15 fps for video sequences with 640x480 pixels resolution. Thus, it can be said that the system is fast enough to be implemented on the real-time video surveillance system. Furthermore, Extracting foreground by background subtraction method is easily affected by environ-ments, such as illumination changing and the weather system fail to detect multiple occluded illegally parked vehicles which are located close to each others.

V. CONCLUSION

In this paper, review on different methods used for detection of illegal parked vehicle in ROI, tracking, recognition tech-niques and segmentation method which is based on the video frame and various tracking technologies are discussed. This approach used towards the illegal parked vehicle detection in ROI with new ideas. We have identified and discussed the limitation/future scope of various methods. A deep learning based framework to detect illegal parking in ROI. It achieves high accuracy and real-time detection results[2]. Even in case of occlusion a new pixel classification method based on GMM is used to detect stationary objects[3]. Another proposed algorithm, which is based on the 1-D projection, can be implemented in real time and is effective even in poor scene conditions[1]. The algorithm benefits greatly from the decreased complexity, allowing to use a more time consuming segmentation and tracking procedure. Dual background model subtraction method is used to extract candidate region of object. The SHOG and SVM-based vehicle detector is also integrated to classify the object into vehicle or other objects[4]. This method detect successfully most parked vehicle, with zero false alarm. Cumulative Dual Foreground Difference method detects successfully most parked vehicles[5]. However, It may fail to detect multiple occluded illegally parked vehicles which are located close each others.However, these all techniques need to concentrate towards handling sudden illumination changes, darker shadows and vehicle occlusions.

VI. REFERENCES

[1]. Jong T. Lee, M. S. Ryoo, Matthew Riley, J. K. Aggarwal,"Real-Time Illegal Parking Detection in Out-door Environments Using 1-D Transformation", IEEE Transactions On Circuits And Systems For Video Tech-nology, VOL. 19, NO. 7, JULY 2009

[2]. Xuemei Xie , Chenye Wang, Shu Chen, Guangming Shi, Zhifu Zhao,"Real-Time Illegal Parking Detection System Based on Deep Learning",Proceedings of the 2017 International Conference on Deep Learning Tech-nologies,Pages 23-27

[3]. W. Hassan, P. Birch, R. Young, and C. Chatwin, "Real-Time Occlusion Tolerant Detection of Illegally Parked Vehicles",International Journal of Control, Automation, and Systems, vol. 10, no. 5, pp. 972982, 2012

[4]. Wahyono, A.Filonenko,K. H. Jo , "Illegally parked vehicle detection using adaptive dual background model",Proc. of IEEE Indutrial Electronic Society Con-ference (IECON), Yokohama, Nov 2015.

[5]. Wahyono, Alexander Filonenko, Kang-Hyun Jo ,"De-tecting illegally parked vehicle based on cumulative dual foreground difference", IEEE 14th International Conference on Industrial Informatics, 2016

[6]. Mukesh Tiwari,Dr. Rakesh Singhai,"A Review of De-tection and Tracking of Object from Image and Video Sequences",International Journal of Computational In-telligence Research ISSN 0973-1873 Volume 13, Num-ber 5 (2017), pp. 745-765 Research India Publications

© 2018 IJSRSET | Volume 4 | Issue 2 | Print ISSN: 2395-1990 | Online ISSN : 2394-4099
National Conference on Advanced Research Trends in Information and Technologies (NCARTICT-2018), Department of IT, L. D. College of Engineering, Ahmedabad, Gujarat, India In association with International Journal of Scientific Research in Science, Engineering and Technology

A Survey of Efficient CCSDS Recommended DWT Decompressor

Shailja M. Maniya[1], Bakul Panchal[2]

[1]M.E. (I.T.) Student, I.T. Department, L.D. College of Engineering Gujarat Technological University, Ahmedabad, Gujarat, India

[2]Assistant Professor, I.T. Department, L.D. College of Engineering Gujarat Technological University, Ahmedabad, Gujarat, India

ABSTRACT

Now a days many satellites are being launched and data retrieval from the satellite on the earth is very expensive task as it requires too much bandwidth. So, we need compression to get the data using less bandwidth and it can also speed up file transfer. For space system there is one committee called Consultative Committee for Space Data System (CCSDS) which publishes the recommended standards on compression methodology for different type of data. CCSDS-122.0-B-1 Standard defines a compression methodology for compression of two dimensional Image Data. Inverse Discrete Wavelet Transform (IDWT) is used for the decompression of two dimensional Image Data which can be considered as time consuming part of the Decompressor. To speed up the decompressor this Inverse Discrete Wavelet Transform (IDWT) can be run parallel using different high-speed hardware. In this paper, we will survey various ways through which efficient CCSDS recommended Discrete Wavelet Decompressor can be made.

Keywords: GPGPU, CCSDS, Discrete Wavelet Transform (DWT), CUDA, NVIDIA, SPIHT, Rice decoding.

I. INTRODUCTION

The CCSDS Image Data Compression is the most widely used particularly for the grayscale image data. This algorithm is very useful for any imaging instruments application. The algorithm is designed in such a way that Its complexity remains sufficiently low so that it can be feasibly implemented on high-speed hardware.

This algorithm is intended to be used for on-board spacecrafts. Compression is used in order to save bandwidth usage & storage space required to save the image data as well as time can also be saved. On the other side, real time need is to decompress the two dimensional image data as soon as it is downlinked on the earth.

Various data Compression technique uses wavelet transformed data for better compression performance.

There are basically two types of compression supported by this CCSDS standard i.e. Lossless Data Compression and Lossy Data Compression. In Lossless Data Compression, data is compressed in such a way that it can be recovered easily on decompression whereas in Lossy data compression technique, data cannot be reproduced without some distortion. Lossy Compression technique uses 9/7 Float Discrete Wavelet Transform, Lossless Compression Technique uses 9/7 Integer Discrete Wavelet Transform. In this paper, we are going to survey for the Lossless Data Decompression technique.

II. CCSDS RECOMMENDED DWT DECOMPRESSOR

In this standard, decompressor consists of two functional parts i.e. bit plane decoding compressed data followed by two dimensional inverse discrete wavelet transform of wavelet transformed data as shown in figure 1.

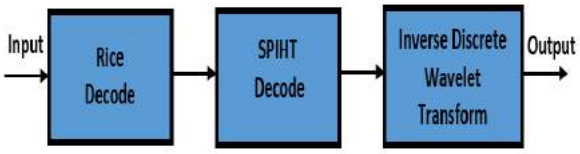


Figure 1. Overview of the Decoding system

A. Rice Decoding

Rice encoding is special case of Golomb coding which can be used to reduce the bits required to represent the lower value numbers. Rice encoding used for encoding of DC Coefficients in the CCSDS-122 standard. An algorithm for rice decoding is as shown below:

Assume Given a Constant C, any Symbol S can be represented using quotient (Q) and reminder (R), where

$$S = Q \times M + R \quad (1)$$

A. Rather than representing both Q and R as binary values, Rice encoding represents Q as a unary value and R as a binary value. Rather than representing both Q and R as binary values, Rice encoding represents Q as a unary value and R as a binary value.

B. For those not familiar with unary notation, a value N may be represented by N 1s followed by a 0.

C. Example: 4 = 11110 and 2 = 110.

If given Bit Length K, Compute the Modulus M, using by given equation (2)

$$M = 2^K \quad (2)$$

Then perform the following steps for each encoded Symbol (S),

1. Determine Q by counting the number of 1s before the first 0.
2. Determine R reading the next K bits as a binary value.
3. Write out S as Q x M + R.

Example: decode the encoded value 100010 when K = (M=16)

1. Q=1
2. R = 0b0010 = 2
3. S = Q x M + R=1 x 16 + 2 = 18

B. SPIHT Decoding

SPIHT (Set Partitioning in Hierarchical Trees) is well known algorithm to encode the image using bit planes and it performs two passes for each bit-plane. While one pass computes sign values and the implicit location information of significant wavelet coefficients, the second pass sends the refined bit values of the significant coefficients which are determined up to the current bit plane.

The decoder rebuilds the wavelet coefficients according to the importance of the information. This process is serial in nature because the importance of information will decide the position of the next node. Thus, decoding process can only finish when all nodes are decoded. Therefore, it is clearly said that SPIHT process should be performed on the CPU.

C. Inverse Discrete Wavelet Transform

As explained in section 1, we are going to do survey for the lossless data compression we have to use Inverse 9/7 Integer Discrete Wavelet Transform in decompressor. Inverse Integer DWT mapping is from two set of wavelet coefficients C_j and D_j back to the signal vector x_i as mentioned in [1] which is given by equations (1) to (6). Here we require special boundary filters to get the data back.

In equation given below, we first compute the even indexed signal values i.e. $x_0, x_2, \ldots, x_{2N-2}$ using DWT coefficients. After computation of these even indexed vector values, odd indexed values can be calculated i.e.$x_1, x_3, \ldots, x_{2N-1}$.

$$x_1 = D_0 + \left\lfloor \frac{9}{16}(x_0+x_2) - \frac{1}{16}(x_2+x_4) + \frac{1}{2} \right\rfloor \quad (1)$$

$$x_{2j+1} = D_j + \left\lfloor \frac{9}{16}(a+b) - \frac{1}{16}(c+d) + \frac{1}{2} \right\rfloor \quad (2)$$

where $a= x_{2j}$, $b= x_{2j+2}$, $c= x_{2j+2}$, $d= x_{2j+4}$ for j=1,…,N-3

$$x_{2N-3} = D_{N-2} + \left\lfloor \frac{9}{16}(e+f) - \frac{1}{16}(g+f) + \frac{1}{2} \right\rfloor \quad (3)$$

where $e= x_{2N-4}$,$f= x_{2N-2}$, $g= x_{2N-6}$

$$x_{2N-1} = D_{N-1} + \left\lfloor \frac{9}{8}x_{2N-2} - \frac{1}{8}x_{2N-4} + \frac{1}{2} \right\rfloor \quad (4)$$

$$x_0 = C_0 + \left\lfloor - \frac{D0}{2} + \frac{1}{2} \right\rfloor \quad (5)$$

$$x_{2j} = C_j + \left\lfloor - \frac{D_{j-1} + Dj}{4} + \frac{1}{2} \right\rfloor \quad \text{for } j = 1,\ldots, N-1 \qquad (6)$$

D. Two dimensional Single Level Inverse DWT

The single-level 2-d DWT transform[1] shall be inverted by repeated application of the 1-d inverse to columns and rows of the transformed data array in the reverse order to that in which the 1-d transforms were applied:

a) each column shall be inverted to produce the intermediate transformed data arrays:

1) The 1-d DWT inverse shall be applied to columns of the LL and LH subbands to obtain the intermediate horizontal low-pass array of figure 3(b),
2) The 1-d DWT inverse shall be applied to columns of the HL and HH subbands to obtain the intermediate horizontal high-pass array of figure 3 (b);

b) the 1-d DWT inverse shall be applied to rows of the intermediate horizontal low-pass and horizontal high-pass arrays to recover the original image array.

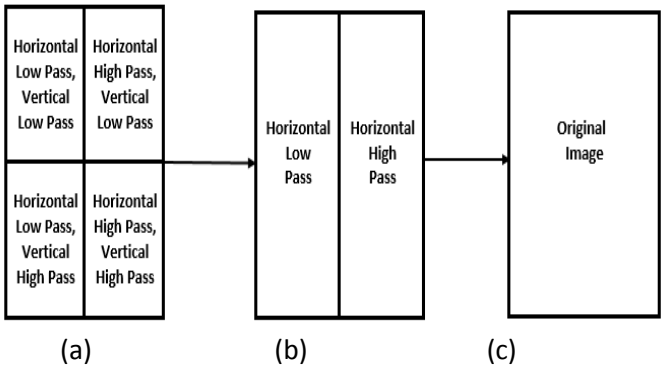


Figure 3. One level 2d Inverse DWT

E. Two dimensional three level Inverse DWT

The inversion process of a multi-level DWT shall be as follows [1]:

a) The four subbands of highest level, LL3, LH3, HL3, HH3, shall be inverted using an inverse single-level 2-d DWT to yield the single subband LL2, which then replaces the higher-level subbands in the transform data matrix;

b) The four subbands LL2, LH2, HL2 and HH2 shall be inverted to yield the single subband LL1, which again replaces the higher-level subbands in the transform data matrix;

c) A final single-level 2-d inverse DWT shall be applied to subbands LL1, LH1, HL1 and HH1 to reproduce the original image.

III. LITERATURE REVIEW

A. Wavelet based decoding system [2]

Changhe Song, Yunsong Li, and Bormin Huang at el. [2] has implemented the decoding system for satellite images. They have implemented a wavelet based decoding system which contains SPIHT with Reed-Solomon decoding. Overview of the decoding system is as shown in below fig. 4.

Figure 4. Overview of decoding system [2]

In this paper, they have used float 9/7 inverse DWT as a wavelet transform with SPIHT and Reed-Solomon decoding.

Table 1.

CPU TIME OF EACH COMPONENT OF THE DECODING SYSTEM

RS decoding	18 ms
SPIHT decoding	62 ms
IDWT	841 ms

As shown in table 1, IDWT is the most time consuming process of the decoding system. They have used General Purpose Graphics Processing Unit (GPGPU) for the faster computation of IDWT. They have also used shared memory for the better time performance of the IDWT computation. GPU based IDWT model is as shown in fig. 5.

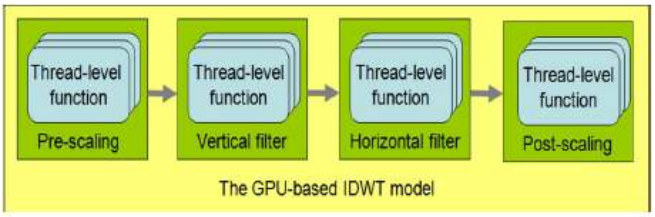


Figure 5. IDWT model on GPU [2]

Fetching steps of horizontal filter and vertical filer are shown in figure 6 & 7 respectively.

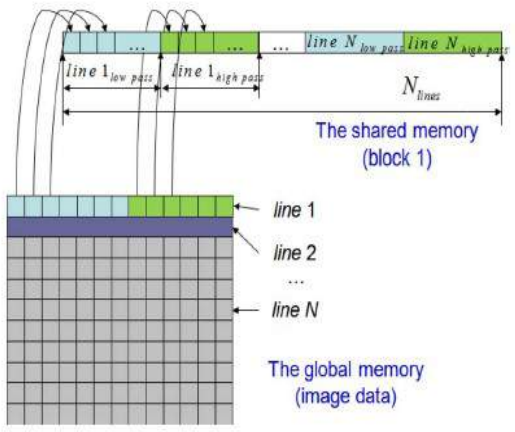


Figure 6. the fetch step of horizontal filter [2]

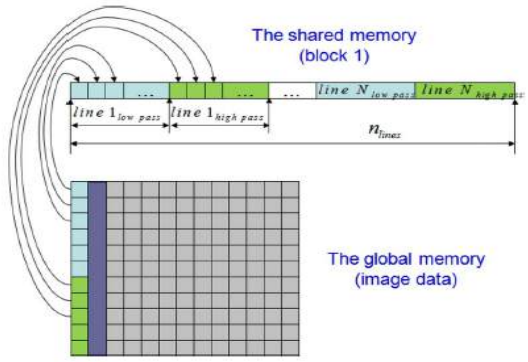


Figure 7. the fetch step of vertical filter [2]

Thus they have prepared CPU-GPU pipelined fashion decoding system as shown in figure 8.

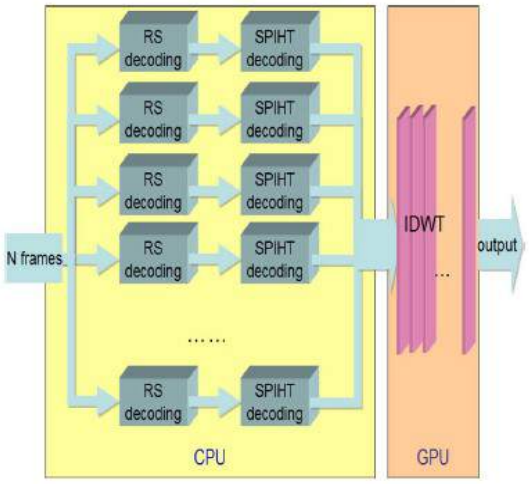


Figure 8. Overview of RS+SPIHT+IDWT decoding system [2]

B. Time efficiency Comparison of Wavelet transform and its inverse on different platforms [3]

Abhishek S. Shetty, Abhijit V. Chitre and Yogesh H. Dandawate at el [3] has done the comparison for the wavelet and Inverse wavelet transform on different platforms. They have used 9/7 Integer Discrete wavelet transform and its inverse for this comparison. Three different platforms are used which are MATLAB, Python using OpenCV and Python using PIL (Python Imaging Library).

As a result the comparison chart is prepared as shown in figure 9. It is observed that Pthon-OpenCV gives better result in terms of time and it is also open source. These results are being obtained using image size starting from 512x512 to 6000x6000.

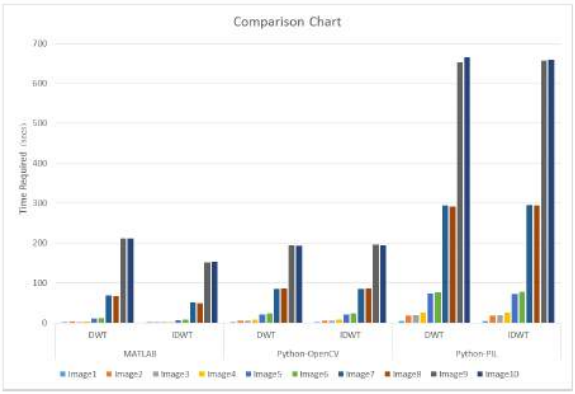


Figure 9. Comparison Chart [3]

C. Efficient parallelization of Discrete Wavelet Transform [4]

Anastasis Keliris, Vasilis Dimitsasy, Olympia Kremmyday, Dimitris Gizopoulosy and Michail Maniatakosz at el [4] has implemented the Discrete wavelet transform on multi core cpu, Many Integrated Cores (MIC) cpu as well as on NVIDIA GPU. In this paper, 9/7 integer DWT is used as a wavelet transform.. They have used the four evaluation systems as shown in table 2.

Table 2.

CONFIGURATION OF THE EVALUATION SYSTEMS

	System 1	System 2
CPU	AMD Phenom II X4 965 4 cores @3.4 GHz 45nm, released: 2009	Intel Xeon E5-2680 8 cores @2.7GHz 32nm, released: 2012
Accelerator	Nvidia Tesla C2070 (Fermi) 448 CUDA cores 40nm, released: 2010	Intel Xeon Phi 5110P 60 cores @1.053GHz 22nm, released: 2012

They have used the algorithm for the transpose of the matrix because this operation is memory-intensive

operation. Algorithm [4] for the matrix transpose is as written below:

Algorithm 1: Cache-oblivious mxn recursive matrix transposition

```
1 in, out: input/output image
2 r off, c off: rows/columns offset
3 Define: BLOCKSIZE (architecture dependent)
4 Function trans (in, out, r off, c off, m, n)
5 if m > BLOCKSIZE or n > BLOCKSIZE
then
6 if n > m then
7 nhalf  n/2;
8 trans (in, out, r off, c off+nhalf, m, nhalf);
9 else
10 mhalf  m/2;
11 trans (in, out, r off+mhalf, c off, mhalf, n);
12 end
13 else
14 rlimit   r off + n;
15 climit   c off + m;
16 for i   r off to rlimit do
17 for j   c off to climit do
18 out[j][i] = in[i][j];
19 end
20 end
21 end
22 return
```

D. Energy consumption analysis of CCSDS image compression running on two different platforms [8]

Christofer Schwartz, Marcelo S. pinho, at el [8] has implemented CCSDS-122 DWT based compressor on the CPU as well as on the GPU. CPU specifications are : Intel Core i7 - 3610QM (third generation) with 4 cores (8 threads) working at a frequency of 2; 241; 003 KHz and GPU specifications are NVIDIA GeForce GT 630M (2.1 Streaming Multiprocessor Capability) — this GPU has two multiprocessor (MP) with 48 cores each (total of 96 cores) working at a clock of 950; 000 KHz.

In this paper, they have implemented bit-plane encoder on the Host side (CPU) and DWT is performed on device (GPU). Timing given by the host and host + device system are analysed in this paper.

IV. COMPARISON

As described in Section III, Changhe Song at el [2] has shown in their results that IDWT part of the decoding system is the most time consuming part in table 1. As per the result of [2], CPU-GPU pipeline system gives better performance than IDWT on GPU and SPIHT + Reed-solomon on CPU as shown in table 3 & 4.

Table 3.

EXECUTION TIME OF THE IMPROVED SYSTEM WITH IDWT ON GPU AND THE 6-FRAME CONCURRENT RS AND SPIHT DECODING CPUS

	Original (ms)	Improved (ms)	Speedup
RS	18	3.7	5
SPIHT	62	12.9	5
IDWT	841	13.3	63
Total	921	29.9	**31**

Table 4.

EXECUTION TIME WITH THE CPU-GPU PIPELINING

	Time (ms)	Speedup
CPU (RS+SPIHT decoding)	16.6	5
GPU (IDWT)	13.3	63
Total	16.6	**55**

As per the result of [3], Python-OpenCV gives better timing results for wavelet transform, further use of GPGPU can still reduce time.

Paper [4] has given results for the DWT algorithm time for different size of images with different types of platforms which is shown in table 5. As shown in the table 5, NVIDIA Tesla gives the best performance from all of the different platforms.

Table 5.

EXECUTION TIMES (IN MS) OF THE DWT ALGORITHM

	256x256	512x512	1024x1024	2048x2048	4096x4096	8192x8192
Phenom ser.	5.7	11.7	66.4	303.4	1348.6	7809.8
Phenom opt.	1.9	4.5	20.4	55.4	197.6	740.6
Tesla	0.2	0.5	1.2	3.9	14.6	58.4
Xeon E5 ser.	1.0	3.6	24.6	108.2	546.3	2589.6
Xeon E5 opt.	1.2	2.3	6.4	16.3	47.2	144.8
Xeon Phi	7.8	28.8	60.1	130.2	263.7	590.1

In terms of energy consumption, the most suitable hardware for the CCSDS-122 Compression

algorithm has been experimented in [8]. The result is shown in fig. 10 using rate-distortion-cost curve.

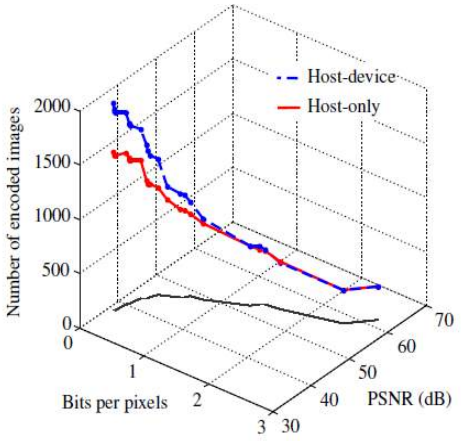


Figure 10. rate-distortion-cost curve [8]

V. CONCLUSION

Different ways to make efficient CCSDS-122 based Decompresssor has been viewed in this paper. This standard Decompressor contains two modules, IDWT and Bit plane Decoder, from which IDWT is most time consuming part of this decoding system. Different ways to do the computation of IDWT in parallel are reviewed, in which NVIDIA, GPU gives best results in terms of time as well as in terms of energy consumption, too. Moreover, CPU-GPU pipeline decoding system can be used to make efficient CCSDS recommended DWT Decompressor for better performance.

VI. REFERENCES

[1]. "Image Data Compression", Recommendation for space data system standards, CCSDS 121.0-B-1. Blue Book,November 2005.

[2]. Changhe Song, Yunsong Li, and Bormin Huang,"A GPU-Accelerated Wavelet Decompression System With SPIHT and Reed-Solomon Decoding for Satellite Images",IEEE journal of selected topics in applied earth observations and remote sensing, vol. 4, no. 3, september 2011.

[3]. Abhishek S. Shetty, Abhijit V. Chitre and Yogesh H. Dandawate, "Time Efficiency Comparison of Wavelet and Inverse Wavelet Transform on Different Platforms",International Conference on Computing Communication Control and automation (ICCUBEA) IEEE-2016.

[4]. Anastasis Keliris, Vasilis Dimitsasy, Olympia Kremmyday, Dimitris Gizopoulosy and Michail Maniatakosz, " Efficient parallelization of the Discrete Wavelet Transform algorithm using memory-oblivious optimizations", 25th International Workshop on Power and Timing Modeling, Optimization and Simulation (PATMOS),pp.25-32, 2015

[5]. John Nickolls,"GPU Parallel Computing Architecture and CUDA Programming Model " , Hot chips 19 Symposium(HCS) IEEE,pp.1-12, 2007

[6]. Khoirudin and Jiang Shun-Liang ,"Gpu application in cuda memory", Advanced Computing: An International Journal (ACIJ), Vol.6, No.2,pp.1-10, March 2015

[7]. NVIDIA, " NVIDIA CUDA Compute Unified Device Architecture", United States, 2007.

[8]. Christofer Schwartz, Marcelo S. pinho,"an energy consumption analysis of ccsds image compressor running in two different platforms", IGARSS-IEEE, pp. 1640-1650,2014.

[9]. William A. Pearlman,Amir Said," Set Partition Coding: Part I of Set Partition Coding and Image Wavelet Coding Systems",foundations and trend in signal processing,Vol. 2,No.2, pp.97-180,2008.

[10]. William A. Pearlman,Amir Said," Image Wavelet Coding Systems: Part II of Set Partition Coding and Image Wavelet Coding Systems",foundations and trend in signal processing,Vol. 2,No.3, pp.181-246,2008.

[11]. William A. Pearlman,Amir Said," A New, Fast, and Efficient Image Codec Based on Set Partitioning in Hierarchical Trees", ieee transactions on circuits and systems for video technology, vol. 6, no. 3,pp.243-249, june 1996

[12]. Jay W. Schwartz, Richard C. Barker," Bit-Plane Encoding: A Technique for Source Encoding", IEEE transactions on aerospace and electronic systems, vol. aes-2, no. 4, pp.385-392, july 1966.

[13].

© 2018 IJSRSET | Volume 4 | Issue 2 | Print ISSN: 2395-1990 | Online ISSN : 2394-4099
National Conference on Advanced Research Trends in Information and g Technologies (NCARTICT-2018), Department of IT, L. D. College of Engineering, Ahmedabad, Gujarat, India In association with International Journal of Scientific Research in Science, Engineering and Technology

Decision Support System for reducing Post-Harvest Loss in Supply Chain Logistics using geographical Information System in Agriculture

Megha K. Patel[1], Hiteishi M. Diwanji[2]
[1]M.E.(I.T.) Student, I.T. Department, L.D. College of Engineering, Ahmedabad, Gujarat, India
[2]Associate Prof., I.T. Department, L.D. College of Engineering, Ahmedabad, Gujarat, India

ABSTRACT

Decision support system (DSS) is a broad concept which caters to the suit of ICT based applications assisting decision makers in order to take appropriate and timely decisions. DSS can help decision makers to make appropriate queries and take particular decision about a particular a particular horticulture crop in order to decide about new areas for plantation as well as supply chain logistics. This system will help in making future projections and consumptions and make scenarios for demand and supply. This system basically works on the premise of spatial network analysis, Site-suitability analysis for cold storages, shortest path analysis, image (raster) operations as geo-spatial layer provided to the system. This system is combination of vector and raster data handling in an integrated way. This DSS software site suitability module is designed for apple and it has been tested with available data for shimla district of Himachal Pradesh and network analysis module is designed potato and it has been tested with the available data for Gautambuddhnagar district of Uttar Pradesh.

Keywords : DSS, AHP, Shortest Path, Advisory, Post-harvest loss, Supply chain management

I. INTRODUCTION

Even though India is second largest country in the production of fruits and vegetables in the all over world after china we need to import fruits from the outside country. It will indirectly affect on our Indian economy. There is post-harvest loss due to lack of improper storage facilities, not using the modern technologies, improper preservation techniques and issues in market availabilities. Farmers of the rural areas do not have access to sufficient storage facilities for storing their produce during post-harvest periods. So, mainly production loss is due to lack of proper infrastructure, supply chain entities like cold storages, market outlets, mandis, transportation entities, proper road network entities especially in hilly areas. To improve production amount, it is extensively required to provide proper infrastructure facilities. Through an integrated management of their supply chains, companies react to increasingly competitive markets and trends towards globalization. Supply chain management (SCM) shown in Figure 1 can be defined as the task of integrating organizational units along the supply chain and coordinating material, information and financial flows in order to fulfill (ultimate) customer demands with the aim of improving competitiveness of a supply chain as a whole. As such, SCM should result in an internally consistent view on how a supply chain should look like in terms of production and distribution processes and their coordination.

Figure 1. Supply chain entities

II. METHODS AND MATERIAL

Decision Support System Characteristics

Decision support system in general is a computer based system that is very useful in making certain important decisions based on criterions to find out solutions to the problems in hand to a decision maker. This type of system can be helpful in such an environment where problems are rapidly changed and not easily specified in advance. Decision support system (DSS) will combine all the possibilities of problem solution so that final result is more accurate than any other methods.

Decision Support System for Supply Chain Logistics will be application which will be very useful to find out the location of best cold storages nearest to the orchards. Finding out the service areas of the cold storages to identify the area to which particular cold storage can serve. Recommend the location of new cold storage so that, the areas where cold storages are not available, new cold storages can be established. Therefore, farmers of that region don't need to go faraway places to store their produce. Post harvesting losses occurring due to lack of proper supply chain entities can be minimized by having proper management plans.

1. Finding out the nearest location of cold storages from the orchard.
2. Calculate service areas for determining the areas to which cold storages can serve.
3. Recommendation of location for establishing new cold storage using site suitability analysis.

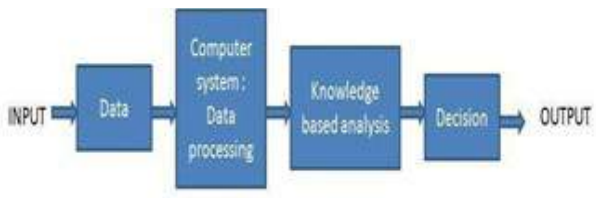


Figure 2. Simple model of Decision Support System

2.1 Site-Suitability Analysis:

To determine site suitability in any region for establishing a location of new cold storage AHP technique is used. Analytical hierarchy process (AHP) is a decision-making technique which can be used to analyze and support decisions which have multiple and even competing objectives. To do this, a complex problem is divided into a number of simpler problems in the form of a decision hierarchy. Once the hierarchy has been established, a pair wise comparison matrix of each element within each level is constructed. Participants can weight each element against each other within each level, which is related to the levels above and below it, and mathematically tie the entire scheme together. AHP is often used to compare the relative suitability of a small number of alternatives concerning the overall goal.

2.2 Network Analysis:

Networks are necessary for the movement of people, transportation of goods, communicate information and control of the flow of matter and energy. Network application is quite vast. Phenomena that are represented and analyzed as networks are roads, railways, cables, and pipelines. A graph is a mathematical abstraction that is useful for solving different networking problems. Finding the shortest paths plays an important role in solving network based systems. In graph theory, a number of algorithms can be applied for finding the shortest path in a graph based network system. This reduces the complexity of the network path, the cost, and the time to build and maintain the network based systems.

2.3 Advisory:

Agriculture sector requires innumerable types of data analytics in various sectors such as crop productivity prediction models, economic models, pest and crop disease prediction models, and so on. Big data analytics in agriculture can be used to analyze huge volumes of structured as well as unstructured data to generate faster and more accurate results that will aid in faster and accurate decision making.

III. RESULTS AND DISCUSSION

3.1 Site-Suitability Analysis:

We proposed the use of GIS based on analytical hierarchy process (AHP) to select sites suitable for building new cold storages. It not only includes environmental factors but also economic factors. We graded every criterion from 1 to 10; 1 being the least suitable and 10 being the most suitable. Thus, our

method combines qualitative and quantitative criteria for establishing location of new cold storage.

3.1.1 Calculating criteria weights by AHP:
Matrix is created according to the size of the parameter.

For example, for seven parameters 7x7 matrix is used.
Step 1. Complete the Matrix.
Sum of all columns.

Table 1. Fill matrix operation

B1	C1	C2	C3	C4	C5	C6	C7
C1	1	1	3	4	5	9	7
C2	1	1	2	3	5	8	6
C3	1/3	1/2	1	1	2	7	3
C4	1/4	1/3	1	1	1	7	9
C5	1/5	1/5	1/2	1	1	1	8
C6	1/9	1/8	1/7	1/7	1	1	1
C7	1/7	1/6	1/3	1/9	1/8	1	1
SUM	**3.03**	**3.28**	**7.75**	**10.25**	**15.125**	**34**	**35**

CR=0.089, B1= Parameters, C1 =Crop, C2 =Wetland, C3 =Road, C4 =DMSP, C5 =Elevation, C6 =Slope, C7 =LULC, W is the weight of C1–C7 to B1.

Step 2. Normalization and weight determination Each cell value is divided by sum of that column value. Average of all the rows is calculated and store in column sum. Average of sum column is 1.

Table 2. Normalization and weight determination

B1	C1	C2	C3	C4	C5	C6	C7	SUM
C1	0.32924	0.32924	0.10974	0.08231	0.06584	0.03658	0.04703	**0.31307**
C2	0.300752	0.30075	0.15037	0.10025	0.06015	0.03759	0.05012	**0.27294**
C3	0.376119	0.25074	0.12537	0.12537	0.06268	0.01791	0.04179	**0.12955**
C4	0.390093	0.29257	0.09752	0.09752	0.09752	0.01393	0.01083	**0.13351**
C5	0.330579	0.33057	0.13223	0.06611	0.06611	0.06611	0.00826	**0.08187**
C6	0.264706	0.23529	0.20588	0.20588	0.02941	0.02941	0.02941	**0.03287**
C7	0.2	0.17142	0.8571	0.25714	0.22857	0.02857	0.02857	**0.03086**

Step 3. Calculation of consistency ration (CR)
Value of consistency ratio must be <= 0.1.

CR = Consistency index(CI) / Random Consistency Index(RI)

Consistency Index (CI) = λ – n/ n-1

λ= Principal Eigen value, n= number of factors
0.31307*1+0.2729*1+0.1295*3+0.133*4+0.081*5+0.0328*9+0.030*7 =2.4166 2.4166/0.31307 = 7.719

Like this all the values of rows calculated and average of **these seven values are considered as a λ.**
CI = 7.707133 – 7/ 7-1
CI = 0.11785
CR = 0.1178/ 1.32 = 0.089 < 0.1

If value of CR is less than 0.1 then the weights are correct for all parameters.

Now, C1 to C7 all parameter's Raster Layer is given as an input according to it's weights generated by AHP. After all this process final output Raster file will generated which is shown in figure 3.

This case study illustrates the process of identifying a single or a few optimal sites. In the end, the best landfill areas were given, and they can be taken as the optimal suitable sites for establishing new cold storages. The better suitable sites for establishing new cold storages areas can be taken as back-up sites.

10 = Most Suitable Sites
1 = Least Suitable Site

Output:

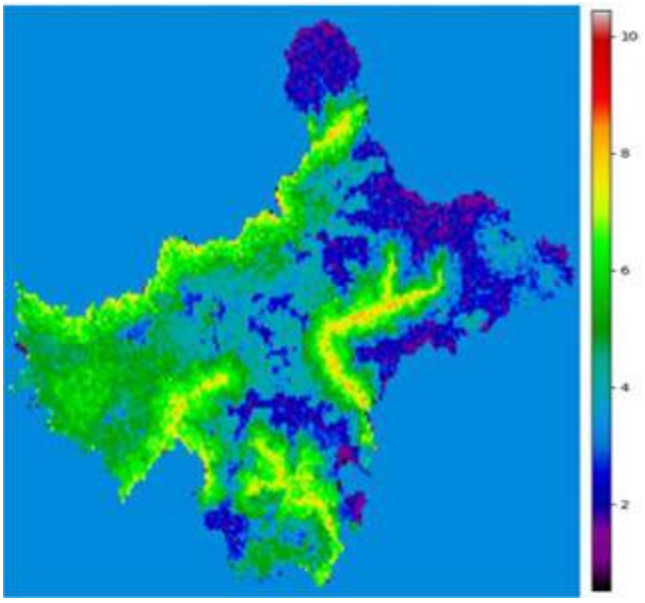


Figure 3. Suitable Sites for establishing new coldstorages in Shimla district

3.2 Network Analysis:
The network analysis is used to find transportation facility from production field to the mandis, processing units, cold chain entities. Network analysis also used for finding out the cost effective path, shortest path in terms of time or distance. Network analysis is needed to find out shortest path from orchard to the nearest cold storage location. Decision support system is a facility management system for supply chain in order to easily help in showing the shortest path. But that shortest path to the cold storage should be based on available capacity of the cold storage. Because sometimes nearest cold storage does not have that much capacity to store the produce of farmers, thereby increasing the cost of transportation unnecessarily. So, it is one kind of facility management system for farmers helping them to get the nearest location of cold storage based on produce capacity and availability of space.

Output:

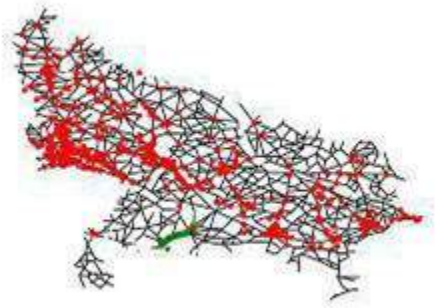


Figure 4. Shortest path from selected source to nearest coldstorages for U.P. State

3.3 Advisory:
An advisory module can be developed using various analytical services like rainfall prediction, crop recommendation, procure agro inputs recommendations, supply chain management, crop disease alert, fertilizer recommendations, etc. based on agricultural and weather data. The role of domain expert/ agricultural scientist is to generate recommendations and constitute decision policies based on analytical results.

IV. CONCLUSION

Decision Support System will help decision makers to make a decision about a particular horticulture crop in order to decide about new areas for the plantation as well as supply chain logistics. This system will help in

making future projections on consumption of scenarios for demand and supply. Mainly post-harvest loss is due to weak supply chain entities like road network, cold chain entities and markets. By improving this type of entities post-harvest loss can be reduced. For that network analysis, proximity analysis, buffer analysis performed.

V. REFERENCES

[1]. A Research on Agricultural Products Supply Chain and Food Safety by Hongling Zheng and Lili Lu

[2]. Landfill site selection using spatial information technologies and AHP: A case study in Beijing, China by Wang Guiqin, Qin Li, Li Guoxue and Chen Lijun

[3]. Big Data Analytics Architecture for Agro Advisory System by Purnima Shah, Deepak Hiremath and Sanjay Chaudhary

[4]. An optimization approach for managing fresh food quality throughout the supply chain by Aiying Rong, Renzo Akkerman and Martin Grunow

[5]. An application of Dijkstra's Algorithm to shortest route problem by Nathaniel Akpofure and Nsikan Paul

[6]. Mapping and Analyses of Post Harvesting Supply Chain Entities of Mango Using GIS by Krishnaveni M. and Kowsalya M.

[7]. Loss in Perishable Food Supply Chain: An Optimization Approach Literature Review by Yared Lemma, Daniel Kitaw and Gulelat Gate

[8]. A Design of Intelligent Route Guidance System Based on Shortest Path Algorithm by Xipeng Zhang, Gang Xiong, Liang Xiao (Correspondence Author), Fenghua Zhu, Xiongguang Yang and Timo R. Nyberg

© 2018 IJSRSET | Volume 4 | Issue 2 | Print ISSN: 2395-1990 | Online ISSN : 2394-4099
National Conference on Advanced Research Trends in Information and Computing Technologies (NCARTICT-2018), Department of IT, L. D. College of Engineering, Ahmedabad, Gujarat, India In association with International Journal of Scientific Research in Science, Engineering and Technology

A Survey on Multiple Image Encryption Using Chaos Based algorithms And DNA Computing

Aarti Patel[1], Dr.Mehul Parikh[2]
[1]M.E(I.T) Student, I.T Department, L.D College Of Engineering, Ahmedabad,Gujarat, India
[2]Associate Prof., I.T Department, L.D College Of Engineering, Ahmedabad, Gujarat, India

ABSTRACT

Due to the development in the field of network technology and multimedia applications, every minute thousands of messages which can be text, images, audios, videos are created and transmitted over wireless network.So encryption is used to provide security. To ensure the security of image transmission,people have proposed many single-image encryption(SIE) algorithms. In the age of big data,although multiple images can be repeatedly encrypted by the SIE algorithm in theory,the encryption effciency is always ineffcient.The encryption algorithm should plaintext sensitive,key sensitive,and lossless.DNA technology has been used with chaotic cryptosystem to double assurance the security of image cryptosystem by chaotic system and DNA biological manipulation.In this paper different multiple image encryption technique based on chaotic map and DNA computing have been studied.
Keywords: Image Encryption,Algorithm,Chaotic Map,DNA Computing.

I. INTRODUCTION

Algorithms, such as DES,AES and RSA are found unsuitable for multimedia data because these algorithms are designed for accurate data. while digital image has some intrinsic features such as bulk data capacity and high redundancy.Fibonacci,Hash,DNA,Chaos,Transform domain and S-box, have been proposed to be applied to image encryption in the past decade.

A. Requirements Of Image Encryption

- Ability to get pixel values from image.
- Create strong encrypted image so that can not easily hacked.
- Faster encyption time so that can easily transfer to person.
- Lossless image which can be get after decrypting it.
- Confusion process in which the pixel positions are permuted to reduce inter-pixel correlation.
- Diffusion process in which consists of some reversible computations that change the pixel values.

B. Parameters Consider For Security Of Image

a. Key Space Analysis

For an image encryption algorithm to have high security, key space should to at least as large as to resist brute force attack.

b. Key Sensitivity

An encryption algorithm should be very sensitive to any secret key. Any trivial change must lead to a different cipher-image or a wrong decrypted image, from the same cipher-image.

c. Plaintext Sensitivity

It means that any tiny change, even just one bit change, in the plain-image could cause a huge difference in the cipher-image.

d. Information Entropy

The information entropy is defined as the degree of uncertainties in the system. The greater the entropy, the more is the randomness in the image, or the image is

more uniform. Thus statistical attacks become difficult.It should be nearer to 8.

$$H(m) = \sum_{i=0}^{2^N-1} p(mi) \times log_2[\frac{1}{p(mi)}]$$

where p(mi) represents the probability of symbol mi , and log2 represents the base 2 logarithm so that the entropy is expressed in bits, N represents the number of bits we use to represent a pixel, and for one colour channel of a pixel, it is clear that N = 8. If an image is ideal random, then for each i, p(mi) = 1/256, and we can easily find that H(m)= 8.

e. NPCR

Number of Pixels Change Rate (NPCR) stands for the number of pixels change rate while one pixel of plain image changed. The NPCR gets closer to 100 to the changing of plain image, and the more effective for the cryptosystem to resist plaintext attack.

$$NPCR = \frac{\Sigma_{i,j} D(i,j)}{M \times N} \times 100\%$$

f. UACI

UACI(Unified Average Changing Intensity) stands for the average intensity of differences between the plain image and ciphered image. The UACI gets closer to 33.333.

$$UACI = (\sum_{i,j} \frac{|C'(i,j) - C(i,j)|}{255})/(M \times N))$$

where M and N are the width and height of the encrypted image, respectively. C and C' are the cipher images, whose corresponding plain images have only one pixel difference. Clearly, in order to with stand the differential attack, the NPCR and UACI values for an ideal cryptosystem should be large enough.

g. Computational Time

It should be less so encryption speed increase.

h. Image Restoration

The cipher-image can be fully recovered by the re-ceiver without loss of data.

i. Robustness

To evaluate robustness of algorithm, attack the encrypted image by salt and pepper noise and block removal. algorithm should robust enough to moderate noise contamination and block missing.

j. Correlation Of Two Adjacent Pixel

It tells us how much there is relation between the same pixels of the original and the encrypted image. The adjacent pixels in plain image are usually highly correlated, which is a weakness to statistical attack. An image encryption should decrease the correlation of two adjacent pixels in the ciphered image. To test the correlation between two vertically adjacent pixels, two horizontally adjacent pixels, and two diagonally adjacent pixels, The result indicates that the correlation coeffcients of the plain image are always nearly equals 1, while that of the ciphered image are greatly reduced to close 0.

$$r_{xy} = \frac{E[(x - \gamma_x)(y - \gamma_y)]}{\eta_x \eta_y}$$

where x and y are the gray values of two adjacent image pixels, and E[.] represents the expectation value, denotes the mean value, and η indicates the standard deviation.

II. PRELIMINARIES

A. Chaos Theory

Chaos is supposed to be that the smallest of changes in a system can result in very large differences in that systems behavior.Chaos is a deterministic, random like process found in nonlinear, dynamical system, which is non-period, nonconverging and bounded.Moreover, it has a very sensitive dependence upon its initial condition and parameter.The chaotic sequences are uncorrelated when their initial values are different and spread over the entire space.A chaotic map is a discrete-time dynamical system, defined as the following Eq. 1:

$$x_{k+1} = f(x_k), x \in (0,1), k = 0,1,2,3..$$

B. DNA Computing

A DNA sequence contains four nucleic acid bases A(adenine),C(cytosine), G(guanine), T(thymine),where A and T are complementary, G and C are complementary.Because 0 and 1 are complementary in the binary, so 00 and 11 are complementary, 01 and 10 are also complementary.By using four bases A, C,G and

T to encode 00; 01; 10 and 11, there are 24 kinds of coding schemes.But there are only 8 kinds of coding schemes that used, which are shown in Figure 1 DNA sequence encoding table.

	1	2	3	4	5	6	7	8
A	00	00	01	01	10	10	11	11
T	11	11	10	10	01	01	00	00
G	01	10	00	11	00	11	01	10
C	10	01	11	00	11	00	10	01

Figure 1. The Encoding And Decoding Rules For DNA Sequences.

C.Benefits Of DNA Computing

- Extraordinary information density,
- Massive parallelism and
- Ultra low energy consumption.

III. LITERATURE REVIEW

A. Multiple-Image Encryption With Bit-Plane Decomposition And Chaotic Maps

Tang proposed algorithm that decomposes input images into bit planes, randomly swaps bit blocks among different bit planes, and conducts XOR operation between the scrambled images and secret matrix controlled by chaotic map. Finally, an encrypted PNG image is obtained by viewing four scrambled grayscale images as its red, green, blue and alpha components.
Some techniques [7,8] can encrypt multiple images, but their decrypted images are not completely the same with the original images. This means that they are lossy algorithms and thus are not suitable for those applications requiring images with good visual quality, such as medical images.The proposed algorithm reaches good performances in security, robustness, and computational time. It can losslessly retrieve original images from the encrypted images.

This algorithm is robust against salt and pepper noise attack and block removal. These techniques ensure that it is difficult to observe useful trace between secret keys and plaintext/ciphertext.But this algorithm encrypt only 4 grayscale images.

B. Multiple Image Encryption Algorithm Based On Mixed Image Element And Chaos

Zhang,Wang proposed algorithm based on the mixed image element and piecewise linear chaotic maps (PWLCM).Firstly,The sender combines original images into a big image, and divides it into many pure image elements; secondly, she scrambles these pure image elements with the chaotic sequence generated by the PWLCM system to get mixed image elements; thirdly, she combines these mixed image elements into a big scrambled image,and segments it into small images with the equal size of original images; finally, these small images, i.e.encrypted images, are named with the filenames generated by another PWLCM system.

This novel algorithm is for k grayscale images without compression technology.Li et al. proposed a MIE algorithm based on the cascaded fractional Fourier transform [7] . most of these algorithms encrypt images in the transform domain and usually combine with the image compression technology,So the decryption images are always with some obvious distortion. Meanwhile, these algorithms require the data conversion between the spatial domain and the transform domain.
Therefore, their encryption effciency is always undesirable. The effciency and the security are contradictory in an encryption algorithm. In Tangs algorithm[1],both the order of image blocks and the content of image blocks are processed. However, only the order of image blocks is scrambled in the new algorithm.Therefore,the security of this algorithm may be a little weaker than Tangs algorithm in theory.

	Tang's algorithm[1]	Zhang's algorithm[2]
Computational time	9.656	0.191
No.of image	4	k
security	more	less
Key size	2^514	10^56

Figure 2. Comparison between Tang's algorithm[1] & Zhang's algorithm[2]

C. Lossless Chaotic Color Image Cryptosystem Based On Dna Encryption And Entropy

The proposed algorithm consists of four processes:
key streams generation process, DNA sequences confusion process, DNA sequences diffusion process and pixel level diffusion.Many DNA-based image encryption algorithms used the DNA sequence

operations such as addition, subtraction and XOR to diffuse the DNA-encoded image with absence of DNA-level confusion, which makes the security of the cryptosystems not high enough.This proposed scheme involves not only DNA-level confusion and diffusion but pixel-level diffusion, which will enhance the security, complexity and sensitivity of the cryptosystem.In this algorithm, the final secret key streams are related to both the chaotic system and the original plain-image, which increases the security level and resistance against known/chosen plaintext attacks of the cryptosystem. In order to reach higher security and sensitivity,ciphertext diffusion method employed in crisscross pattern which encrypts two equal sub-images in parallel.

D. A Chaotic Color Image Encryption Using Integrated Bit-Level Permutation

Author has proposed algorithm which convert the color image into three bit-level images (R, G, B components) and combine them to one bit-level image.Then, only use bit-level permutation architecture based on chaotic system to encrypt the integrated image.When diffuse the position of the integrated binary image, the value of the gray pixel is changed as well, so this architecture can achieve similar security to permutation diffusion architecture.

For color image encryption, most of the previous algorithms used multi-round permutation diffusion architecture and encrypt their three color component respectively which is time consuming. The encryption and decryption speed of our proposed method is 16.97 MB/s while the speed of AES with 128 bit key, AES with 192 bit key, AES with 256 bit key are 11.23 MB/s, 9.25 MB/s, 9.19 MB/s.

E. A Light Weight Secure Image Encryption Scheme Based On Chaos And DNA Computing

In most of the schemes the authors considered the statistical tests like key-space analysis, histogram analysis, correlation of two adjacent pixels, differential attack analysis, information entropy analysis, known plain-text and cipher-text only attack etc. and over all complexity but they have not given enough emphasis on memory uses and energy consumption, throughput of the algorithms.

In the proposed scheme chaotic logistic map is used which will generate a highly randomized number sequence.The chaotic logistic map runs on low com putational overhead, so it becomes an light weight PRNG. The permuted data are converted to DNA sequence.Same PRNG is again used to generate a random bit sequence. For this purpose, this binary sequence is also converted to its DNA sequence .The DNA sequences C and D are adde together which results in a new DNA sequence E. Eis again converted back to sequence of 8 bit (integer) F form. XORing of each element of the sequence is done with the elements previous to that index on F which gives the final encrypted image.

	Xiangjun Wu's algorithm[3]	Teng's algorithm[4]	Bhaskar'salgorithm[5]
Key Size	2^299	2^128	2^133
NPCR	99.6074	99.623	99.7570
UACI	33.4570	33.32	39.12
SPEED	35.24MBITS/SEC		
CORRELATION COEFFICIENT	FOR LENA RED(HORIZONTAL)=-0.0124 (VERTICAL)=-0.0001 (DIAGONAL)=-0.005 GREEN(HORIZONTAL)=-0.0038 (VERTICAL)=-0.0059 (DIAGONAL)=-0.0086 BLUE(HORIZONTAL)=-0.0075 (VERTICAL)=-0.0062 (DIAGONAL)=-0.0006	FOR LENA RED=-0.010889 GREEN=-0.018110 BLUE=-0.006140	0.001178542895092

Figure 3. Comparison between Xiangjun Wu's algorithm[3], Teng's algorithm[4] & Bhaskar's algorithm[5]

IV. CONCLUSION

In this literature Survey ,It is concluded that image encryption is more secure if confusion and diffusion process is more complex.An encryption method should more dependent on plain image to resist against plaintext attack and reduce correlation between adjacent pixels to enhace resistance against stastical attack.So algorithm proposed in Multiple image encryption based on mixed image element and chaos is less secure.

V. REFERENCES

[1] Zhenjun Tang, Juan Song, Xianquan Zhang, Ronghai Sun.Multiple-image encryption with bit plane decomposition and chaotic maps.Optics and Lasers in Engineering(2016).

[2] Xiaoqiang Zhang , Xuesong Wang . Multiple image encryption algorithm based on mixed image element

and chaos, Computers and Electrical Engineering (2017).
[3] Xiangjun Wu ,Kunshu Wang ,Xingyuan Wang Haibin Kan. Lossless chaotic color image cryptosystem based on DNA encryption and entropy. Springer Science+Business Media B.V. (2017).
[4] Lin Teng,& Xingyuan Wang & Juan Meng.A chaotic color image encryption using integrated bit-level permutation. Springer Science+Business Media New York (2017).
[5] Bhaskar Mondal ,Tarni Mandal .A light weight secure image encryption scheme based on chaos & DNA computing. Journal of King Saud University Computer and Information Sciences(2016).
[6] Li Y , Zhang F , Li Y , Tao R . Asymmetric multiple-image encryption based on the cascaded fractional Fourier transform. Opt Lasers Eng (2015).
[7] Zhang G, Liu Q. A novel image encryption method based on total shuffling scheme. Opt Commun (2011).
[8] Liu H,Wang X. Color image encryption using spatial bit-level permutation and high dimension chaotic system. Opt Commun(2011).

© 2018 IJSRSET | Volume 4 | Issue 2 | Print ISSN: 2395-1990 | Online ISSN : 2394-4099
National Conference on Advanced Research Trends in Information and Technologies (NCARTICT-2018), Department of IT, L. D. College of Engineering, Ahmedabad, Gujarat, India In association with
International Journal of Scientific Research in Science, Engineering and Technology

Service Level Agreements & its Impact on Cloud

Pranav Tank, Darshan Upadhyay, Kunal Khimani
Asst. Professor, I.T. Department, VVP Engineering College, Rajkot, Gujarat, India

ABSTRACT

With the growth of technology, there has been scalable increase in use of Cloud services in real world applications. Cloud becomes efficient as it works on Pay-as-you-use basis. So with increasing number of cloud service, Service Level Agreement between Service Provider (seller) and Service Consumer (customer) becomes more important which will guarantee service type and its quality. Not just withstanding the point of quality of service, SLA also works on monitoring the service and refining the parameters which can increase the level of service. The objective of this paper is to present comprehensive study about how SLA works in a cloud environment and how it can impact in real world applications.

Keywords: Cloud, Pay-as-you-use, SLA, Quality.

I. INTRODUCTION

Cloud computing is term used in numerous definitions by many research scholars. But in general terms it is defined for delivery of hosted service over the internet. Cloud means internet and computing means type of internet-based computing, where services such as storage, servers, applications, etc are delivered to consumer through internet. Basically usage of Cloud has been increased due to pay-as-you-use service implemented in applications. Service level agreement is more unified version established through negotiation between service consumer and provider.

1.1 Types of Cloud

Cloud is specified in mainly 4 categories: Public, Private, Hybrid and Community. The Community Cloud is rarely used and so in this paper we will be over viewing previous three versions only.

1.1.1 Public Cloud

Public Cloud is implemented for general use. Data remains publicly available to all users. Users are charged for the time duration they use the services, like usage of CPU cycles, bandwidth or storage consumed, etc [1]. Usually, a third party cloud service delivers cloud service over the internet. They are more vulnerable to security threats than other cloud models. Example of Public Clouds is Amazon Web Service (AWS), Microsoft Azure, IBM Websphere, etc.

Figure 1. Public Cloud

1.1.2 Private Cloud

Private Cloud services are delivered between business data center to internal users. Users may not be charged for the services. This is more secure and controlled model of cloud. Example of Private Clouds includes VMware, OpenStack, etc.

Figure 2. Private Cloud

1.1.3 Hybrid Cloud

Hybrid Cloud services is combination of Public cloud and on-premises Private Cloud. Goal of Hybrid cloud is to use infrastructure of Public Cloud and still maintain vulnerability of the data.

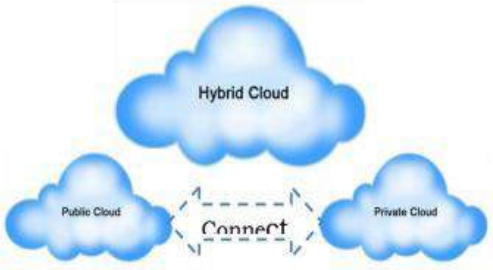


Figure 3. Hybrid Cloud

1.2 Characteristics of Cloud

On-Demand Service: These are services required by the customer to manage their resources. Service is provided over internet by provider to customer so that they can manage their own computing resources.

Resource Pooling: Cloud service provides pool of resources that can rapidly provisioned and be released with minimal efforts. Consumers can access resource directly from remote data centre.

Elasticity: Cloud services can be requested or managed from cloud providers as per customers request. Elasticity means service can be scaled-up or scaled-down as per customer requirement.

Measured Service: Services are billed according to customers requirements. Customers demand defines services required.

1.3 Service Models of Cloud

Cloud computing is computational process in which service are delivered over network using computing resources. There are three main types of service models:

- ✓ Software as a Service (SaaS)
- ✓ Platform as a Service (PaaS)
- ✓ Infrastructure as a Service (IaaS)

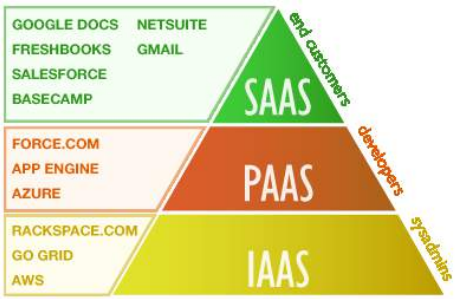


Figure 4. Cloud Service Models

1.3.1 Software as a Service (SaaS)

In this service model, consumer uses application running on a cloud infrastructure. Few popular examples for SaaS, as shown above in Fig. are Salesforce, Netsuite, Basecamp etc.

1.3.2 Platform as a Service (PaaS)

In this service model consumer uses platform which includes all system and environments consisting all phase of SDLC life cycle. Few popular examples of PaaS, as shown in Fig. are Azure, App Engine etc.

1.3.3 Infrastructure as a Service (IaaS)

In this service model provider delivers the user the infrastructure over the internet. User is able to deploy and run various softwares including application or system softwares. Few popular examples of IaaS, as shown in Figure are Rackspace.com, AWS etc.

II. GENERAL OVERVIEW OF SERVICE LEVEL AGREEMENT (SLA)

Main components of SLA include SLA Definition, SLA components, Customer and Service Provider Definition and Qualities & Parameters determined by SLA.

2.1 SLA Definition

"Service Level Agreement" is defined as format that contains explanation of the agreed service, parameters of the level of service, the guarantees regarding the Quality of Service, and arrangements for all cases of violations. In short, SLA is the tool for protecting the stability of the service[3]. The SLA is very significant as a contract that is held between the provider of the service and another party who could be one of following; consumer of the service, broker negotiator, or monitoring negotiator. SLA can be considered as quantifiable contract, the service that service provider will provide

and what sanction service provider will pay if the dedicated objective is not met.

Parameters that SLA includes for Cloud service are Web services, Networking, Internet, Data centre management, etc. Most summarized SLA contains terms like: business endings, pricing strategy etc.

2.2 SLA Components

The more refined SLA improves customer acceptance level, as it supports provider to focus and work on needs prioritized by customers. It also helps to check whether process is on the right path. Consumer can even check the quality and examine the service based on Service Level Objective (SLO) determined in SLA.

Quality of Service (QoS) is one of the Enhanced terms which is measured with Key Performance Indicator (KPI) component in SLA. Customers check whether in the agreed contract whether or not these indicators match Service Level Objective (SLO). SLO contains levels of parameters, specific value, behaviour of services etc as a target to achieve. Actual parameter values are compared with stated ones to evaluate the performance. These indicators are used to check and test these parameters and are used later in determining the violation.

2.3 Consumer and Provider definition

The word Service Provider refers to firm or business which provides service. Service supplier may be network, worker, transporter, Internet Service Provider (ISP) or an Application Service Provider (ASP)[4]. The word Service Customer refers to firm or organization that employs the service supplied by the service provider. Service providers exploit this base to enhance their usage of infrastructure to achieve signed conditions of services. Service consumers exploit the SLA to reach the stage of quality of service they require and to keep appropriate business models for a long period.

2.4 Parameters determined by SLA

There are two groups of qualities that are identified by SLA: measurable and un-measurable qualities. Measurable qualities could be measured in metrics; whereas Un-measured qualities could be measured from specified estimation. Metrics are exploited in monitoring procedure, software procedure enhancement, business policy employment, and mainly any area where information has to be gathered to confirm whether objectives are being achieved. These metrics decides the amounts needs to be collected to confirm whether SLA parameters are achieved[6]. Note that, if the parameters are declared and negotiated, the changes needs to be applied to both consumer and provider side[5]. Following are the parameters:

- Reasonable, which allows contributor to encourage to perform well. For example, SLA parameter identifies cost depending on amount of service used.
- Achievable, all the parameters needs to be included in the metrics. For example, unexpected kind of communication over the internet.
- Quantifiable, metrics must be measurable and permit for measurements. For example, if gathering metrics exploits important resources, it might not deserve to spend any effort.

III. LIFE CYCLE OF SLA

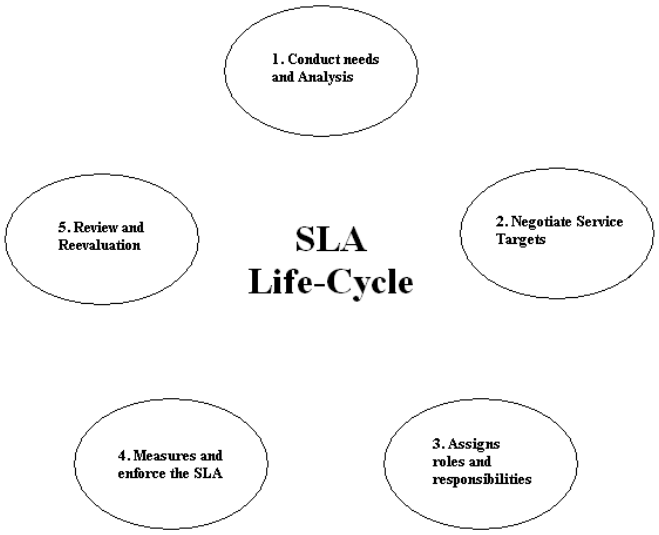


Figure 5. SLA Life Cycle

SLA includes 5 stages to be completed. These stages consists of; Needs and Analysis, Negotiation, Assigning roles and responsibilities, Measures and enforcement and Review[3].

3.1 Needs and Analysis:

This stage includes the identification of customer requirements and needs, the network capabilities, the identification of the suitable service features and parameters, service's levels, service executional environment, and the implementation of the standard of SLA templates.

3.2 Negotiate Services:

This stage includes the negotiation of an SLA with the consumer to select the values of SLA parameters related to specific services, the costs gained from the service customer after signing the SLA, the costs incurred by the service provider when the SLA is violated, the definition and at last periodicity of the reports associated with service to be delivered to the service customer.

3.3 Assign Roles and Responsibilities:

This stage include the service's resource provisioning, where the service is enabled and prepared for the consumption, configuration of the network which might be to achieve specific requirements in the service, or to support the service network overall, and service activation. Service provisioning and deployment stage may need the reconfiguration of the service resources to support the executional stage which will lead to a successful achievement of the SLA parameters. Thus it becomes very important to assign roles of each user and make them aware of the strategy they need to work on.

3.4 Measure and enforce SLA:

This stage includes measurement and assessment of the SLA. Assessment is done in two phase, assessment with the individual customer and overall service assessment. In first phase reviewing QoS, possible enhancement and altering requirements are examined. Whereas, in second phase, readjustment of goals, support problems and establishing service levels are measured. At this phase we would propose our work.

3.5 Review and Evaluation of SLA:

Once the SLA is signed, however, if job isn't over; SLA has to enforce the guidelines both (consumer + provider) agreed to. To evaluate the SLA one should also include Master Service Agreement (MSA) to compare each parameter. MSA defines services and liabilities in general terms, which will be same for all the consumers. Moreover, SLA is specific for the client and will do exactly that is expected.

IV. PROPOSED WORK

Generally the work proposed till now includes changes in SLA parameters, reducing the factors affecting violation of SLA or maintaining/upgrading the threshold value of SLA. But in our work, we would like to propose an automation system which focuses on MSA rather than on SLA. To upgrade the parameters, majorly parameters included in SLA are compared with MSA. So before comparison we can develop a automation system which compares the parameters predefined in MSA with the parameters included in SLA. The changes can be marked directly between them. If the difference is large the system will upgrade the required parameter/s directly so as to build a more effective SLA. This will also help in reducing the violation ratio.

Following are the parameters which are predefined in MSA and also required while designing a SLA.

Table 1. Parameters included for designing SLA in Cloud Environment.

Object	No. Of metrics Parameters	General Measurement (Unit)
Hardware	11	Time/Number
Software	05	Time
Network	10	Time/Percent
Storage	10	Time/Percent/Number
Service Desk	09	Percent/No. ff Language

4.1 Working of Metric Parameters

As shown in Table-1, Hardware includes parameters like Availability, Response Time, Failure Frequency, Processor Time, and Instruction per second, Number of Workstation etc. Software includes parameters like Service Time, Solution Time, and Number of Licenses etc. Network includes, Availability LAN, Availability WAN, Access RAS, Latency times, etc. Storage includes, Max. Down Time, Backup, Memory Size, Periods of Operation, etc and Service desk includes parameters like, Failure forwarding degree, Language Variety, Self Solution rate etc.

These objects and metrics are already included in Master Service Agreement (MSA). During SLA design, which differs from consumer to consumer depending on requirement, these parameters are taken into consideration. Thus, a well-versed SLA is formed, which will provide required QoS. Ultimately we look for better SLA which provides best Quality of Service. A way of increasing QoS is by changing the Key Performance Indicator (KPI) value of the parameter. But this will not be that effective as it points to the same

problem of changing the parameters after/on SLA violation.

Existing system compares the parameters only after/on SLA violation. So far seen system works on Metrics and Non-metrics parameters of SLA. While in the proposed system parameters are included depending on the customer's requirement. Depending on this requirement Service Level Objective (SLO) responsible for designing a SLA, decides whether to include the parameter or to upgrade the parameter. If the Service Level Objective is achieved above/just at the threshold value, parameter is included in SLA. If the Service Level Objective is achieved below the threshold value, parameter is considered for update and rejected.

So, we focus on parameters included in High-Level Metrics and Low-Level Metrics, for designing a SLA. Therefore, parameters that are required for designing are included in SLA using SLO mechanism. If required, parameters are upgraded/changed first and are included, this reduces the probability of violation of SLA to approx 0%. So it becomes easy for the system to compare/review the threshold for designing better SLA. This system will be beneficial for not only preventing SLA violation but also in achieving required QoS pre-defined in SLA.

V. CONCLUSION AND FUTURE WORK

Our studies on number of SLAs currently used throughout have revealed that today's prevailing contract focuses on QoS metrics, namely Availability, Response Time etc. Other parameters are mostly never mentioned and so the steps are taken for reducing the violation of SLA.

This proposed work will focus on the basics of SLA and not on reducing the violation of SLA but preventing the violation of SLA. The collection of Metrics, both high-level and low-level, will assist for developing this automation system. These metrics are easy to automate and commonly used in IT process and service. Categorization scheme is populated by updating categories, metrics, addition of parameters, etc. In summary, the proposed categories are useful to find an initial problem and possible metrics which can be automated and used as measure of performance. With the varying taxonomical structure of an enterprise, level of categories / subcategories tends to change in future.

VI. REFERENCES

[1]. Pranali Gajjar, Brona Shah: An Efficient Scalable Framework for Auto Scaling Services in Cloud Computing Environment, IJIRT, Volume 2 Issue 12, May 2016.

[2]. Hayes, I.: Five Principles for Selecting SLA Metrics, http://www.clarityconsulting.

[3]. com/five_principles.htm, Clarity Consulting, 2004.

[4]. Paschke, A., M. Bichler, Jens Dietrich ContractLog: An Approach to Rule Based

[5]. Monitoring and Execution of Service Level Agreements. Int. Con. on Rules and Rule

[6]. Markup Languages for the Semantic Web (RuleML 2005), Galway, Ireland, 2005.

[7]. Lijun Mei, W.K. Chan and T.H. Tse, "A Tale of Clouds: Paradigm Comparisons and Some Thoughts on Research Issues", IEEE Asia-Pacific Services Computing Conference, 2008, pp 464-469.

[8]. R. Pragaladan, P. Suganthi, "A Study on Challenges of Cloud Computing in Enterprise Perspective", July 2004.

[9]. M.Kriushanth, L. Arockiam, and G. Justy Mirobi, "Auto Scaling in Cloud Computing: An Overview", July 2013.

© 2018 IJSRSET | Volume 4 | Issue 2 | Print ISSN: 2395-1990 | Online ISSN : 2394-4099
National Conference on Advanced Research Trends in Information and .ting Technologies (NCARTICT-2018), Department of IT, L. D. College of Engineering, Ahmedabad, Gujarat, India In association with International Journal of Scientific Research in Science, Engineering and Technology

A Survey of Authentication of RFID Devices Using Elliptic Curve Cryptography

Suthar Monali, Prof Alka J Patel
*[1]ME , IT Department, LDCE, Ahmedabad, Gujarat, India
[2]Assistant professor , IT Department, LDCE, Ahmedabad, Gujarat, India

ABSTRACT

RFID is a wireless technology for automatic identification and data capture and it's the core technology to implement the internet of things. Because of that, the security issue of RFID is becoming more important. In past , simple mathematical and logical method, hash based schemas and simple PKI schemas are introduce RFID authentication . In this paper I illustrate about the possible security attack on RFID and three different authentication algorithm of RFID based on ECC and I also describe about why ECC is the best method among them .

Keywords: RFID, Reader, Tag, Backend Sever , Authentication , Security , ECC.

I. INTRODUCTION

RFID system is composed of tags, readers, backend sever, and antennas. RFID tags are available in affordable charges, wireless devices which can be communicate with RFID readers [1].RFID architecture shown in below figure which consist tag, reader and back-end server. Tag consist EPC (electronic product code) which store details about tag. Reader is responsible for reading and writing tag information. Back end sever will save all data about tag which are in one group .Communication in RFID network will start on reader broadcast message or query. Communication between tag-reader and reader-server is in insecure channel.

In this paper we first analyse security attack possible on RFID in section II, we discuss RFID device performance measurement in section III, three authentication method or protocol discussed in section IV.

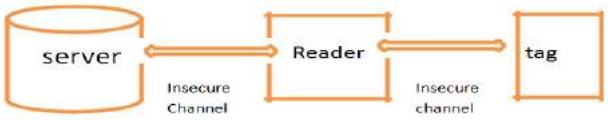


Figure 1. RFID architecture

Security Attck On Rfid

Denial of Service (DOS) : In both of wireless and wired communication, there are Denial of Service (DOS) . Once attackers control a large number of fake readers and tags, they can make the data connection to abuse computational resources, and even use up the resources and network bandwidth.[1]

Eavesdropping: The communication channel between the tag and the reader can be eavesdropped, because the radio frequency channel is not secure communication channel .[2]

User privacy: The attacker can monitor the tag using the tag identifier in order to know the user's behaviour, when the user identity is linked to a certain tag. Also, the attacker can trace the user location with the tag identifier, when the output of the tag such as the tag identifier is unchangeable.[2]

Replay attack: The attacker obtains messages between the tag and the reader by eavesdropping and reuses the message in order to impersonate a legitimate tag or a legitimate reader.[2]

Spoofing attack : The attacker impersonates a reader, sends a query to a tag, and then obtains the response of the tag. When the legitimate reader queries the tag, the attacker will send the obtained response to reader in order to impersonate the tag.[2]

Cloning attack: An attacker can build a cloned tag which will be interpreted by the reader as the legitimate tag, due to the fact that most tags are not tamper-proof.[2]

Performance

RFID schemes cannot use computationally intensive cryptographic algorithms for privacy and security because tight tag cost requirements make tag-side resources (such as processing power and storage) scarce .[3]

- Capacity minimisation: The volume of data stored in a tag should be minimised because of the limited size of tag memory

- Computation minimization: Tag-side computations should be minimized because of the very limited power available to a tag.

- Communication compression: The volume of data that each tag can transmit per second is limited by the bandwidth available for RFID tags [3]

- Scalability: The server should be able to handle growing amounts of work in a large tag population. It should be able to identify multiple tags using the same radio channel [3] Performing an exhaustive search to identify individual tags could be difficult when the tag population is large [3].

II. LITERATURAL SURVEY

Authenticity can be achieved by a secure protocol running between RFID tag and reader [4]. To achieve authentication public key cryptography (PKC), non-public key cryptography (NPKC), hash function, hash with random number, simple bitwise operation, AES, HMAC schema can be used. The suitability of PKC for RFID is an open research problem due to the limitation in tag cost, gate area and power consumption. Among PKC algorithms, ECC based algorithms would be the best choice for RFID system due to their small size key and efficient computation. So, ECC is very attractive for small devices like RFID with limited computational capacity, memory and low bandwidth network. In this paper we will discuss three ECC based RFID authentication algorithms.

a) A secure ECC based RFID authentication protocol with ID verifier

This paper[Liao's schema] proposes an ECC based mutual authentication algorithm that satsfies the essiential requirements in RFID system [2].

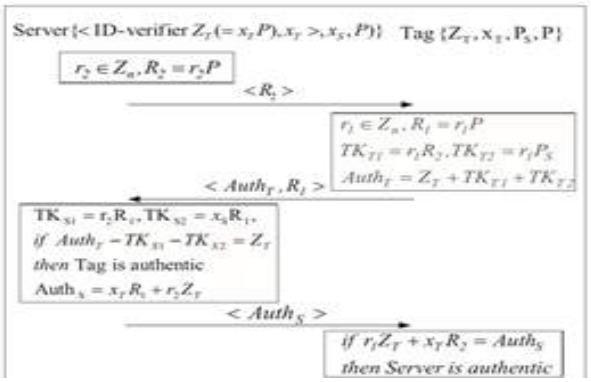


Figure 2. Liao's schema

In this algorithm tag believes that the ID verifier Zt is securely transmitted to the server and vice versa. This algorithm provide Mutual authentication, confidentiality, forward security, scalability. This algorithm resisting replay attack, tag masquerade attack, server spoofing attack, location attack, cloning attack [2].To implement this schema successfully a powerful server device needed [2]. There are also some other schema [3-6] which are more efficient then this schema in tag computational time [2].

b) Cryptanalysis and improvement of an efficient mutual authentication RFID scheme based on elliptic curve cryptography

This paper is improved version of Chou's protocol [7] based on ECC which is failed to provide mutual authentication and cloning attack. [8]

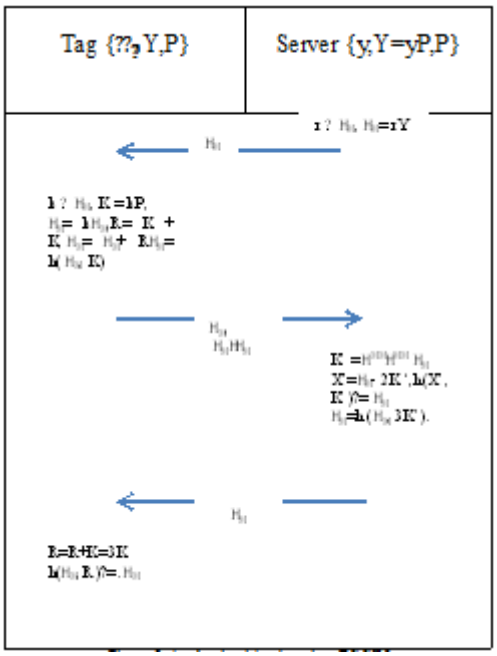


Figure 3. chou's algorithm based on ECC[9]

Weaknesses of Chou's schema: Lack of tag privacy, forward privacy and mutual authentication. Farash introduces improved version of Chou's algorithm to achieve mutual authentication and tag privacy [figure 4]. Farash's schema has proof against mutual authentication, tag privacy. Computational cost of Frash's schema is same as Chou's schema [8].The total computation of schema is very high. To improve this some pre computing technique should use [8].

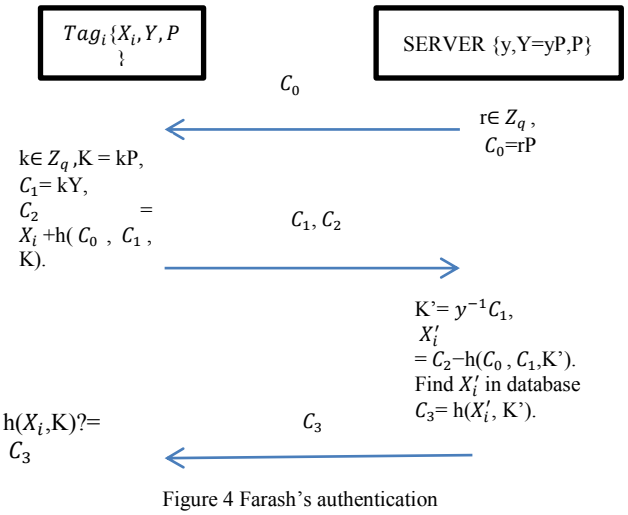


Figure 4 Farash's authentication

c) Elliptic Curve Cryptography Based Mutual Authentication Protocol for Low Computational Capacity RFID Systems - Performance Analysis by Simulations

In this paper no reader communication is only happen between back end server and tag. In this group key is use to perform authentication instead of individual key [9].

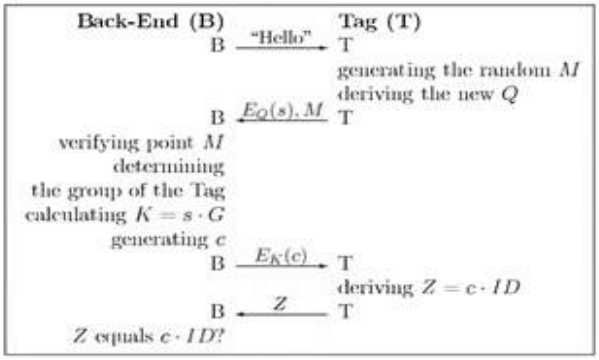


Figure 5. Godor's schema[9]

While in use of 160 bit elliptic curve might be very big and required high computational capacity and strong back-end sever[9] .This paper have implementation in OMNET ++.

Godor's schema is acceptable for every attack except DOS . This schema also prove that computational time for 112 bit and 160 bit are almost near.

III. COMPARISON

Attacks	Liao's	Farash's	Godor's
Mutual authentication	Yes	Yes	Yes
Scalability	-	-	-
DOS	-	-	No
Cloning	-	-	-
Server spoofing	-	-	-
Replay	-	Yes	-

Comparison factor	Liao's	Farash's	Godor's
Computational time	.32 sec	Not measured	.1006(160bit)

IV. CONCLUSION AND FUTURE WORK

In the conclusion three of them are ECC based authentication schema; all are proved more efficient against simple PKI, simple HASH, AES and RSA. All three have much computational time to perform authentication and all three

needed high capacity back-end sever. So in future some pre-computational method can implement with ECC.

V. REFERENCES

[1]. Xiao Nie, Xiong Zhong "Security In the Internet of Things Based on RFID: Issues and Current Countermeasures" Proceedings of the Published by Atlantis Press, Paris, France.

[2]. Yi-Pin Liao , Chih-Ming Hsiao "A secure ECC-based RFID authentication scheme integrated with ID-verifier transfer protocol " Department of computer science and information engineering , St.John's university,Taipei,ROC (2013), published by ELSEVIER, http://dx.doi.org/10.1016/j.adhoc.2013.02.004

[3]. C.P. Schnorr," Efficient identification and signatures for smart cards", in: Gilles Brassard (Ed.), Advances in Cryptology – CRYPTO'89, Lecture Notes in Computer Science, 435, Springer-Verlag, 1989, pp. 239–252.

[4]. T. Okamoto, "Provably secure and practical identification schemes and corresponding signature schemes", in: E.F. Brickell (Ed.), Advances in Cryptology – CRYPTO'92, Lecture Notes in Computer Science, 740, Springer-Verlag, 1992, pp. 31–53.

[5]. Y.K. Lee, L. Batina, I. Verbauwhede, "EC-RAC (ECDLP Based Randomized Access Control): Provably Secure RFID Authentication Protocol", IEEE International Conference on RFID, 2008, pp. 97–104.

[6]. Xinglei Zhang, Linsen Li, Yue Wu, Quanhai Zhang, "An ECDLP-Based Randomized Key RFID Authentication Protocol", 2011 International Conference on Network Computing and Information Security.

[7]. ChouJS(2013)" An efficient mutual authentication RFID scheme based on elliptic curve cryptography" .J Supercomput. doi:10.1007/s11227-013-1073-x

[8]. Mohammad Sabzinejad Farash " Cryptanalysis and improvement of an efficient mutual authentication RFID scheme based on elliptic curve cryptograph" published by Springer Science+Business Media New York 2014, DOI: 10.1007/s11227-014-1272-0

[9]. sGy6z6 Godor, Norbert Giczi, Sandor Imre Dr."Elliptic Curve Cryptography Based Mutual Authentication Protocol for Low Computational Capacity RFID Systems -Performance Analysis by Simulations", Department of Telecommunications, Budapest University of Technology and Economics Magyar Tud6sok korutja 2., Budapest, Hungary H-ll17, 978-1-4244-5849-3/10/$26.00 2010 IEEE

© 2018 IJSRSET | Volume 4 | Issue 2 | Print ISSN: 2395-1990 | Online ISSN : 2394-4099
National Conference on Advanced Research Trends in Information and Technologies (NCARTICT-2018), Department of IT, L. D. College of Engineering, Ahmedabad, Gujarat, India In association with
International Journal of Scientific Research in Science, Engineering and Technology

Design and Implementation of OCR to identify English Characters and Numbers

Rachit Adhvaryu*, Rachana Parikh, Komil Vora
Information Technology, V.V.P. Engineering College, Rajkot, Gujarat, India

ABSTRACT

Optical character recognition has become one of the most successful applications of technology in the field of pattern recognition and artificial intelligence. At present, there is great demand of such softwares which can identify the characters from scanned documents or images. Optical Character Recognition deals with the problem of recognizing optically processed characters. Optical recognition is performed off-line after the writing or printing has been completed, as opposed to on-line recognition where the computer recognizes the characters as they are drawn. This paper presents the system which identifies characters from the images. The objective of this system prototype are to develop a prototype for the Optical Character Recognition (OCR) system and to implement the Template Matching algorithm in developing the system prototype.

Keywords: OCR, Template Matching, Gray Scale Images, Color Images, Fixed Size Templates

I. INTRODUCTION

Optical character recognition (OCR) is the process of classification of optical patterns contained in a digital image corresponding to alphanumeric or other characters [1]. OCR has attained much popularity in the both academics as well as in industry. Over the last few years, machine reading has grown rapidly through the development of much more sophisticated and easy OCR systems. OCR Technology allows us to convert scanned documents, pdf files and images from digital camera to editable and readable form [1]. Optical character recognition belongs to the family of techniques performing automatic identification [2]. OCR has become most prominent and successful technological applications in the field of pattern recognition and artificial intelligence [2]. Below we discuss these different techniques and define OCR's position among them.

Automatic Identification

The traditional way of entering data into a computer is through the keyboard. However, this is not always the best nor the most efficient solution. In many cases automatic identification may be an alternative. Various technologies for automatic identification exist, and they cover needs for different areas of application. Below a brief overview of the different technologies and their applications is given [1].

a. Speech recognition

In systems for speech recognition, spoken inputs from predefined library of words are recognized. Such systems should be speaker-independent and may be used for instance for reservations or ordering of goods by telephone. Another kind of such systems are those used to recognize the speaker, rather than the words, for identification [1].

b. Radio frequency

This kind of identification is used for instance in connection with toll roads for identification of cars. Special equipment on the car emits the information. The identification is efficient, but special equipment is needed both to send and to read the information [2]. The information is also inaccessible to humans.

c. Vision systems

By the use of a TV-camera objects may be identified by their shape or size. This approach may for instance be used in automatons for recirculation of bottles [2]. The

type of bottle must be recognized, as the amount reimbursed for a bottle depends on its type.

d. Magnetic stripe

Information contained in magnetic stripes is widely used on credit cards etc. Quite a large amount of information can be stored on the magnetic stripe, but specially designed readers are required and the information cannot be read by humans [2].

e. Bar code

The bar code consists of several dark and light lines representing a binary code for an eleven digit number, ten of which identify the particular product. The bar code is read optically, when the product moves over a glass window, by a focused laser beam of weak intensity which is swept across the glass window in a specially designed scanning pattern. [2] The reflected light is measured and analysed by a computer.

f. Magnetic ink

Printing in magnetic ink is mainly used within bank applications. The characters are written in ink that contains finely ground magnetic material and they are written in stylized fonts which are specifically designed for the application. Before the characters are read, the ink is exposed to a magnetic field. This process accentuates each character and helps Simplify the detection. The characters are read by interpreting the waveform obtained when scanning the characters horizontally [2]. Each character is designed to have its own specific waveform. Although designed for machine reading, the characters are still readable to humans. However, the reading is dependent on the characters being printed with magnetic ink.

g. Optical Mark Reading

This technology is used to register location of marks. It may be used to read forms where the information is given by marking predefined alternatives. Such forms will also be readable to humans and this approach may be efficient when the input is constrained and may be predefined and there is a fixed number of alternatives [2].

h. OCR

Optical character recognition is needed when the information should be readable both to humans and to a machine and alternative inputs can't be predefined. In comparison with the other techniques for automatic identification, optical character recognition is unique in that it does not require control of the process that produces the information [1].

II. OPTICAL CHARACTER RECOGNITION

Optical Character Recognition deals with the problem of recognizing optically processed characters. Optical recognition is performed off-line after the writing or printing has been completed, as opposed to on-line recognition where the computer recognizes the characters as they are drawn. Both hand printed and printed characters may be recognized, but the performance is directly dependent upon the quality of the input documents [2, 3].

The more constrained the input is, the better will the performance of the OCR system be. However, when it comes to totally unconstrained handwriting, OCR machines are still a long way from reading as well as humans. However, the computer reads fast and technical advances are continually bringing the technology closer to its ideal.

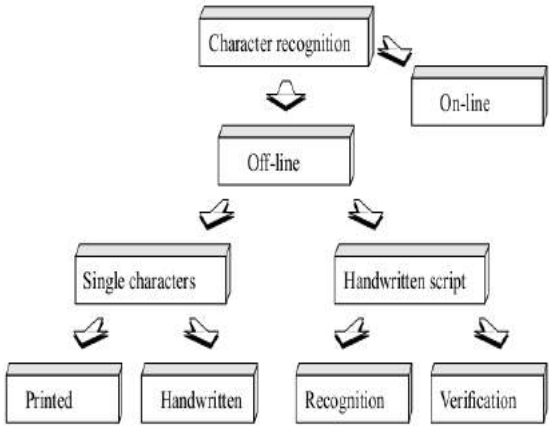


Figure 1. Types of character recognition

A. Online Character Recognition

Today, many tools are available over internet which identifies the given characters in the fraction of time [3]. These tools are mainly used in Multiple Choice Questions examinations conducted using OMR Sheets.

B. Offline Character Recognition

Offline character recognition deals with developing the systems as per the requirement of the user or as per the requirement of desired output. This type of character recognition can be done on the following:

i. Single Character

In this, the system is created in such a way that it identifies one single character at one time [3]. The single characters can be in 2 ways:

a) Printed Characters

In this, the system tries to recognize the printed characters either in the form of text file or an image.

b) Handwritten Characters

In this, the system tries to recognize the handwritten characters which is in the form of scanned document and image.

ii. Handwritten Scripts

The system developed for handwritten scripts are mainly used for 2 purposes:

a) Recognition

This type of system mainly recognizes a particular set of characters from the given handwritten scripts. These systems split the recognized characters in a single character or the group of characters depending on the requirement of system [3, 4].

b) Verification

Once the characters are recognized, the correctness of the characters is required to be validated. The verification deals with the correctness of shapes, diagonals and curves of characters [3]. Also using various Natural Language Processing techniques, these characters once verified can classified in many different classes.

III. METHODS OF OCR

There are 2 methods of OCR:

A. Matrix Matching

The Matrix Matching is technique in which the library of characters is created. The system then compares the scanned characters with the library character matrices. This system works best when the characters to be scanned and the library characters have very little or no variation in style [4].

B. Feature Extraction

Feature Extraction generally deals with the features of characters like shape, closed areas, diagonal lines, line interaction and curves. It is more effective and flexible methods as it has a wide scope to identify the same character with different shapes and dimensions [4].

In this paper, we are focus on Matrix Matching (Template Matching) method.

IV. COMPONENTS OF OCR

A typical OCR system consists of several components. In figure 2 a common setup is illustrated. The first step in the process is to digitize the analog document using an optical scanner. When the regions containing text are located, each symbol is extracted through a segmentation process [4, 5]. The extracted symbols may then be pre-processed, eliminating noise, to facilitate the extraction of features in the next step.

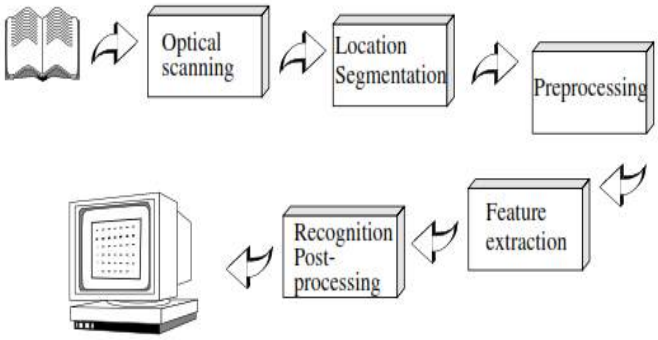


Fig.2 Components of OCR

A. Optical Scanning

It converts multilevel colored image to bi-level image. It uses "Thresholding" method. The quality of recognition depends on bi-level image. Thresholding can be fixed or Variable depends on contrast and brightness of bi-level image [5].

B. Location and Segmentation

It deals with isolation of characters or words. It segments the lines into words or characters. A problem occurs if the characters are connected or fragmented and consists of several parts. It is hard to distinguish noise from text [5, 7].

C. Pre-processing

Image resulting from scanning may contain noise. Pre-processing smoothens the image of characters. It uses Filling and Thinning and also uses Normalization [5, 8].

D. Feature Extraction

It captures essential characteristics of characters. Extraction of features divided into three groups [5, 6]:

i. Distribution of Points
ii. Based on statistical distribution of points.
iii. Includes Zoning, Moments, Crossing and Distance

E. Recognition Post Processing

This phase reduces the dimensionality of feature vector. It extracts feature by deformation like rotation and translation. Transformation can be Fourier, Walsh, Haar, Hough or Karhunen-Loeve and based on the curve describing the contour of the characters. It also deals with Error detection and correction [6].

The identity of each symbol is found by comparing the extracted features with descriptions of the symbol classes obtained through a previous learning phase. Finally contextual information is used to reconstruct the words and numbers of

the original text. In the next sections these steps and some of the methods involved are described in more detail.

V. IMPLEMENTATION

A. Implementation Steps

The template-matching algorithm implements the following steps [10]:

a. Firstly, the character image from the detected string is selected.
b. After that, the image to the size of the first template is rescaled.
c. After rescale the image to the size of the first template (original) image, the matching metric is computed.
d. Then the highest match found is stored. If the image is not match repeat again the third step.
e. The index of the best match is stored as the recognized character.

B. Flow control of Template Matching

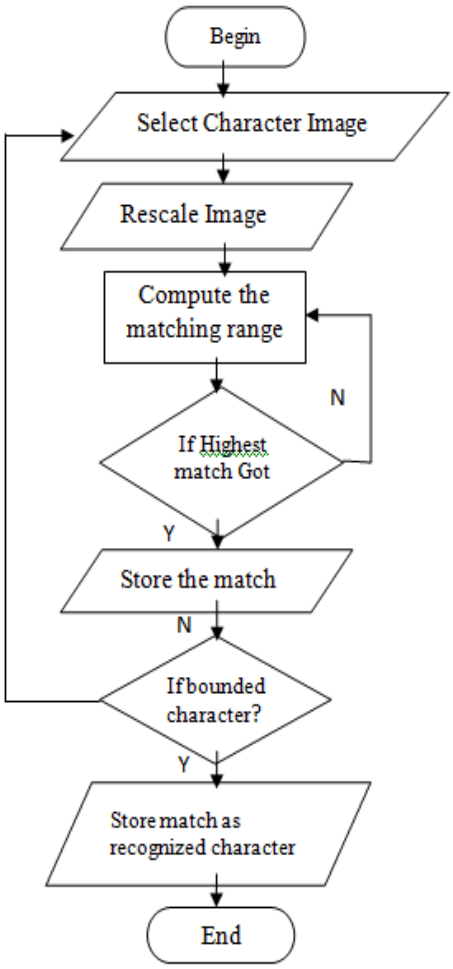


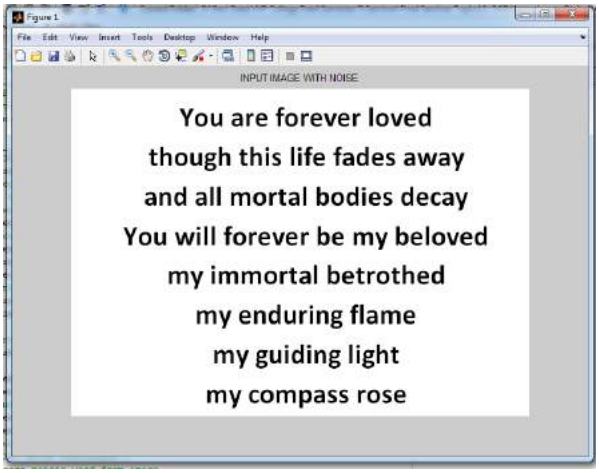
Figure 4.2.1. Shows the Flowchart.

VI. RESULTS

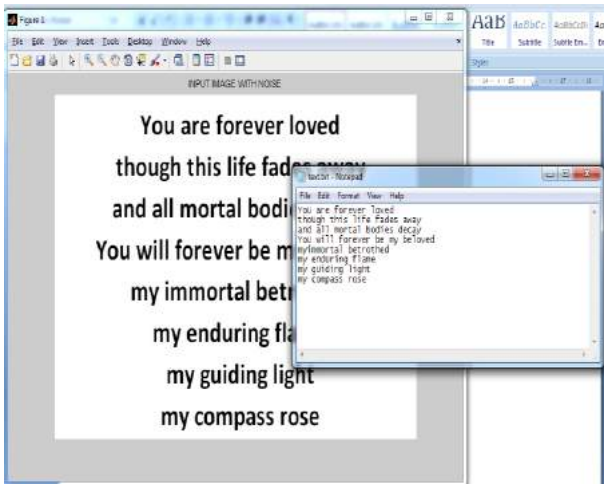
The system was developed using Template Matching method of OCR. 2 Gray Scale and 1 Color Images were taken as an input [9] and the following results were found:

A. Input 1:

Gray Scale Image with medium sized characters were given input to the system.



Output 1:

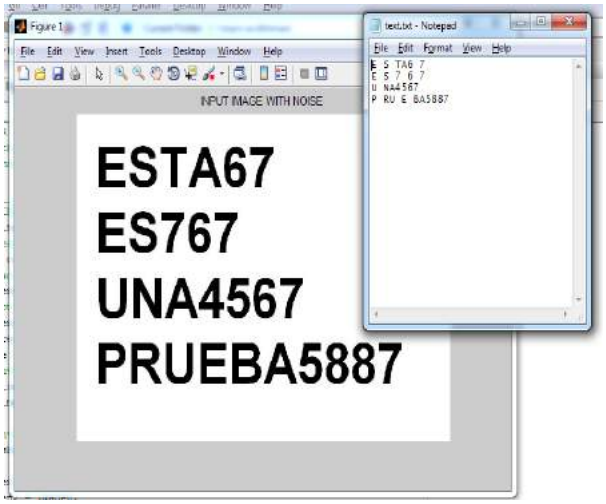


All the texts are clearly visible and are as per the size of library characters, so each character of whole string is recognized properly and generates output without any error.

B. Input 2:

Gray Scale Image with large sized characters was given as an input:

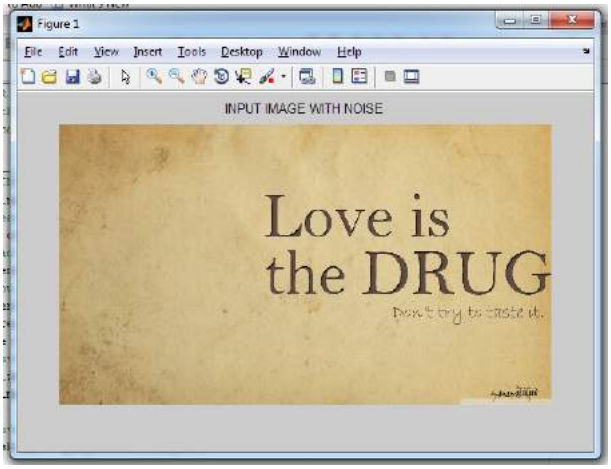
Output 2:



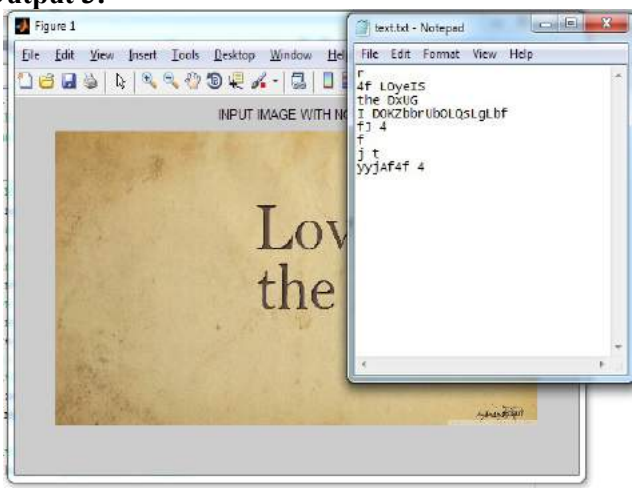
OCR recognizes all the character and digits but the additional spaces are padded in between in resultant text because the library characters are smaller in size as compared to the given input characters.

C. Input 3:

A colored image with some text was given as an input



Output 3:



It fails to identify the characters from the image, and generates in-correct output as the system is developed for Gray Scale Image only.

VII. CONCLUSIONS & FUTURE WORK

As an overall view of the system prototype, it could be conclude that this system prototype has been developed by using the technique that has mentioned and elaborated which is the Template Matching approach to recognize the character image. Besides, the interface of the system prototype looks user-friendly and makes the user of this system prototype easier to use it. As a result, the recognition process of this system become smoothly because of the steps that used in this system while recognizing the character. Even though this system prototype could gives several advantages to the users, but this system prototype are still facing a number of limitations like:

1. The system prototype has some limitation related to performance.
2. It works only with stored templates of alphabets and numbers with fixed sized templates.
3. It works only for gray scale image only.

In future, an effective system can be developed which works on Feature Extraction method and also colored image. This new system may be a motivation for enormous and fruitful researches in future.

VIII. REFERENCES

[1]. Maninder Kaur, "A Brief Review on Optical Character Recognition Techniques", International Journal of Computer Science and Mobile Computing, Vol.6 Issue.2, February-2017, pg. 95-100.

[2]. A. Chaudhuri, "Optical Character Recognition Systems for Different

[3]. Languages with Soft Computing", Studies in Fuzziness and Soft Computing 352, Springer International Publishing AG 2017.
[4]. Chowdhury Md Mizan, Tridib Chakraborty and Suparna Karmakar, "Text Recognition using Image Processing", International Journal of Advanced Research in Computer Science, ISSN No. 0976-5697, Volume 8, No. 5, May-June 2017.
[5]. Ashima Singh and Swapnil Desai, "Optical character recognition using template matching and back propagation algorithm", 2016 International Conference on Inventive Computation Technologies (ICICT) Vol-3.
[6]. M. Shen, "Improving OCR Performance with Background Image Elimination", 2015 12th Int. Conf. Fuzzy Syst. Knowl. Discov., pp. 1566-1570, 2015
[7]. E. N. Bhatia, "Optical Character Recognition Techniques : A Review", IJARCSSE, vol. 4, no. 5, pp. 1219-1223, 2014
[8]. Kartar Singh Siddharth, Mahesh Jangid, Renu Dhir, Rajneesh Rani (2011), "Handwritten Gurmukhi Character Recognition Using Statistical and Background Directional Distribution Features", International Journal on Computer Sceince and Engineering, Vol. 3(6), Pages 2332-2345.
[9]. Kirill Safronov, Dr.-Ing. Igor Tchouchenkov, Dr.-Ing Heniz Worn (2007), "Optical Character Recognition Using Optimisation Algorithm", Workshop on Computer Science and Information Technology, Pages 1-5.
[10]. Aparna Vara Lakshmi Vemuri, T.V.Sai Krishna, Atul Negi, "Dataset Generation for OCR" [OCR_Datasheet_Generation.pdf]
[11]. Jesse Hansen, "A Matlab Project in Optical Character Recognition (OCR)".

© 2018 IJSRSET | Volume 4 | Issue 2 | Print ISSN: 2395-1990 | Online ISSN : 2394-4099
National Conference on Advanced Research Trends in Information and Technologies (NCARTICT-2018), Department of IT, L. D. College of Engineering, Ahmedabad, Gujarat, India In association with International Journal of Scientific Research in Science, Engineering and Technology

Research Trends in Web and Web Usage Mining

Rachana Parikh*, Rachit Adhvaryu, Komil Vora
Information Technology, V.V.P. Engineering College, Rajkot, Gujarat, India

ABSTRACT

Web Technology is evolving very fast and Internet Users are growing much faster than estimated. The website users are using a wide range of websites leaving back a variety of information. This information must be used by the websites administrator to manipulate their websites according to the users of the websites. This actually is Data Mining. Web Mining is one of the applications of Data Mining. Web mining plays an important role in the decision making in the corporate, education and research environment. Modern developments in digital media technologies have evolved a huge amount of data transmitting over the web and with this huge data storage is required for easy and feasible access. Web Usage Mining is a mining of usage of websites and the information used and delivered on the websites. It is a technique to extract information from the web which includes web documents, hyperlinks between the documents and web usage logs. In this paper we describe the detailed survey of web mining, different techniques of web usage mining and its importance.

Keywords: Data Mining, Internet, Web, Web Mining, Web Usage Mining

I. INTRODUCTION

The World Wide Web (WWW) has lots of information and this information is increasing in a large amount daily. It is a very complex task to filter such information. A web or a website is a collection of web pages generally made using HTML and some programming languages which contains images, text, hyperlinks and similar digital data. The need to understand and use large, complex, information-rich data sets is common to virtually all fields of business, science, and engineering [1]. The technique to extract useful knowledge residing in these data and to act on that knowledge is becoming increasingly important in today's competitive world. Web Mining is the application of Data Mining techniques to retrieve information from the web which includes web documents, hyperlinks between the documents and web usage logs. Also the entire process of applying a computer-based methodology for discovering and retrieving knowledge from web documents is a web mining [1]. As the web data is updated every second, it is not compulsory that every user will get the same data whenever it is retrieved.

II. WEB MINING

Web Mining is classified into 3 main mining techniques [2]. The taxonomy of web mining is as follows:

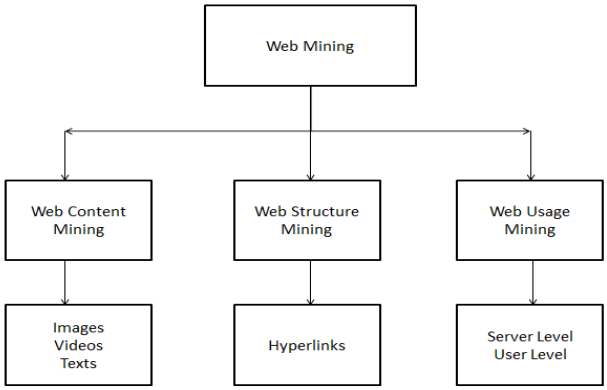


Figure 1. Web Mining Taxonomy

A. WEB CONTENT MINING

Web Content Mining is the way of retrieving important information from the web documents. The information can be in the form of images, videos, audios, texts and hyperlinks. A text mining and developing application for the same is one of the favourite topics for the researchers. Research in web content mining consists resource discovery from the web, categorizing and clustering of documents and extraction of informations from the webpages [2]. Here is the brief detail of web content mining techniques.

a) Image Mining

Image Mining is the technique which is used to detect unusual patterns and extract important and useful information from the images stored on the web or the large database. Thus image mining mainly deals with defining relationships between different images from the web and large databases [4]. Image mining is used in various fields like medical diagnosis, space research, remote sensing, agriculture and industries. The images include maps, geological structures and even it is used in the field of education.

b) Video Mining

Mining video is more complex than mining image data. Video is the collection of moving images like animation. There are 3 types of videos: 1) the produced (movies, news videos etc), 2) the raw (traffic, surveillance etc) and 3) the medical video (x-ray, cardiogram etc). The information from the video can be a) detecting trigger events (movement of vehicle and people), b) determining typical patterns of activity, generating person or object centric views of activity, and c) classifying activities into named categories (walking, sleeping etc), clustering and determination of interactions between two entities[5]. Video mining can also be classified into pixel based, statistics based, feature based and histogram based.

c) Audio Mining

As audio is the continuous media like video, the techniques and tools for audio processing and mining information is similar to video mining. Audio data can be in the form of radio, speech or spoken languages. Also television news has audio which are integrated with the videos [6]. Mining audio data requires conversion of audio to speech for better processing. Very few works has been carried out in this field.

d) Text Mining

The most trending research in the field of web content mining is text mining. The text mining refers to the text representation, classification, clustering, information extraction and search for hidden patterns. Text mining is the process of extracting useful information from the text and converting to automated discovery of knowledge [2]. It is natural extension of data mining or applying data mining techniques on a specific domain.

B. WEB STRUCURE MINING

This type of mining focuses on the data which describes the structure of the content of the web page. It is classified into two types: 1) intra-page structure: links within the page, 2) inter-page structure: links between 2 web pages [2]. This can be classified into two types based on structure of information:

a) Hyperlinks

A hyperlink is a structural unit that connects one web page with other web page either within same location or different location. A hyperlink connecting web pages in the same location is called intra-document hyperlink and a hyperlink connecting web pages at different locations is called inter-document hyperlink [2].

b) Document Structure

The content within a web page can also be organized in a tree structure based on HTML and XML Tags used to create a web page. Mining can be done to identify document object model (DOM) structures automatically from the documents [2].

C. WEB USAGE MINING

Web usage mining is one techniques of data mining to retrieve interesting and useful patterns from the web logs [3]. Web logs stores the identity or the origin of web users along with the browsing behaviour on the web site. Web Usage Mining can be grouped based on the type of usage logs:

a) Web Server Data

User logs are collected by the web server which includes IP address, page references and the time accessed by the user [3].

b) Application Server Data

Application servers are used to track various types of business events which can be used to improve the performance for any business firms [3]. For e.g. E-commerce websites uses such servers to know the events, business policies developed by their competitors.

c) User Level Data

User level data is the software developed using the information available from the web server and application server data [3]. It is an end user application which is used for various purposes.

D. Tools for Web Mining

Web mining tools helps the users to download essential information from the web. It collects the accurate and necessary information for the user which can be helpful in mining. The different tools are:

a) Automation Anywhere

It is a tool which is used to find web data very easily;. It is unique Intelligent and Smart Automation Application used for quick automation of any complex tasks [3].

b) Web Info Extractor

These tools is used to collect web content, constantly updating data and analysing data like images, videos, texts etc. [3].

c) Screen-Scraper

It is a tool used in searching databases and document structure. It provides a graphical interface allowing the user to navigate through URLs, data elements and hyperlinks and extract useful information from it [3].

d) Mozenda

This tool is used to extract and manage web data. User is allowed to setup tools at different places which can store and publish data at a regular interval of time [3].

e) Web Content Extractor

This tool is used to retrieve information from various websites like online auctions, online shopping, business directories, financial sites etc. The data can be represented in the form of excel, HTML, XML or any other script [5].

III. WEB USAGE MINING

Number of internet users in India is growing by 150K every month or 1.8 Million new users every year. India is the fastest growing online market in the world with 75% of the users being below the age of 35. The recent survey of Internet Users all over the world is described in image below: [15]

TOP 20 COUNTRIES WITH HIGHEST NUMBER OF INTERNET USERS - JUNE 30, 2017

#	Country or Region	Population, 2017 Est.	Internet Users 30 June 2017	Internet Penetration	Growth (*) 2000 - 2017	Facebook 30 June 2017
1	China	1,388,232,693	738,539,792	53.2 %	3,182.4 %	1,800,000
2	India	1,342,512,706	462,124,989	34.4 %	9,142.5 %	241,000,000
3	United States	326,474,013	286,942,362	87.9 %	200.9 %	240,000,000
4	Brazil	211,243,220	139,111,185	65.9 %	2,682.2 %	139,000,000
5	Indonesia	263,510,146	132,700,000	50.4 %	6,535.0 %	126,000,000
6	Japan	126,045,211	118,453,595	94.0 %	151.6 %	26,000,000
7	Russia	143,375,006	109,552,842	76.4 %	3,434.0 %	12,000,000
8	Nigeria	191,835,936	91,598,757	47.7 %	45,699.4 %	16,000,000
9	Mexico	130,222,815	85,000,000	65.3 %	3,033.8 %	85,000,000
10	Bangladesh	164,827,718	73,347,000	44.5 %	73,247.0 %	21,000,000
11	Germany	80,636,124	72,290,285	89.6 %	201.2 %	31,000,000
12	Vietnam	95,414,640	64,000,000	67.1 %	31,900.0 %	64,000,000
13	United Kingdom	65,511,098	62,091,419	94.8 %	303.2 %	44,000,000
14	Philippines	103,796,832	57,607,242	55.5 %	2,780.4 %	69,000,000
15	Thailand	68,297,547	57,000,000	83.5 %	2,378.3 %	57,000,000
16	Iran	80,945,718	56,700,000	70.0 %	22,580.0 %	17,200,000
17	France	64,938,716	56,367,330	86.8 %	563.1 %	33,000,000
18	Turkey	80,417,526	56,000,000	69.6 %	2,700.0 %	56,000,000
19	Italy	59,797,978	51,836,798	86.7 %	292.7 %	30,000,000
20	Korea, South	50,704,971	47,013,649	92.7 %	146.9 %	17,000,000

Figure 2. Country wise Internet Users

The above image clearly shows that India ranks 2nd and the internet users in India are increasing gradually. With this, the data and usage logs are getting larger and complex to manage and analyze. Web usage mining is the application of data mining techniques to discover usage pattern from Web data, in order to understand and better serve the needs of Web-based applications. Web usage mining consists of three phases: 1. preprocessing, 2. pattern discovery, 3. pattern analysis. Figure 3. represents these 3 phases of web usage mining.

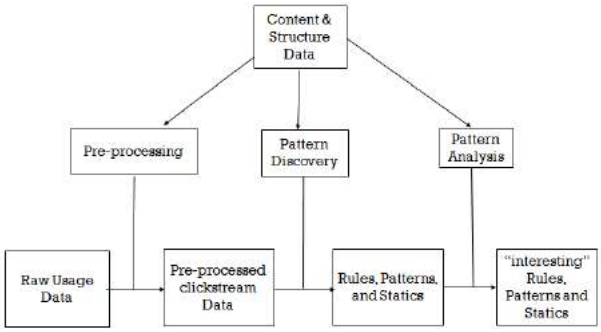


Figure 3. Web Usage Mining

The first is preprocessing state in which user sessions are inferred from log data. The second searches for patterns in the data by making use of standard data mining techniques, such as association rules or mining for sequential patterns. In the third stage an information filter bases on domain knowledge and the web site structures is applied to the mining patterns in search for the interesting patterns [7]. The web usage mining process can be classified into two parts. The first part includes transforming the Web data into suitable transaction form. This includes Data Preprocessing and Data Integration components. The second part includes some Data Mining and Pattern Discovery [7].

A. Data Processing and Data Integration

Data preprocessing consists of data filtering, user identification, session/transaction identification, and topology extraction. Data filtering filters out some noise, i.e., unsuccessful requests, automatically downloaded graphics, or requests from robots, to get more compact training data [10]. People use some heuristic rules to identify user, such as IP address, cookies, etc. Data Preprocessing converts the available data sources into the data abstractions.

a) Usage preprocessing:

Usage pre-processing consists of Web pages, such as IP addresses, page references, and the date and time of accesses. Typically, the usage data comes from an Extended Common Log Format (ECLF) Server log [11].

b) Content Preprocessing:

Content pre-processing consists of converting the text, images, scripts and multimedia data into forms that are useful for the web usage mining process [11].

c) Structure Preprocessing:

Web structure mining analyses the link structure of the web in order to identify relevant documents. The structure of a site is created by the hypertext links between page views. The Google Search engine makes use of the web link structure in the process of determining the relevance of a page. The Google search engine achieves good results because while the keyword similarity analysis ensures high precision the use of a probability measure ensures high quality of the pages returned [11].

The usage data collected at the different sources such as Server level, Client Level and Proxy Level represent the navigation patterns of different segments of the overall Web traffic.

i) Server-level Collection:

A Web server log records the browsing behaviour of site visitors. The data recorded in server logs reflect the concurrent and interleaved access of a Web site by multiple users. These log files can be stored in various formats such as Common Log Format (CLF) or Extended Common Log Format (ECLF). ECLF contains client IP address, User ID, time/date, request, status, bytes, referrer, and agent.[10] Tracking of individual users is not an easy task due to the stateless connection model of the HTTP protocol. In order to handle this problem, Web servers can also store other kind of usage information such as cookies in separate logs, or appended to the CLF or ECLF logs. Packet sniffing technology (also referred to as "network monitors") is an alternative method for collecting usage data through server logs. [10]

ii) Client level collection:

Client-side collection can be implemented by using a remote agent (such as Java scripts or Java applets) or by modifying the source code of an existing browser (such as Mosaic or Mozilla) to enhance its data collection capabilities.[9]

iii) Proxy Level Collection:

The Internet Service Provider (ISP) machine that users connect to through a model is a common form of proxy server. A web proxy acts as an intermediary between client browsers and Web servers. Proxy-level caching can be used to reduce the loading of time of a Web page experienced by users as well as the network traffic load at the server and client sides. [10]

B. Data Mining and Pattern Discovery

Data Mining comprises of the following data techniques which are as follows:

a) Association Rules:

Association rule generation can be used to relate pages that are most often referenced together in a single server sessions. In the context of web usage mining, association rules refer to sets of pages that are accessed together with a support value exceeding some specified threshold.[10] Association rule mining has been well studied in Data Mining, especially for basket transaction data analysis. Aside from being applicable for e-Commerce, business intelligence and marketing applications, it can help web designers to restructure their web site. The association rules may also serve as heuristic for pre fetching documents in order to reduce user-perceived latency when loading a page from a remote site [7].

b) Clustering:

Clustering is a technique to group together a set of items having similar characteristics. Clustering can be performed on either the users or the page views. Clustering analysis in web usage mining intends to find the cluster of user, page, or sessions from web log file, where each cluster represents a group of objects with

common interesting or characteristic. User clustering is designed to find user groups that have common interests based on their behaviours, and it is critical for user community construction. This information is useful for the Internet search engines and Web assistance providers. [7]

c) Deviation/Outlier Detection:

It contains techniques aimed at detecting unusual changes in the data relatively to the expected values. Such techniques are useful, for example, in fraud detection, where the inconsistent use of credit cards can identify situations where a card is stolen. The inconsistent use of credit card could be noted if there were transactions performed in different geographic locations within a given time window. [12]

d) Statistical Analysis:

Statistical techniques are the most common method to extract knowledge about visitors to a web site. By analyzing the session file, one can perform different kinds of descriptive statistical analyses (frequency, mean, median, etc) on variables such as page views, viewing time and length of a navigational path. Many web traffic analysis tools produce a periodic report containing statistical information such as the most frequently accessed pages, average view time of a page or average length of a path through a site. This information can be useful for improving the system performance. [12]

Pattern discovery uses methods and algorithms developed from several fields such as statistics, data mining, machine learning and pattern recognition. The knowledge that can be discovered is represented in the form of rules, tables, charts, graphs, and other visual presentation forms for characterizing, comparing, predicting, or classifying data from the web access log [13]. Several pattern discovery techniques are as follows:

a) Sequential Patterns:

The technique of sequential pattern discovery attempts to find inter-session patterns such that the presence of a set of items is followed by another item in a time-ordered set of sessions or episodes. A new algorithm MiDAS (Mining Internet data for Associative Sequences) for discovering sequential patterns from web log files has been proposed that provides behavioural marketing intelligence for e-commerce scenarios.[13] MiDAS contains three phases: 1. A priori phase is the input data preparation, which consists of data reduction and data type substitution. 2. Discovery Phase discovers the sequences of hits and generates the pattern tree. 3. A posteriori Phase filters out all sequences that do not fulfil the criteria laid in the specified navigation templates and topology network and also pruning is done in this phase. By using this approach, Web marketers can predict future visit patterns, which will be helpful in placing advertisements aimed at certain user groups [13].

b) Dependency modeling:

Dependency modelling is another useful pattern discovery task in web mining. The goal here is to develop a model capable of representing significant dependencies among the various variables in the web domain. As an example, one may be interested to build a model representing the different stages a visitor undergoes while shopping in an online store based on the actions chosen (i.e., from a casual visitor to a serious potential buyer. [14]

c) Pattern analysis:

Pattern analysis is the last step in the overall Web Usage mining process. The motivation behind pattern analysis is to filter out uninteresting rules or patterns from the set found in the pattern discovery phase. The exact analysis methodology is usually governed by the application for which Web mining is done [13]. The most common form of pattern analysis consists of a knowledge query mechanism such as SQL. Another method is to load usage data into a data cube in order to perform OLAP operations. Content and structure information can be used to filter out patterns containing pages of a certain usage type, content type, or pages that match a certain hyperlink structure [14].

IV. APPLICATIONS OF WEB MINING

There are many applications of web mining. Most dominating applications of web mining are as follows:

a) In the world of online shopping, it is very important to know the customer behavior and experience with the website. The feedback of the experience given by the user helps the website owner to improve their content in an efficient way. The main target of the website owner is that once a customer is visiting the website for a purchase, the user should not move to the other websites [9].

b) Google is the best and widely used search engines. It provides the users to access information from over billions of web pages indexed on its server. The quickness and quality of information provided by the search engines make them the most successful search engines. Web mining helps the search engines to know the behavior of the user, the keywords they search and based on this they give a Page Rank [9].

c) Web usage mining offers users the ability to analyze massive volumes of click stream or click flow data, integrate the data seamlessly with transaction and demographic data from offline sources and apply sophisticated analytics for web personalization, e-CRM and other interactive marketing programs [8].

d) Personalization for a user can be achieved by keeping track of previously accessed pages. These pages can be used to identify the typical browsing behavior of a user and subsequently to predict desired pages [8].

e) Web usage patterns can be used to gather business intelligence to improve Customer attraction, Customer retention, sales, marketing and advertisement, cross sales [8].

f) Mining of web usage patterns can help in the study of how browsers are used and the user's interaction with a browser interface [8].

V. CONCLUSION

As the web data and its usage increases day by day, it is very important to analyse the web data and retrieve the information. Thus web mining, web usage mining and its techniques play an important role in information extraction from the web. In this paper, we have briefly discussed various web and its usage mining techniques, tools and applications. By the techniques and the tools described in the paper, one can use in his/her research and new techniques can be developed for more effective, efficient and faster results. We hope that this primary discussion is the beginning for a fruitful researches and results in future.

VI. REFERENCES

[1]. Arvind Kumar Sharma, P.C. Gupta, -"Exploration of efficient methodologies for the improvement in web mining techniques-A survey", International Journal of Research in IT & Management (ISSN 2231-4334) Vol.1, Issue 3, July 2011.

[2]. G. Srivastava, K. Sharma, V. Kumar,"Web Mining: Today and Tomorrow", in the Proceedings of 2011 3rd International Conference on Electronics Computer Technology (ICECT), pp.399-403, April 2011.

[3]. S. K. Madria, S. S. Bhowmick, W. K. Ng,and E. P. Lim, "Research Issues in Web Mining", in proceding of data mining and knowledge discovery, 1st International conference, DaWk 99, pp 303-312, 1999.

[4]. Ji Zhang, Wynne Hsu and Mong Li Lee "An Information-Driven Framework for Image Mining" Database and Expert Systems Applications in Computer Science, 2001, Volume 2113/2001, 232-242, DOI: 10.1007/3-540-44759-8_24

[5]. Boresczky J. S. and L. A. Rowe, "A Comparison of Video Shot Boundary Detection Techniques", Storage & Retrieval for Image and Video Databases IV, Proc. SPIE 2670, 1996, pp.170-179.

[6]. A. Czyzewski, "Mining Knowledge in Noisy Audio Data", in Proc. 2nd Int. Conf. on Knowledge Discovery and Data Mining, pages 220-225, 1996.

[7]. J.W. Han, M .Kamber, Data Mining-Concepts and Techniques, Elsevier Science & Technology Books, 2006.

[8]. J. Srivastava, R. Cooley, M. Deshpande, P. Tan, "Web Usage Mining: Discovery and Applications of Usage Patterns from Web Data", SIGKDD Explorations, 1(2), 20001, pp.2-23.

[9]. Jaideep Srivastava and Robert Cooley "Usage Mining: Discovery and Applications of Usage Patterns from Web Data", ACM SIGKDD Explorations Newsletter .2000,1(2):12-23.

[10]. Bamshad Mobasher, "Web Usage Mining", Springer Berlin Heidelberg, 2007, 449-483.

[11]. Navin Kumar Tyagi, A.K. Solanki and Sanjay Tyagi: "An Algorithmic Approach to Data Preprocessing in Web Usage Mining". International Journal of Information Technology and Knowledge Management, Volume 2, No. 2, July-December 2010, pp. 279-283.

[12]. Jose Roberto de Freitas Boullosa. "An Architecture for Web Usage Mining".

[13]. Yan Wang." Web Mining and Knowledge Discovery of Usage Patterns". CS 748T Project. February, 2000.

[14]. Cyrus Shahabi, Amir M. Zarkesh, Jafar Adibi, and Vishal Shah, "Knowledge Discovery from Users Web-page Navigation", IEEE RIDE 1997.

[15]. http://www.internetworldstats.com/top20.htm

© 2018 IJSRSET | Volume 4 | Issue 2 | Print ISSN: 2395-1990 | Online ISSN : 2394-4099
National Conference on Advanced Research Trends in Information and Computing Technologies (NCARTICT-2018), Department of IT, L. D. College of Engineering, Ahmedabad, Gujarat, India. In association with International Journal of Scientific Research in Science, Engineering and Technology

Survey on Different Method to Improve Performance of The Round Robin Scheduling Algorithm

Bhavin Fataniya[1], Manoj Patel[2]
[1]M.E(I.T) Student, I.T Department, L.D College Of Engineering, Ahmedabad, Gujarat, India
[2]Assistant Prof., I.T Department, L.D College Of Engineering, Ahmedabad, Gujarat, India

ABSTRACT

Round Robin (RR) scheduling algorithm is the widely used scheduling algorithm in multitasking and real time environment. Its performance highly depends on a Time Quantum, which is a predefined amount of time assigned by CPU to every task to be executed. If we chose the less time quantum then context switch is high and if we chose the high time quantum then its leads to First-Come-First-Serve (FCFS). So, the performance of the system totally depends upon the choice of optimal time quantum. In this paper I will survey a different method to improve performance of the round robin scheduling algorithm using dynamic time quantum and also compare different method to each other to see the performance of the algorithm with respect to turn around time, waiting time, number of context switch.

Keywords: Round Robin,Dynamic Time Quantum.

I. INTRODUCTION

CPU scheduling is one of the fundamental features that an Operating System (OS) needs to perform multitasking. CPU scheduling means allocation of sufficient resources to a process being executed. The policy of CPU scheduling is to maximize the efficiency of the CPU. This efficiency includes maximizing the throughput, minimizing context switches, Average Waiting Time (AWT) and Average Turnaround Time (ATT) of the processes in ready queue. Throughput can be defined as the number of processes executed per unit cycle. The process waiting time of a process is the time it spends in ready queue before its execution.Turn around time is the sum of waiting and the execution time of a process.CPU scheduling algorithms are subdivided in to two main categories. A non-preemptive algorithm is the one in which CPU is allocated to a process till completion of its execution while in a preemptive algorithm, a higher priority process can block currently running process. Or we may simply conclude that if there are n processes in a ready queue, a non-preemptive process can have only n − 1 context switches. While a preemptive process can have more context switches. Some of the algorithms that have been developed and implemented are as follows;

First Come First Serve (FCFS) [9]: As obvious from the name,FCFS processes the jobs in the order in which they arrive in the ready queue.

Shortest Job First (SJF) [9]: In this algorithm, the ready queue is first sorted on ascending order of burst times of the processes arrived in it. Then the processes are assigned CPU in sequence.This algorithm is optimal than others in most cases. But in SJF,we must be able to have a pre-knowledge of burst times of all the processes which is difficult practically. Moreover, it can cause starvation for longer processes.

Priority Scheduling (PS) [9]: In priority scheduling, the processes come in ready queue along with assigned priorities to them. OS assigns CPU to the process with highest priority. Then the CPU is assigned to a process with lower priority and so on. Priority scheduling algorithm respects the priority of a process rather than

focusing on efficiency constraints of CPU. So, it may cause best and worst cases depending on burst times of corresponding processes. FCFS and SJF, both are non-preemptive algorithms. But PS can be of both typei.e. preemptive or non-preemptive. If the currently executing process can continue its execution on arrival of a higher priority process, then it is a non-preemptive priority scheduling .On the other hand, if the process is blocked on arrival of a higher priority process, it is a preemptive priority scheduling.

Round Robin (RR) [9]: This algorithm allocates CPU to all processes for an equal time interval. A process is blocked and put at the end of ready queue after a constant time slice, known as Time Quantum. This process is assigned CPU time again once the execution of all other processes in their respective time quanta. The efficiency of Round Robin depends entirely on the time quantum selected. If the time quantum is selected too large, the algorithm will become FCFS. While on the other hand, if the quantum is much small, it will yield much overhead and larger average waiting and turnaround times .With advancement in technology, many Round Robin algorithms based on dynamic time quantum have been developed. In this, a dynamic time quantum is chosen instead of a constant time quantum. It may be changed after a cycle or just after an arrival of next process in ready queue .In next section we will see the original round robin algorithms, some of the problem of round robin algorithms and survey of different method to improve performance of the round robin scheduling algorithm using dynamic time quantum and also compare different method to each other to see the performance of the algorithm with respect to turn around time, waiting time ,number of context switch.

II. ROUND ROBIN ALGORITHM

It is one of the oldest, simplest, fairest and most widely used scheduling algorithms, designed especially for time-sharing systems. A small unit of time, called time slice or quantum is defined. All run able processes are kept in a circular queue. The CPU scheduler goes around this queue, allocating the CPU to each process for a time interval of quantum. New processes are added to the tail of the queue. The CPU scheduler picks the first process from the queue, sets a timer to interrupt after time quantum, and dispatches the process. If the process is still running at the end of the quantum, the CPU is pre-empted and the process is added to the tail of the queue. If the process finishes before the end of the quantum, the process itself releases the CPU voluntarily. In either case, the CPU scheduler assigns the CPU to the next process in the ready queue. Every time a process is granted the CPU, a context switch occurs, which adds overhead to the process execution time.
An example of round robin algorithm is given below:

Process	AT	BT
P1	0	40
P2	0	55
P3	0	60
P4	0	90
P5	0	102

Time quantum=25.

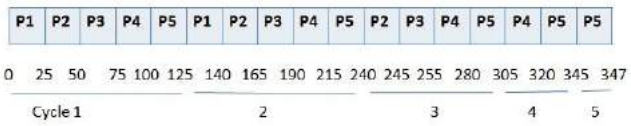


Figure 1. Gantt chart for round robin

Turn-around time and waiting time for above example is calculated in below table.

PROCESS	T.A.T	W.T
P1	140	100
P2	245	190
P3	255	195
P4	320	230
P5	347	245

III. ROUND ROBIN PROBLEM

The performance of RR depends on the size of the time quantum.In some case, additional context-switch is required even if only small amount of BT is remaining.For example in above Gantt chart at last only p5 process is remaining and remaining B.T is 2 m.s so for execution of only 2 m.s ,additional context switch is required. Turn-around time also depends on time

quantum.Increase in quantum time - cause less context-switch,lesser turn-around time but can cause high response time, and waiting time. Decrease in quantum time-cause less waiting time and response time but can cause high turn-around time and high no of context-switch Hence a medium sized time quantum can reduce all the factors to some extent. medium sized time quantum can be achieved by using dynamic time quantum method. now we will see different method based on dynamic time quantum.

IV. LITERATURE REVIEW OF DIFFERENT ROUND ROBIN ALGORITHM BASED ON DYNAMIC TIME QUANTUM.

Assumptions

AT=arrival time, BT=burst time, T.A.T=turn-around time, W.T=waiting time, CS= context switch, C.T=completion time.

During analysis we have considered CPU bound processes only. In each test case 5 independent processes are analyzed in uni-processor environment. Corresponding burst time and arrival time of processes are known before execution. all the arrival time and burst time is in M.S(milli second).Arrival time and burst time of different case is below.

Case 1. arrival time is same [8]

Process	AT	BT
P1	0	105
P2	0	60
P3	0	120
P4	0	48
P5	0	75

Case 2. arrival time is different [8]

Process	AT	BT
P1	0	45
P2	5	90
P3	8	70
P4	15	38
P5	20	55

In our analysis for various algorithm we have used above two table(case 1,case 2) to see the performance of particular algorithms.Turn-around time and waiting time can be calculated using below formula.

T.A.T=C.T-A.T.

W.T=TA.T-B.T.

A. Modified Round Robin Algorithm for Resource Allocation based on average.[3]

a) Methodology

Pandaba Pradhan,Prafulla Ku. Behera, B N B Ray [3] proposes an algorithm MRRA to improve the performance of Round Robin. In MRRA for each cycle the average of burst time of the processes is calculated and used as time quantum.

Step-1:process arrive in ready queue based on A.T.

Step-2:sort the process based on B.T of process found on ready queue.

Step-3:find time quantum using average of B.T of process which is placed on ready queue.

Step-4:now assign this time quantum to all process which is loaded in ready queue.

Step-5:if process B.T-time quantum=0 then terminate that process.

Step-6:IF process B.T-TIME QUANTUM!=0 then put that process at end of ready queue. Continue with step:2 to step:6 until all process are completed.

Step-6:calculate T.A.T,W.T,NO OF C.S.

Case 1. arrival time is same

The Gantt chart shown as in Figure 2

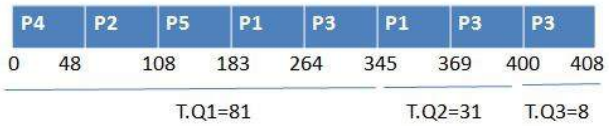


Figure 2. Gantt chart for MRRA [3] for same arrival time of process.

Turn-around time and waiting time for above example is calculated in below table.

PROCESS	T.A.T	W.T
P1	369	264
P2	108	48
P3	408	288
P4	48	0
P5	183	108

Case 2. arrival time is different

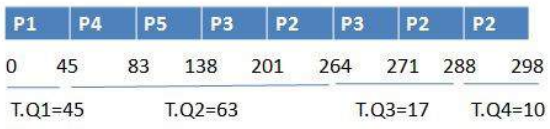


Figure 3. Gantt chart for MRRA [3] for different arrival time of process.

Turn-around time and waiting time for above example is calculated in below table

PROCESS	T.A.T	W.T
P1	**45**	**0**
P2	**293**	**203**
P3	**263**	**193**
P4	**68**	**30**
P5	**118**	**63**

B. Self-Adjustment Time Quantum in Round Robin Algorithm Depending on Burst Time of the Now Running Processes (SRBRR ALGO).[2]

b) Methodology

Rami J. Matarneh [2] proposes an algorithm SRBRR to improve the performance of Round Robin. In SRBRR for each cycle the median of burst time of the processes is calculated and used as time quantum.In other word time quantum is calculated using TQ=median(BT of process in ready queue).Median is calculated using below formula.

$$Q = \tilde{x} \equiv \begin{cases} Y_{(N+1)/2} & \text{if N is odd} \\ \frac{1}{2}(Y_{N/2}) + (Y_{1+N/2}) & \text{if N is even} \end{cases}$$

where, Y is the number located in the middle of a group of numbers arranged in ascending order.

Consider below example

Step-1:process arrive in ready queue based on A.T.

Step-2:find time quantum using median of B.T of process which is placed on ready queue.

Step-3:now assign this time quantum to all process which is loaded in ready queue.

Step-4:if process B.T-time quantum=0 then terminate that process.

Step-5:IF process B.T-TIME QUANTUM!=0 then put that process at end of ready queue. Continue with step:2 to step:6 until all process completed.

Step-6:calculate T.A.T,W.T,NO OF C.S.

Case 1. arrival time is same

The Gantt chart shown as in Figure 4.

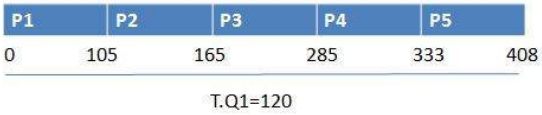


Figure 4. Gantt chart for SRBRR [2] for same arrival time of process.

Turn-around time and waiting time for above example is calculated in below table.

PROCESS	T.A.T	W.T
P1	**105**	**0**
P2	**165**	**105**
P3	**285**	**165**
P4	**333**	**285**
P5	**408**	**333**

Case 2. arrival time is different.

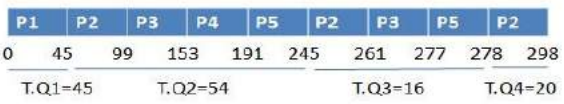


Figure 5. Gantt chart for SRBRR [2] for different arrival time of process

Turn-around time and waiting time for above example is calculated in below table

PROCESS	T.A.T	W.T
P1	**45**	**0**

P2	293	203
P3	269	199
P4	176	138
P5	258	203

C. Implementation of Alternating Median Based Round Robin Scheduling Algorithm[5]

c) Methodology

The proposed algorithm uses a set of two time quanta for scheduling the ready queue. These time quanta are used alternately in cycles of scheduling. Before the first cycle of scheduling starts, one of the time quanta is set equal to the median of burst time of processes present in the ready queue while the other one is calculated as the difference between the highest value of burst times and the median of burst times of processes present in the ready queue.The two-time quanta remain fixed throughout scheduling and are not recalculated until the architecture of the ready queue gets changed by the arrival of a new process.The larger time quantum is used if even number of cycles of scheduling are completed while the smaller time quantum is used if the odd number of cycles of scheduling are completed. Therefore, the two-time quanta are used alternately in cycles of scheduling**.**

Step-1:process arrive in ready queue based on A.T.

Step-2:sort the process based on B.T of process found on ready queue.

Step-3:now calculate time quantum 1=median of B.T of process which is present in ready queue and time quantum 2=high B.T- time quantum 1.

Step-4:now assign this time quantum 1 if even number of cycle is completed and assign time quantum 2 if odd number of cycle completed to all process which is loaded in ready queue.

Step-5:if process B.T-time quantum=0 then terminate that process.

Step-6:IF process B.T-TIME QUANTUM!=0 then put that process at end of ready queue.Continue with step:2 to step:7 until all process completed.

Step-7:calculate T.A.T,W.T,NO OF C.S.

Case 1. arrival time is same

The Gantt chart shown as in Figure 6.

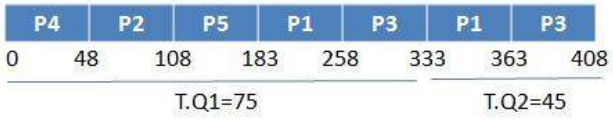


Figure 6. Gantt chart for AMBRR [5] for same arrival time of process

Turn-around time and waiting time for above example is calculated in below table.

PROCESS	T.A.T	W.T
P1	366	261
P2	108	48
P3	408	288
P4	48	0
P5	183	108

D. Design of A Modulus Based Round Robin Scheduling Algorithm(MB ALGO)[4]

d) Methodology

The Modulus Based (MB) algorithm has been devised on the basis of two scheduling algorithms namely MRR(AVG) algorithm and SRBRR(median) algorithm. This algorithm gives the results intermediate between both of its parent algorithms. if we encounter scenarios where the MRR(Avg) algorithm behaves more efficiently than the SRBRR algorithm or vice versa, then we can select the given MB algorithm in order to get more stable results. The formula of calculation of time quantum Q is given in equation.

$$Q = \sqrt{\frac{Median^2 + Average^2}{2}}$$

Step-1:process arrive in ready queue based on A.T.

Step-2:sort the process based on B.T of process found on ready queue.

Step-3: using above formula find time quantum of B.T of process which is placed on ready queue.

Step-4:now assign this time quantum to all process which is loaded in ready queue.

Step-5:if process B.T-time quantum=0 then terminate that process.

Step-6:IF process B.T-TIME QUANTUM!=0 then put that process at end of ready queue.Continue with step:2 to step:5 until all process completed.
Step-6:calculate T.A.T,W.T,NO OF C.S.

Case 1. arrival time is same
The Gantt chart shown as in Figure 7.

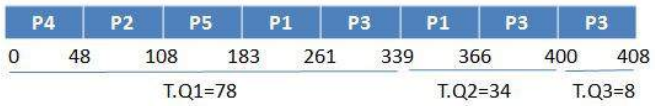


Figure 7:Gantt chart for MB algorithm[4] for same arrival time of process

Turn-around time and waiting time for above example is calculated in below table

PROCESS	T.A.T	W.T
P1	366	261
P2	108	48
P3	408	288
P4	48	0
P5	183	108

Case 2. arrival time is different.

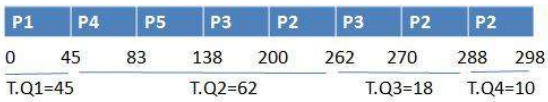


Figure 8. Gantt chart for MB [4] for different arrival time of process

Turn-around time and waiting time for above example is calculated in below table

PROCESS	T.A.T	W.T
P1	45	0
P2	293	203
P3	262	192
P4	68	30
P5	118	63

E. An Improved Dynamic Round Robin Scheduling Algorithm Based on a Variant Quantum Time(DRR ALGORITHM)[1]

e) Methodology

Step-1:process arrive in ready queue based on A.T.
Step-2:sort the process based on B.T of process found on ready queue.
Step-3: if the arrival time of all process is same then first time quantum is calculated using below formula
T.Q1=Σ two high B.T/2.
Then second time quantum is calculated using below formula.
T.Q2=Σ of high remaining B.T and T.Q1 /2.
Step-4:now assign this time quantum 1(T.Q1) to first cycle and for remaining cycle used T.Q2.
Step-5:if process B.T-time quantum=0 then terminate that process.
Step-6:IF process B.T-TIME QUANTUM!=0 then put that process at end of ready queue.Continue with step:2 to step:7 until all process completed.
Step-7:calculate T.A.T,W.T,NO OF C.S.
Step-8 : if the arrival time of all process is different than time quantum is calculated using below formula.
T.Q1=B.T OF FIRST ARRIVED PROCESS
T.Q2= (Σ of high remaining B.T and T.Q1 /2) – (Σ OF THE TWO LOOWEST REMAINING AT/2) (used only once time)
T.Q3== (Σ of high remaining B.T and T.Q1 /2) – (LOOWEST REMAINING AT/2)
Step-9:now assign this time quantum 1(T.Q1) to first cycle and for second cycle used T.Q2 and for remaining cycle used T.Q3.
Step-10:if process B.T-time quantum=0 then terminate that process.
Step-11:IF process B.T-TIME QUANTUM!=0 then put that process at end of ready queue.Continue with step:8 to step:12 until all process completed.
Step-12:calculate T.A.T,W.T,NO OF C.S.

Case 1. arrival time is same
The Gantt chart shown as in Figure 9.

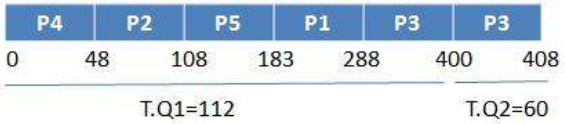


Figure 9. Gantt chart for DRR ALGORITHM [1] for Same arrival time of process

Turn-around time and waiting time for above example is calculated in below table.

PROCESS	T.A.T	W.T
P1	288	183
P2	108	48
P3	408	288
P4	48	0
P5	183	108

Case 2. arrival time is different

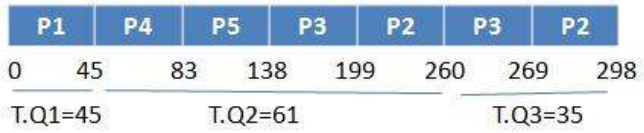


Figure 10. Gantt chart for DRR ALGORITHM [1] for different arrival time of process

Turn-around time and waiting time for above example is calculated in below table.

PROCESS	T.A.T	W.T
P1	45	0
P2	293	203
P3	261	191
P4	68	30
P5	118	63

F. Improved Round Robin Scheduling using Dynamic Time Quantum.(IRR ALGORITHM)[6]

f) Methodology

Debashree Nayak [6] proposes an algorithm IRR to improve the performance of Round Robin. In IRR,for each cycle the median of burst time of the processes is calculated.Median is calculated using below formula.

$$Q = \tilde{x} \equiv \begin{cases} Y_{(N+1)/2} & \text{if N is odd} \\ \frac{1}{2}(Y_{N/2}) + (Y_{1+N/2}) & \text{if N is even} \end{cases}$$

where, Y is the number located in the middle of a group of numbers arranged in ascending order.
Then time quantum is calculated using below formula
T.Q= (MEDIAN + HIGH B.T) / 2

Step-1:process arrive in ready queue based on A.T.

Step-2:sort the process which found in ready queue.

Step-3:using above formula find time quantum of B.T of process which is placed on ready queue.

Step-4:now assign this time quantum to all process which is loaded in ready queue.

Step-5:if process B.T-time quantum=0 then terminate that process.

Step-6:IF process B.T-TIME QUANTUM!=0 then put that process at end of ready queue.Continue with step:2 to step:7 until all process completed.

Step-7:calculate T.A.T,W.T,NO OF C.S.

Case 1. arrival time is same
The Gantt chart shown as in Figure 11.

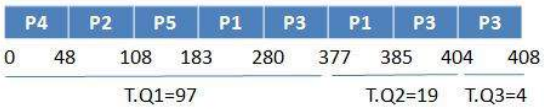


Figure 11. Gantt chart for IRR ALGORITHM [6] for same arrival time of process.

Turn-around time and waiting time for above example is calculated in below table.

PROCESS	T.A.T	W.T
P1	385	280
P2	108	48
P3	408	288
P4	48	0
P5	183	108

Case 2. arrival time is different

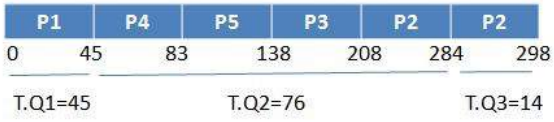


Figure 12. Gantt chart for IRR ALGORITHM [6] for different arrival time of process.

Turn-around time and waiting time for above example is calculated in below table.

PROCESS	T.A.T	W.T
P1	45	0
P2	293	203
P3	200	130
P4	68	30
P5	118	63

V. COMPARISON

Section IV describe the all dynamic quantum based round robin algorithm All the above described algorithms except SRBRR[2], first sort processes on ascending order of their burst times.This tends their algorithm towards SJF and most of algorithms take advantage of optimization of SJF rather than their own algorithms. Moreover,the sorting on ascending order is itself problam.It is important to note that the time quantum calculated by the MRR(Avg)[3] algorithm which varies greatly due to possibility of smaller or larger differences in execution times of tasks.Hence this average may be larger than median of burst times of all the tasks giving smaller T.A.T and less number of context switches but larger response time and W.T.Hence SRBRR algorithm is more stable than the MRR(AVG) algorithm if sorting of process is used in SRBRR[10].but in original SRBRR sorting of process is not used so its leads to F.C.F.S in case of process arrival time is same.SRBRR is perfectly work if arrival time of process is different. The calculation of the time quantum of MB[4] algorithm is a computationally intensive task that consumes a bulk of time for the calculation of Median and the Average of burst times of tasks in each cycle..In this section we will see the Comparison of each algorithm based on average turn-around time(A.T.A.T),average waiting time(A.W.T) and No of C.S.

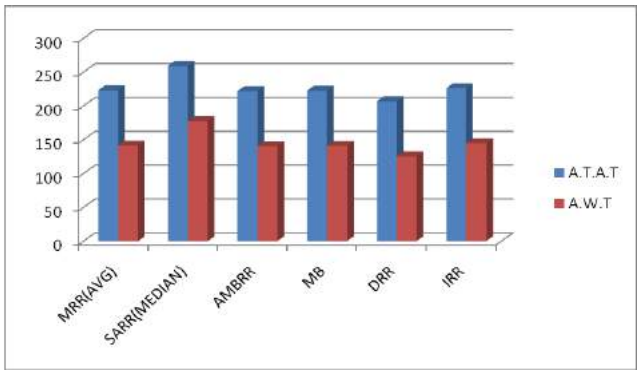


Figure 14. A.T.A.T and A.W.T of when all process arrived at same time(case:1)

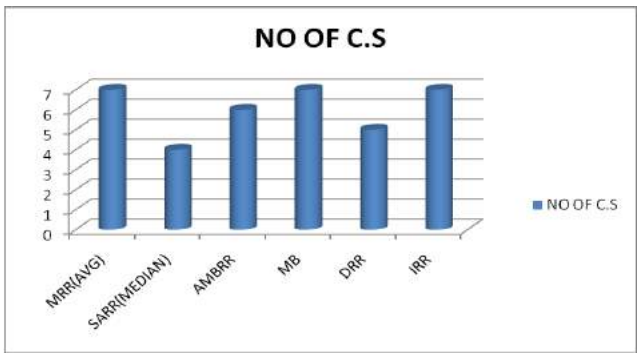


Figure 15. number of context switch when all process arrived at same time.(case:1).

Above Figure 14 and 15 show the comparison between A.T.A.T ,A.W.T and NO of C.S for all algorithms which was describe in section IV.From above graph we will conclude that D.R.R algorithm give the best result compare to all other algorithms. It will give best average turn-around time and best average waiting time.No of C.S is less in SARR algorithm but its leads to FCFS so this is not useful. D.R.R[1] algorithm give less number of context switch compare to remaining algorithm.

Figure 16 and 17 show the comparison between A.T.A.T ,A.W.T and NO of C.S for all algorithms which was describe in section IV.From above graph we will conclude that I.R.R algorithm give the best result compare to all other algorithms. It will give best average turn-around time and best average waiting time.I.R.R algorithm give less number of context switch compare to remaining algorithm.

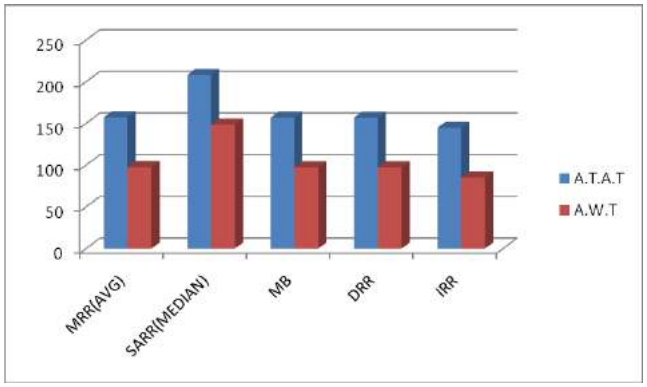


Figure 16. A.T.A.T and A.W.T when all process arrived at different time(case:2)

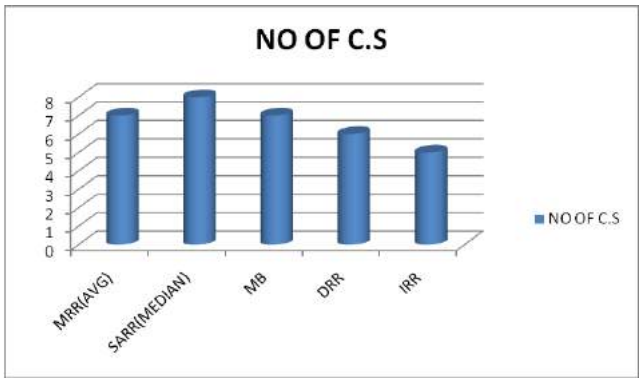


Figure 17. number of context switch when all process arrived at different time.(case:2)

VI. CONCLUSION

Different scheduling algorithms are reviewed in this paper.section IV describe the overall working of all dynamic quantum based round robin algorithm.section V describe the comparison of all algorithm. At last we conclude that D.R.R[1] algorithm is more stable than all other algorithm in case of arrival time of all process is same and I.R.R[6] algorithm is more accurate than all other algorithm in case of different arrival time of process.

VII. REFERENCES

[1]. A. Alsheikhy, R. Ammar, and R. Elfouly, "An improved dynamic Round Robin scheduling algorithm based on a variant quantum time",IEEE, 2015 11th International Computer Engineering Conference (ICENCO), 2015, pp. 98-104.

[2]. Rami J. Matarneh,"Self-Adjustment Time Quantum in Round Robin Algorithm Depending on Burst Time of Now Running Processes", American J. of Applied Sciences 6(10): 1831 - 1837,2009.

[3]. Pandaba Pradhan, Prafulla Ku. Behera and B N B Ray, "Modified Round Robin Algorithm for Resource Allocation in Cloud Computing ", ScienceDirect, Procedia Computer Science 85 (2016) 878 – 890,2016.

[4]. Salman Arif, Saad Rehman and Farhan Riaz "Design of A Modulus Based Round Robin Scheduling Algorithm",IEEE, 9th Malaysian Software Engineering Conference, Dec. 2015

[5]. Salman Arif,Naveed Ghaffar,Ali Javed "Implementation of Alternating Median Based Round Robin Scheduling Algorithm",IEEE, International Conference on Computer and Information Technology,2016.

[6]. Debashree Nayak, Sanjeev Kumar Malla and Debashree Debadarshini "Improved Round Robin Scheduling using Dynamic Time Quantum", International Journal of Computer Applications (0975 – 8887) Volume 38– No.5, January 2012

[7]. A. Noon,A. Kalakech and S. Kadry,"A New Round Robin Based Scheduling Algorithm for Operating Systems: Dynamic Quantum Using the Mean Average", International Journal of Computer Science Issues (T.TSI),Vol.3,Issue 3,No. 1,pp. 224-229,May 2011.

[8]. Amar Ranjan Dash, Sandipta kumar Sahu and Sanjay Kumar Samantra"AN OPTIMIZED ROUND ROBIN CPU SCHEDULING ALGORITHM WITH DYNAMIC TIME QUANTUM" International Journal of Computer Science, Engineering and Information Technology (IJCSEIT), Vol. 5,No.1, February 2015.

[9]. Sonia Zauaoui,lotfi Boussaid and Abdellatif Mtibaa, "CPU scheduling algorithms case and comparative study",17th international conference on sciences and techniques of automatic control & computer engineering=STA 2016,sousse,Tunisia,Decmber,2016.

[10]. H.S.Behera, R. Mohanty and Debashree Nayak "A New Proposed Dynamic Quantum with Re-Adjusted Round Robin Scheduling Algorithm and Its Performance.

[11]. Analysis", International Journal of Computer Applications (0975 – 8887) Volume 5– No.5, August 2010.

© 2018 IJSRSET | Volume 4 | Issue 2 | Print ISSN: 2395-1990 | Online ISSN : 2394-4099
National Conference on Advanced Research Trends in Information and Computing Technologies (NCARTICT-2018), Department of IT, L. D. College of Engineering, Ahmedabad, Gujarat, India In association with International Journal of Scientific Research in Science, Engineering and Technology

A Survey on Efficient Frequent Pattern Mining Techniques

Janak Thakkar[*1], Dr. Mehul Parikh[2]
[*1]IT Department, R C Technical Institute, Ahmedabad, Gujarat, India
[2]Associate Professor, IT Department, LDCE, Ahmedabad, Gujarat, India

ABSTRACT

Data Mining is the technique to abstract the useful data from the large dataset for different perspectives. Frequent pattern mining has become an important data mining technique to find the frequent patterns from the data set that appears frequently. Frequent Pattern Technique is widely used in financial, retail, telecommunication and many more. The major concern of these industries is faster processing of a very large amount of data. Various techniques and algorithms have been proposed for this purpose. Apriori, FP-tree are the pioneer techniques among them. In this paper, we have analysed algorithms for finding frequent patterns with the purpose of discovering how these algorithms can be used to obtain frequent patterns over large transactional databases with most efficient way in various aspects. This has been presented in the form of a comparative study of the following algorithms: Apriori, Frequent Pattern (FP) Growth, dNC-ECPM Algorithm, OCFP–growth, IA-TJ-FGTT(Important Attributes - Transaction Joining - Frequency Gathering Table Technique).

Keywords: Data Mining, Frequent Pattern Mining, Apriori, FP-Growth, Erasable Patterns, Close Patterns

I. INTRODUCTION

Frequent patterns are the itemsets that occur with frequency which is greater than a user-specified threshold. Frequent patterns plays an significant role in mining associations, correlations, and many other interesting relationships among data in datasets. It also helps in data indexing, classification, clustering, and other data mining tasks as well. Thus, FPM is an important data mining technique and topic for the research.

FPM was first introduced by Agrawal etal.[1,2] for market basket data analysis in the form of association rule mining, which is useful for discovering interesting relationships hidden in large data sets. For example, we may find a strong relationship, which can be represented in the form of association rules or sets of frequent items, exists between the sale of diapers and beer because many customers who buy Bread also buy Butter.

A. Apriori algorithm

Agrawal and Srikant [2] noticed an interesting downward closure property, called Apriori, among frequent k-itemsets: A k-itemset is frequent if all of its sub-itemsets are frequent. For this, first scanning the database to give the frequent 1-itemsets, then using the frequent 1-itemsets to generate candidate frequent 2-itemsets, by this process frequent k-itemsets can be generated for some k. This is the essence of the Apriori algorithm[7].Though in many cases, the Apriori algorithm reduces the size of candidate sets. However, it has two significant disadvantages : (1) generating a huge number of candidate sets, and (2) Scanning of dataset more number of times.

B. FP-growth algorithm

Han et al. [3] Proposed an FP-growth method that mines the complete set of frequent itemsets without candidate generation in less number of Scans. FP-growth follows divide-and-conquer method[8]. In first scan of the database, it derives a list of frequent items in which items are ordered by frequency-descending order. According to the frequency- ascending list, the database is compressed into a frequent-pattern tree (FP-tree),

which retains the itemset association information. The FP-tree is mined by starting from each frequent length-1 pattern (as an initial suffix pattern), which constructs the conditional pattern base, then constructing its conditional FP-tree, and performing mining recursively on such a tree. The pattern growth is achieved by the joining of the suffix pattern with the frequent patterns generated from a conditional FP-tree. The FP-growth algorithm transforms the problem of finding long frequent patterns to searching for shorter ones recursively and then joining the suffix. It uses the least frequent items as a suffix, offering good selectivity. Performance studies shows that FP-Growth method significantly reduces search time.

The remaining document is organised as follow: In Section 2 will represent three different algorithms for the mining of frequent patterns in more efficient way. In section 3, Comparative study of all this algorithms will be represented in tabular format which covers some basic characteristics of algorithms with significant advantages and disadvantages. Section 4 will cover the conclusion with some limitations and future work. In section 5, all references will be listed.

II. LITURATURE SURVEY

As Apriori and FP - Tree have some disadvantages, to overcome this paper proposed some novel algorithms and techniques for efficient pattern mining.

C. dNC-ECPM Algorithm

Tuong Le, Giang Nguyen, Tzung-Pei Hong [4] proposed an efficient algorithm for erasable closed pattern from dataset to reduce Memory Usage and Mining Time. Paper mainly focus on the Erasable Closed Patterns (ESP) from the dataset. In algorithm first step is to identify the ECPs from the dataset in such a way that which doesn't affect the information loss. In second step dNC-ECPM algorithm will be implemented for mining the ECPs generated in first step. The dNC-ECPM algorithm follow the systemic approach to discard the erasable close patterns. Initially dNC_set structure is defined which consists set of diff. Node Codes Based on WPPS–tree(Weight, Pre order, Post Order, Childnodes). After successful generation of dNC_set structure, algorithm is implemented. The NC_diff algorithm is to determine the dNC_set and the relationship of dNCs(X) and dNCs(Y). A node code Ci is an ancestor of another node code Cj if and only if Ci.pre-order ≤ Cj.pre-order and Ci.post-order ≥ Cj.post-order. The dNC_set algorithm to construct WPPC-tree with threshold value and determine E1 and generate dNCset of erasable patterns from E1. Significant advantage of the dNC_ECPM is less memory usage and mining time.

D. OCFP – Growth Algorithm

Another approach for efficient frequent pattern mining was proposed by team of Hsiao-Wei Hu, Hao Chen Chang, Wen-Shiu Lin [5]. To overcome the limitations of Apriori algorithm OCFP–growth (Optimized Close Frequent Pattern) algorithm was proposed. The algorithm is based on the OMIS – tree (Optimized Minimum Item Support Trees). In this algorithm for Minimum support on the individual items will be counted and after that minimum support on the itemset. Definition 2: A subset of frequent itemset may be not frequent. Definition 3: A subset containing the item with the lowest MIS value in a frequent itemset must be frequent. The process of to give input OMIS-tree, which is a set of frequent item F,MIS of each item in F, where the value of k = 2. Output will be the complete set of all f's conditional frequent k-itemset (patterns). After input is taken the step by step functions will be implemented. In first step OCFP-growth algorithm will be called with MIS-tree by declaring root as a null. Second step is to generating the patterns with MIS value. In third step, set of the conditional patterns will be constructed. Fourth and final step is to add frequent patterns will be added into the MIS Tree and run the programs till all infrequent patterns are discarded.

E. IA-TJ-FGTT Technique

Another technique is developed by the Saravanan. Suba, Dr.T. Christopher[6] to mine the frequent patterns from the large datasets. Algorithm is IA-TJ-FGTT which stands for Important Attributes -Transaction Joining-Frequency Gathering Table Technique. The technique is based on the association rule[7][8] mining method which perform pruning, joining operations on the attributes. In techniques first important attributes are selected from the all attributes. After selecting the important attributes duplicate transactions are discarded and dataset is reduced. Then transactions are joined. After joining the truncation Frequency Gathering Table(FGT) is generated which resulting in the Frequent Patterns.

III. COMPARITIVE STUDY

After evaluating individually now below listed table shows the comparative analysis of the all the approaches discussed in the paper. comparison is focused on the selected fields like, Technique used by the algorithm, Advantages, Disadvantages, Storage structure, Mining Time and Accuracy.

Table 1. Comparative Analysis of various Pattern Mining Algorithms

	Apriori	**FP -Growth**	**dNC-ECPM**	**OCFP**	**IA-TJ-FGTT**
Technique	Breadth First Search	Divide And Conquer	Divide And Conquer	Closed Frequent Pattern	Association Rule Mining
Advantages	- Easy to Implement	- Scanning of Dataset for Two times only	- Memory Usage is low	Execution time is less	- get perfect patterns
Disadvantages	- Too many scans of dataset - Require Large Memory Space	- expensive then the Apriori	- not efficient in case of Large Datasets	- generating the MIS tree is difficult	- mining time is more compare to others because of more operations
Storage Structure	Array	Tree	Tree	Tree	Array
Mining Time	More	Less Compare to Apriori	Less than Apriori and FP -Growth	Reduced Mining Time and Execution	More as compare to dNC-ECPM and OCFP
Accuracy	Less Compare to others	Less Compare to others	High in Small Dataset Low in Large Data Set	Better than other	High than all the Techniques

IV. CONCLUSION

In this paper, we gave the brief overview of some of the existing algorithms and new approaches for the frequent pattern mining. Than each individual algorithm is compared and analyzed with the remaining algorithms by various fields. The comparison shows that all the approaches are efficient in reducing mining time but up to some extent not effective in the accuracy. More focused research can be possible in the direction of getting accurate and concrete frequent patterns which can be utilised for the planning and forecasting.

V. REFRENCES

[1] First Agrawal R, Imielinski T, Swami A. Mining association rules between sets of items in large databases. In Proceedings of the 1993 ACMSIGMOD international conference on management of data(SIGMOD'93), pages207-216.1993

[2] Agrawal R, Srikant R. Fast algorithms for mining association rules. InProceedings of the 1994 international conference on very large databases(VLDB'94), pages487–499.1994

[3] Han J, Pei J, Yin Y. Mining frequent patterns without candidate generation. In Proceeding of the 2000 ACM-SIGMOD international conference on management of data (SIGMOD'00) pages1–12.2000

[4] Efficient algorithms for mining erasable closed patterns from product datasets" 10.1109/ACCESS.2017.2676803, IEEE Access

[5] "An optimized frequent pattern mining algorithm with multiple minimum supports" 2016 IEEE International Conference on Big Data (Big Data) Hsiao-Wei Hu, Hao-Chen Chang, Wen-Shiu Lin

[6] Saravanan. Suba, Dr.T. Christopher. "An Improved and Efficient Frequent Pattern Mining Approach to Discover Frequent Patterns among Important Attributes in Large Data set Using IA-TJFGTT" 2016 IEEE International Conference on Advances in Computer Applications (ICACA)

[7] Goswami.D, Chaturvedi.A, Raghuvanshi.C "An Algorithm for Frequent Pattern Mining Based On Apriori" IJCSE 2010.

[8] Mabroukeh.N and Ezeife.C "A Taxonomy of Sequential Pattern Mining Algorithms" ACM Computing Surveys, Vol. 43, No. 1, Article 3, Publication date, November 2010

© 2018 IJSRSET | Volume 4 | Issue 2 | Print ISSN: 2395-1990 | Online ISSN : 2394-4099
National Conference on Advanced Research Trends in Information and
ıg Technologies(NCARTICT-2018), Department of IT,L. D. College of Engineering,
Ahmedabad, Gujarat, India In association with
International Journal of Scientific Research in Science, Engineering and Technology

A Survey on Phishing Detection based on Visual Similarity of web pages

Ms Niyati Raj[1], Prof. Jahnavi Vithalpura[2]
[1]PG Student, IT Department, L.D. college of Engineering, Ahmedabad, Gujarat, India
[2]Assistant Professor, IT Department, L.D. College of Engineering, Ahmedabad, Gujarat, India

ABSTRACT

Phishing attack uses scam web pages which pretending to be an important website and takes user's personal information such as credit card number, passwords and other sensitive details. Anti Phishing is very important for online transactions and user privacy protection.
In this paper, I have done survey on different methods of phishing detection based on visual similarity and also compared them to see better accuracy and correctness with law performance head.
Keywords: Phishing detection, Visual similarity, Privacy protection

I. INTRODUCTION

Phishing is a form of social engineering attack in which attacker steals the sensitive information such as credit card numbers and passwords with spoofed web pages. Such communications are usually done through emails that trick users to visit fraudulent websites that collects users' private information. Users use social networks to communicate and share. User's privacy protection has become one of the most research issues.

Phishing pages need to lure users by their visual appearance. Page contents and page layouts are visually similar to the target pages. In a web based phishing attack, attacker sets up phishing web pages to lure users to input their sensitive information. The attacker sends emails or publishes web links on social networks that trick users to visit phishing pages. As social networks become a convenient platform to initial social engineering attacks.

Phishing can be detected by analysis of URLs of phishing pages and by analysis of page content similarity. Attackers have flexibility in changing URL features to evade detection. One key feature of phishing pages is that they usually maintain the similar visual appearance as their target pages. The software classification approaches can automatically detect the phishing messages by using white list/blacklist, URL based and Content based.

The black/white-list method is the most widely deployed anti phishing techniques used in browsers. The black/white list methods utilize a blacklist consisting of previously detected phishing URLs, IP addresses or keywords to classify the web page being visited as legitimate or phishing. White list can also be used to filter the famous legitimate web pages.

The most widespread blacklists are the Google safe browsing API[4] and the PhishTank blacklist[5]. Though the blacklist and whitelists are frequently updated, they can not deal with zero-hour phishing attacks[6] because the new zero-hour phishing site can not be added to the blacklist before it is submitted by a victim. The heuristics based methods explore some heuristics that exist in phishing attacks in reality.
In content based detection scheme based on the visual similarity between a page and other target pages. The features used include: text and styles,

images in the page, and the overall visual appearance of the page. Content based approaches generally extract content features of web pages to identify suspicious websites. To deal with such evasion attempts, some solutions compare images of rendered pages to evaluate their visual similarity.

II. LITERATURE REVIEW

A. Visual Similarity based Anti-phishing with the combination of Local and Global Features[1]

The algorithm is proposed by the Yu Zhou, Yongzheng Zhang , Jun Xiao, Yipeng Wang, Weiyao Lin, year 2014.

In this paper, they proposed a novel visual similarity based phishing detection method purely on image level by combining global and local features of the Web page image pair.

The global image feature is extracted only in the visible region of the whole Web page, not in the overall Web page.

The flowchart of their proposed approach is illustrated in Figure 1, which includes two steps.

The first step is logo detection. First, the snapshot of the suspected Web page and the logo image of the protected Web page are input.

In each image, the Speeded Up Robust Features (SURF) [7] detector is used to detect key points which represent the characteristics of the corresponding image. Then, the SURF descriptor is generated for each image. These two sets of key points are matched according to the Euclidean distance. The matched key point pairs are then filtered, and good matched points are reserved.

Based on the good matched key points, if the suspected Web page contains the target logo, a homography matrix can be found and the region that the logo locates can be extracted. The second step is the global similarity computation. The suspected Web page snapshot and the protected Web page snap shot are cut to the visible regions, and two images correspond to the visible regions are obtained. For each result image, they follow the work to extract signature, and the EMD distance between two signatures is taken as the global similarity score.

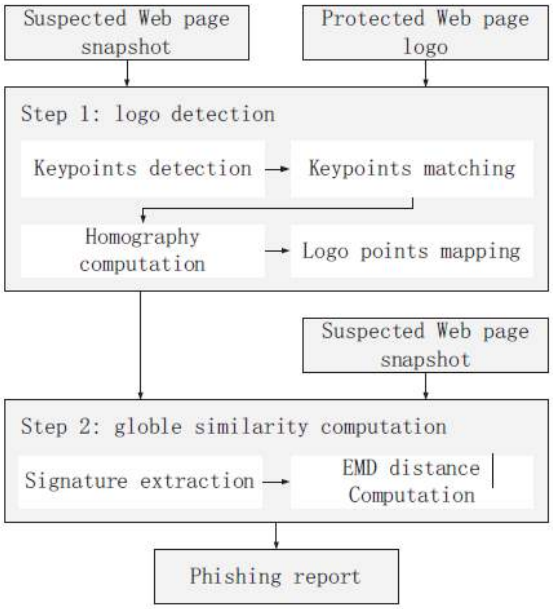


Figure 1. Flow chart of proposed approach [1]

If the suspected Web page snapshot contains the logo of the protected Web page and the global similarity score is beyond to the threshold, the suspected Web page is classified as the phishing Web page. In other words, the local and the global similarities are combined sequentially. In the next two Sections, the logo detection and the global similarity computation are respectively introduced in detail.

B. Bait Alarm: Detecting Phishing sites using similarity in fundamental visual features[2]

This algorithm is proposed by Jian Mao, Pei Li, Kun Li, Tao Wei, and Zhenkai Liang, year 2013.

In this paper, they proposed a solution, Bait Alarm, to efficiently detect phishing web pages. Page layouts and contents are fundamental feature of web pages' appearance. Since the standard way to specify page layouts is through the style sheet (CSS), they developed an algorithm to detect similarities in key elements related to CSS.

They implemented Bait Alarm in a Google Chrome extension.
The overall architecture of the Bait Alarm extension is shown in Figure 3.

Bait Alarm includes three main components: Pre-Processor, Layout Monitor and Network Library.

The Pre-Processor consists of Page Filter, DOM, and HTML Parser. After a web page is loaded, the Page Filter checks it over. If the web page has been loaded before, it does not need further analysis. If the loaded page is new and contains some specific UI (e.g., login form), the Page Filter triggers the detecting process. The HTML Parser and the DOM extract the layout information of the suspicious page. When the user inputs personal information, such as Login ID, the browser holds the page and the Pre-Processor sends the layout information to the Layout Monitor. The Layout Monitor consists of a Layout Model Builder and a Similarity Checker. When the Layout Monitor gets the layout information of the suspicious page from the Pre-Processor, the Layout Model Builder models them into "comparison-unit" and sent them to the Similarity Checker, together with additional page features (e.g., page domain, etc.). After the Similarity Checker gets the comparison unit of the suspicious page, it searches the Network Library for the victim pages feature model (comparison unit) indexed by the same personal information that has been inputted by the user before.

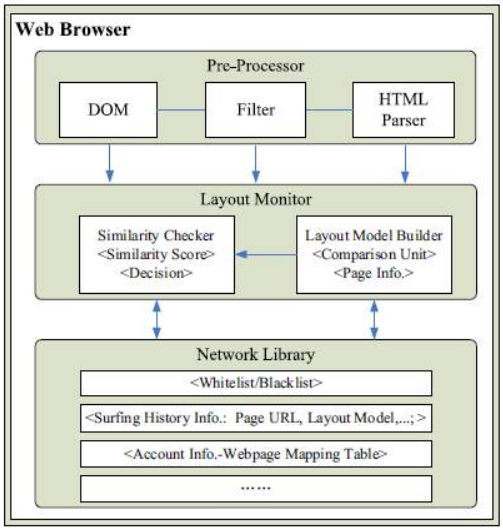


Figure 2. Architecture of Bait Alarm

If the Similarity Checker does not find the matched page, then it informs the browser to release the page and treat it as a new registering web site. The Similarity Checker reports the page information and its layout model to the Network Library.

If the Similarity Checker finds the matched page and gets layout model and additional page information. The checker calculates the similarity score of the pages and outputs the decision based on their similarity score and additional page information. In this scheme, if a page's similarity score is less than the preset threshold, the page is innocent. Then browser releases the page and the Similarity Checker reports the page information and its layout model to the Network Library. Otherwise, the Similarity Checker checks additional page information to make the decision.

The checker will submit the related information to the Network Library and inform the browser to pop up a warning page. The Network Library maintains the user's surfing history information (e.g., URL, layout model, etc.), Whitelist/Blacklist and a "Personal Info-Historical Page Mapping Table". The table is used to search for the victim pages based on users' information captured by the browser.

C. Utilisation of website logo for phishing detection[3]

This algorithm is proposed by Kang Leng Chiew, Ee Hung Chang, San Nah Sze,Wei King Tiong, year 2015.

Even though anti-phishing methods based on textual elements receive more attention, it has some limitations. Using a graphical element, especially the logo, is important. This will compensate for the limitations faced in textual-based methods, and will make the detection more robust. They use a logo image to determine the identity consistency between the real and the portrayed identity of a website. Consistent identity indicates a legitimate website and inconsistent identity indicates a phishing website.

The proposed method consists of two processes, namely logo extraction and identity verification. The first process will detect and extract the logo image from all the downloaded image resources of a webpage. In order to detect the right logo image, they utilise a machine learning technique. Based on the extracted logo image, the second process will employ the Google image search to retrieve the portrayed identity.

Since the relationship between the logo and domain name is exclusive, it is reasonable to treat the domain name as the identity. Hence, a comparison between the domain names returned by Google with the one from the query website will enable them to differentiate a phishing from a legitimate website.

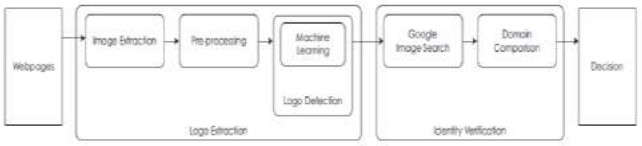


Figure 3. Framework of System [2]

The proposed method involves two main processes: logo extraction and identity verification. A logo extraction process will extract the logo from the query website. Based on the extracted logo, the identity verification process will evaluate the consistency between the real identity and portrayed identity of the query website. If the identity is consistent, the query website is legitimate, and vice versa.

Consistent identity means that the real identity and the portrayed identity are identical. The real identity can be obtained from the domain name of the query website. Whereas the portrayed identity can be retrieved from a database which entry has logo matches to the extracted logo.

Since the relationship between the logo and domain name of a website is exclusive, any mismatch is an indication to the phishing attack. Clearly, a complete and up-to-date database of different website logos with the corresponding domain names is needed. Maintaining this database effectively alone is impossible. Hence, they utilised Google Images as a source of the knowledge database. The first page of the search result. Next, a domain comparison sub process will parse from the first and third elements of the search result. After that, this sub process will extract only the domain name from each of the URLs and compare them to the domain name of the query website. They have taken the liberty to refer to the name which excludes the TLD (top level domain) and any sub domain as the domain name. For example, the domain name for http://www.mydomain.com is my domain. If the comparisons return at least one match, this method will classify the query website as legitimate. Otherwise, it is classified as a phishing website.

As for the limitation in the logo detection sub process mentioned above (i.e., when multiple images of logo and non-logo are returned), this method will repeat the identity verification process for each image, and aggregate the comparison results. Similarly, this method will classify the query website as legitimate if the aggregated comparisons return at least one match.

To fully utilise the Google Images database, authors employed the content-based image retrieval feature from the Google image search facility. It allowed them to retrieve the portrayed identity of a query website from the vast image database. This is depicted as a Google image search sub process. The output from the Google image search sub process is the search result which includes elements.

D. Use of HOG Descriptors in Phishing detection[4]

This method is proposed by ahmet Selmen Bozkir and Ebru akcapinar Sezer in year 2016.

This paper proposed to evaluate and solve this problem by leveraging a pure computer vision based method in the concept of web page layout similarity. Proposed approach employs histogram of oriented gradients descriptor in order to capture cuse of page layout without the need of time consuming intermediate stage of segmentation.

This system was designed to detect zero – day phishing attacks.

For the following reasons, HOG descriptors were preferred in this study:

(i) HOG descriptors are able to capture visual cues of overall page layout

(ii) They are able to provide a certain degree of rotation and translation invariance.

Extracting HOG descriptors require three main steps:

(i) Gradient computation:
Grid of equal sized cells is obtained by dividing the image.

(ii) Orientation Binning
For each pixel, gradient vector is converted to an angle and orientation bins are built according to angle ranges.

(iii) Block normalization
Normalized histograms are concatenated and final descriptor is formed.

This system consists of two modules. The first module so called “wrapper”, was designed and implemented in order to find out effective page boundaries and taking a screenshot of web page.

Second module called as "Hogger" was implemented in order to take JPEG file and output a concatenated HOG feature vector.

a. Identifying Region of Interest:

The "wrapper" window was precisely set for taking 1024 pixel wide screen shots. At next stage, crop the portion below 1024 pixels. For the cases where height of web page is lower than 1024 pixels, apply a dominant color detection method for filling the empty lowest part in order to have full square input image. In this way, input mages were generated concerning the existing dominant color in web page. Finally the output image will be converted to grayscale in order to increase the gradient computation accuracy.

b. Revealing the Cues of Page Layout via HOG Descriptors

In order to reveal the appropriate cell size this system applied two different grid configurations. In first configuration (HOG128), the input image consisting of 1024×1024 pixels was divided into 8×8 cells having side length of 128 pixels.

For the second configuration (HOG64), the side length of square sized cells is reduced to 64 pixels which totally results 16×16 grid.

By use of these two types of grid configuration, it aimed to understand and evaluate the levels of details.

c. Use Case Scenario

First collect URLs of legitimate pages LPi which have potential phishing risk and the layout signature of the LPi is stored in legitimate corpus database along with its root domain. Once all the pages which need phishing detection were loaded to the central corpus, a suspicious page SPj can be checked against the legitimate corpus in order to verify whether it has a high similar legitimate target. During the verification process, Histogram Intersection Kernel (HIK) is employed as a similarity metric.

E. A Computer vision technique to detect Phishing Attacks[5]

The algorithm is proposed by Routhu Srinivasa Rao, Syed Taqi Ali, year 2015, India.

In this paper, they proposed a novel solution to defend zero-day phishing attacks. Their proposed approach is a combination of white list and visual similarity based techniques. They used computer vision technique called SURF detector to extract discriminative key point features from both suspicious and targeted websites. Then they are used for computing similarity degree between the legitimate and suspicious pages.

The basic idea of their proposed solution is described below.

1) Maintain a legitimate image database consisting of all popular website screenshots along with their URLs (whitelist).
2) Obtain the accessed URL and do comparing with whitelist of URLs.
3) If comparison is successful URL will be considered as Innocent and no further checking will be required. This removes the extra overhead of comparison of legitimate website display.
4) If the given URI is not found in the white-list then SURF algorithm is applied on the suspicious website screenshot and legitimate image database.
5) Extract the SURF features from both suspicious website screenshot, image database and compare for similarity check.
   a) If similarity score is greater than the threshold, webpage is considered as suspected.
   b) If similarity score is less than threshold, URL is considered as innocent. The domain will be the part of white-list in next update.

III. COMPARISION

Research Paper No.	Merits	Demerits
1.	Purely work on image level Can achieve over 90% TP rate and 97% TN rate.	Some phishing web pages do not contain the official logo which results in the failure of logo detection and then these pages are classified as normal.
2.	Better than white list based techniques.	Needs to be enhanced.
3.	Use logo images to determine the identity consistent between the real identity and the portrayed identity of a website. The captured screenshot is the actual rendered web content, which means there is no other hidden image.	Need enhancement of the logo extraction process with a more effective logo detection algorithm.
4.	An efficient and fast phishing page detection scheme	Can be enhanced by providing image content invariance.
5.	Combination of white list and visual similarity based technique Used SURF detector	Needs to improve the computational cost and accuracy cost.

IV. CONCLUSION

Phishing is a most popular attack used by attackers to collect sensitive information from users. We surveyed paper on phishing attack based on visual similarities and from that we conclude that CSS based phishing detection are more efficient and faster in compare of other methods.

V. REFERENCES

[1] Yu Zhou, Yongzheng Zhang , Jun Xiao, Yipeng Wang, Weiyao -"Visual Similarity based Anti-Phishing with the Combination of Local and Global Features" in IEEE 13th Conference on Trust, Security and Privacy in Computing and Communications, 2014.

[2] Jian Mao, Pei Li , Kun Li, Tao Wei, and Zhenkai Liang ―"BaitAlarm: Detecting Phishing Sites Using Similarity in Fundamental Visual Features" in fifth conference on Intelligent Networking and Collaborative Sysytems, 2013.

[3] Kang Leng Chiew, Ee Hung Chang, San Nah Sze,Wei King Tiong "Utilisation of website logo for phishing detection" in Elsevier 2015.

[4] ahmet Selmen Bozkir and Ebru akcapinar Sezer - "Use of HOG Descriptors in Phishing detection" in 4th international symposium on digital forensics and security (ISDFS'16), 2016.

[5] Routhu Srinivasa Rao, Syed Taqi Ali "A computer vision technique to detect Phishing Attacks" 5th international conference on communication systems and network technologies, 2015

[6] Google, https://developers.google.com/safe-browsing.

[7] PhishTank, https://www.phishtank.com.

[8] S. Sheng, B. Wardman, G. Warner, L. Cranor, J. Hong, and C. Zhang, ―An empirical analysis of phishing blacklists,‖ in Sixth Conference on Email and Anti-Spam, 2009.

[9] H. Bay, A. Ess, T. Tuytelaars, and L. Van Gool, ―Speeded-up robust features (SURF),‖ CVIU, vol. 110, no. 3, pp. 346–359, 2008.

© 2018 IJSRSET | Volume 4 | Issue 2 | Print ISSN: 2395-1990 | Online ISSN : 2394-4099
National Conference on Advanced Research Trends in Information and Technologies (NCARTICT-2018), Department of IT, L. D. College of Engineering, Ahmedabad, Gujarat, India. In association with
International Journal of Scientific Research in Science, Engineering and Technology

RFID based Security in Internet of Things: A Study

Thomson Christain*[1], Lalit Sengunthar[2], Oneel Christian[3], Urja Mankad[4]
[*1]PG Scholar, L. J. Institute of Computer Applications, Ahmedabad, Gujarat, India
[2]PG Scholar, L. J. Institute of Computer Applications, Ahmedabad, Gujarat, India
[3]PG Scholar, L. J. Institute of Computer Applications, Ahmedabad, Gujarat, India
[4]Assistant Professor, L. J. Institute of Computer Applications, Ahmedabad, Gujarat, India

ABSTRACT

The Internet of Things (IoT) provides connectivity for everything and everyone. The IoT can be considered as the future evaluation of the Internet that realizes machine to machine (M2M) learning. Security and privacy are the main issues for IoT applications and facing some enormous challenges. This paper analyses the security architecture and security requirements of IoT as a whole and privacy protection, trust management is being discussed with the help of RFID which is explained in detail with all its consequences and also shown its application in different areas.
Keywords: Internet of Things (IoT), Security, Radio Frequency Identification (RFID), machine to machine (M2M).

I. INTRODUCTION

Internet of Things (IoT) is an ecosystem of connected physical device that are accessible through the internet, i.e. objects that have been assigned an IP address and have the ability to collect and transfer data over a network with automatic assistance or intervention. Nowadays, IoT is widely applied to social life applications such as smart grid, intelligent transportation, smart security, and smart home [1]. The IoT is considered as the future evaluation of the Internet that realizes machine-to-machine (M2M) learning [2]. The basic idea of IoT is to allow autonomous and secure connection and exchange of data between real world devices and applications [3]. The IoT links real life and physical activities with the virtual world [4]. The numbers of Internet connected devices are increasing at the rapid rate. These devices include personal computers, laptops, tablets, smart phones, PDAs and other hand-held embedded devices. Most of the mobile devices embed different sensors and actuators that can sense, perform computation, take intelligent decisions and transmit useful collected information over the Internet [5]. Using a network of such devices with different sensors can give birth to enormous amazing applications and services that can bring significant personal, professional and economic benefits [6]. The IoT consists of objects, sensor devices, communication infrastructure, computational and processing unit that may be placed on cloud, decision making and action invoking system [7]. The objects have certain unique features and are uniquely identifiable and accessible to the Internet. These physical objects are equipped with Radio-Frequency Identification (RFID) tags or other identification bar-codes that can be sensed by the smart sensor devices [6]. The IoT is a hot research topic that is getting increasing popularity for academia, industry as well as government. Many European and American organizations and multinational companies are involved in the design and development of IoT to achieve different type of useful and powerful automated services [1].

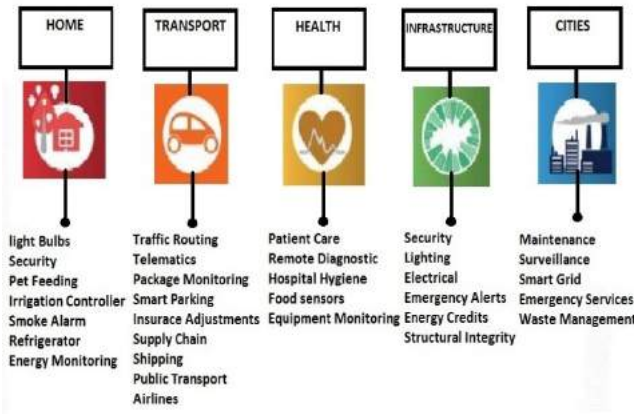


Figure 1. Applications of IoT [27]

Smart city, smart energy, smart transportation, smart health, smart agriculture, and many more areas of IoT are generating curiosity among world's population as shown in the figure 1. Smart surveillance, automated driving, smarter electricity management systems, urban security and environmental monitoring all are examples of internet of things applications for smart cities [28].

RFID

Radio frequency identification system (RFID) is an automatic technology and aids machines or computers to identity objects, record metadata or control individual target through radio waves. The RFID technology was first appeared in 1945, as an espionage tool for the Soviet Union. Main advantage of the RFID is the automated identification and data capture that promises wholesale changes across a broad spectrum of business activities and aims to reduce the cost of the already used systems such as bar codes [10].

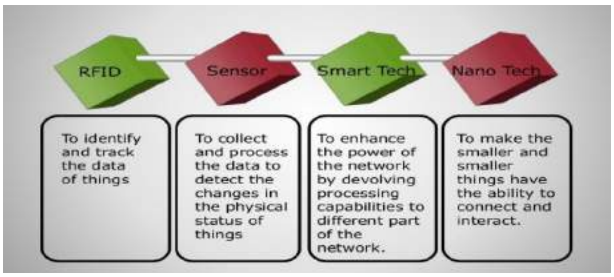


Figure 2. Perceptual Layer Devices [15]

The IoT can be viewed as a huge network consisting of networks of devices as shown in the figure 2 and computers connected through a series of intermediate technologies where numerous technology like RFID's, wireless connections may act as enablers of this connectivity [8].

1) Tagging Things: RFID trace and addresses the real-time item.
2) Feeling Things: Sensors collects data from the environment
3) Shrinking Things: Miniaturization and Nanotechnology has provoked the ability of smaller things to interact and connect within the smart devices or "things".
4) Thinking Things: Embedded intelligence in devices thorough sensors has formed the network connection to the Internet.

This paper will cover the security architecture of IoT (section II), security requirements in IoT in context of RFID (section III) and conclusion (section IV).

II. SECURITY ARCHITECTURE OF IoT

In general, the IoT can be divided into four key levels [9]. Figure 3 shows that the level architecture of the IoT [34].

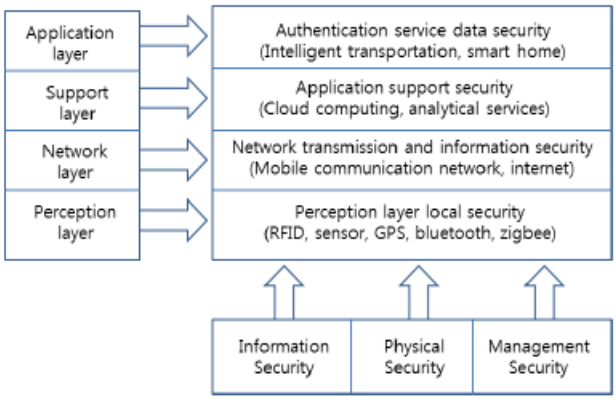


Figure 3. Layered Architecture of IoT

A. Perceptual Layer: The most fundamental level is the perceptual layer (otherwise called recognition layer), which gathers a wide range of data through physical gear and distinguishes the physical world, the data incorporates object properties, environmental condition and so on; and physical equipment incorporate RFID per user, a wide range of sensors, GPS and different types of equipment. The key segment in this layer is sensors for catching and representing the physical world in the digital world.

B. Network Layer: The second level is network layer. Network layer is responsible for the transmission of information from perceptual layer, classification, initial processing of information and polymerization. In this layer the information transmission is depend on several basic networks, which are the International Conference on Computer Science and Electronics Engineering. Internet, satellite nets, wireless network, mobile communication network, network infrastructure and communication protocols are also essential to the information exchange between devices.

C. Support Layer: The third level is support layer. Support layer will set up a reliable platform for the application layer, on this platform all kind of intelligent computing powers will be organized through cloud computing and network grid. It plays the role of combining application layer upward and network layer downward.

D. Application Layer: The application layer is the topmost and terminal level. Application layer provides the personalized services according to the user's

requirements. Users can access to the internet of thing through the application layer interface using of personal computer, television, or mobile equipment and so on. Management and network security play an important role in above each level.

III. SECURITY REQUIREMENTS

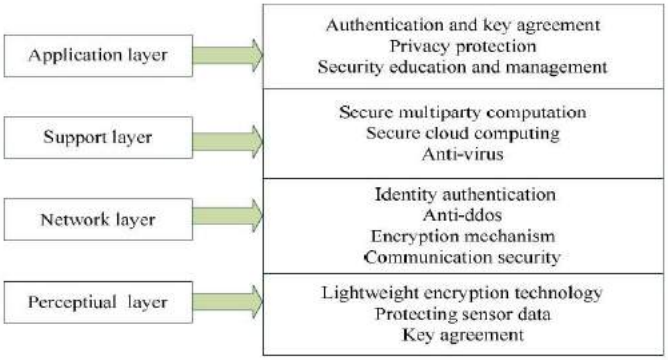


Figure 4. Security requirements in each level [34]

We can summarize the security requirements for each level in the following layers, as shown in Figure 4.

A. Perceptual Layer: At first node authentication is necessary to prevent illegal node access, secondly to protect the confidentiality of information transmission between the nodes, before the data encryption key agreement is an important process in advance and data encryption is absolute necessity. The stronger are the safety measures the more is consumption of resources to solve this problem lightweight encryption technology becomes important, which includes Lightweight cryptographic algorithm and lightweight cryptographic protocol. At the same time the authenticity and integrity of sensor data is becoming research focus.

B. Network Layer: In this layer existing communication security mechanisms are difficult to be applied. Identity authentication is a kind of mechanism to prevent the illegal nodes, confidentiality and internality are of equal importance, and thus we also need to establish internality mechanism and data confidentiality. Besides distributed denial of service (DDoS) in the network and is particularly severe in the internet of thing, it is necessary to prevent the DDoS attack for the vulnerable node is another problem to be solved in this layer [14].

C. Support Layer: Support layer needs a lot of the application security architecture such as secure multiparty computation and cloud computing, almost all of the strong encryption algorithm and stronger system security technology, anti-virus and encryption protocol.

D. Application Layer: To solve the security problem of application layer, we need two aspects. One is the authentication and key agreement across the heterogeneous network, the other is user's privacy protection. In addition, education and management are very important to information security, especially password management [25].

A. RFID Components For IoT

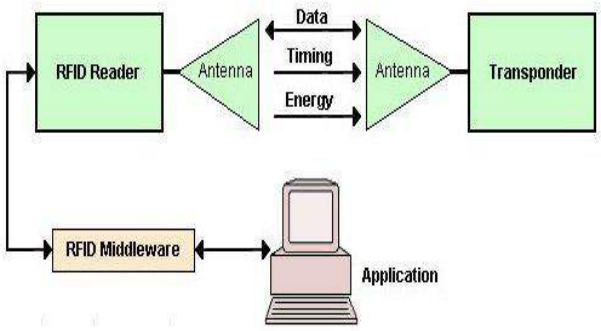


Figure 5. Components of IoT system [26]

RFID consist of different components like tags, readers and application, as shown in Fig. 5.

1. Transponder/Tags (transmitters/responders): The tag is a microchip connected with an antenna, which can be attached to an object as the identifier of the object.
2. Readers (transmitters/receivers): The RFID reader communicates with the RFID tag using radio waves.
3. Application system

B. SECURITY ISSUES OF RFID TECHNOLOGY AND SOLUTIONS

RFID (Radio Frequency Identification) is a non-contact automatic identification technology, which can automatically identify the target tag signal to obtain relevant data, identifying the process does not require manual intervention, and can work in harsh environments [11][38]. While RFID is widely used, it exposes a lot of problems as follows [36]

a. Uniform coding currently there is no uniform International encoding standard for RFID tag. The EPC (Electronic Product Code) standard supported by European and The most influential standards are the UID (Universal Identification)

standards supported by Japan. It may cause problems that errors may occur in the reading process or the reader cannot obtain access to the tag information.

b. Conflict collision as multiple RFID tags may also transmit data information to the reader at the same time, which may cause the reader not able to get data correctly [12]. Using the anti-collision technique can prevent multiple tags from transmitting information to the reader simultaneously. RFID conflict collision can be divided into two categories: tags' collision and readers' collision [15]. When a large number of labels are in the reader's working scope, and the reader cannot access to data correctly, this is called tags' collision. IoT requires wide range of RFID sensor coverage, and multiple readers' cooperative work is particularly important, but the working scope of reader overlaps. So information may become redundant which increases the burden on the transmission of network. This is called readers' collision. Different collisions have different solutions. Currently, tag anti-collision algorithm has been studied adequately, but research for reader anti-collision algorithm is not enough. Reader anti-collision algorithm is mainly divided into solutions based on the scope-based and time-base solutions [16, 17]. The main idea of the scope-based anti-collision algorithm is to try to avoid overlapping of reader work scope to achieve the purpose of reducing the conflict zone, but this solution requires an additional central control area to calculate the working scope between the readers, which increases the complexity and cost [18].

C. RFID PRIVACY PROTECTION

Low cost tags led to RFID's limited resources, such as weak computational capabilities and low storage capacity, thus it requires lightweight solutions for privacy protection, which includes location privacy and data privacy.

a. Data Privacy: RFID security and privacy technologies can be divided into two categories: physical-based schemes and password-based schemes, the former sends deactivation kill command, block tags, signal interference etc. The later includes schemes such as hash locks [19], random hash lock [20], anonymous ID [21], re-encryption [22]. Different organization styles for IoT require different ways of privacy protection agreement. A compromise solution for data privacy issues is to store less important information in RFID tag, and store important information in the up level service.

b. Location privacy: Although RFID tags do not store important information, but hackers can still get the tag ID information for the purpose of tracking the position of the tag [23]. For example, when a reader equipped with vehicle GNSS information reads a tag's information, it can easily obtain the approximate location information of the tag according to its effective operational range.

D. TRUST MANAGEMENT

Trust management in IoT, we must take node privacy more seriously. So we need to introduce trust management into IoT RFID system. Trust management exists not only just between the RFID tags and readers, but also between the base stations and the readers. In trust management field digital signature technology is of great usage. It has been used for device authentication, data authentication and data exchange between different applications for a long time. Protocols and cryptographic algorithms play important roles for digital signature technology. While standard cryptographic algorithms and protocols require storage space and computing resources more than the available resources of RFID tags, so RFID authentication algorithm must not only take into account security and privacy issues, but also consider the tag storage and computing power. Limited resources of RFID tags would be the focus of ongoing research and complexity of security. Above all, uniform encoding, conflict collision, privacy protection and trust management are four typical technologies for the security issues of RFID. With uniform encoding standard, we encode tag information uniformly, which can maximize information exchange. With a good conflict collision resolution technology, minimize potential data interference and we can make RFID readers read information correctly. With a good lightweight data privacy protection, we have helped protect data privacy and location privacy. Finally, with appropriate trust management algorithms, we can enable trust management for reader's/ RFID tags, base stations and readers [14].

.

IV.CONCLUSION

We have focused on the security architecture and security requirements of IoT, We concisely reviewed security in the IoT and requirements from four layers' perceptual layer, network layer, support layer and application layer. For perceptual layer RFID technology is of great importance, we analyzed the security issues of RFID technologies and their corresponding solutions including: Uniform Coding, Conflict Collision, RFID Privacy Protection, and Trust Management. This paper will help to analyse the IoT major security issues for research and development.

V. REFERENCES

[1] Sundmaeker, H., Guillemin, P., Friess, P., &Woelffle´, S. (2010). Vision and challenges for realising the internet of things.Cluster of European Research Projects on the Internet of Things—CERP IoT.

[2] Y. Huang and G. Li, "Descriptive Models for Internet of Things," in IEEE International Conference on Intelligent Control and Information Processing (ICICIP), August 2010.
[3] T. Fan and Y. Chen, "A Scheme of Data Management in the Internet of Things," in 2nd IEEE International Conference on Network Infrastructure and Digital Content, Sept. 2010.
[4] Y. Huang and G. Li, "A Semantic Analysis for Internet of Things," in International Conference on Intelligent Computation Technology and Automation (ICICTA), May 2010.
[5] Q. Zhou and J. Zhang, "Research Prospect of Internet of Things Geography," in 19th International Conference on Geoinformatics, June 2011.
[6] J. Li, Z. Huang, and X. Wang, "Countermeasure Research about Developing Internet of Things Economy," in International Conference on E - Business and E -Government (ICEE), May 2011.
[7] Y. Yu, J. Wang, and G. Zhou, "The Exploration in the Education of Professionals in Applied Internet of Things Engineering," in 4th International Conference on Distance Learning and Education (ICDLE), October 2010.
[8] https://www.slideshare.net/MohanKumarG/internetofthings-iot-aseminar-ppt-by-mohankumarg
[9] G. Yang, J. Xu, W. Chen, Z. H. Qi, and H. Y. Wang, "Security characteristic and technology in the internet of things," Journal of Nanjing University of Posts and Telecommunications (Natural Science), vol. 30, no. 4, Aug 2010.
[10] https://www.slideshare.net/swethak36/iot-rfid-finalppt
[11] Hu, F., & Wang, F. (2010). Study of recent development about privacy and security of the internet of things. In Proceedings of the international conference on web information systems and mining (pp. 91–95).
[12] Lv, B. Y., Pan, J. X., Ma, Q., & Xiao, Z. H. (2008). Research progress and application of RFID anti-collision algorithm. In Proceedings of the international conference on telecommunication engineering (vo1. 48, no. 7, pp. 124–128).
[13] http://assets.devx.com/articlefigs/15504.jpg
[14] https://www.researchgate.net/profile/Jiafu_Wan2/publication/254029342_Security_in_the_Internet_of_Things_A_Review/links/550b9adf0cf290bdc111f796/Security-in-the-Internet-of-Things-A-Review.pdf
[15] Finkenzeller, K. (2003). RFID handbook fundamentals and applications in contactless smart cards and identification (2nd Ed.). West Sussex: Wiley.
[16] Wang, D., Wang, J. W., & Zhao, Y. P. (2006). A novel solution to the reader collision problem in RFID system. In Proceedingof the IEEE wireless communications, networking and mobile computing (WiCOM 06) (pp. 1–4).
[17] Song, I. C., Hong, S. H., & Chang, K. H. (2009). An improved reader anti-collision algorithm based on pulse protocol with lotoccupied probability in dense reader mode. In Proceeding of the IEEE 69th vehicular technology conference (pp. 1–5).
[18] Kim, J., Lee, W., Yu, J., Myung, J., Kim, E., & Lee, C. (2005). Effect of localized optimal clustering for reader anti-collision inRFID networks: Fairness aspects to the readers. In Proceeding of the IEEE international conference on computer communications and networks (pp. 497–502).
[19] Juels, A., Rivest, R. L., &Szydlo, M. (2003). The blocker tag: Selective blocking of RFID tags for consumer privacy. In Proceedings of the 10th ACM conference on computer and communications security (CCS 2003), (pp. 103–111).
[20] Ohkubo, M., Suzuki, K., & Kinoshita, S. (2003). Cryptographic approach to privacy- friendly tags. RFID privacy workshop (p.82). Cambridge, MA: MIT.
[21] Duels, A., Pappu, R., & Euros, S. (2003). Privacy protection RFID-enabled banknotes. In Proceedings of seventh internationalfinancial cryptography conference (pp. 103–121).
[22] T2TIT Research Group. (2006). The T2TIT—Thing to thing in the internet of things-project. ANR.
[23] Lakafosis, V., Traille, A., & Lee, H. (2011). RFID-CoA: The RFID tags as certificates of authenticity. In Proceedings of theIEEE international conference on RFID (pp. 207–214).
[24] https://www.slideshare.net/indravi/will-internet-of-things-iot-be-secure-enough

[25] Hachem, S., Teixeira, T., &Issarny, V. (2011). Ontologies for the internet of things (pp. 1–6). New York: ACM.
[26] https://image.slidesharecdn.com/internet-of-things-iota-seminar-ppt-by-mohan-kumar-g-160122172302/95/internetofthings-iot-aseminar-ppt-by-mohankumarg-7-638.jpg?cb=1453483528/
[27] https://www.slideshare.net/MohanKumarG/internetofthings-iot-aseminar-ppt-by-mohankumarg
[28] https://www.analyticsvidhya.com/blog/2016/08/10-youtube-videos-explaining-the-real-world-applications-of-internet-of-things-iot/
[29] https://techcrunch.com/2015/10/24/why-iot-security-is-so-critical/
[30] https://www.forbes.com/sites/gilpress/2017/03/20/6-hot-internet-of-things-iot-security-technologies/#6ab252291b49
[31] https://www.cisco.com/c/en/us/about/security-center/secure-iot-proposed-framework.html
[32] https://www.ericsson.com/en/.../ericsson.../end-to-end-security-management-for-the-iot...
[33] www.collaberatact.com/IoT/Security
[34] Suo, Hui, et al. "Security in the internet of things: a review." Computer Science and Electronics Engineering (ICCSEE), 2012 international conference on. Vol. 3. IEEE, 2012.
[35] Tamboli, Jinita, et al. "Security in the Internet of Things." Communication, Cloud and Big Data: Proceedings of CCB 2014 (2014).
[36] Kamath, Srikanth H., Suyashi Pandey, and Kar Tanisha. "Security Issues in Internet of Things." International Journal of Emerging Research in Management &Technology 6.5 (2017): 260-264.
[37] Dezhgir, Hamid, And Haniyeh Hooshmand. "Security On The Internet Of Things." (2017).
[38] Sayana, Laxman Singh, And Bineet Kumar Joshi. "Security Issues In Internet Of Things."

© 2018 IJSRSET | Volume 4 | Issue 2 | Print ISSN: 2395-1990 | Online ISSN : 2394-4099
National Conference on Advanced Research Trends in Information and Technologies (NCARTICT-2018), Department of IT, L. D. College of Engineering, Ahmedabad, Gujarat, India. In association with
International Journal of Scientific Research in Science, Engineering and Technology

Survey of Machine Learning Algorithms For Dynamic Resource Pricing In Cloud

Meetu Kandpal[1], Dr. Kalyani Patel[2]
[1]MCA, Gujarat Technological University, Ahmedabad, Gujarat, India
[2]MSc(IT), Gujarat University, Ahmedabad, Gujarat , India

ABSTRACT

The paper provides insights of various machines learning algorithm that could be helpful in deriving the dynamic pricing of resources in cloud. Currently machine learning has impact on many IT and non IT sectors. At the same time because of great change in computing from on premise to cloud computing many big companies has opted cloud computing in which resources are provided on demand basis via internet. On the basis of resource usage machine learning algorithm help to predict the future demand and also help in deciding the price of resource at the time of request (spot request).

Keywords: Cloud Computing, Cloud Pricing, Machine Learning Algorithms

I. INTRODUCTION

One of the aims of machine learning algorithms is prediction. Like linear regression, multiple regression algorithms help various sectors to increase profitable opportunities by forecasting demand. Forecasting can be based on the past data sets. As cloud services are delivered on demand basis the future demand of recourses can be forecast by using these machine learning algorithms and new pricing model can be derived.

The paper has started with cloud computing and its pricing policies(1.1), brief about machine learning algorithm and its categories(1.2), literature surveyed on machine learning algorithm is represented under heading(2), conclusion(3).

1.1Cloud Computing and pricing policies: cloud computing are the services provided over the internet, like other utility consumer needs to pay the bills as per the consumption. There are various layers of services provided by cloud like software as a service (SaaS), platform as a service (PaaS), infrastructure as a service (IaaS) etc are presented in Table-1.

The major reasons for the companies opting cloud is security, scalability, fast data recovery etc. Big cloud providers are Amazon (AWS), Google (Google cloud), Microsoft, IBM etc.

Table 1. Service model

Service Model Name	Description	Providers	Services
Software as a Service (SaaS)	Cloud provider delivers its software as a service to be used by the customer on demand.	Google	Google doc
		Microsoft	Microsoft Office
		Salesforce	CRM
		in	LinkedIN
Platform as a service (PaaS)	Provider's delivers application development tools in addition to services for testing, deploying, hosting and maintaining applications.	Google	AppEngine
		Microsoft	Azure
		Salesforce	Heroku
		Amazon web services	Elastic Beanstalk
		IBM	Bluemix
		RedHat	OpenShift

Infrastructure as a Service (IaaS)	Providers provides users with physical and virtual resources that satisfies the requirement of the user applications in terms of CPU, memory, operating systems and storage	Amazon	EC2,S3
		Google	Google compute engine
		Microsoft	Windows Azure
		IBM	IBM SmartCloud Enterprise

The cloud providers have their own policies to charge for the services like Amazon has on demand pricing, reserved pricing and auction based.

With On-Demand instances, user pays for compute capacity by the hour with no long-term commitments or upfront payments. User can increases or decreases their compute capacity depending on the demands of the application and only pay the specified hourly rate for the instances they use. [1]

Reserved Instances provide you with a significant discount (up to 75%) compared to On-Demand instance pricing. In addition, when Reserved Instances are assigned to a specific Availability Zone, they provide a capacity reservation, giving you additional confidence in your ability to launch instances when you need them. [2]

Auction pricing was proposed by Amazon EC2 which whether or not a subscriber obtains the VM depends on the price it offers. It is also known as dynamic pricing. Auction based pricing is fair for both customer and service provider because the price is set as per the level of demand and supply [3]. In auction pricing there are two terms first spot price and second bid price.

For using spot instance the customer can request for instance (spot request). Request contains the bid price also. If the bid price is greater than the spot price the resource will be allocated to the customer immediately otherwise it will be allocated when price of spot become less than the bid price.

Table 2. Cloud Resources

Sr#	Name of family	Instance Name	Description
1	General purpose	M4,M3,T2	Memory to CPU ratio suitable for most general purpose applications.
2	Compute Optimized	C4,C3	More CPU than RAM for compute sensitive applications.
3	Memory Optimized	R3,R4	Larger size for memory intensive applications.
4	GPU	P2	Parallel processing for 3D graphic capability using OPENGL.
5	Storage	L3,L2	For data warehousing

Above Table 2 represents some example of resources provided by AWS.

1.2 Machine Learning-Machine learning is the branch of computer science that concerned with generating such algorithms that takes input from the historic data(training data set) and make prediction based on it. One of the examples is Amazon's algorithms that recommend item based on items you have purchased before

There are mainly three types of machine learning algorithm as shown in Fig-1. Supervised machine learning, unsupervised learning, reinforcement learning. Supervised learning is algorithm in which there are training data labelled data is given.

Currently most industry uses the supervised learning to achieve the work well known algorithm in this category

are artificial neural network (ANN), linear regression, support vector machine (SVM),deep learning based CNN (convolutional neural network) and RNN(Recurrent neural network) architectures .Unsupervised is same without labelled data. K-means clustering is an example of unsupervised algorithm still most of the unsupervised algorithms are under the area of research. Reinforcement learning is the algorithm in which machines are forced to learn and take appropriate action contextually. It is mostly used in area of AI.

In recent years a London based start up which is later acquired by the Google, Deepmind achieve the major breakthrough in the area of AI by using the reinforcement learning algorithm. The company developed the algorithm which beats the highest score by human in 50 plus games. The summary of various machine algorithms is presented in Table: 3 (Machine learning algorithms).

These algorithms help the cloud provider to predict the price of the resources at the time of bidding.

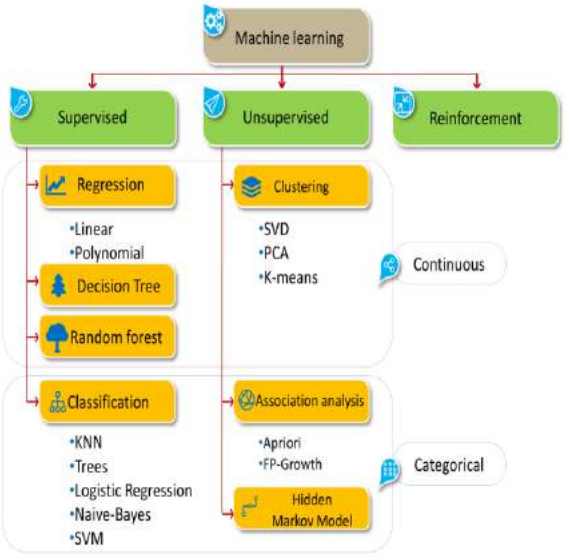


Figure 1. Categories of Machine learning algorithm [4]

II. METHODS AND MATERIAL

Table 3. Machine learning algorithms

Sr No.	Name of Algorithm	Description	Mathematical model	Applications	Advantages	Limitations
1.	Linear Regression [5]	Supervised learning technique to predict the outcomes in quantitative way. Its work on the basis of error minimization technique to fit the linear model.	$h_\theta(x) = \theta_0 + \theta_1 x$ θ_0, θ_1 are parameters Cost function $J(\theta_0, \theta_1) = \frac{1}{2m}\sum_{i=1}^{m}(h_\theta(x^{(i)} - y^{(i)})^2$	Evaluate the trend in the business and help in forecasting future demand.	It is useful when there is a liner relationship between dependent and independent variable.	It is limited to linear relationship only. It is predicting numeric output only.
2.	Multiple Regression [6]	It is extended version of linear regression model to predict the outcomes of more than one variables	$h_\theta(x) = \theta_0 + \theta_1 x$ θ_0, θ_1 are parameters Cost function $J(\theta_0, \theta_1) = \frac{1}{2m}\sum_{i=1}^{m}(h_\theta(x^{(i)} - y^{(i)})^2$	Predict the multivariable outcomes/multi classification problems. Used in market research.	It is useful when there is a liner relationship between dependent and independent variable.	It is limited to linear relationship only. It is predicting numeric output only.

3.	Decision Tree[7]	In decision tree outcome is based on the combination of conditions.	$Entropy = -p\log_2 p - q\log_2 q$ Entropy for each branch is calculated to get the gain or decrease in entropy. $Gain(T,X) = Entropy(T) - Entropy(T,X)$	Used in bank for loan processing, in finance for option pricing	It finds the most important feature in data sets, easy to understand.	Complex for the large data set.
4.	Naive Bayes classifier [8]	Naive Bayes classifier works on the basis of bayes theorem to predict the outcome of posterior probability from the prior probability.	$P(c \mid x) = P(x \mid c)P(c) / P(x)$	Sentiment analysis, document categorization-mail spam filtering	It is well when the input variables are categorical, converges faster, requiring relatively little training data than logistic regression	It works under the assumption of Markov chain that means the current outcome depends on only the previous outcome. y(t) outcome depends on y(t-1)
5.	Artificial Neural Network (ANN) [9]	ANN is method which has input unit hidden units and output unit's weights to predict the outcomes. Mainly work on the basis of back propagation algorithm and gradient decent optimizer.	derivative = output * (1.0 - output) error = (weight_k * error_j) * transfer_derivative(output) weight = weight + learning_rate * error * input	Very good to solve the classification problem and pattern analysis	Easy to understand and produces good classification results.	-
6.	K-means clustering [10]	It is well known unsupervised learning algorithm to make the cluster/ label from the data. It's a generative algorithm which works iteratively.	-	It is used by the most of the search engine like Google to cluster WebPages by similarities of their content.	It does not required labelled data.	Missing variate values cannot be handled by this algorithm.
7.	Support vector machine (SVM) [11]	It is a supervised algorithm, it can be helpful in analysing the data for regression as well as for classification	Depends on the selection of kernel.	Recognition of hand written characters, image classification.	It represents the data in latent space. It performs well for image data.	Computationally complex as compare to ANN.
8.	Hidden Markov Model (HMM) [12]	Hidden Markov model work on the basis of Markov chain assumption, state transition probability and emission transition probability.	θ=sequence of state o=sequence of observation $q_0 q_1$ =(start, final) A=state transition probability B=emission probability C=initial state probability compute probability model (A,B,C)	Language translation, pattern recognition, language modelling	It reveals the hidden features of given sequential data set.	Complex to understand

9.	Deep learning [13]	Deep learning algorithms are deep neural networks which work on the raw data sets. DNN automatically select the best suited features to perform the task.	There are different architectures to deal with different types of problems.	Classification Problems, Pattern recognitions, Speech recognition, Natural language processing, Object detection, etc.	No need of manual feature engineering. It selects the feature automatically from data sets and performs the task.	It is computationally very expensive and need lot of training data set as well.

III. CONCLUSION

As there are different cloud service provider in the market and the boom of the cloud, IT companies has opted cloud. According to the behaviour of the company adoption of cloud the above machine learning algorithm can predict the future demand of cloud resources and help the provider to drive the new pricing policies. Each algorithm has its advantage and limitations; it can be found which will give the best result by implementing WEKA (Waikato Environment for Knowledge Analysis) [14] and other tools. Authors [15] compared various supervised learning algorithms.

IV. REFERENCES

[1]. https://aws.amazon.com/pricing/

[2]. https://aws.amazon.com/ec2/pricing/

[3]. Al-Roomi, May, et al. "Cloud computing pricing models: a survey." International Journal of Grid & Distributed Computing 6.5 (2013): 93-106.

[4]. https://yourfreetemplates.com/free-machine-learning-diagram/

[5]. Montgomery, Douglas C., Elizabeth A. Peck, and G.

[6]. Geoffrey Vining. Introduction to linear regression analysis. Vol. 821. John Wiley & Sons, 2012.

[7]. Woodside, Arch G. "Moving beyond multiple regression analysis to algorithms: Calling for adoption of a paradigm shift from symmetric to asymmetric thinking in data analysis and crafting theory." (2013): 463-472.

[8]. Jin, Chen, Luo De-Lin, and Mu Fen-Xiang. "An improved ID3 decision tree algorithm." Computer Science & Education, 2009. ICCSE'09. 4th International Conference on. IEEE, 2009.

[9]. Wu, Xindong, et al. "Top 10 algorithms in data mining." Knowledge and information systems 14.1 (2008): 1-37.

[10]. Picatoste Ruilope, Álvaro. Energy demand forecasting in smart buildings. MS thesis. Universitat Politècnica de Catalunya, 2017.

[11]. J. A. Hartigan and M. A. Wong Journal of the Royal Statistical Society. Series C (Applied Statistics) Vol. 28, No. 1 (1979), pp. 100-108

[12]. Tsochantaridis, Ioannis, et al. "Support vector machine learning for interdependent and structured output spaces." Proceedings of the twenty-first international conference on Machine learning. ACM, 2004.

[13]. Tumilaar, Kezia, Yohanes Langi, and Altien Rindengan. "Hidden Markov Model." de CARTESIAN 4.1 (2015): 86-94.

[14]. LeCun, Yann, Yoshua Bengio, and Geoffrey Hinton. "Deep learning." Nature 521.7553 (2015): 436-444.

[15]. Domingos, Pedro. "A few useful things to know about machine learning." Communications of the ACM 55.10 (2012): 78-87.

[16]. Caruana, Rich, and Alexandru Niculescu-Mizil. "An empirical comparison of supervised learning algorithms." Proceedings of the 23rd international conference on Machine learning. ACM, 2006.

© 2018 IJSRSET | Volume 4 | Issue 2 | Print ISSN: 2395-1990 | Online ISSN : 2394-4099
National Conference on Advanced Research Trends in Information and Technologies (NCARTICT-2018), Department of IT, L. D. College of Engineering, Ahmedabad, Gujarat, India In association with International Journal of Scientific Research in Science, Engineering and Technology

Recommendation Engine for Mobile-Commerce Product using Heuristic Algorithm

Shanti Verma[*1], Kalyani Patel[2]
[*1]Assistant Professor, L.J. Institute of Computer Applications, Ahmedabad, Gujarat, India
[2]Assistant Professor, K.S. School of Business Management, Ahmedabad, Gujarat, India

ABSTRACT

Now days every age group people use online shopping to complete their daily needs. In this research proposal researcher tries to improve TAM model of Mobile commerce for searching of product. For that researcher first conduct an online survey using Google forms to find the factors associated with mobile commerce. By performing statistical analysis researcher calculate load factors associated with these factors which will used to update the class relationship of defined ontology of product for mobile commerce. Researcher will tries to compare various heuristic search algorithms and find out the best algorithm for defined ontology. Finally researcher will try to apply query based search for available ontology and updated ontology and prove that search results are better in updated ontology. The research proposal finally provides output as a recommendation engine for mobile commerce based on heuristic approach.

Keywords: Online Shopping, Survey, Analysis, Mobile Commerce, Ontology, Heuristic Search

I. INTRODUCTION

I am a Full time PhD. Candidate of Gujarat Law Society (GLS) university, Ahmedabad, India. I have started my PhD work in July 2015 with primary data collection for smart phone and Mobile commerce usage in India using Google forms. I have collected 335 responses in major states of India like Uttar Pradesh, Gujarat, Maharashtra, Karnataka and Madhya Pradesh. The primary data collection method is questionnaire. The mode of filling this questionnaire is through Google Forms. The target audiences are smart phone users who do mobile shopping.

Several cross-tabulations have been used in the questionnaires in order to simultaneously record the responses across more than one variable/response sets for meaningful analysis of the concerned issues. As the data available for various mobile commerce websites are in binary form, there is a strong need to collect primary data.

There are two types of search algorithms: Blind and uninformed, used for ontology based product structure. I have done a comparative study and found that Heuristic search algorithms which are part of uninformed algorithms provide optimized results than blind search algorithms. There are various heuristic algorithms available in market. Most cited heuristic algorithms are Travelling salesman Problem (TSP), A*, Ant Colony optimization and Genetic algorithms (GA).

Currently I try to compare above four algorithms for my training data set/ontology and find which best suited algorithm for my work is. After choosing the Algorithm, I try to add some relationship in the available product ontology and check is there is any improvement in search results after updating in ontology. I have finalized my research topic. My research proposal topic is "Heuristic Search based recommendation engine for product of mobile commerce".

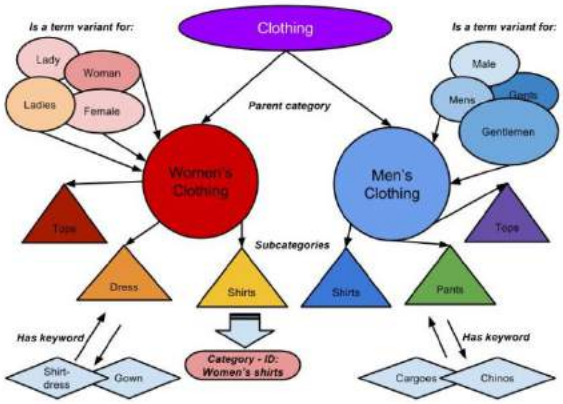


Figure 1. Example of M-commerce product (Cloth) Ontology

II. CONTEXT AND MOTIVATION

In India number of mobile phone users grows with rapid rate due to low cost mobile phones with better configuration and fast data access in low cost. Due to availability, personalization and location based services mobile commerce is used in India by every age people. In India younger age people (below 35) is 75% which plays important role in the growth of online shopping users. There are various key factors which plays important role in success of Mobile commerce. These are accessibility, entertainment, reliability, mobility, externality and reciprocity [11]. By the end of year 2017 it will be expected that 3 billion Smartphone's and 1 billion tablet users Worldwide [15]. This implies that Mobile commerce is hot topic for future research.

The research on Mobile-commerce buying behavior trends aims to contribute in the following areas: The research provides findings and suggestions useful for online shopping applications providers for providing recommendation of product. The research tries to establish the relationship between Mobile users buying behavior and demographics. The main objective of this research is tried to update the class relationship of product ontology which is helpful to provide better search in mobile-commerce environment.

III. LITERATURE REVIEW

3.1 Related Work- Mobile-Commerce

Mobile commerce is a platform where user can purchase product with the use of mobile device connected through wireless data connection [11]. The difference between online shopping and mobile shopping is that in mobile shopping customer use handheld (mobile) devices while in online shopping they can use both mobile and fixed devices like pc [12]. In Smartphone's we see that technology is upgraded day by day with minimum cost. With effect that mobile commerce market is now in boom and major companies are investing in that [13]. Traditional shopping model are applicable only for personal computer users. For mobile shopping different models are used such as Technology acceptance model (TAM) [14]. In mobile shopping users get better offers on product as compared to traditional shopping with effect number of users who uses

3.2 Related Work- Ontology

As the number of internet users grows exponentially, size of database of World Wide Web (WWW) also grows exponentially. It is very difficult to search in huge database manually [5]. So organization of database is major concern as one query has different meaning but we want the correct result. Patil,

Aradhana R., and Amrita A. Manjrekar was combined the concepts of Ontology, text feature extraction and hierarchical clustering to provide proper information retrieval [6]. Now a day's web pages are the major sources of text documents for example: Digital Library. To apply searching there is a need of homogenous cluster of text. Harmony search is a algorithm which is based on heuristic function which produce optimized result of given query [7]. Kuznetsov, V. A., V. A. Mochalov, and A. V. Mochalova proposed an algorithm which automatically updates the ontology basing using some expert rules and logical inference. They use the PHREK algorithm for frequent pattern matching [8]. If we want to create ontology we must aware about rules, relationships and attributes of classes very precisely. Cheng, Bo, Shengda Zhong, and Junliang Chen proposed an expert system for multimedia conferencing management [9].

3.3 Related Work- Heuristic search algorithms

Heuristic search is a technique of artificial intelligence which is used for problem solving. Traditional search algorithms like BFS and DFS produce very slow results, while heuristic algorithms provide quick and precise result of given query. Heuristic algorithms are generally used for recommendation system which used collaborative filtering to provide precise result.

Peñaranda, Cristian, et al. proposed recommendation system for online cooking recipe based on ontology [10]. There are various researchers who work on extracting the text and converting it in ontology based representation. Then there is a need to apply heuristic algorithm like harmony search, A* and simulated Annealing for the problem defined to achieve precise result of query.

IV. STATEMENT OF THESIS/PROBLEM

The research is related to the computer science branch. In the thesis researcher try to satisfy the following hypothesis for the primary data analysis to show the relationship between different attributes of mobile shopping:

H1: There is no significant relationship between Gender and expenditure done in mobile shopping.

H2: There is a significant relationship between category of item purchased and payment method opted by Indian consumer in mobile shopping.

H3: There is no significant difference between population means of gender.

H4: There is no significant difference between population means of yearly income.

H5: There is no interaction between gender and yearly income.

After completing the primary analysis researcher try to develop ontology of product for mobile shopping. Using MATLAB tool they try to identify the best heuristic search algorithm for the defined ontology of product. Generally heuristic search algorithms used Euclidean distance and Manahan distance to find the goal in graph. In this research, researcher tries to modify heuristic function and associate load factor to find the result of defined query

V. TEST OF HYPOTHESIS

To test the hypothesis authors collect the 335 responses of mobile shopping customers using Google forms and apply various statistical methods like Chi-square test, ANOVA, correlation etc.

H1: There is no significant relationship between Gender and expenditure done in mobile shopping [2]

Authors take expenditure done in mobile shopping and Gender of consumer to test the H1 and get the below results of Chi-square test.

```
chisq.test(tab, correct=T)
Pearson's Chi-squared test      data:  tab
X-squared = 4.2519, df = 4,   p-value = 0.373
```

The results of study show that Gender is dependent on the expenditure done in mobile shopping at 95% significance.

H2: There is a no significant relationship between category of item purchased and payment method opted by Indian consumer in mobile shopping [3].

Authors take two questions from questionnaire.

1) Type of product purchased
2) Payment method opted to purchase that product.

To test H2 authors apply chi-square test of independence and get the below results.

```
Pearson's Chi-squared test      data:  tab
X-squared = 50.862, df = 28,   p-value = 0.005174
```

The results of test have P-value 0.005174 which is less than 0.05. It shows the rejection of H2.

These results conclude that there is a relationship between category of item purchased and payment method opted for that product.

H3: There is no significant difference between population means of gender [4].

To test H3 authors apply two way ANOVA to find association between demographic factor gender with income of consumer in mobile shopping. The following results come in the form of ANOVA table given below

SOURCE	SUM OF SQUARES	DF	MEAN SQUARES	F	SIG	H_0
Gender	8.46	1	8.4585	6.0690	0.01427	Accepted H1

The results of above table shows the significance value of 0.01427 which is less than 0.05, means acceptance of H3. We can conclude from above that the there is no impact of gender on yearly income.

H4: There is no significant difference between population means of yearly income [4].

To test H4 authors apply two way ANOVA to find association between demographic factor gender with

income of consumer in mobile shopping. The following results come in the form of ANOVA table given below

SOURCE	SUM OF SQUARES	DF	MEAN SQUARES	F	SIG	H_0
yearly income	84.44	4	21.10890	15.1458	2.294e-11	Accepted H4

Results from above table we may conclude that there is a there is no impact of yearly income on gender of mobile commerce consumers.

H5: There is no interaction between gender and yearly income.

SOURCE	SUM OF SQUARES	DF	MEAN SQUARES	F	SIG	H_0
gender *yearly income	2.88	4	0.7209	0.5173	0.72310	Rejected H5

From the above table we conclude that there is no influence of gender and yearly income towards shopping habit

VI. RESEARCH GOALS & METHODS

To collect primary data for satisfy the defined hypothesis in previous section researcher conduct an online survey about "usage and perception of *Indian* customer for Mobile commerce". Main aim to conduct survey is to find out some important parameters that should be included in Mobile commerce model. From the survey we conclude that page navigation, payment method, organization of product, image based search should be improved. Researcher selects the parameter organization of product that should be improved in TAM of m-commerce. For that researcher want to update ontology of product To apply the search mechanism in ontology, from literature I found that Heuristic search is better .I decide to first develop my heuristic function and apply various heuristic algorithm such as best first search, Beam search and A* and compare the results. After comparison I want to find out the best suited algorithm for my heuristic function using tool MATLAB. Then I try to propose an improved TAM model for M-commerce. To complete the thesis researcher proposed a flow of research as a flow chart representation. In this flow chart researcher present a stepwise approach to complete thesis. There is always probability that our defined hypothesis not satisfied. For that researcher provides alternatives to satisfy the defined hypothesis.

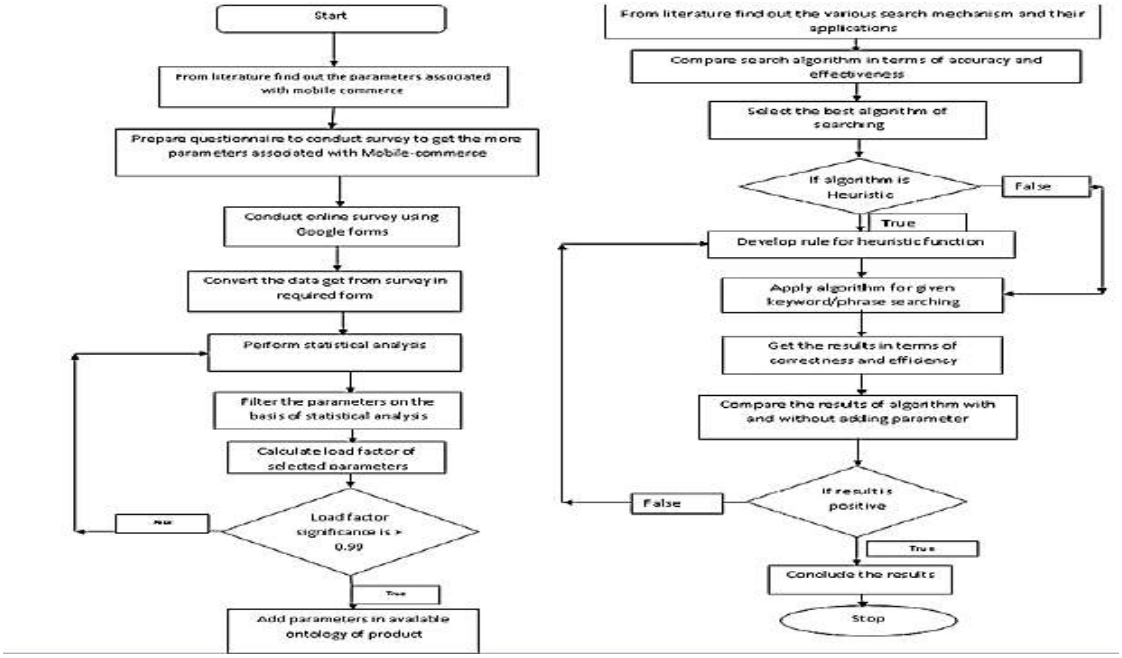


Figure 2. proposed flow of research

VII. DISSERTATION STATUS

Now a day's Indian online market is very competitive. To sustain in market it is very important to satisfy customer needs. As a part to analyze this researcher conduct an online survey to know the customer perception about mobile shopping and smart phone usage. From the collected data researcher published contribute the following research paper which are published in IEEE/Springer and presented in international conference.

1. Survey of smart phone and mobile commerce usage in India [1].
2. Impact of consumer gender on expenditure done in Mobile shopping using test of independence [2].
3. Dependency between type of product purchased and payment method opted by Indian consumer in mobile shopping [3].

Association between shopping habit and demographics of mobile commerce users in India using two-way ANOVA [4].

To provide the better recommendation of product in mobile shopping, researcher find out the heuristic algorithms provide optimal result than blind search algorithms.

Researcher now try to do a comparative study to heuristic search algorithms like TSP, A*, Genetic algorithm and try to find out the best searching algorithm for ontology of product for mobile commerce. After finalize the algorithm, researcher try to update the ontology class relationship and prove that after changing the relationship the search algorithm gives better results than previous ontology of product.

This result is helpful to update the technology acceptance model (TAM) of mobile commerce. Till now I have completed draft version of three chapters of my thesis, which are 1) Introduction to research 2) Mobile-Commerce and user behavior analysis and 3) Role of Recommendation system in Mobile shopping.

VIII. EXPECTED CONTRIBUTIONS

The expected outcome researcher wants from this research is to develop a heuristic approach based search engine for mobile commerce product. For searching researcher tries to use statistical approach i.e. load factor based search.

IX. REFERENCES

[1]. Prof Shanti verma, Dr. Kalyani patel. "Survey on smart phone and mobile commerce usage in India." IEEE International advance computing conference (IACC-2017). IEEE, 2017.

[2]. Kalyani patel, Shanti verma. "Impact of Consumer Gender on Expenditure done in Mobile shopping using Test of Independence." Second International Conference on ICT forSustainable Development (ICT4SD-2016). Springer Proceedings Book-LNNS Series, 2016.

[3]. Verma, Shanti, and Kalyani Patel. "Dependency between type of product purchased and payment method opted by Indian consumer in Mobile shopping." Cloud Computing, Data Science & Engineering-Confluence, 2017 7th International Conference on. IEEE,2017.

[4]. Kalyani patel Shanti verma "Impact of Consumer Gender on Expenditure done in Mobile shopping using Test of Independence" , Second International Conference on ICT for Sustainable Development (ICT4SD-2016. Springer ProceedingsBook-LNNS Series, 2016.

[5]. Muthukumar, S., and N. Muthu. "The Indian kaleidoscope: emerging trends in M-Commerce."

[6]. Sanjivani Mahor, Sourabh Jain "M-Commerce with Process of Knowledge Discovery" International Journal of Computer Technology and Electronics Engineering (IJCTEE) Volume 1 , Issue 2 page 150-154

[7]. Gan Teck Wei, Shirly Kho, Wahidah Husain, Zurinahni Zainol, " A Study of Customer Behaviour Through Web Mining" Journal of Information Sciences and Computing Technologies Volume 2, Issue 1 2015 page 103-107

[8]. Gupta, Rajan, and Chaitanya Pathak. "A Machine Learning Framework for Predicting Purchase by Online Customers based on Dynamic Pricing."Procedia Computer Science 36 (2014): 599-605.

[9]. Turkyilmaz, Ceyda Aysuna, Sakir Erdem, and Aypar Uslu. "The Effects of Personality Traits and Website Quality on Online Impulse Buying." Procedia-Social and Behavioral Sciences 175 (2015): 98-105.

[10]. Gohary, Ali, and Kambiz Heidarzadeh Hanzaee. "Personality Traits as Predictors of Shopping Motivations and Behaviors: A Canonical Correlation Analysis." Arab Economic and Business Journal 9.2 (2014): 166-174.
[11]. Chiang, I-ping, and Yi-Shiang Liao. "Exploring the key success factors of mobile commerce in Taiwan." Advanced Information Networking and Applications Workshops (WAINA), 2012 26th International Conferenceon. IEEE, 2012.
[12]. Wu, Jen-Her, et al. "Falling in love with online shopping carnival on singles' day in China: An uses and Gratifications perspective." Computer and Information Science (ICIS), 2016 IEEE/ACIS 15th International Conference on. IEEE, 2016.
[13]. Naili, Mohamed, Abdelhak Boubetra, and Abdelkamel Tari. "Mobile Commerce Websites Ranking." NGMAST. 2014.
[14]. Li, Qiang, et al. "Multi-Classes Feature Engineering with Sliding Window for Purchase Prediction in Mobile Commerce." 2015 IEEE International Conference on Data Mining Workshop (ICDMW). IEEE, 2015.
[15]. Seth, Earley. "Mobile Commerce: A Broader Perspective." IT Professional 16.3 (2014): 61-65.

© 2018 IJSRSET | Volume 4 | Issue 2 | Print ISSN: 2395-1990 | Online ISSN : 2394-4099
National Conference on Advanced Research Trends in Information and Technologies (NCARTICT-2018), Department of IT, L. D. College of Engineering, Ahmedabad, Gujarat, India In association with
International Journal of Scientific Research in Science, Engineering and Technology

Survey of IoT Frameworks for Smart Water Metering

Urja Mankad[*1], Harshal A. Arolkar[2]

[*1]Assistant Professor, L. J. Institute of computer Applications, Ahmedabad, Gujarat, India
[2]Associate Professor,GLS Institute of Computer Technology, Ahmedabad, Gujarat, India

ABSTRACT

Applications designed for Internet of Things (IoT) needs support of interoperable physical infrastructure, communication protocols, data storing and managing techniques and user-friendly interface. IoT is not as simple as the TCP/IP suite where protocols and communication structure are fixed.The key challenge for IoT is to have a standardized framework to connect any 'thing' to an IoT network. Here the major challenge is of heterogeneity of objects. Many researchers have tried to propose different frameworks to support IoT applications. This paper surveys different types of IoT frameworks proposed to help get insight about IoT trends and future scope for smart water metering system.

Keywords: IoT; framework; Smart Water Metering; Architecture

I. INTRODUCTION

Use of Internet has increased exponentially in the last decade or so. Approximately half of the world's population uses Internet 0. Similarly deployment of IoT based applications is increasing rapidly now. The idea of IoT is to form an Internet of computing devices or objects with sensors which are connected with computers or mobile phones, operated or monitored by human beings. IoT can be expressed as a wireless network layer on top of the Internet where millions of things are continuously tracked, monitored and managed by people all over the world. It is said that by the year 2020 more than 50 billion objects will be joining this network [34].

IoT technology has been widely used in different sectors like, Healthcare, Agriculture, Education, Industrial and many more [1][12]. Different applications of IoT like Smart building, Smart board teaching, Smart Meters, Smart Parking, Smart Tracking, Smart Health, Intelligent Traffic Management, Environment Monitoring, Waste Management, Air Quality Monitoring, Smart Lighting and others have been proposed and implemented. Smart water metering is one of the upcoming applications of IoT[3].

There are many different frameworks available in the market for IoT applications as well as researchers are trying to propose more and more frameworks suitable for IoT applications[9][22][29][29][31][32][33]. In this paper we have done survey of different frameworks with the focus of suitability for smart water metering system. Section II contains related work, Section III contains present situation for the deployment of smart water meters in India. Section IV concludes with the future thoughts and upcoming challenges.

In today's hi-tech era people don't have time to do all the tracking and monitoring tasks on daily basis. IoT connects different types of devices, objects and people together to communicate and exchange data or information to solve many problems which make theroutine easy and simple. Planning and scheduling of task can also be done easily if our objects are on internet. We can synchronize daily activities without stressing our mind. As a result there are technological developments going on for Things Identification Mechanism (Tagging Objects / Addressing Things), Wireless sensing mechanism (Sensors / RFID), Embedded Systems (Microprocessors / OS), Data Management Technology (Data Mining / Big Data) [24][26]

Basic Architecture of IoT:

Internet of Things can be referred as machine-to-machine interaction with secure connectivity, appropriate infrastructure, large amount of data transfer, computation, storage and analytics for converting data to information for people and businesses [23].Layered model for IoT is similar to the one shown in fig.1.

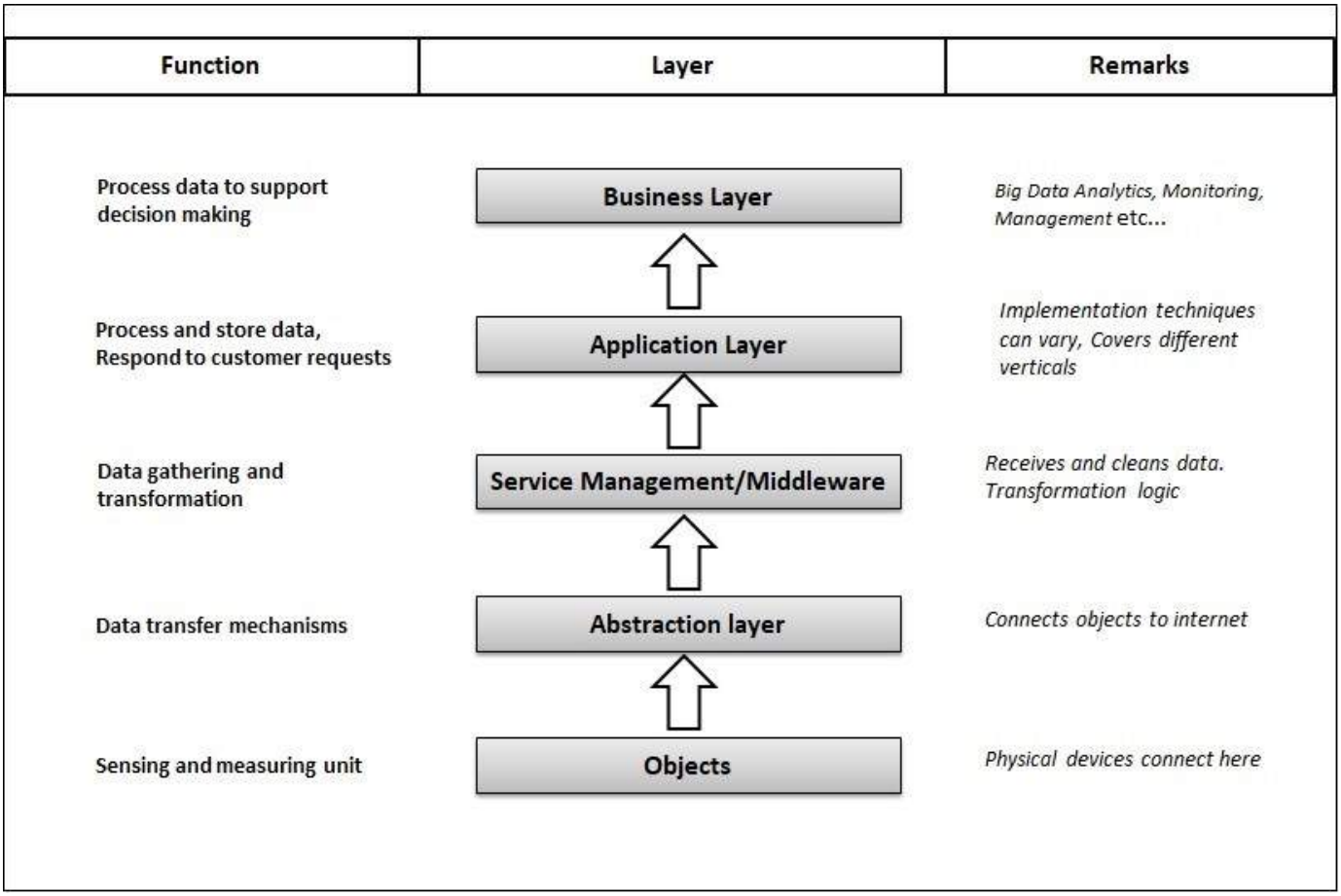


Figure 1. Layered Model for IoT

Description of layers Error! Reference source not found.[5]:

1. Object Layer: Physical objects / sensors, of IoT are connected at this layer. Large amount of data are generated at this layer and from there digitized data are send to the abstraction layer using a secure channels.
2. Abstraction Layer: Data gathered from the objects are sent to the service management layer using different connectivity/communication services of the internet.
3. Service Management Layer: Provides a middleware for facilitating IoT applications to run on different hardware platforms.
4. Application Layer: This layer provides high quality smart services to customers depending on their request.
5. Business Layer: This layer manages system activities and services. It is responsible to build abusiness model to support decision support system based on Big Data analysis.

II. METHODS AND MATERIAL

When we talk about Internet we have different types of networks like Wired Network, Mobile Network, Sensor network and so on but there the nodes are of same type and are using same methods and protocols to exchange data to communicate in between. But when we say Internet of Things, the main challenge is of heterogeneity of objects and their communication techniques and the physical connections[27].

Due to an incredible advancement in IoT services, different frameworks based on interoperability hardware, data exchange methods, low power communication with cloud services, data compression and security have been proposed by researchers[6][7][28].

For the IoT implementations a conceptual layered structure is needed to serve as a support for building applications [2]. The IoT framework should provide interrelation between objects at physical layer, operating system / mainframe support, communication mechanism and application programming interface standards etc.
The Internet of Things Middleware framework joins different components together for enabling smooth communication between objects or devices or people in the network. The IoT interface provides the interaction between the Internet and Things. [14][17]

Many enterprises and alliances have come up with different IoT frameworks or Platforms to increase the connectivity of the devices / objects into the cloud/ network services. These frameworks and platforms provide interoperability among devices and powerful analytic power for the applications to build on [11]. Some frameworks also provide value added services like security and authentication to access the data from different applications of IoT [17][21]. Here the survey is carried out from two perspectives that are commercial frameworks / open source frameworks and frameworks proposed by researchers.
Commercial frameworks used by some of the organizations are proprietary solutions. They provide their components and technologies for a specific problem.

Maven Systems [39] uses a device called WiArt which is used as a wireless connectivity module and collects data from the smart meters using GSM. This data are sent to the data concentrator using GPRS. They provide cloud storage for collected data and different management reports can be generated. Maven systems have their own mesh topology in which communication occur using ISM bands.

Sensus[40] have SensusRF as a LPRF wireless connectivity module. Other solutions are Connit[41][42], IoTSens[43], WebNMS[44], Smarter Homes [45]. Many proprietary solutions are available in the market. Table 1 shows comparison of few of them.

Table 1. Comparison of Commercial solutions for smart water metering

	Maven	**Sensus**	**Connit**	**IoTSens**	**WebNMS**
Data collection/ connectivity devices	WiART, Wireless connectivity module (LPRF), GSM based smart metering modul	Sensus RF, Wireless connectivity module (LPRF), GSM/GPRS/3G	Blue Pulse Mini, Blue Pulse XL2i, LPRF (SigFox), GSM/GPRF communication	Watchmeter Data Logger, Limnimeter, Wireless connectivity (2.4 GHz), GSM Based connectivity	Devices are 3rd party (flow meter, pressure meter), Connect to gateway over RS232, MODBUS, GSM/Ethernet connectivity from gateways
Data concentrator units	Wireless (GPRS)/Wired (Ethernet) Data Concentrator Units, Walk by/Hand held data collector units also available	GPRS / Ethernet based data collection, Hand held collection units also available	GPRS/3G/Ethernet	Not specified by provider	Not specified by provider
Meter data management (cloud software)	Close storage, Reporting	Close storage, Reporting	Close storage, API export, Reporting	Not specified by provider	Cloud storage, Alerts, notification, analytics/reporting, Rule based process automation

Meter types/ integration	Analog and Digital inputs	Not specified by provider	Pulse meters	Not specified by provider	Not specified by provider
Communication Technologies	Maven SmartMesh (Mesh topology) on ISM Band - Works on 433 MHz, 865-868 MHz (802.15.4g) SmartMesh using wired connectivity - UART, TTL, RS232, RS485 ZigBee (on 2.4 GHz)	Mesh topology on ISM Band - Works on 433 MHz, 868 MHz (902 MHz in USA) ZigBee (2.4 GHz)	Star topology using SigFox on ISM Band - 868 MHz (902 MHz in US)	SigFox, LoRa/LoRaWAN - Works on 433 MHz, 868 MHz (902 MHz in US) ZigBee (2.4 GHz), WiFi (2.4 GHz), Bluetooth (2.4 - 2.485 GHz), NB-IoT (Narrow Band IoT), M2M (Data, SMS, Voice) using GSM network	REST, SNMP, MQTT, SOAP, XML

Limited Open source frameworks are available in the market [18]such as Open Source Grid Platform, Open Meter Project etc. but they do have some loopholes which are under research by the researchers [15][16]

OSGP (Open Smart Grid Protocol) is supported and maintained by OSGP Alliance. They have developed a generic, scalable and independent 'Internet of Things' open source platform. Using this platform, smart objects can be easily connected, controlled and monitored. The platform can also help simulate development of smart solutions like Smart Water Metering. As it is open source, it can be used for our own applications and devices with reduced time and development cost.

Weakness of OSGP[15][16]:

1. RC4 algorithm is used for encryption in which stream cipher have known IVs so attackers can easily break WEP within seconds
2. Weak Digest function that may allow meet-in-the-middle attack by sending NACK for wrong digest
3. Sending unsecure broadcast message for firmware updates
4. Session key is the same as master key for encryption, so it is fairly exposed to a number of attacks, and a compromise of this key is possible

Jan Sliwa[26] has proposed a generalized framework for data exchange between objects of IoT system. Here he has analyses different types of data and provided classification for IoT data exchange schemes. There is a general structure and a support platform structure where a Data Broker, the main system element which will manage data among different types of recipients.

Kevin Laubhan et.al.[17]have presented a framework for low power data transmission from sensors to cloud. Here the node architecture has a low power microcontroller which will collect data from the sensors and the regional hub will be collecting data from the node architecture. The cloud architecture will provide efficient data processing and accessibility. The author still has to find different methods of node deployment to maximize coverage area, and optimize scheduling algorithm for minimizing energy usage.

Neeharika Cherukutota et.al.[10]has proposed an IoT framework for smart water meter reading system. They have used a Mediatek cloud sandbox and RESTfull web services to communicate between water meters and cloud.

Enrique Carrillo et.al [8]has proposed a framework for smart systems where computational power needed by the system is provided by the cloud services. The proposed framework uses Arduino integrated development environment for managing and monitoring things connected in the network. Raspberry PI work as a main gateway controller for all Arduino andMicrosoft Azure cloud services are used for computational power.

Mohammed ShahanasK., andDr.BagavathiSivakumar P. [25] have proposeda framework for a smart water management system in the context of smart city initiatives inIndia. They have tried to provide cloud based cost effective solution for data transmission and management.

Sanchez, Tomas, et al [24] have implemented an Adhoc network with clustering of objects and proposed a framework for improving network lifetime in collecting data.
Lloret, Jaime, et al[18] has also proposed a three layered communication architecture. They also says that proprietary solutions for meter communications cannot be suitable for all types of smart meters as there is a lack of standardization in communication protocols which can provide better interoperability among smart meters.

III. DISCUSSION

In this section we have a discussion related to the present situation for the deployment of smart waters in India. Water scarcity has become a serious threat to the globe. So we need some smart systems that will continuously measure the water usage and will help the utility companies in proper decision making. Manual meter reading is not just waste of human resources but also inconvenient and inaccurate. Many researchers have given solutions for Automated Meter Reading Systems in many countries. We would like to connect all water meters into the internet through the IoT as a global solution towards smart city. By the year2020 life will become easier with the help of all automated things. Almost all the things will be interacting with each other and human interaction will be negligible for the task of daily routine. Using smart water meter we can minimize the human load by taking automatic meter reading and sending it to the utility office for further decision making and generating bill. It will also helpful to the customers to know their real time water usage and can get the actual bill in time.

Smart water metering is defined in[29] as the system where smart meters provide meter reading data by monitoring water usage and transfer the consumption data frequently to the utility office, where it is integrated with the business system, like Billing System and the information is also shared with the customers via SMS or web portal. Using smart meters the water consumption data can be gathered automatically without physical visit by a human. Smart water meters can be used to find the solution for the problems like, Water losses, water quality, disasters and drought. Basic Components of Smart Water Metering System are defined by[12][13]. Fig 2 shows the basic components of smart water metering.

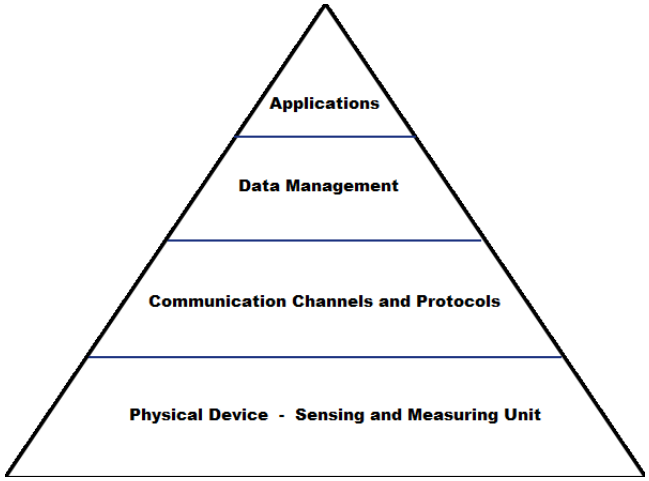


Figure 2. Smart Water Metering System Components

Smart Water Meter components are used in such a way that they can solve many problem of routine life related to water. Like Waste water management, Prevention of water losses due to leakage problem, Managing water quality and water level etc.. To achieve this goal we require a generalized framework architecture which can help easy and efficient implementation of smart water metering system.Many researchers have tried to give a generalized framework for IoT applications. Where Neeharika[10] has proposed an architectural framework for IoT based water meter.

Shahanas[25] has proposed an architecture for implementing smart water monitoring system. It includes basic three units . 1. Collecting water level from the tanks using sensors, 2. Transmit the data to a centralized server using Arduino and Raspberry Pi. 3. Visualizing the data through a web interface and also in the cloud using Ubidots cloud Platform. SMS and email alert system are also included as a part of user interface.
Ankith S et.al.[4]have proposed a design of smart water flow meter system for India. They have designed a prototype using HTTP Compatible CoAP based monitoring and control system. They have compared their system with the other [18][19][20] smart water meter reading systems and automated meter reading systems which are using ZigBee wireless technology for communication between water meters and base stations and data concentrators. In their proposal they have used 6LoWPAN protocol for wireless communication and they have tried to prove that it has more advantages over ZigBee. They stated that 6LoWPAN is more suitable for IoT based applications because using 6LoWPAN protocol sensor networks can be directly connected to the Internet without IP conversion. Proposed system's

prototype is simulated using ContikiOS, MySQL database and visual studio for web application which provides GUI to the users for the interpretation of usage data in the graphical way.

Table 2 shows comparison of the frameworks proposed by various researchers with respect to metering components and major IoT challenges.

The comparison shows that most of the proposals are all vertically defined according to the situation and need. Some are talking about communication protocols; some are talking about data gathering techniques. Further there are different solutions for water flow measuring, leakage management, billing patterns, cloud computations etc…but not a single solution provides flexible, generalized framework which can be suited for any scenario at any location for the implementation of smart water metering system. Hence it is required to set a framework in which any type of smart water metering system implementation becomes easy.

Table 2. Comparison of Frameworks

	Smart Water Metering Components			IoT Major Challenges		
Proposals	**Sensing Technique**	**Channel Acquisition & Communication Mechanism**	**Data Gathering and Analyzing**	**Interoperability**	**Scalability**	**Security**
3	X	✓	✓	X	X	X
4	✓	✓	✓	X	X	X
5	✓	✓	✓	X	X	X
6	✓	✓	X	X	X	✓
7	X	X	✓	X	X	X
29	✓	X	✓	X	✓	X
31	X	X	✓	✓	X	X
32	X	X	✓	✓	X	X
33	X	X	X	✓	X	X

As shown in Table 2, major IoT Implementation challenges are Interoperability Scalability and security. When we talk about smart water metering system, there may be different water meters communicating with each other as well as may send or receive data to other devices I the Internet, so Interoperability requirement is must. Apart from that, as the residential and commercial area will increase, smart metering range must be increasing with same communication speed and accuracy so that will be a big challenge. And last but not the least is the security. When user's personal device information is shared in the Internet, security is the main concern of an individual.

IV.CONCLUSION

The study of IoT framework concludes that no common framework can be used to develop any type of IoT applications or to deploy any system under IoT network. The main issue of the available IoT architectures is lack of full interoperability, privacy and security. Due to heterogeneity in IoT frameworks they have many network related problems like, channel allocation, routing mechanism, interface design and so on, many organizations and researchers are trying to build a uniform architectural framework that can fulfil the technical need for IoT. So, there is a strong requirement of a generalized framework to deploy a smart water metering system under IoT.

V. REFERENCES

[1] Agrawal, Shashank, and Dario Vieira. "A survey on Internet of Things-DOI 10.5752/P. 2316-9451.2013 v1n2p78." Abakós 1.2 (2013): 78-95

[2] Al-Fuqaha, Ala, et al. "Internet of things: A survey on enabling technologies, protocols, and applications." IEEE Communications Surveys & Tutorials 17.4 (2015): 2347-2376.

[3] Andrea Zanella et al. Senior Member, IEEE, "Internet of Things for Smart Cities", IEEE Internet of Things Journal, vol. 1, no. 1, Feb. 2014.DOI: 10.1109/JIOT.2014.2306328

[4] Ankith, S., et al. "Design of IPv6 Network enabled Smart Water Flow Meter System for India."

Proceedings of the Asia-Pacific Advanced Network 40 (2015): 115-120.

[5] Atzori, Luigi, Antonio Iera, and Giacomo Morabito. "The internet of things: A survey." Computer networks 54.15 (2010): 2787-2805.

[6] Baoding Zhang, Jialei Liu, "A kind of design schema of wireless smart water meter reading system based on ZigBee technology" in E-Product E-Service and E-Entertainment (ICEEE), 2010 International Conference on, Nov. 2010, pp 1-4

[7] Beal, Cara, and Joe Flynn. "The 2014 Review of Smart Metering and Intelligent Water Networks in Australia and New Zealand." Report prepared for WSAA by the Smart Water Research Centre, Griffith University, 2014.

[8] Carrillo, Enrique, et al. "IoT framework for smart buildings with cloud computing." Smart Cities Conference (ISC2), 2015 IEEE First International. IEEE, 2015

[9] Chauhan, Saurabh, et al. "IoTSuite: a framework to design, implement, and deploy IoT applications: demonstration abstract." Proceedings of the 15th International Conference on Information Processing in Sensor Networks. IEEE Press, 2016

[10] Cherukutota, Neeharika, and Shraddha Jadhav. "Architectural framework of smart water meter reading system in IoT environment." Communication and Signal Processing (ICCSP), 2016 International Conference on. IEEE, 2016

[11] Datta, Soumya Kanti, et al. "oneM2M architecture based IoT framework for mobile crowd sensing in smart cities." Networks and Communications (EuCNC), 2016 European Conference on. IEEE, 2016.

[12] Dlodlo, Nomusa, et al. "The state of affairs in internet of things research." Electronic Journal of Information Systems Evaluation 15.3 (2012): 244-259

[13] Guest Editorial, "Special Issue on Internet of Things (IoT): Architecture, Protocols and Services" in IEEE Sensors Journal, vol. 13, no. 10, October 2013

[14] Joseph Boman et al., "Flexible IoT Middleware for Integration of Things and Applications" in 10th IEEE International Conference on Collaborative Computing: Networking, Applications and Worksharing, 2014

[15] Jovanovic, Philipp, and Samuel Neves. "Practical cryptanalysis of the open smart grid protocol." International Workshop on Fast Software Encryption. Springer, Berlin, Heidelberg, 2015.

[16] Kursawe, Klaus, and Christiane Peters. "Structural weaknesses in the open smart grid protocol." Availability, Reliability and Security (ARES), 2015 10th International Conference on. IEEE, 2015.

[17] Laubhan, Kevin, et al. "A low-power IoT framework: From sensors to the cloud." Electro Information Technology (EIT), 2016 IEEE International Conference on. IEEE, 2016.

[18] Lloret, Jaime, et al. "An integrated iot architecture for smart metering." IEEE Communications Magazine 54.12 (2016): 50-57.

[19] Miorandi, Daniele, et al. "Internet of things: Vision, applications and research challenges." Ad Hoc Networks 10.7 (2012): 1497-1516

[20] Mudumbe, Mduduzi John, and Adnan M. Abu-Mahfouz. "Smart water meter system for user-centric consumption measurement." Industrial Informatics (INDIN), 2015 IEEE 13th International Conference IEEE, 2015.

[21] Nimmermark, Erik, and Alexander Larsson. "Comparison of IoT frameworks for the smart home." (2016)

[22] Park, Kil-Sung, and Sun-Hyung Kim. "Design of Mainframe for IoT Framework." 2015 8th International Conference on u-and e-Service, Science and Technology (UNESST). IEEE, 2015

[23] Ray, P. P. "A survey on Internet of Things architectures." Journal of King Saud University-Computer and Information Sciences (2016).

[24] Sanchez, Tomas, et al. "Adding sense to the internet of things—an architecture framework for smart object systems." Pers Ubiquitous Comput 16.3 (2012): 291-308

[25] Shahanas, K. Mohammed, and P. Bagavathi Sivakumar. "Framework for a Smart Water Management System in the Context of Smart City Initiatives in India." Procedia Computer Science 92 (2016): 142-147

[26] Sliwa, Jan. "A Generalized Framework for Multi-party Data Exchange for IoT Systems." 2016 30th International Conference on Advanced Information Networking and Applications Workshops (WAINA). IEEE, 2016

[27] Stankovic, John A. "Research directions for the internet of things." IEEE Internet of Things Journal 1.1 (2014): 3-9
[28] Wesnarat, Akekaluck, and Yodyium Tipsuwan. "A power efficient algorithm for data gathering from wireless water meter networks." Industrial Informatics, 2006 IEEE International Conference IEEE, 2006.
[29] White paper by Sierra Wireless, Inc. – "Cellular Communications and the Future of Smart Metering", September, 2010. 2 2010
[30] http://www.iotleague.com/list-of-10-iot-frameworks-and-platforms/ Accessed on 4th Dec 2016
[31] https://developer.ibm.com/architecture/webapp, Accessed on 3rd Jan 2017
[32] https://blog.persistent.com/index.php/2015/08/21/an-end-to-end-iot-platform-the-parts-that-make-the-whole-whole/, Accessed on 3rd Jan 2017
[33] http://www.oracle.com/technetwork/articles/entarch/shrivastava-automated-frameworks-1692936.html, Accessed on 4th Dec 2016
[34] http://spectrum.ieee.org/tech-talk/telecom/internet/popular-internet-of-things-forecast-of-50-billion-devices-by-2020-is-outdated. Accessed on 8th Jan 2017
[35] https://www.tizen.org/.../ballroom1_31st_1100-1200_iotivity_-_connecting_things-in-iot-ashok-subash.html
[36] www.http://events.linuxfoundation.org/sites/events/files/slides/IoTivityCoreFramework.pdf
[37] www.oracle.com/us/industries/utilities/046596.pdf
[38] https://dzone.com/articles/iot-software-platform-comparison
[39] www.mavensystems.com/monitormymeter/index.html
[40] https://sensus.com/smart-water-network/
[41] www.connit.com/en/
[42] www.connit.com/en/iot-smart-water-metering-sigfox-lora-water-management-solution/
[43] www.iotsens.com/smart-water-metering/
[44] https://www.webnms.com/iot/smart-metering.html
[45] https://smarterhomes.com/

© 2018 IJSRSET | Volume 4 | Issue 2 | Print ISSN: 2395-1990 | Online ISSN : 2394-4099
National Conference on Advanced Research Trends in Information and Technologies (NCARTICT-2018), Department of IT, L. D. College of Engineering, Ahmedabad, Gujarat, India In association with International Journal of Scientific Research in Science, Engineering and Technology

Attribute and Energy Aware Tree Formation in Wireless Sensor Network

Vrunda Popat*[1], Harshit Vakil[2], Vishakha Sanghavi[3]
[1]IT Department, V.V.P. Engineering College, Rajkot, Gujarat, India
[2]IT Department, Government Polytechnic, Rajkot, Gujarat, India
[3]IT Department, V.V.P. Engineering College, Rajkot, Gujarat, India

ABSTRACT

Data Aggregation is a vital technology to save energy in Wireless Sensor Network (WSN). Data aggregation has been recognized as an competent method to reduce energy consumption by dropping the number of packets sent. For data aggregation solutions, different topologies are created to aggregate the data. For e.g. line, chain, cluster and tree. In this paper, our focus is on tree based data aggregation. There are different methods of tree formation for tree based data aggregation techniques. Most of them do data aggregation efficiently in homogeneous networks but not in heterogeneous networks. Many different applications can be deployed together in WSNs and packets generated by different applications have different attributes. These packets from different applications cannot be aggregated. To make data aggregation more efficient in heterogeneous networks, Attribute and Energy aware tree formation is introduced. In this routing approach, any node will select parent of same type and having higher residual energy and there are more chances of aggregation between same type of packets so that data aggregation can be efficient even in heterogeneous networks.

Keywords: Wireless Sensor Networks, Data Aggregation, Attribute Aware, Tree based Data Aggregation, Heterogeneous Networks

I. INTRODUCTION

A. Wireless Sensor Network

A Wireless Sensor Network is having a no. of wireless sensor nodes which form a sensor field and a sink node. These sensor nodes, having the abilities to sense their surroundings and perform some computation and communicate wirelessly [1].

There is a limitation of size and battery power, these devices typically have limited storage capacity, limited energy resources, and limited network bandwidth. Data propagates via wireless links in the network and wireless transmission is extremely expensive when compared to local processing of data.

Many different types of sensors available like thermal, visual, infrared, acoustic and they monitor a wide variety of environment conditions like humidity, temperature, lightning condition, pressure, soil makeup, noise levels and characteristics such as speed, direction, and size of an object. WSN applications are in many areas like emergence surveillance [2], environment monitoring [3], target tracking [4], military, health, home, space exploration, chemical processing, disaster relief, and other commercial areas [6].

B. Data Aggregation in Wireless Sensor Network

Data aggregation refers to acquiring data from the sensor nodes to the gateway node. Data aggregation reduces the power consumed during data transmission between the sensor nodes [7].

The sink node is secure as it has unlimited energy available and the sensor nodes are having limited energy and are unsecured. The sensory information collected by these nodes is sent to the Gateway node through Wireless hop by hop transmissions and it is aggregated at intermediate sensor nodes to conserve energy by using suitable aggregation function on the data received. The

reduction in network traffic and energy consumption on sensor nodes is achieved by data aggregation as in [1, 9].

To make easy deployment it is required that sensor devices are inexpensive, small in size but having long lifetime. So, protocols for sensor networks need to be designed vigilantly and limited resources like energy, storage and computation should be used efficiently. The Data Aggregation in WSNs can be classified into five basic types: 1) Centralized Data Aggregation 2) Cluster Based Data Aggregation 3) Multipath Data Aggregation 4) Tree Based Data Aggregation 5) In Network Data Aggregation [12].

C. Tree based Data Aggregation

In the tree-based approach, aggregation is performed by constructing an aggregation tree as shown in Figure 1, which is mostly a minimum spanning tree, where root is the sink node and leaves are the source nodes.

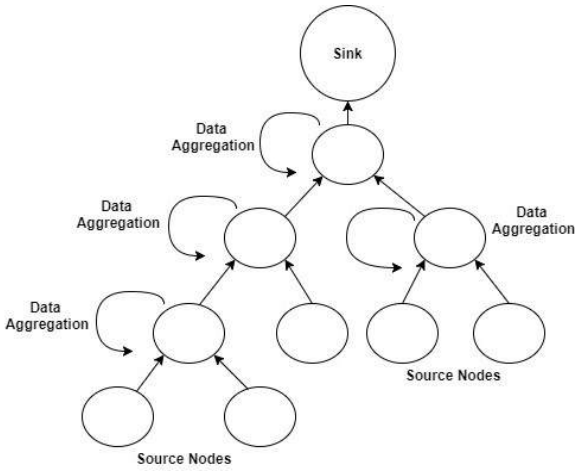


Figure 1. Tree based data aggregation

Each node has a parent node to forward its data. The data flow starts with leaves nodes up to the sink and the aggregation done by parent nodes [5].

Here Tree based data aggregation techniques are explored for further study and in tree based approach, most of the available techniques works with homogeneous networks only, which consists of only one type of sensor nodes. Research work has been started in attribute aware tree based approach for data aggregation in heterogeneous networks in wireless sensor networks [17].

II. MOTIVATION AND RELATED WORK

A. Motivation

In Wireless Sensor Networks, sensor nodes first collect the data, aggregate the data and then transmit it to the gateway node directly or indirectly with the help of other sensor nodes. Tree based data aggregation protocol works with homogeneous networks only while in real life, many times WSN is a heterogeneous network. Current schemes do not provide efficient data aggregation in heterogeneous environment. Here heterogeneous environment represent two or more application running on same network.

It is found that attribute aware aggregation is very much required for heterogeneous networks which should be explored more for research work. As here when there are two applications which are sensing different type of data are always finding the same path, which is statically predefined. In this case, if any node tries to select the parent node of its type then aggregation can be performed efficiently and overall network lifetime can be increased.

B. Related Work

The data sampled by the same type of sensors have more redundancy since the sensor nodes are usually quite dense in wireless sensor networks. To make data aggregation more efficient, the packets with the same attribute, defined as the identifier of different data sampled by different sensors, such as temperature sensors, humidity sensors, etc., should be grouped together [8].

Data aggregation is performed at intermediate nodes in tree-based routing protocols, and a brief representation of data is transmitted to the root node, i.e. sink. One of main tasks for tree-based scheme is to construct an energy efficient data aggregation tree. Since the tree constructed in advance is static, most tree-based schemes can only be suitable for applications in which source nodes are known.

In the following section, we focus on some of tree based routing protocols separately which works well in homogeneous networks by reviewing the main concepts and briefly commenting on the pros and cons of each scheme.

1) **TAG -** The Tiny Aggregation Approach: TAG [13] is a data-centric protocol. Tiny Aggregation algorithm uses the routing scheme which consists of two phases: 1) The distribution phase, where queries are scattered to the sensor nodes 2) The collection phase, where the aggregated sensor readings are routed up the aggregation tree. After the tree is constructed, the queries are sent along with the structure to all nodes in the network. TAG uses the selection and aggregation database query languages (SQL). During the data collection phase, because of the tree structure, each parent has to wait for data from all of its children before it send the aggregated data. TAG may be ineffective for dynamic topologies or link/device failures. As the topology changes, TAG has to rearrange the tree structure, that means it is expensive in terms of energy consumption. It only supports homogenous networks.

2) **EADAT -** Energy Aware Distributed Heuristic: EADAT is an energy-aware distributed heuristic algorithm [15] to construct Data Aggregation Tree in WSNs and also maintain it. In this approach, the node having higher residual energy level has the higher probability to become non-leaf and that is parent tree node, as the number of alive nodes are more the network lifetime. It does not supports aggregation for different type of sensor reading packets.

3) **PEDAP -** Power Efficient Data gathering and Aggregation Protocol: PEDAP [16] constructs a minimum spanning tree considering transmission overhead as the link cost, and thus each communication round consumes less energy. It is expensive to reconstruct the spanning tree for every communication round. Does not support aggregation in heterogeneous environments.

4) **ADA -** Attribute aware Data Aggregation scheme: Nearly all of the existing tree based data aggregation scheme work suitably in homogeneous environment but rarely considers the impact of diversity, including diverse sensors or different applications in the same WSN [16]. This scheme provides routing based on type or attribute of nodes.

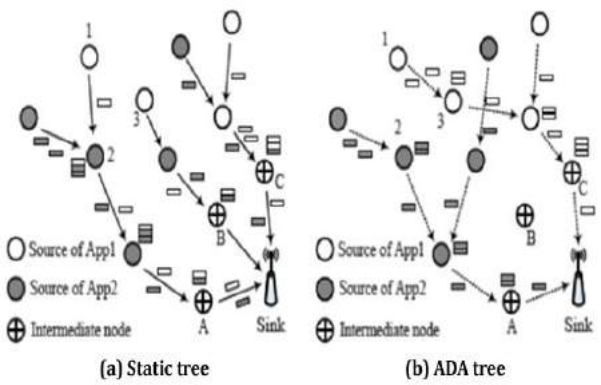


Figure 2. Basic idea of ADA scheme[17]

As shown in Figure 2, it illustrates a typical tree-based routing protocol consists of the shortest path tree rooted at the sink with metric such as hop-count. The static and predetermined routing protocol hardly adapts to heterogeneous environment. Routing in ADA improves aggregation efficiency even in heterogeneous environment. Here Energy of node is not considered while selecting parent for maintaining a tree.

III. IMPLEMENTATION AND RESULTS

A. Flow of work done

For research work, we developed two tree formation techniques for using as performance measures of data aggregation efficiency.

a) Shortest path tree : This scheme will select any parent node from all possible parent nodes which has one lesser level than its level in tree. So it is a static tree formation technique which takes in consideration only distance of parent node.

b)Attribute and Energy aware tree : This scheme will select any parent node of its similar type from all possible parent nodes which has one lesser level than its level in tree and it also considers residual energy while selecting parent node from all possible parent of similar types and which is enhanced ADA (Attribute and Energy aware Data Aggregation scheme).

This ADA program decides parents of each node and make tree based routing possible. ADA scheme gives text file as output which contains list of parent node for every node in network. And it is given to NS2 as input. So that NS2 routing protocol will know about parent

node of every node and also children of every parent node and so that finally it works as a Tree based routing protocol.

There are some modification performed in DSR routing protocol so that it can be used as tree based routing protocol and also aggregation function is added in DSR routing protocol.

1)Flow Chart of Attribute and Energy aware Tree formation Scheme (ADA): In Figure 3, the flow chart of ADA program is shown.

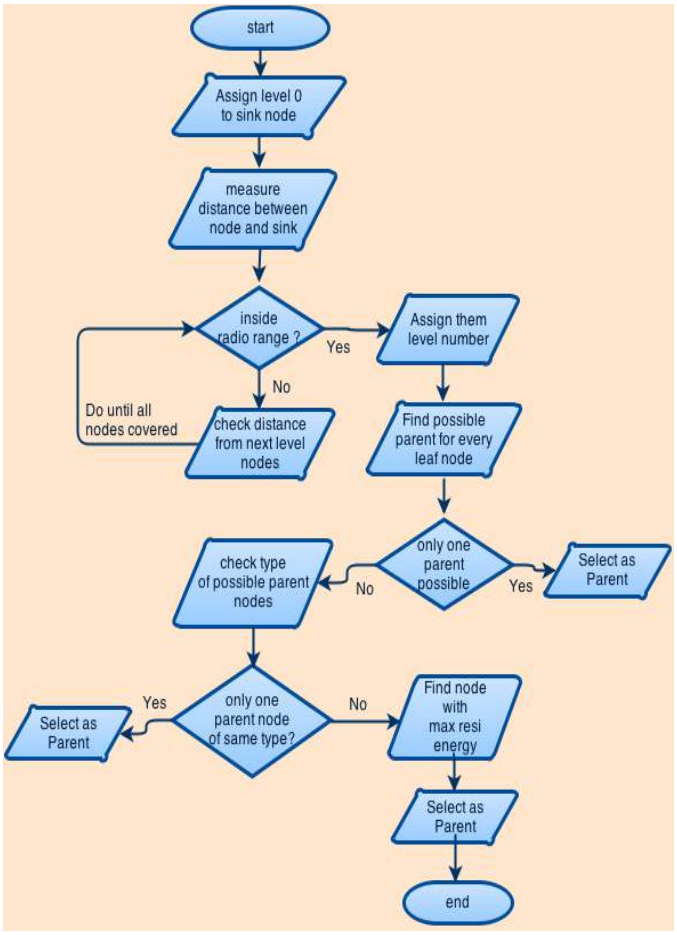


Figure 3. Flow chart of ADA program

2) Output of ADA program : Output Text file as shown in Figure 4 is given to NS2 as input. Here first column represents parent node id, second column represents child node id.

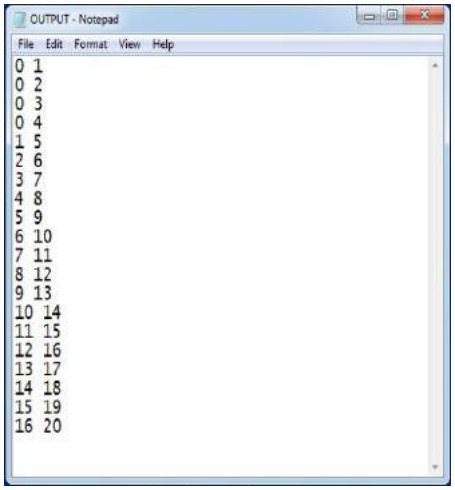


Figure 4. Output text file of ADA program

3) Implementation done in NS2: For generation of tree based topology for data aggregation, some changes are made in DSR routing protocol. Some functions added in DSR routing protocol. Output text file generated by ADA program is given to NS2 as input and this file has list of Parent nodes in first column and child nodes in second column.

Get Parent and child list for tree topology based routing and using this information every node will know if it is a parent node or not and every parent node will know list of all child nodes.

When node receive any packet and timer's current state is on then node does not forward that packet if it is of same type with first packet received, and it is of different type from first packet received then node will forward that packet. If any node receive more packets of same type then more aggregation is possible.

Using energy and attribute aware scheme there are more chances that same type of nodes are children of a single parent of same type and so that more data aggregation can be performed.

B. Simulation Results

All simulation results are observed considering the parameters shown in table I and II. Some source nodes starts initially and send packets to sink node and total max 1000 packets sent at the rate of 1 packet per seconds.

Table 1. simulation parameters

Parameter	Value
Number of Nodes	50, 100, 150, 200, 250, 300, 350, 400, 450
Simulation Time	1000 seconds
Simulation Area	1000 x 1000 meters
Radio Range	250 meters
MAC	802.11

Table 2. energy model parameters

Parameter	Value
Idle Power	0.843 W
Transmit Power	1.650W
Receive Power	1.40W
Radio Range	250 meters
Initial energy of nodes	140J
Packet Size	512 bytes
Idle Power	0.843 W

All graphs are generated for two tree formation schemes: 1) Shortest Path Tree and 2) Attribute and Energy aware Data Aggregation Tree, considering two types of application running in heterogeneous network.

1) Graph of No. of nodes Versus Aggregation Ratio: As shown in Figure 5, the X - axis of graph represents No. of nodes and Y - axis of graph represents Aggregation Ratio. As no. of nodes increases network become more dense and chances of aggregation are more in dense network.

Formula of Aggregation Ratio is as below: Aggregation Ratio = avg [(Total forwarded packets at node 1/ Total received packets at node 1) + (Total forwarded packets at node 2/ Total received packets at node 2)+..+ (Total forwarded packets at node n/ Total received packets at node n)]

So, as node density increases, Aggregation Ratio decreases which shows Aggregation is performed more efficiently. Graph shows Aggregation Ratio is lesser in ADA scheme with compare to SPT scheme. SPT has less chances of aggregation than ADA scheme.

In SPT scheme : There are less chance of Aggregation so that any node will forward more packets out of total received packets. It may send 7 out of 10.

In ADA scheme : There are more chance of Aggregation so that any node will forward less packets out of total received packets. It may send 4 out of 10.

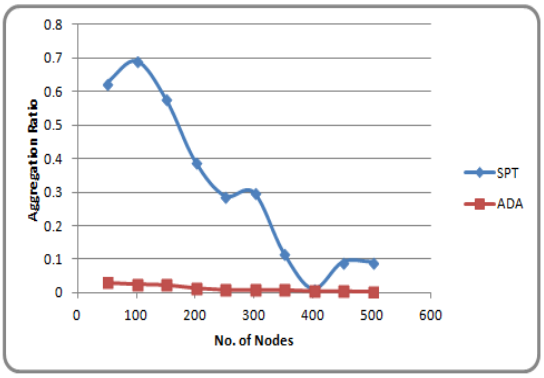


Figure 5. No. of nodes Versus Aggregation Ratio

From this graph, we can say ADA performs Aggregation efficiently than SPT scheme.

2) Graph of No. of nodes Versus PDR: As shown in Figure 6, X - axis represents No. of nodes and Y - axis represents Packet Delivery Ratio. As node density increases Packet Delivery Ratio decreases which shows Aggregation is performed more efficiently.

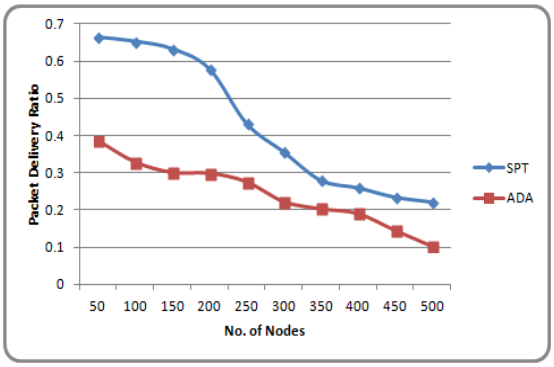


Figure 6. Graph of No. of nodes Versus PDR

Graph shows Packet Delivery Ratio is lesser in ADA scheme with compare to SPT scheme. So that in this scenario, we can say ADA performs Aggregation efficiently than SPT scheme in heterogeneous network.

3) Graph of No. of nodes Versus avg. Residual Energy: As shown in Figure 7, X - axis represents No. of nodes and Y - axis represents Avg. Residual Energy. As node density increases Avg. Residual energy decreases much slowly in case of ADA than that of case when SPT scheme.

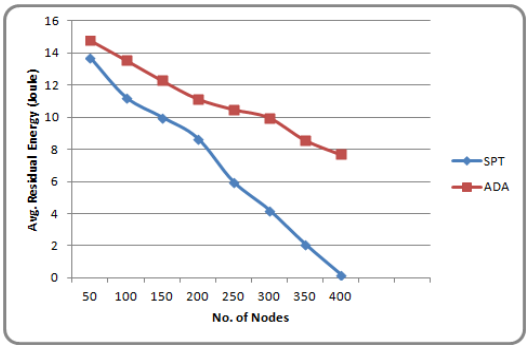


Figure 7. Graph of No. of Nodes Versus Avg. Residual Energy

So that in this scenario, less energy is used in ADA because it forwards less packets. In other way, we can say ADA performs Aggregation efficiently than SPT scheme. Thus, ADA scheme is Energy aware also.

4) Graph of No. of applications Versus Aggregation Ratio: As shown in Figure 8, X - axis represents No. of applications and Y - axis represents Aggregation Ratio.

Graph shows that aggregation efficiency decreases with increase in no. of application so ADA performance is vary with more no. of application. But it still performs better than SPT scheme in all ways ADA is better in heterogeneous networks. Here SPT increase aggregation ratio faster than ADA. This graph is for fixed number of nodes 200.

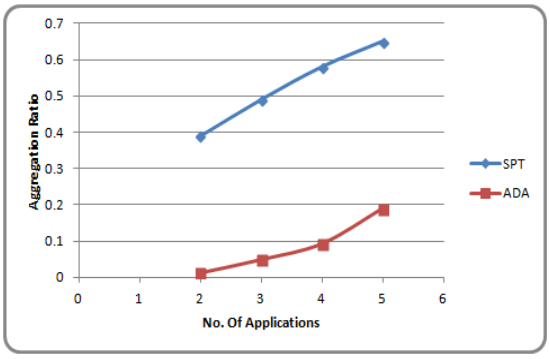


Figure 8. Graph of No. of applications Versus Aggregation Ratio

IV. CONCLUSION

ADA program is developed for tree formation which considers type of parent node and also residual energy while selection of parent which performs better in heterogeneous networks compare to other tree formation techniques like SPT (shortest path tree formation scheme). Performance comparison is done between SPT and ADA for parameters like Packet Delivery Ratio, Aggregation Ratio and Avg. Residual Energy Versus node density and Aggregation Ratio Versus No. of application. Results justify that ADA performs far better aggregation than SPT in all parameters. Overall residual energy is saved more in ADA scheme so that network lifetime is increased. Finally, it is observed that Attribute aware tree formation is an efficient method for data aggregation in all aspects for heterogeneous networks.

V. FUTURE WORK

ADA can perform better even if tree maintenance part is managed. If tree is maintained properly, no. of alive nodes can be more and network lifetime increases. If any parent node dies, all children of that node also become disconnected which make very poor performance of ADA. For that, if any node's energy go below pre defined threshold then it's children should select new parent so that they can continue sending data in the network and overall network lifetime can be increased.

ADA's performance can be measured in multiple sink scenario.

VI. REFERENCES

[1] Anindita Ray, Debashis De. 2012. "Data Aggregation Techniques in Wireless Sensor Network: A Survey" International Journal of Engineering Innovation and Research Volume 1, Issue 2, ISSN: 2277 – 5668, 2012.

[2] Sert S A, Onur E, Yazici A., "Security attacks and counter measures in Surveillance Wireless Sensor Networks" , Application of Information and Communication Technologies (AICT), 9th International Conference on. IEEE, pp. 201-205, 2015.

[3] Guo X, Wang Y, Wei X., "Design of WSN-based environment monitoring system on repair of gas leakage", Control and Decision Conference (CCDC), 27th Chinese. IEEE, pp. 3340- 3344, 2015.

[4] Demigha O, Hidouci W K, Ahmed T., "Energy efficiency in collaborative target tracking in wireless sensor network: a review", Communications Surveys & Tutorials, IEEE, Volume 15, Issue 3, pp. 1210-1222, 2013.

[5] Rashmi M. R, "Analysis Of Data Aggregation Techniques In Wireless Sensor Networks" International Journal of Applied Engineering Research and Development (IJAERD) ISSN(P): 2250-1584; ISSN(E): 2278-9383 Volume 5, Issue 1, pp. 21-30, 2015.

[6] Jingchen Zhang a, Huazhong Zhang b, Xibei Jia b "A Data-aggregation-centric Wireless Sensor Network Routing Architecture" Journal of Information and Computational Science , Volume 9, Issue 3, pp. 635-642, 2012.

[7] Amrutha mohanan.k, P. Vijayalakshmi, "Trust based Data Aggregation in Wireless Sensor Networks" in Journal International Journal of Computer Applications ISSN:0975 – 8887, Volume 73, Issue 22, pp. 8-12, 2013.

[8] E. Fasolo, M. Rossi, J. Widmer, M.Zorzi "In-network Aggregation Techniques for Wireless Sensor Network: A Survey" IEEE Wireless Communications, Volume 14, Issue 2, pp. 70-87, 2007.

[9] Jamal N. Al-Karaki Ahmed E. Kamal, "Routing Techniques in Wireless Sensor Networks: A Survey", IEEE Wireless Commun.Mag., Volume 11, Issue 6, pp. 6–28, 2004.

[10] S.Lindsey and C.Raghavendra, "PEGASIS: Power Efficient Gathering in Sensor Information Systems", Proc. IEEE. Aerospace Conference, pp.1-6, 2002.

[11] Ahmad Sardouk, LeIla Merghem-Boulahia,Rana Rahim-Amoud and Dominique Gaiti, "Data Aggregation in WSNS: A Survey", Materials Science Research Journal, ISSN : 1935 – 2441, 2010.

[12] Abhijith H V, Dr M Dakshayini, "Comparative Study On Data Aggregation Techniques for Wireless Sensor Networks", ISSN: 2321-8169, Volume 2, Issue 4, pp. 838- 845, 2015.

[13] S. Madden et al., "TAG: a Tiny Aggregation Service for Ad Hoc Sensor Networks" , OSDI 2002, Boston, MA, 2002.

[14] C. Intanagonwiwat et al., "Directed Diffusion for Wireless Sensor Networking" IEEE/ACM Trans. Net., Volume 11, Issue 1, pp. 2-16, 2002.

[15] M. Ding, X. Cheng, and G. Xue, "Aggregation Tree Construction in Sensor Networks," IEEE 58th Vehic. Tech. Conf., pp. 2168-2172, 2003.

[16] H. O. Tan and I. Korpeoglu, "Power Efficient Data Gathering and Aggregation in Wireless Sensor Networks," SIGMOD Record, Volume 32, Issue 4, pp. 66-71, 2009.

[17] Fengyuan Ren, Jiao Zhang,Yongwei Wu, Tao He, Canfeng Chen and Chuang Lin, "Attribute-Aware Data Aggregation Using Potential-Based Dynamic Routing in Wireless Sensor Networks", Journal IEEE Transactions on Parallel and Distributed Systems, Volume 24, Issue 5, pp. 881-892, 2013.

© 2018 IJSRSET | Volume 4 | Issue 2 | Print ISSN: 2395-1990 | Online ISSN : 2394-4099
National Conference on Advanced Research Trends in Information and Technologies (NCARTICT-2018), Department of IT, L. D. College of Engineering, Ahmedabad, Gujarat, India In association with
International Journal of Scientific Research in Science, Engineering and Technology

Cloud Management using Network Function Virtualization to increase Service agility

Poonam Rangpariya[1], Dr. Chirag Suryakant Thaker[2]
[1]Department of Computer Science, Ahmedabad, Gujarat, India
[2]Asst. prof. Information Technology Department, L.D Eng College, Ahmedabad, Gujarat, India

ABSTRACT

Network Function Virtualization (NFV) aims to provide high-performance network services through cloud computing and virtualization technologies and its changing the way networks are managed by providing more scalability and flexibility. NFV promises to increase service agility and reduce the overall OPEX and CAPEX experienced by the network operators. In this paper, discussing NFV and NFV and SDN bring together a promise for service agility and OPEX reduction through extreme automation of IT and networking processes in a Service Provider environment.

Keywords: Network Function Virtualization, Software Defined Networking, cloud computing.

I. INTRODUCTION

Internet has become an integral part of everyone's life in present days. The services are delivered by the service providers, as requested by the end user. The evolution of cloud computing has provided an opportunity for the end user to customize the use of the internet as required and pay only for the requested services. However, to provide all these simplicities and customization, there are a few complex systems that work behind the scenes. Every vendor providing services to the customers has a data centre to host the service and manage them. There are severs, routers, switches and middle boxes running in a data centre environment. In traditional networks, there are physical dedicated hardware resources for compute, network and storage components, separately performing their respective tasks [2].

Increased complexity of these networks leads to increase in cost. Administration and management of the physical devices in these networks increases the capital expenditure and operational expenditure for a service provider[4]. More the number of physical devices, higher the CAPEX and OPEX. Moreover, network resource usage is highly unpredictable, leading to uncertainty among the service providers.

Not all network functions or the internetwork devices are used round the clock. Some of the network functions and devices are unused or not fully utilized. These devices and functions can be managed appropriately by sharing the resources with other users or suspending it when not in use.

This paper presents the difference in service agility between traditional network and the cloud network using NFV. On the cloud computing platform the virtualization of network function using NFV is carried out. Orchestration and management of resources, using NFV and OpenStack is demonstrated. Rest of this paper is organized as follows. Section II Methods and material Section III proposed architecture, Section IV Conclusion paper.

II. METHODS AND MATERIAL

1) Dynamic scaling

The network function and their functionality decoupling into insatiable software components provides greater flexibility to scale the actual VNF performance in a more dynamic way and with finer granularity, for instance, according to the actual

traffic for which the network operator needs to provision capacity[1].

In Figures, we use an example of a Customer Premises Equipment to illustrate the economies of scale that may be achieved by NFV. Traditional CPE shows a typical (current network service)

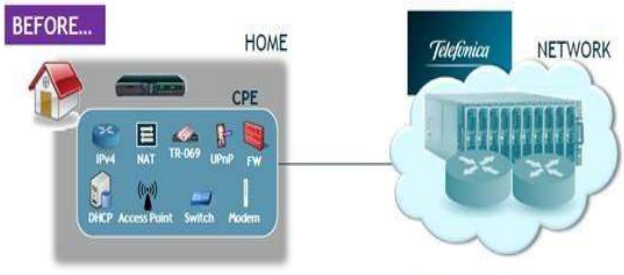


Figure 1. Traditional CPE

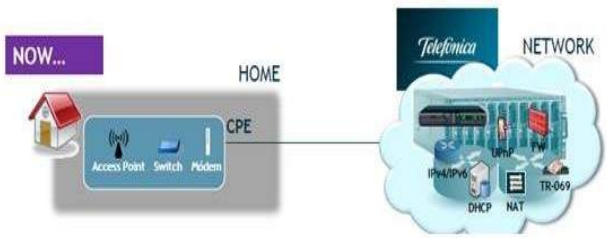


Figure 2. CPE with NFV
(Customer Premises Equipment)

The function made up the implementation of the CPE and their function like: Dynamic Host Configuration Protocol (DHCP), Network Address Translation (NAT), routing, Universal Plug and Play (UPnP), Firewall, Modem, radio and switching[1]. In this example, a single service (the CPE) is made up of eight functions. These functions may have precedence requirements.

For example, if the functions are part of a service chain2, it may be required to perform firewall functions before NAT. Currently, it is necessary to have these functions in a physical device located at the premises of each of the customers both. With such an implementation, if there is a need to make changes to the CPE, say, by adding, removing or updating a function, it may be necessary for a technician from the ISP to individually talk to or go to each of the customers. It may even require a complete change of the device in case of additions. This is not only expensive for the ISPs, but also for the customers.

In Capital Premises Equipment with NFV we show a possible implementation based on NFV in which some of the function software the CPE are transferred to a shared infrastructure at the ISP, which could also be a data center.

This makes the changes described above easier since, for example, updating the DHCP for all customers would only involve changes at the ISP. In the same way, adding another function such as parental controls for all or a subset of customers can be done at once. In addition to saving on operational costs for the ISP, this potentially leads to cheaper CPEs if considered on a large scale.

2)NFV Architecture

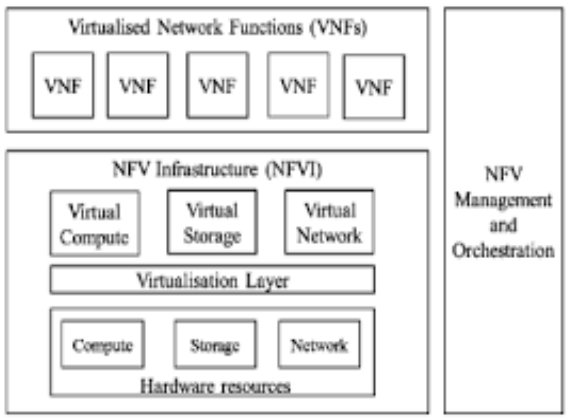


Figure 3. NFV Architecture

A. NFV Infrastructure (NFVI)

IN NFVI combination of both software and hardware resources make up the environment in which VNF are deployed. Two type of resources like physical resources include commercial of shelf computing hardware, storage and network that provide processing &connectivity To VNFs. second resource virtual are abstractions of the computing, storage and network resources. this is achieved using a virtualization layer hypervisor based ,which decouples the resources from the underlying physical resources[3].

In a data center environment, the storage and computing resources may be represented in terms of more virtual machines (VMs), while virtual networks are made up virtual nodes and links.

B. VNFS – virtualized Network Functions

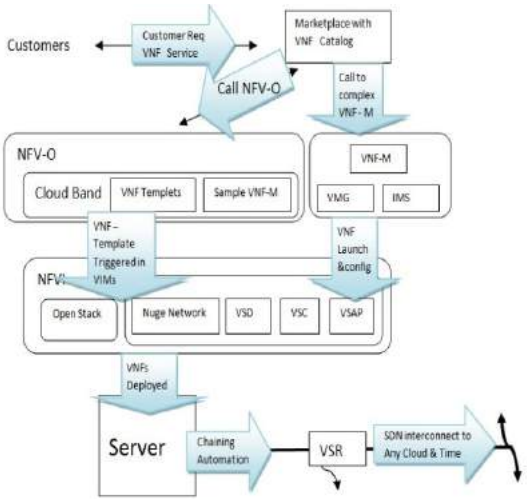
A Network function is well define functional interface and well define functional behaviour of the NFV. example of the NF is that the residential gateway in home network, DHCP server and firewall ,instruction detection etc. VNF is an implementation of an NF that is deployed on virtual resources such as a VM. A service is an offering provided by a TSP that is composed of one or more NFs[1]. In the case of NFV, the NFs that make up the service are virtualized and deployed on virtual resources such as a VM.

C. NFV Management and Orchestration (NFV MANO)

The provisioning of VNFs required the functionality that provides the NFV MANO, and related operation and VNFs configuration, function run on these infrastructure. It include the orchestration and the management of lifecycle of software and physical resources the virtualization infrastructure, support to that and also the lifecycle of VNF management . it also includes databases that are used to stores and data models the information which defined both deployment as well as lifecycle properties of services ,resources and functions[2].

In the NFV framework NFV MANO focuses on all virtualization –specific management tasks. in addition the framework define interfaces that can be used to communications between the different components, as well as coordination with traditional network system such as a OSS and BSS so as to allow for management of both VNFs as well as functions running on legacy equipment.

III. PROPOSED ARCHITECTURE



NFV- Service Chaining

NFV and SDN bring together a promise for service agility and operational and capital expenditure reduction through extreme automation of networking processes in a Service Provider environment. Shown in fig two type of customers (Internal and external) requests to the marketplace with VNF catalog for VNF service. M.K with VNF Catalog need to the NFV-orchestration (NFV-O) so they request to the NFV-O. Cloud based NFV-O provides to the VNF Templates like firewall and catalog type. also provide sample VNF-M.

The VNF Templates triggered in OpenStack and nuge network. OpenStack Provide more expansive networking automation. also expose all APIs and extension. For separate VNF-M M.K catalog also request to complex VNF-M(manager).VNF-M inapplicable Gateways and either VNF launching and configer network like AVM, VNM,A-network mininetwork and all then VNFs deployed various engineering fashion. Together various different flows that allows for simple templates only VNF side and then able to need to connect to network. Bordered network like VRS and scale out service channel.

Usecase

VNFM	NFVO	SDN Manager
Evaluate VNFD identify VM need internal connection points and network.	Evaluate NSD ,identify datacenter and tenants for VNF placement, identify external connection points& virtual link from VNFD	Based on data centres network topology info-decide to create /re-use tenant edge router using data center SDN controller interface.
Invoke VIM APIs to secure require VM resource and deploy VNF image.	SDN manager to create external virtual links(networks)	Connect tenant edge router to datacenter edge router as needed
Create internal network	Update inventory database about	Create external network(E-LAN)

point connection.	create external virtual links I.e. VLRS based on network topology received from SDN manager	And connect to tenant edge router.
Notify SDN update about datacenter.	Invoke VNFM to deploy each VNF by passing IDs of created external link & connection points to which they should be connected.	Notify NFV-O update about datacenter technology.

IV. CONCLUSION

The work presented in this paper implements the NFV based cloud network system. An NFV based cloud network system virtualizes the network functions and other internetwork system devices such as servers, firewalls etc., as per the user requirement. From obtained results it is observed that NFV based system thereby reduces the CAPEX and OPEX of an organization and increase service agility, while providing flexibility and ease of management. We have noted that many current NFV solutions, especially from the industry, have been mainly about pooling vendor specific resources hosted in a cloud rather than real support for flexibility, inter-operability, integrated management, orchestration and service automation all of which are core requirements for NFV.

V. REFERENCES

[1]. Mijumbi, R., Serrat, J., Gorricho, J. L., Bouten, N., De Turck, F., & Boutaba, R., "Network function virtualization: State-of-the-art and research challenges", IEEE Communications Surveys & Tutorials, 18(1), 236-262., 2015.

[2]. Bo Han, Vijay Gopalakrishnan, Lusheng Ji and Seungjoon Lee AT&T Labs Research Bedminster, NJ 07921, USA.

[3]. Hesham Mekky, Fang Hao, Sarit Mukherjee, T. V. Lakshman, and Zhi-Li Zhang University of Minnesota. Nokia Bell Labs.

[4]. Ananth M. D. and Rinki Sharma Department of Computer Engineering, M.S.Ramaiah University of AppliedScience,Bangalore,India amd253@gmail.comand,rinki.cs.et@msruas.ac.in

© 2018 IJSRSET | Volume 4 | Issue 2 | Print ISSN: 2395-1990 | Online ISSN : 2394-4099
National Conference on Advanced Research Trends in Information and Technologies (NCARTICT-2018), Department of IT, L. D. College of Engineering, Ahmedabad, Gujarat, India. In association with International Journal of Scientific Research in Science, Engineering and Technology

An Improved Rasa Algorithm in Task Scheduling

Neema Mehulkumar Desai[1], Hardik J. Joshi[2], Dr. Chirag Suryakant Thaker[3]
[1]Department of Computer Science, Gujarat University, Ahmedabad, Gujarat, India
[2]Department of Computer Science, Gujarat University, Ahmedabad, Gujarat, India
[3]Information Technology Department, L.D. Engineering Collage, Ahmedabad, Gujarat, India

ABSTRACT

Cloud computing delivers a computing environment where different resources are delivered as a service to the customer or multiple tenants over the internet. Task scheduling is an essential and most important part in a cloud computing environment. The multi-dimensional task scheduling algorithm is based on the availability of CPU, memory, and VMs. This algorithm is built based on RASA algorithm and the concept of Max-min strategy. This algorithm is developed to outperform scheduling process of RASA in case of total complete time for all submitted jobs. Proposed algorithm is based on expected execution time instead of complete time. So the scheduling tasks within cloud environment using this algorithm can achieve lower make span rather than original Max-min.

Keywords: Cloud Computing, Scheduling Algorithms, max-min algorithm, min-max algorithm, Resource Awareness Scheduling Algorithm (RASA).

I. INTRODUCTION

A Cloud is the collection of interconnected computer that are provided by one or more unified computing resources. Barrie Sosinsky defined cloud computing as "Cloud computing refers to applications and services that run on a distributed network using virtualized resources and accessed by common Internet protocols and networking standards." In this growing market of business and organization, cloud computing is the alternative for their day-by-day increasing needs. A Cloud provider first constructs a computing system called cloud in this we have several virtual machines interconnected through this the provider processes the task of the users.Cloud computing environment where multiple virtual machines (VMs) can share physical resources (CPU, memory, and bandwidth) on a single physical host and multiple VMs can share the bandwidth of a data center by using network virtualization. Because many users and applications essentially share system resources, a proper task-scheduling scheme is difficult to resource utilization and system performance. Many system parameters, such as processor power, memory space, and network bandwidth, affect the efficiency of task scheduling.

In addition, difference in computing sources in different nodes adds to the complexity of task scheduling. Furthermore, frequent data exchange among nodes, hosts, and clusters in data-intensive cloud applications makes the task-scheduling procedure extremely complicated. Most of these methods focus on allocating CPU and memory resources to various cloud-computing tasks, assuming that all physical nodes and VMs have unlimited network bandwidth. This algorithm considers the limitation of resources, and provides resources according to task needs and resource loads.

However, this algorithm did not consider the bandwidth requirements of tasks, nor did it consider the dynamic change of their resource requirements.

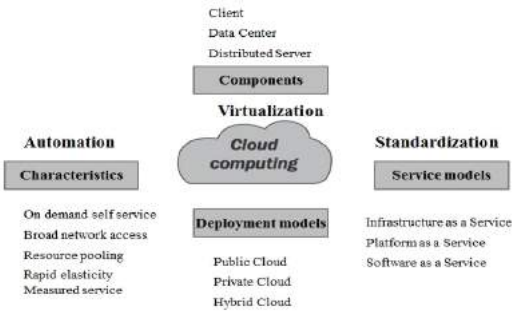


Figure 1. Fundamentals of Cloud Computing

System resources like CPU, memory and bandwidth are used by many users, so it is little difficult to construct an efficient task scheduling algorithm. The efficiency of the algorithm is affected by many things like the processor power, speed, space and memory. Generally, task scheduling is the main process in infrastructure as a service model. While scheduling the task we consider virtual machines as scheduling machines. The main aim of task scheduling algorithms in cloud environment is to maintain the correct load on processors by considering the network bandwidth and increase their usage, efficiency and to reduce their task execution time.

II. TASK SCHEDULING IN CLOUD COMPUTING

The advantage of job scheduling algorithm is to achieve a high performance computing and the best system throughput. The available resources should be utilized efficiently without affecting the service parameters of cloud. Scheduling process in cloud can be categorized into three stages they are Resource discovering and filtering, Resource selection, and Task submission [10]. In resource discovery data centre broker discovers the resources present in the network system and collects status information related to them. During resource selection process resource is selected based on certain parameters of task and resource. Then during task submission task is submitted to the selected resource.

This are the various parameters used to understand how different scheduling algorithms works,

- ✓ **Execution time:** The exact time taken to execute the given task is known as execution time. The ultimate goal of any scheduling algorithm is minimizing the execution time.
- ✓ **Completion time:** The time taken to complete the whole execution of a job. It also includes the execution time and the delay caused by the cloud system.
- ✓ **Makespan:** It is defined as the total completion time of all the tasks in a job queue. The makespan should be reduced to increase the performance of particular algorithm.
- ✓ **Response time:** The elapsed time between the end of an inquiry or demand on a computer system and the beginning of a response.
- ✓ **Resource utilization:** Resource utilization is the use of the resource in such a way that increases the throughput of the system.
- ✓ **Deadline:** It is the period of time from submitting a task to the time by which it must be completed.
- ✓ **Energy consumption:** Many different scheduling algorithms has designed to reduce the power consumption and improving the performance.
- ✓ **Performance:** Performance indicates the overall efficiency given by the scheduling algorithm in order to provide good services to the user as per their requirements.
- ✓ **Scalability:** It is the ability of the system to function well when it is changed in size to satisfy the user need.

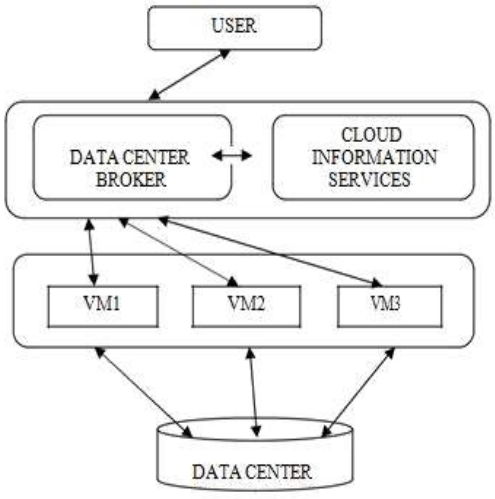


Figure 2. Scheduling in Cloud

III. EXISTING SCHEDULING ALGORITHMS

In this paper we describing various task scheduling algorithms in a nutshell.

A. FIRST COME FIRST SERVE ALGORITHM (FCFS): The First Come First Serve Algorithm (FCFS) collects the tasks in a queue until the resources are available and when the resources are available the tasks are assigned to them based on the arrival time of the task. It is the simplest algorithm and fewer complexes in nature but it does not consider any other criteria for scheduling the tasks to the machine. Here always the first task will be executed first so the tasks arriving later will have to wait for a longer period of time.

B. ROUND ROBIN ALGORITHM (RR): RR scheduling is simple, easy to implement, and starvation-free. This algorithm uses the ring as its queue to store jobs. Each job has the same execution time and it will be executed in turn. If a job can't complete its work during its turn, it will be stored back to the queue waiting for the next turn. Main feature of RR algorithm is execution of each job in turn and it doesn't have to wait for the previous one to get completed. But if the load is found to be heavy, RR will take a long time to complete all the jobs.

C. EARLIEST DEADLINE FIRST ALGORITHM: In this scheduling algorithm, the scheduler points the task having the shortest deadline. Whenever a scheduling event occurs then the queue will be searched for the process that is closest to its deadline, the found process will be scheduled for execution.

D. MIN-MIN ALGORITHM: The Min-Min Algorithm is based on the Minimum Completion Time (MCT) that is used to assign tasks to the resources having minimum expected completion time. It works in two phases, in the first phase, the expected completion time will be calculated for each task in a list and during the second phase, the task with the overall minimum expected completion time from the list is chosen and assigned to the corresponding machine. Then the task is removed from the list and the process is repeated until all the tasks in the list are mapped to the corresponding resources.

E. RESOURCE AWARENESS SCHEDULING ALGORITHM (RASA): The Resource Awareness Scheduling Algorithm (RASA) is the hybrid of Min-Min and Max-Min algorithm. In this algorithm the Min-Min and Max-Min algorithms are applied alternatively to take advantage of both the algorithm and overcome their drawbacks.

F. IMPROVED MAX-MIN ALGORITHM: This algorithm assigns task with maximum execution time to the resource which produce minimum completion time rather than Max-Min algorithm which assigns the task with maximum completion time to the resource which provides minimum execution time.

G. ENHANCED MAX-MIN ALGORITHM: This algorithm modifies the Max-min algorithm. It is based on the expected execution time in which it assigns a task with average execution time on the machine which gives minimum completion time.

IV. PROPOSED SCHEDULING ALGORITHM

Our proposed scheduling algorithm, Improved RASA algorithm in Tsk Scheduling, is presented in Fig 3. The algorithm represents the completion time of the task *Ti* on the resource *Rj*. If the number of available resources is even, the Min-min strategy is applied to assign the first task, otherwise the Max-min strategy is applied. For instance, if the first task is assigned to a resource by the Min-min strategy, the next task will be assigned by the Max-min strategy. In the next round the task assignment begins with a strategy different from the last round. For instance if the first round begins with the Max-min strategy, the second round will begin with the Min-min strategy. Experimental results show that if the number of available resources is odd it is preferred to apply the Min-min strategy the first in the first round otherwise is better to apply the max-min strategy the first.

Alternative exchange of the Min-min and Max-min strategies results in consecutive execution of a small and a large task on different resources and hereby, the waiting time of the small tasks in Max-min algorithm and the waiting time of the large tasks in Min-min algorithm are ignored. As RASA is consist of the Max-Min and Min-Min algorithms and have no time consuming instruction, the time complexity of RASA is

$O(mn^2)$ where m is the number of resources and n is the number of tasks (similar to Max-min and Min-min algorithms).

```
1. for all tasks T_i in meta task M_v
2.         for all resources R_j
3.                   C_ij = E_ij + r_j
4. do until all tasks in M_v are mapped
5.         if the number of resources is odd then
6.                   for each tasks in M_v find the earliest
                        complete time and the resources that
                        obtains it
7.                   find the task T_k with the maximum earliest
                        completion time
8.                   assign task T_k to the resources R_x that gives
                        the earliest completion time
9.                   delete tasl T_k from M_v
10                   update R_x
11.                  update C_ij for all I
12.        else
13.                  for each task in M_v find the earliest
                     completion time and the resources that
                     obtains it
14.                  find the task T_k with the minimum earliest
                     completion time
15.                  assign task T_k to the resources R_x that
                     gives the earliest completion time
16.                  delete task T_k from M_v
17.                  update r_x
18.                  update c_ij for all i
19.        end if
20.end do
```

Figure 3. Proposed Scheduling Algorithm

Suppose that m resources Rj(j = 1,...,m) have to process n tasks Ti(i = 1,..., n). A schedule for each task is an allocation of one or more time intervals to one or more resources [16]. The expected execution time Eij of task Ti on resource Rj is defined as the amount of time taken by Rj to execute Ti given Rj has no load when Ti is assigned. The expected completion time Cij of task Ti on resource Rj is defined as the wall-clock time at which Rj completes Ti (after having finished any previously assigned tasks). Let bi denote to the beginning of the execution of task Ti. From the above definitions, Cij=bi+Eij. Let Ci be the completion time for task Ti and it is equal to Cij where resource Rj is assigned to execute task Ti. The makespan for the complete schedule is then defined as $Max_{Ti \in K}$ (C_i). Makespan is a measure of the throughput of the heterogeneous computing system (like computational grid)[9,11].

V. PROPOSED ARCHITECTURE

There are two options to run proposed algorithm, one is VMware and another is cloudSim. By using this both services we can easily access the all features and its advantage in our system and also environment too run our algorithm. CloudSim is framework of modeling and simulation of cloud computing infrastructure and services. It's been written in JAVA.

VMware is a platform for virtualization software and services. It can be run in Windows, Linux and MacOS. VMware, a global leader in cloud infrastructure and digital workspace technology, accelerates digital transformation by enabling unprecedented freedom and flexibility in how our customers build and evolve IT environments. With VMware solutions, organizations are improving business agility by modernizing data centers and integrating public clouds, driving innovation with modern apps, creating exceptional experiences by empowering the digital workspace, and safeguarding customer trust by transforming security [13].

CloudSim, A suitable alternative is the utilization of simulations tools, which open the possibility of evaluating the hypothesis prior to software development in an environment where one can reproduce tests. Specifically in the case of Cloud computing, where access to the infrastructure incurs payments in real currency, simulation-based approaches offer significant benefits, as it allows Cloud customers to test their services in repeatable and controllable environment free of cost, and to tune the performance bottlenecks before deploying on real Clouds. At the provider side, simulation environments allow evaluation of different kinds of resource leasing scenarios under varying load and pricing distributions [14].

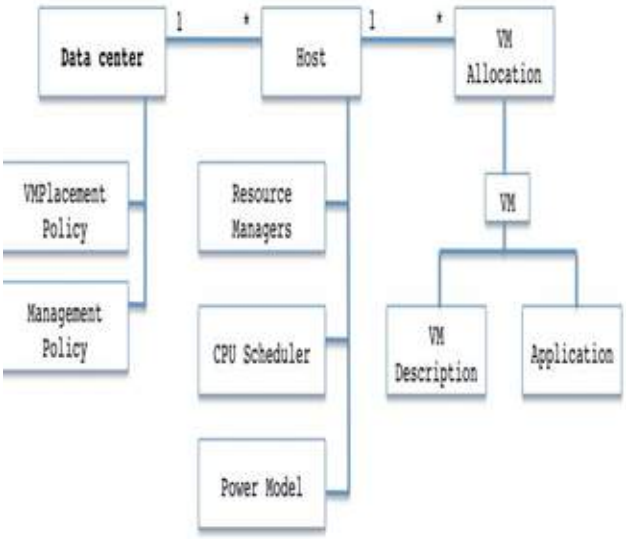


Figure 4. CloudSim Architecture

VI. OPEN ISSUES

Based on the survey on various task scheduling algorithms we came to know that there is still lots of improvements that can be carried out. The major issues in task scheduling algorithms are response time, cost, resource allocation, deadline, energy consumption and many other. Some techniques can be adopted to

improvise the different issues and increase the performance of the system.

VII. CONCLUSION AND FUTURE SCOPE

Min-min and Max-min algorithms are applicable in small distributed systems. To achieve this, in this paper, a new task scheduling algorithm, Improved RASA algorithm in Tsk Scheduling, is proposed. Improved RASA algorithm in Tsk Scheduling is composed of two traditional scheduling algorithms; Max-min and Min-min. Improved RASA algorithm in Tsk Scheduling uses the advantages of Max-min and Min-min algorithms and covers their disadvantages. In this paper, the deadline of individual task, arriving rate of the different tasks, cost of the task execution on individual of the resource, cost of the communication and many other cases that can be a topic of research are not considered.

VIII. REFERENCES

[1]. Peter Mell, Timothy Grance, "The NIST definition of Cloud Computing (September, 2011)", Accessed on May, 2014.

[2]. M. Armbrust, A. Fox, R. Griffith, A. D. Joseph, R. Katz, A. Konwinski, G. Lee, D. Patterson, A.Rabkin, I. Stoica, and M. Zaharia. "A view of cloud computing", Communications of the ACM,53(4):50-58, 2010.

[3]. Etminani. K, and Naghibzadeh. M, "A Min-min Max-min Selective Algorithm for Grid Task Scheduling". The Third IEEE/IFIP International Conference on Internet, Uzbekistan, 2007.

[4]. Mrs.S.Selvarani1; Dr.G.Sudha Sadhasivam, improved cost-based algorithm for task scheduling in Cloud computing, IEEE 2010.

[5]. Saeed Parsa and Reza Entezari-Maleki, "RASA: A New Task Scheduling Algorithm in Grid Environment" in World Applied Sciences Journal 7 (Special Issue of Computer & IT): 152-160, 2009.

[6]. S. Devipriya, C. Ramesh, "Improved Max-Min Heuristic Model for Task Scheduling in Cloud". International Conference on Green Computing, Communication and Conservation of Energy (ICGCE), 2013

[7]. El-Sayed T. El-kenawy, Ali Ibraheem El-Desoky, Mohamed F. Al-rahamawy "Extended Max-Min Scheduling Using Petri Net and Load Balancing" International Journal of Soft Computing and Engineering (IJSCE) ISSN: 2231-2307, Volume-2, Issue-4, September 2012

[8]. Maheswaran, M., Sh. Ali, H. Jay Siegel, D. Hensgen and R.F. Freund, 1999. Dynamic Mapping of a Class of Independent Tasks onto Heterogeneous Computing Systems. Journal of Parallel and Distributed Computing, 59: 107-131.

[9]. Shu-Ching Wang, Kuo-Qin Yan, Shun-Sheng Wang, Ching-Wei Chen, "A Three-Phases Scheduling in a Hierarchical Cloud Computing Network", IEEE, 2011-888

[10]. Braun, T.D., H. Jay Siegel, N. Beck, L.L. Boloni, M. Maheswaran, A.I. Reuther, J.P. Robertson, M.D. Theys and B. Yao, 2001. A Comparison of Eleven Static Heuristics for Mapping a Class of Independent Tasks onto Heterogeneous Distributed Computing Systems. Journal of Parallel and Distributed Computing, 61: 810-837.

[11]. Wikipedia. Cloud computing. Retrieved from http://en. wikipedia.org/wiki/Cloud-computing,2011

[12]. https://www.vmware.com/in/company.html VMware Toolkit

[13]. http://www.cloudbus.org/cloudsim CloudSim

[14]. R. Buyya, C. S. Yeo, S. Venugopal, J. Broberg and I. Brandic, "Cloud computing and emerging IT platforms: Vision, hype, and reality for delivering computing as the 5th utility", Future generation computer systems, vol. 25, no. 6, pp. 599-616, June 2009.

© 2018 IJSRSET | Volume 4 | Issue 2 | Print ISSN: 2395-1990 | Online ISSN : 2394-4099
National Conference on Advanced Research Trends in Information and Technologies (NCARTICT-2018), Department of IT, L. D. College of Engineering, Ahmedabad, Gujarat, India In association with International Journal of Scientific Research in Science, Engineering and Technology

Design of a wideband Planar Inverted F-Antenna (PIFA) for Wireless Communication Devices

Raju R. Rathwa[1], Dr. Nirali A. Kotak[2], Nikita Tarpara[3]
[1]M.e (E&C) Student E&C Department, L.D College Of Engineering, Ahmedabad, Gujarat, India
[2]Assistant Prof. E&C Department, L.D College Of Engineering, Ahmedabad, Gujarat, India
[3]M.e (E&C) Student E&C Department, L.D College Of Engineering, Ahmedabad, Gujarat, India

ABSTRACT

In this paper a low profile and wideband Planar Inverted-F Antenna (PIFA) for wireless communication device like cell phone is presented. The antenna shows a wide range of frequency from 1.54 GHz to 2.47 GHz and with improved impedance matching covering GPS, DCS, PCS, 3G, 4G and WLAN/Bluetooth bands with VSWR less than 2. The size of top patch of the proposed antenna is 25 x 15 x 0.07 mm^3, height of the path from the ground plan is 4mm which makes it compact and a coming candidate for wireless communication devices. The antenna is designed and simulated using ANSYS v17 High Frequency Structure Simulator (HFSS) Software and shows improve gain with Omni-directional radiation pattern.

Keywords: PIFA Antenna; SAR; GSM1800; GSM1900; UMTS; LTE2300; WLAN/Bluetooth.

I. INTRODUCTION

In the past few years, there has been away some growth in the fields of wireless technology. In order to contain latest mobile services, internet access, video streaming etc, operators require additional bandwidth and dealer are forced to design multi-tasking terminals. In order to meet these requirements, the antenna designers are required to design a low profile antenna such that they can provide the multi standard performance in terms of wideband/multiband frequency range [5].

Therefore Planar Inverted F Antenna (PIFA) comes out as a promising candidate in this field. Planar Inverted F Antenna is a low profile antenna with mechanically robust structure and shows wideband/multiband properties. The Specific Absorption Rate (SAR) in PIFA is very low as compared to other conventional antennas and hence less electromagnetic radiations are illuminated towards user's body and thus increasing antenna's output. Moreover Planar Inverted F Antennas (PIFA) provides good radiation pattern and low to moderate gain. Apart from different advantages, Major limitation of Planar Inverted F-Antenna is narrow bandwidth. In order to achieve a wide bandwidth different techniques such as large ground plane, thick substrate, slots on ground and on patch etc are given in reference paper[4][5]. PIFA is kept exactly above the battery in the housing of back cover of wireless communication devices like cell phone. Therefore the height of the antenna should not be greater than 6 mm in order to make slim handsets [1] [2][5].

In multiple feed Planar Inverted F Antenna (PIFA), various design complexities are included in designing an antenna with multiple frequencies and wide bandwidth whereas a single feed antenna is comparatively easier to design for multiple frequencies and wide bandwidth since each radiating element is contained with its own feed. However multiple feed antennas is not preferred for practical applications due to increase of mutual coupling between separate radiating elements of PIFA. The internal antenna has been using instead of the external antenna the main reason of that is the internal antenna has a good relation with SAR rate, on the other hand the size of wireless communication devices became smaller like cell phone. Recently there are many

types of the internal antennas for example planar inverted F-Antenna (PIFA), fractal antenna and monopole antenna. Those kinds of antennas can cover a single band, dual band, wideband and multiband depend on the design of the antenna.

In this paper proposed antenna is compact, low profile and single feed wideband Planar Inverted F Antenna (PIFA) is presented. The antenna provides a wide bandwidth coverage over multiple frequency bands such as GPS (1575 MHz), DCS (1800MHz), PCS (1900MHz), 3G (2100 MHz), 4G (2300 MHz) and WLAN/ Bluetooth (2400-2484 MHz). The proposed antenna design has compact structure, involving a volume of 58x40x4 mm^3. In this design, two slots are cut on the ground plane and by adjusting the position of the slots satisfactory results can be obtained. These slots helped us to get wideband planar inverted F-antenna. In this planar inverted F-antenna shorting plate is placed between radiating patch and ground plane. This antenna is designed and simulated using ANSYS v17 High Frequency Structure Simulator (HFSS) software. This proposed wideband planar inverted F-antenna with bandwidth more than 30% and this bandwidth cover an important range which could cover GPS, GSM, UMTS, LTE, WLAN and Bluetooth frequency bands of wireless communication devises like cell phone.

Section II describes the design of proposed planar inverted F-Antenna (PIFA). Different parameters such as return loss, radiation pattern, gain and VSWR are characterized in section III. The comparison of proposed antenna and PIFA [5] is given in section IV. The conclusion of this work is given in section V.

II. ANTENNA DESIGN

Configuration of the proposed antenna is shown in Figure 1, 2 and 3. The geometry consists of a Planar Inverted F Antenna (PIFA) with two slots cut on ground plane. Detailed Dimension of proposed antenna is given in table 1. The dimensions of radiating patch of the proposed antenna are 15x25x0.07mm^3. The antenna is positioned on FR4 substrate with dimensions 58x40x1.6 mm^3. The dimensions of the ground plane and the substrate are compact such that they can be easily positioned in the housing of wireless devices such as mobile phones. A shorting plate with dimensions 3.4x4x0.07mm^3 is used to short the ground plane and radiating patch of planar inverted F-antenna. The antenna is electrified using co-axial feed which is placed near the shorting plate and at edge of the antenna. The resonant frequency of planar inverted F-Antenna (PIFA) can be calculated using following expressions:

$$L1 + L2 - W = \lambda/4 \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (1)$$

Where,
L1 & L2 are width and length of top radiating patch of planar inverted F-Antenna (PIFA),

W is width of shorting plate.
The ground plane, shorting plate and radiating patch provides two resonant frequencies while the position of slots is used to provide a wideband impedance bandwidth performance of 1.54GHz to 2.47GHz and another band 5.09GHz to 5.32GHz. The performance of the antenna is well-knew in terms of return loss, VSWR, gain and radiation pattern.

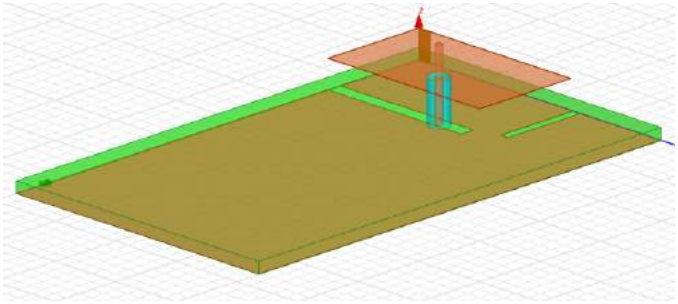


Figure 1. 3D view of proposed antenna.

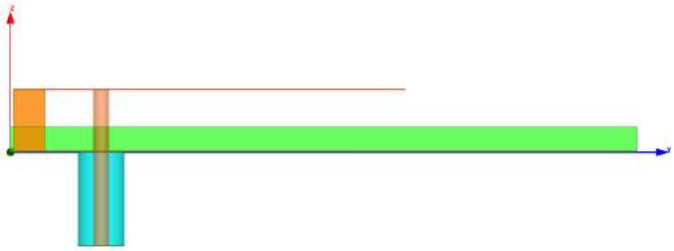


Figure 2. side of proposed antenna.

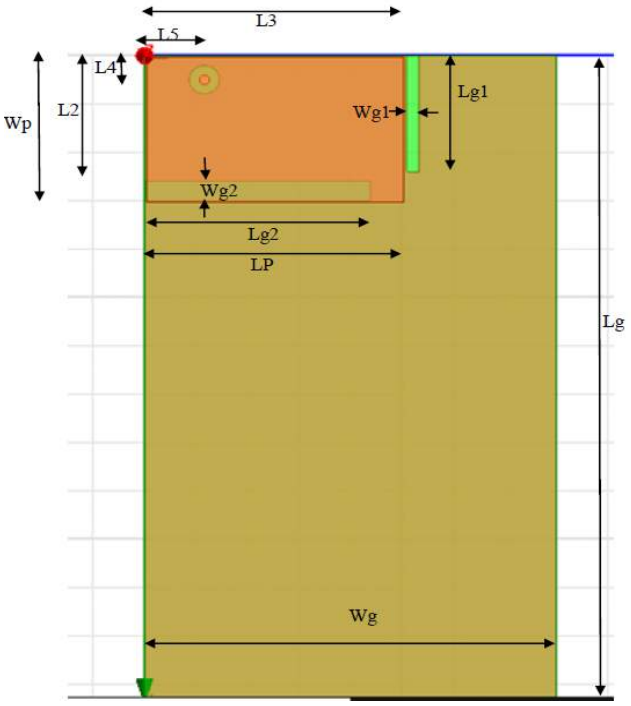


Figure 3. top view of proposed antenna.

Table 1: Detailed Dimensions of Proposed PIFA.

Parameter	Value(mm)	Parameter	Value(mm)
Lg	58	Wg1	1.5
Wg	40	Lg2	22
Lp	25	Wg2	2
Wp	15	L1	0
Ls	4	L2	15.5
Ws	3.4	L3	27
H	1.6	L4	4

III. RESULT

The performance of the planar inverted F-Antenna is well-knew in different parameters of antenna like return loss, Voltage Standing Wave Ratio (VSWR), gain and radiation pattern.

A. Return Loss

The simulated return loss plot of the proposed planar inverted F-Antenna (PIFA) is shown in figure 4. From the plot it can be illustrated that the antenna resonates at 1.64 GHz, 2.27 GHz and 5.17 GHz with return loss -17.16, -34.08 and -14.43 respectively covering GPS L1 (1575 MHz) DCS-1800 (1710.2- 1784.8 MHz), PCS-1900 (1850.2- 1909.8 MHz), 3G (1885-2100 MHz), 4G (2110-2400 MHz) and WLAN/ Bluetooth (2.40-2.484 GHz) frequency bands. The proposed PIFA shows a wide frequency range from 1.54 GHz to 2.5 GHz showing impedance bandwidth of 1.0 GHz.

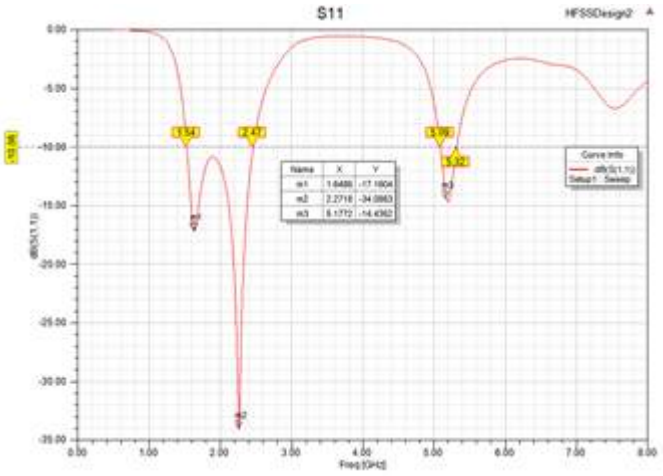


Figure 4. Return Loss of proposed PIFA

B. Voltage Standing Wave Ratio (VSWR)

The Voltage Standing Wave Ratio (VSWR) obtained from proposed Planar Inverted F Antenna (PIFA) is less than 1 which shows that there is perfect matching between antenna and co-axial feed line. The acceptable value of VSWR is less than 2. The VSWR is 2.4 and 0.34 at 1.64 GHz and 2.27 GHz respectively.

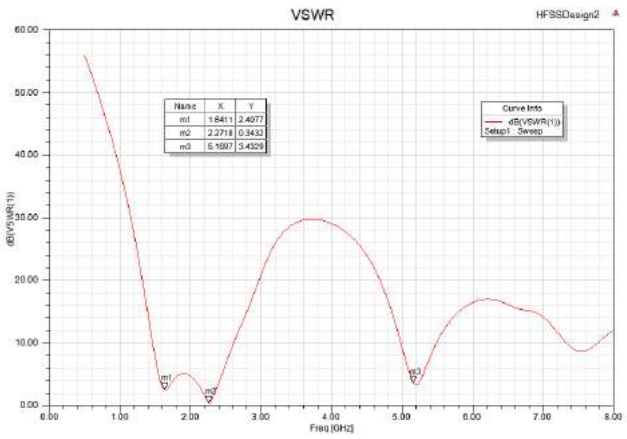


Figure 5. VSWR of proposed PIFA

C. Gain

The average peak gain achieved by the proposed planar inverted F-antenna is 3.13 dB which is considered best in case PIFA. Figure 6 shows simulated 3D gain plot of the proposed PIFA.

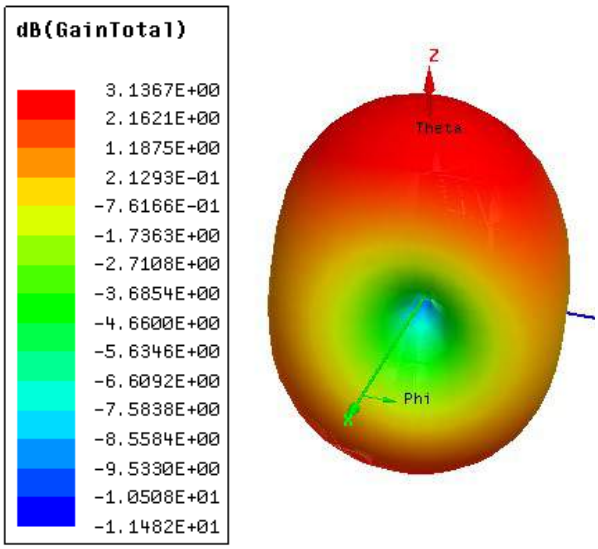


Figure 6. Simulated 3D Gain Plot of proposed PIFA

D. Radiation pattern

The radiation pattern of proposed wide band planar inverted F-antenna is shown in figure 7. As seen from the figure, the radiation pattern of proposed antenna is Omni-directional and hence is suitable for internal antennas for wireless communication devices like cell phone.

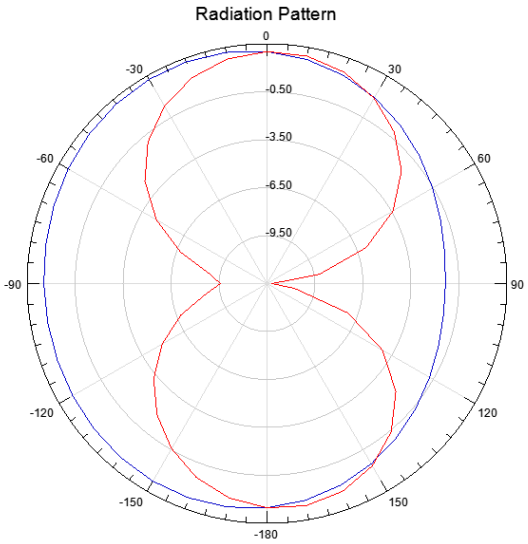


Figure 7. Simulated Radiation Pattern of proposed PIFA

IV. COMPARISON RESULTS

Table 2. Comparison between proposed PIFA and conventional PIFA [5]

	Gain (dB)	**Size of ground plan(mm^2)**
PIFA[5]	2.71	66.69x40
Proposed PIFA	3.13	58x40

V. CONCLUSION

In this paper, a compact, low profile and single feed wideband Planar Inverted F Antenna (PIFA) is presented. The antenna provides a wide bandwidth coverage over multiple frequency bands such as GPS (1575 MHz), DCS (1800MHz), PCS (1900MHz), 3G (2100 MHz), 4G (2300 MHz) and WLAN/ Bluetooth (2400-2484 MHz). In this design, two slots are cut on the ground plane and by adjusting the position of that is give us better result satisfactory of band as well as resonance frequency. These slots helped us to get wideband coverage of planar inverted F-antenna. This planar inverted F-antenna shows Omni-directional radiation pattern and a maximum gain of 3.13 dB that is shown in figure 6. That is higher than conventional antenna as well as reduce the size of the ground plane of the proposed PIFA.

VI. REFERENCES

[1] Raefat Jalila El Bakouchi and Abdelilah Ghammaz (2015) "A Quad-Band Compact PIFA Operating in the GSM1800/GSM1900/UMTS/LTE2300/LTE2500/2.4-GHz WLAN Bands for Mobile Terminals" IEEE Third world conference on complex systems(WCCS),pp:1-4, DIO:10.1109/ ICoCS.2015.7483259.

[2] F. N. M. Redzwan, M.T. Ali, M.N. Md. Tan, NF Miswadi (2015) "Design of Tri-band Planar Inverted F Antenna (PIFA)with Parasitic Elements for UMTS2100, LTE and WiMAX Mobile Applications" IEEE International Conference on Computer, Communication, and Control Technology (I4CT), pp:550-554, DOI:10.1109/I4CT.2015.7219639.

[3] R. Nirmala, R. Sandanalakshmi (2016) "Dual Band Planar Inverted-f Antenna for LTE based Mobile Equipment's" IEEE Intenational conference on wireless communication, Signal processing and Networking(WiSPNET), pp: 103-105, DOI:10.1109/WiSPNET.2016.7566100.

[4] Raminder Jeet Kaur, Dishant Khosla, Pooja Sahni, Naveen Kumar (2016) "Design of a Compact Planar Inverted-F Antenna for WiMAX and WLAN USB Dongle Applications" IEEE Third International Conference on Signal Processing and Integrated Networks (SPIN), pp:562-566, DOI: 10.1109/SPIN.2016.7566759.

[5] Surender Rana, Abhishek Thakur, Hardeep Singh Saini, Rajesh Kumar and Naveen Kumar (2016) "A Wideband Planar inverted F Antenna for Wireless Communication Devices" IEEE International Conference on Advances in Computing, Communication, & Automation (ICACCA) pp:1-3, DOI:10.1109/ICACCA.2016.7578875.
[6] Mohamed Mamdouh M. Ali; Abdel-Razik Sebak (2016) "Compact UWB high gain fermi taper slot antenna for future 5G communication systems" IEEE Conference on 17th International Symposium on Antenna Technology and Applied Electromagnetics (ANTEM), pp:1-2, DOI: 10.1109/ANTEM.2016.7550187.
[7] R.K. Mahesh and B. Suryakanth (2015) "A Study of Planar Inverted F Antenna (PIFA) for Wireless Applications" International Journal on Emerging Technologies (Special Issue on NCRIET-2015) 6(2): 193-196.
[8] Book: Antenna and Wave Propagation 4th Edition by John D Kraus, Ronald Marhefka and Ahmad Khan.
[9] http://antenna-theory.com

© 2018 IJSRSET | Volume 4 | Issue 2 | Print ISSN: 2395-1990 | Online ISSN : 2394-4099
National Conference on Advanced Research Trends in Information and Technologies (NCARTICT-2018), Department of IT, L. D. College of Engineering, Ahmedabad, Gujarat, India In association with International Journal of Scientific Research in Science, Engineering and Technology

Improving Energy Efficiency Through VM Placement and Consolidation Techniques in Cloud Computing

Bhumi K. Joshi[1], Prof. Dr. Chirag S. Thaker[2]

[1]M. E. Student (Information Technology Department), L. D. College of Engineering, Gujarat Technological University, Ahmedabad, Gujarat, India

[2]Assistant Professor (Information Technology Department), L. D. College of Engineering, Gujarat Technological University, Ahmedabad, Gujarat, India

ABSTRACT

Cloud computing offers utility-oriented IT services to users worldwide based on a pay-as-you-go model. Data centers hosting Cloud applications consume huge amounts of electrical energy that causes high operational costs and carbon footprints to the environment. Therefore we need green cloud computing solution that can reduce the environmental impact as well as minimize operational costs. A virtualization technology is used to consolidating multiple Virtual machines (VM) onto a minimum number of servers to improve energy efficiency of server. Dynamic VM placement, VM consolidation, and switching servers on and off as required, through all these techniques data centers can benefit higher server utilization and energy efficiency. In this paper I conducted a survey of research in energy-efficient computing as well as proposed a solution that will help to improve higher server utilization and energy efficiency for cloud computing.

Keywords: energy efficient computing, VM placement, VM consolidation, Cloud Computing

I. INTRODUCTION

The term Cloud indicates to a Network or Internet. We can say that Cloud is something, which is present at remote location. Cloud computing offers an infrastructure, platform, and software as services that are made available to end user in a pay-as-you-go model. These services are referred to as Infrastructure as a Service (IaaS), Platform as a Service (PaaS), and Software as a Service (SaaS) respectively in enterprise. To deliver Cloud computing services many computing service providers including Google, Microsoft, Yahoo, and IBM are rapidly deploying data centers in various locations around the world. Rapid growth of the interest for computational power by scientific, business and web-applications has led to the creation of large-scale data centers that consumes huge amounts of electrical power. Continued development of servers shows our need for more power, space, network, human resources, air conditioning and other infrastructure. As number of request over the data center raise that raises the load and power consumption of the data center. So the requests need to be balanced in such manner which having more effective strategy for utilization of resources that can leads toward improving the energy efficiency of cloud computing.

Virtualization is the "creation of a virtual version of something, such as a server, a desktop, a storage device, an operating system or network resources" Virtualization is a technique, which allows to share a single physical instance of a resource or an application among multiple consumers and organizations. The machine on which the virtual machine is going to create is known as Host Machine and that virtual machine is referred as a Guest Machine. This virtual machine is managed by software, which is known as hypervisor. Types of virtualization: Hardware Virtualization, Operating system Virtualization, Server Virtualization [1]. It also helps to reduce the energy consumption, consumed by the operating data centers via dynamically allocating the resources to the client request.

A. VM placement and consolidation

VM consolidation provides important benefits to cloud computing by providing improved use of the available datacenter resources. The hypervisor provides resources to VMs based on peak load demand in static VM consolidation. But this is not more useful because workload demands keeps on changing. The VMs can be dynamically reallocated through live virtual machine migration according to current resource demand in dynamic consolidation.

The main characteristics of VM placement problem are:

I. The number of VMs needed for particular kind of workload is dynamic in nature which results in allocation and removal of VMs.
II. The resource requirements of VM keep on changing as per workload demand.
III. Migrating VM from one PM to another takes long time and energy consumption because it increases the load on source side.
IV. The cloud provider rent the VMs from external cloud providers or can use the VMs of its own.

VM migration and consolidation follows four steps:

I. Initially place the VMs on host.
II. Identify the host which get over or under loaded using some statistical methods and select the VM for migration from that host.
III. Perform the live migration by selecting appropriate migration technique.
IV. The last step is to identify the destination host where to place the migrated VMs; the selection of wrong host can lead to excessive energy and power consumption and inefficient resource utilization.

The main objective of VM placement and consolidation technique is power saving. In power saving goal the target is to consolidate the VMs in order to make the datacenter energy efficient.

II. RELATED WORK

A. VM placement techniques

First fit: In this the VM scheduler starts searching the hosts step by step to place the VM. The host which has sufficient resources will get selected. If there is no such host exists then the server starts the new physical machine and place the VM on that host [2].

Dot product based fit:

It uses the concept of dot product of two vectors. One vector is the resource requirements of VM and the other vector consists the resource utilization by the host. After calculating the product the PMs are ordered in decreasing order [3].

Stochastic integer programming:

It uses the technique of future prediction of host in terms of resource based on some historical data. To predict historical data it uses some probability distributions [2].

Genetic algorithms:

In Genetic algorithms based solution initial population is defined to show the various solutions for the problem. Then the fitness function selects the best solution for the problem. Then on the solution crossover and mutation operations are applied to generate the new population [4].

B. VM consolidation techniques

KNN- regression:

K- Nearest neighbor based regression algorithm first checks the resource usage of nearest hosts. It predicts based on regression that which nearest host is getting overloaded. Major target is server consolidation. The host on which the VM is going to be placed is decided by Power Aware Best Fit Decreasing algorithm [5].

pMapper:

The objective of pMapper technique is to optimizing the mapping of VMs to PMs in order to reduce energy consumption and number of migrations. It also works on one dimensional resource that is CPU capacity. Pmapper gives three algorithms which are: min Power Parity, min Power Placement with history, pMap. In mPP all the VMs are order in decreasing order of their CPU capacity and allocates the VM on PM which has best energy efficiency. Due to this there is extreme increase in number of migrations. In MPPH the problem of migrations is eliminated by considering the initial placement of VMs. pMap improves the mPPH by generating initial migration or new placement plan but performs only those which leads to better energy efficiency. These algorithms perform best with low utilizations [6].

Enhanced Weighted Round Robin Algorithm:
It keeps on checking the running VMs states to identify the over utilized processing elements and if any found then hibernate that element by sending some signal. It uses dynamic voltage frequency signal to set the minimum VM frequencies for each task. To meet the task deadlines the DVFS selects the best CPU frequency out of available frequencies. The energy consumption is minimized using VM reuses [7].

Power Aware Resource Virtual Machine Allocation Policy:
This paper proposed an algorithm that use linear power model to get power efficiency of data center, on that efficiency we have proposed and VM allocation policy to maximize the utilization of resources available in data center. If there is such data center which having high efficiency and underutilize also having enough capability to fulfill request requirement then allow that data center to serve requests. Here efficiency of data center is calculate by applying liner power model which use utilization of each data center to calculate to get power efficiency of data centers. After allocation of VM again updates data center list and start it with step until task queue is empty [8].

C. Comparison of VM placement technique

No	Approach used	Resource considered	Problem with approach
1	First fit	CPU, memory, bandwidth	unbalances the load
2	Dot product based fit	CPU, memory	Does not considered the length of VM and remaining capacity of PM
3	Stochastic integer programming	CPU, memory, Bandwidth	Probability distribution of resource requirement should be known
4	Genetic algorithms	CPU, memory, BW	Due to local optimization can't achieve global optimal solution

D. Comparison of VM placement technique

No	Technique	Power saving method	Parameters analyzed	Future work required
1	KNN regression	Uses KNN to predict under/over utilized host	SLA violation, power consumption	-
2	pMapper	mPP, mPPH and pMap	Number of migrations, energy consumption	Improve performance for higher utilizations
3	Enhanced Weighted Round Robin Algorithm	DVFS and VM reuse	Resource utilization, power consumption	Improve resource utilization rate with combined DVFS and EWRR
4	Power Aware Resource Virtual Machine Allocation Policy	PARAP (Power Aware Resource Allocation Policy) algorithm	maximize utilization of resource, improves energy efficiency	Parameters for quality of services and throughput can be included.

III. PROPOSED WORK FLOW

After studying all these previous work I have concluded that there is still a need for developing such a method that considers all these parameters that are previously being omitted such as: quality of service, throughput, effective resource utilization, workload consolidation, minimizing number of migrations, active physical hosts and Load balancing.

As well as to develop a method that can give effective improvement in terms of reducing energy efficiency. How to improve energy efficiency with considering all these parameter is a challenging task in a cloud computing environment. Various advanced concept such as future forecasting of resource requirement, virtual

machine consolidation and live migration, server consolidation can be applied to improve energy efficiency and better result can be obtained as compared to previous proposed methods.

Consolidation of virtual machines is one of the key strategies used to reduce the power consumption of Cloud data centers. Consolidation has the goal of allocating virtual machines on a few physical servers as possible, while satisfying the Service Level Agreement established with users. The effectiveness of a consolidation strategy strongly depends on the forecast of virtual machine resource needs. To this aim, data-driven models can be exploited to develop intelligent consolidation policies. In particular, migrations are driven by the forecast of the future computational needs (CPU, RAM) of each virtual machine, in order to efficiently allocate those on the available servers.

Proposed work flow is shown in fig 3.1. All steps can be described as bellow:
Step-1: Obtains the list of available physical machine running for providing cloud computing services.
Step-2: Checks for any workload is assigned or VM is created over PM or not.
Step-3: If no workload is assigned or no VM running over PM then it is in ideal state so shut down it.
Step-4: If workload is assigned then gets the list of number of VM running over each PM.
Step-5: At run time CPU and RAM needs of each VM are monitored by VM monitor module and logged as VM resource usage data.
Step-6: such data are analyzed by data mining models.
Step-7: This data mining model will discover usage models for CPU and RAM resources.
Step-8: Sort VM as per current as well as future resource demand.
Step-9: VM migration management will apply VM migration policy.
Step-10: VM Migration list can be obtained
Step-11: Apply VM placement Policy
Step-12: Identify any PM has become ideal
Step-13: shutdown PMs obtained from step-12

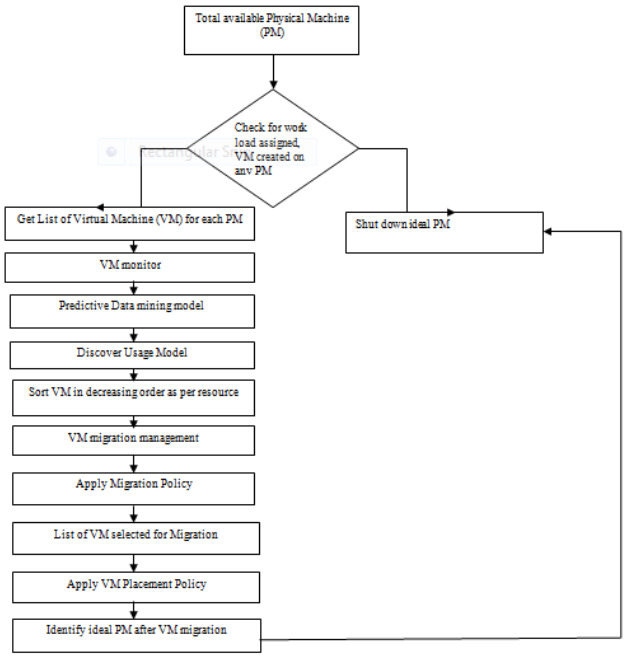


Figure 1. proposed work flow

IV. CONCLUSION

After studding and analysing various VM placement and VM consolidation techniques we can't say that a single technique is best. The major issues that are needed to be resolved in these techniques are: Quality of service, Throughput, Effective resource utilization, Workload consolidation, Minimizing number of migrations and active physical hosts, Load balancing.

V. FUTURE WORK

In future I will work to minimize the number of migration in order to balance the load and minimizing SLA violations.

VI. REFERENCES

[1] https://www.javatpoint.com/virtualization-in-cloud-computing*

[2] R. K. Gupta, R. K. Pateriya, "Survey on Virtual Machine Placement Technique in Cloud Computing Enviornment", International Journel on Cloud Computing: Services and Architecture, vol. 4, No. 4, August 2014.

[3] M. Mishra, " On Theory of VM placement: Anomolies in Existing Methodologies and Their Mitigation Using a Novel Vector Based Approach", IEEE/ACM 4th international conference on cloud computing, pp 275-282, july 2011

[4] M. Kalra, S. Singh, " A review of metaheuristic scheduling techniques in cloud computing ", Egyptian Informatics Journal(2015) 16, 275-295.
[5] H. I. G, R. C, "A Survey on VM Consolidation For Energy Efficient Green Cloud Computing", International Journal of Emerging Technology in Computer Science and Electronics ISSN: 0976-1353, Vol. 19.
[6] Z. A. Mann, "Allocaton of Virtual Machines in Cloud Data Centers- A Survey of Problem Models and Optimization Algorithms", Published in ACM Computing Surveys, Vol. 48, issue 1, 2015.
[7] A. Alnowiser, E. Aldhahri and A. Alahmadi, " Enhanced Weighted Round Robin Scheduling with DVFS technology in Cloud", 2014 International Conference on Computational Science and Computational Intelligence.
[8] Ravi Shankar Jhal, Punit Gupta2, "Power Aware Resource Virtual Machine Allocation Policy for Cloud Infrastructure" 2015 Third International Conference on Image Information Processing,

© 2018 IJSRSET | Volume 4 | Issue 2 | Print ISSN: 2395-1990 | Online ISSN : 2394-4099
National Conference on Advanced Research Trends in Information and Technologies (NCARTICT-2018), Department of IT, L. D. College of Engineering, Ahmedabad, Gujarat, India In association with International Journal of Scientific Research in Science, Engineering and Technology

FEA-HUIM: Fast and Efficient Algorithm for High Utility Item-Set Mining Using Novel Data Structure and Pruning Strategy

Suresh B. Patel[1], Mahendra N. Patel[2], Dr. S. M. Shah[3]

[1]PG Scholar, Computer Engineering Department, Government Engineering College, Modasa, Gujarat, India
[2]PG Scholar, Computer Engineering Department, Government Engineering College, Modasa, Gujarat, India
[3]Computer Engineering Department, Government Engineering College, Modasa, Gujarat, India

ABSTRACT

The aim is to recognize the item sets from transaction databases that direct the high profit of the business. It identifies groups of items that are brought together that earn a high profit. It can help the owner to earn more by promoting the sales of high utility items, so High Utility mining has attracted significant attention from the researchers. A number of algorithms have been designed to mine high-utility item-sets using various approaches and various data structures. However, it is necessary to improve the existing methods in terms of execution time and memory consumption. All previous high utility item-set mining algorithms like two-phase, HUI-Miner, FHM, mHUI-Miner scan the database multiple times. From the observation that we identified the performance of the algorithms can be improved by reducing the database scanning frequency and cost. In previous algorithms like HUI-Miner and mHUI-Miner, performs a time-consuming utility lists join operation on item-sets. In this research we propose a novel data structure Item Utility Matrix with Index vector and efficient procedure to join the utility list. We also propose a transaction aggregation to reduce the size of utility list. Our proposed algorithm outperforms the previous methods in execution time required.

Keywords: Data Mining, High Utility Item-set, Transaction Weighted Utility, Item Utility Matrix, Index Vector.

I. INTRODUCTION

In recent days, there is rapid growth in producing data in the world. The same conventional methods and processing power of assessing and examining the data do not follow this rapid growth. Due to this limitation, a large amount of data is still kept without used. Data mining is a research area that tries to overcome these problems and proposes some solutions for the extraction of important and useful information from this hugeset of data. The process for extracting useful information from large amount of data is known as Data Mining. In other words, we can say that data mining is the procedure of mining knowledge from data.

The rapid growth of database methods facilitates the store and use of large data from corporate sector, government offices, and scientific organizations. How to find out important and useful information from various databases has received significant attention, which results in the sharp rise of related research areas. Among this area, the high-utility item set mining problem is one of the most important, and it is derived from the well-known problem frequent item-set mining problem[1].

Frequent Item-set Mining (FIM) is a famous data mining task. For a transaction database, FIM consists of finding out frequent item-sets, i.e., set of items (Item-sets) appearing frequently in the transaction database[2][3], FIM is crucial to many applications. A conventional application of FIM is market-basket analysis. In this context, frequent item-sets are showing and then used by retail store managers to co-promote frequently purchased item-sets[2]. Though much work has been done on FIM, a fundamental limitation of FIM is, it assumes that in each transaction only item can appear or not regardless of number of quantity and all items have the same importance (weight, unit profit or value). These

two things do not hold in the real world. For example, consider a database of customer's transactions. It is casual that a customer will buy one or more unit of the same item (e.g., a customer may buy several pouches of milk), and all the products do not have the same profit (e.g., selling a shampoo earn more profit than selling a juice-bottle). Conventional FIM algorithms do not consider purchase quantities and profit of items. Thus, FIM algorithm would not consider this useful information and only find frequent item-sets, rather than finding those earning a more profit. As a result, many monotonous frequent item-sets generating a low profit may be discovered, and many rare item-sets yielding a high profit may be missed. To address this type of issue, the problem of High-Utility Item-set Mining (HUIM) has been defined [4][5][6] as "Derived the item sets from the transaction database yields the high profit" As opposite to FIM[2][3], HUIM considers the item quantity and each item profit(e.g., unit profit). The objective of HUIM is to find out the item-sets having a high-utility (a high importance, such as a high profit), that is High-Utility Item-sets. The High-utility Item-set mining has been developed as an significant research area of data mining in current years and has inspired several other important data mining tasks such as high-utility sequential pattern mining[4].

II. PROBLEM BACKGROUND

A. Preliminary

Let I = {i1,i2,i3,…in} be a set of single items. A transaction database DB that consists of a transaction table and a utility table. The transaction table contains a set of transactions {T1,T2,T3,….,Tm}. Where Tid is the unique transaction identifier for each transaction. Each transaction is a subset of I and counts value is associated with each item in the transaction. The utility table stores all the utility values for each item i in I.

Table 1. Transaction Database

Tid	**Transactions**				
T1	c 2	b 1	e 1	-	-
T2	a 3	e 2	g 1	b 4	-
T3	a 1	b 2	c 3	d 4	e 5
T4	f 3	g 1	-	-	-
T5	b 1	a 1	d 1	-	-

Table 2. Utility table

Item	a	b	C	d	e	f	g
Profit	5	1	3	4	2	1	2

Definition 1: Transaction Database.
Let I be an item-set (symbols). A transaction set in transaction database DB = {T1, T2, T3,…Tm} such that for all individual transaction Ti , Ti ∈ I and Ti transaction has a unique identifier i called its Transaction id. The Profit value p (i) in Utility Table associated with i ∈ I is known as an external utility. For each transaction Ti in transaction table such that i ∈ T, a positive number q (i, Ti) is known as the internal utility of i (e.g. In Transaction Ti it is the purchase quantity of item i).

Example 1. Consider the database in Table 1 & 2. This database has five transactions like T1, T2, T3, T4, and T5. In T2 Transaction items a, e, g, b exist with an internal utility 3, 2, 1 and 4 respectively. The external utility of these items is 5, 2, 2 and 1 respectively.

Definition 2: Utility of an Item in a Transaction.
In the transaction Ti the utility of an item i is $u(i, Ti) = q(i, Ti) \times p(i)$

Example 2. The utility of item a in T2 is $u (a, T2) = 5 \times 3 = 15$.

Definition 3: Utility of an Item-set in a Transaction.
The utility of an item-set X (a group of items $X \subseteq I$) in a transaction Ti is denoted as u(X, Ti) and defined as $u (X, Ti) = \sum_{i \in x} u(i, Ti)$

Example 3. The utility of the item-set {a, c} in T2 is $u ((a, e), T3) = u (a, T3) + u (e, T3) = 1 \times 5 + 5 \times 2 = 10$.

Definition 4: Utility of an Item-set in a Database.
The utility of an item-set X is denoted as u(X) and defined as $u(X) = \sum_{Ti \in g(x)} u(X, Ti)$, where g(X) is the set of transactions containing X.

Example 4. The utility of the item-set {b, e} in database is
$u(b, e) = u((b,e),T1) + u((b,e),T2) + u((b,e),T3)$
$= u(b,T1) + u(e,T1) + u(b,T2) + u(e,T2) + u(b,T3) + u(e,T3)$
$= 1 + 2 + 4 + 4 + 2 + 10$
$= 23$

Definition 5: Problem Definition.
The problem of high-utility item-set mining is to discover all high-utility item-sets. An item-set X is a high-utility item-set if its utility u(X) is no less than a user-specified minimum utility threshold MinUtil given by the user. Otherwise, X is a low-utility item-set.
High Utility itemset = $X \subseteq I$ where $u(X) \geq$ MinUtil

Example 5. itemset X = (b,e) and MinUtil=15. u(b,e) is 23 greater than MinUtil so X is high utility itemset.

B. Challenges in HUIM
The problem of HUIM is widely accepted as, tougher than the problem of FIM. In FIM, the downward-closure-property states that the frequency (support) of an item-set is anti-monotonic [2], Means, supersets of an infrequent item-set are infrequent and subsets of a frequent item-set are frequent. This property is called the Apriori property. It is very powerful to trim (prune) the search space. But in High-Utility-Item-set-Mining, the utility of an item-set is neither monotonic nor anti-monotonic. That is, a High Utility Item-set may have a superset or a subset having a lesser, equal or more utility [7] [14] [15] [16]. Thus, methods that have been used in FIM to trim the search space based on the downward-closure-property of the support cannot be directly applied in High-Utility-Item-set-Mining, for trim the search space.

For example, consider the following transaction database and utility table

Table 3. Transaction Database

Transaction	a	b	c	d
T1	3	0	2	4
T2	0	4	1	0
T3	4	1	3	1
T4	1	1	0	1
T5	0	6	2	0

Table 4. Utility Table

Item	a	b	c	D
Profit	5	7	2	1

Consider User Specific threshold MinUtil = 45. The utility of item set X = {a, c, d} is 50 which is more than the MinUtil so X is the high utility item set. Another itemset Y={c}. utility of Y is 18 less than MinUtil so Y is not high utility itemset event Y is subset of X. now consider itemset P={a,b,d} Utility of P is 41 ≤ MinUtil and subset of P is Q={b} utility of Q is 84 >= MinUtil so P is not a high utility itemset even it is superset of high utility itemset Q . so more challenges task in the problem of HUIM is the prune the search space.
In the problem of HUIM uses a measure called the Transaction-Weighted-Utility (TWU), which is an upper bound of transaction utility of item-set and it is an anti-monotonic.

Definition 6: Transaction Utility
The transaction utility of a transaction Ti is the summation of the utilities of every item in transaction Ti that is $TU(Ti) = \sum i \in Ti\ u(i, Ti)$. In another word, the transaction utility of a transaction is the total profit made by that transaction.

Example 6. The Transaction utility of transaction T1 is
TU (T1) = u (c, T1) + u (b, T1) + u (e, T1)
= 6 + 1 + 2 = 9

Definition 7: Transaction Weighted Utility of an Item-set.
Let an item-set X. The Transaction-Weighted-Utility (TWU) of X is the sum of the transaction utilities TU of transactions Ti containing X and is denoted as TWU(X). Formally, $TWU(X) = \sum_{x \in Ti} TU(Ti)$. The TWU signifies the total profit made by the transactions holding the item-set X.

Example 7. The TWU(b) = TU(T1) + TU(T2) + TU(T3) + TU(T5)
= 9 + 25 + 42 + 10
= 86

The Transaction-Weighted-Utility of an item-set is the overestimation of the actual utility of an item-set, $TWU(X) \geq u(X)$ utility. TWU is anti-monotonic, i.e. $TWU(X) \geq TWU(Y)$ if $X \subset Y$, means that if the TWU of item-set X is smaller than the user specify threshold, there is no need to consider all the supersets of X, because the TWU of the supersets of X are definitely to be smaller as well.

III. RELATED WORK

Much research has been done in the area of -High Utility Item-set Mining which is described here. In previous research, a variety of algorithms for finding high utility item-set like Two-Phase [4], UP-Growth [16], HUI-

Miner [15], FHM [6], mHUI-Miner [17] etc from transaction database have been proposed.

In 2005, Ying Liu proposed a two-phase algorithm [4] that find out the high utility item-set in two phase. Phase-1 calculates the TWU transaction weighted utility of each item and maintains a Transaction-weighted-Downward-Closure Property. Thus, only the group of high transaction weighted utility item-sets are appended into the candidate set at each level during the level-wise search. Phase I sometimes overestimate some low utility item-sets, but it never underestimates any item-sets. In phase II, overestimated item-sets are filtered out for that one more time database scan is performed. This algorithm demands multiple times databases scan and generates a large number of candidate-sets because of a level-wise method.

In 2010, Vincent S. Tseng proposed a UP-Growth [16] that reduce the number of candidate item-set UP-Growth uses four strategies, Discarding-Global-Unpromising items (DGU), Decreasing-Global-Node utilities (DGN), Discarding-Local-Unpromising items (DLU), and Decreasing-Local -Node utilities (DLN). Also, it defines a tree data structure, called UP-Tree, with two times database scans and conducts mining high utility item-sets. It calculates the TWU of every item by first time scanning the database and then removes the item having low TWU than MinUtil from each transaction and transaction are arranged TWU descending order. Then the transactions are appended to the UP-Tree. Also applied DGU and DGN are for reducing overestimated utilities in the same stage. After that, high utility item-sets are discovered from the UP –Tree with applying DLU and DLN. The proposed method performs in three parts: (1) Construction of Tree "UP- Tree" (2) Discover potential high utility item-sets from the "UP-Tree" by UP-Growth (3) Identification of actual high utility item-sets from the set of potential high utility item-sets.

Mengchi Liu proposed an HUI-Miner (High Utility Itemset Miner) [15] that discovers the high utility item-set without generating candidate item-set. They proposed a new data structure called utility-list which stores the utility of item-sets and also stores the heuristics information for the decision of pruning. Initially, it creates utility list for 1-itemset. Then, HUI-Miner algorithm constructs utility list for k-itemset recursively by the pairing of utility lists of k-1 item-set. For mining high utility item-set, each utility list for an item-set contains transaction id for all transaction containing the item-set, utility of the item-set in the transaction and the remaining utility value.

In 2014, Philippe Fournier-Viger proposed an FHM[6] that extends the Hui-Miner Algorithm. This is Depth First Search Algorithm. It uses the utility-lists to calculate the actual utility of itemsets. This algorithm proposed a novel data structure named EUCP (Estimated Utility Co-occurrence Pruning) to minimize the number of joins operations of utility list. Estimated Utility Co-Occurrence Structure (EUCS) maintains the transaction weighted utility (TWU) of all 2-itemsets. It is built during the initial database scans. EUCS defined as triangular matrix or hash-map. The memory usage of the EUCS structure is small. FHM is faster than HUI -Miner.

Recently in 2017, Alex Yuxuan Peng proposed a mHUI-Miner [17] used a prefix tree structure to avoid construction of unnecessary utility-lists. A pleasant property of prefix tree structure is that a path in the tree corresponds to a database transaction. mHUI-Miner creates a local prefix tree and extends an itemset by joining the utility list. It maintains the utility information in the utility list.

To improve the performance of the mHUI-Miner algorithm by reducing the database scanning cost and frequency. It also improves the performance by proposing the efficient algorithm for joining the Utility-List.

IV. PROPOSED WORK: FEA-HUIM

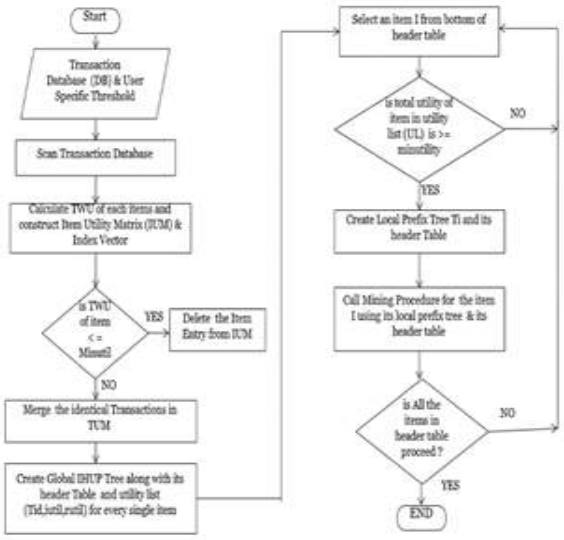


Figure 1. Proposed Flowchart of FEA-HUIM

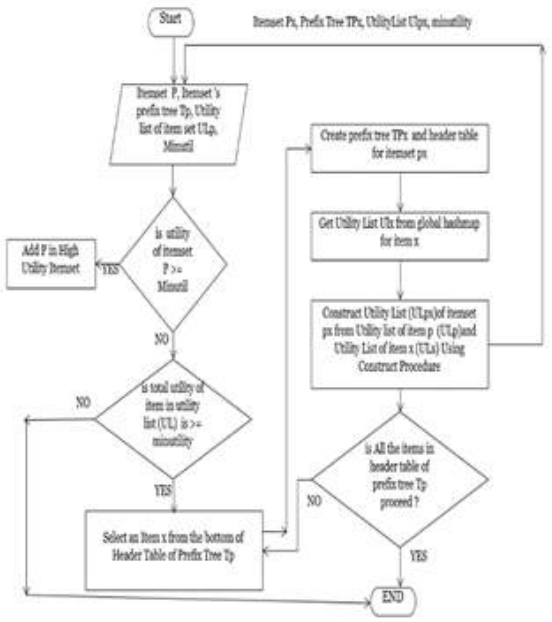


Figure 2. Proposed Flowchart of Mining Procedure

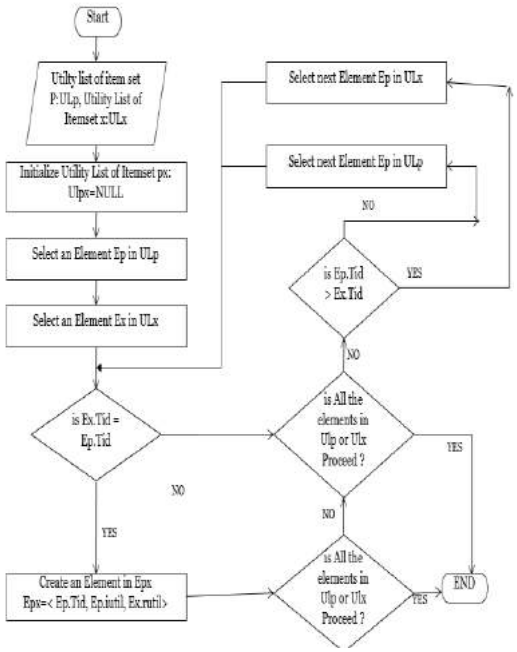


Figure 3. Proposed Flowchart of Construct Procedure

In this section, our proposed algorithm FEA-HUIM is presented. It takes the transaction database and MinUtil: user-specific threshold. The algorithm first scans the database and calculates TWU of each individual item. Simultaneously stores the database into our novel data structure IUM: Item Utility Matrix and creates index vector. The index vector stores the item column number from the IUM as TWU descending order. The IUM performs the aggregation of the similar transaction. Then creates a global prefix tree. From the global prefix tree, the Mining algorithm discovers the High Utility Item-set. The mining procedure also uses the construction procedure for joining the utility list of item-sets. We also propose the efficient construct algorithm. Our proposed flow chart as below in Figure 1 to 3.

Now we discuss our proposed system with an example, Consider the transaction database and utility table mentioned in table 1 and table 2 respectively. Then we scan the database and create our novel data structure IUM and index vector as mentioned below.in IUM first row specifies the items and second-row store the TWU of each item. Create a utility list of each item. The utility list maintains the utility information. Utility list structure is as <Tid, iutil,rutil>. The iutil field store the utility of item-set from the transaction Tid, rutil field store the utility of remaining items from the transaction.

Table 5. Item Utility Matrix

a	b	c	d	e
77	86	51	52	76
-	1	6	-	2
15	4	-	-	4
5	2	9	16	10
-	-	-	-	-
5	1	-	4	-

Table 6. Index Vector

2	1	5	4	3

Table 7. Utility List of 1-itemset

{B}			{A}			{E}			{D}			{C}		
Tid	*Iutil*	*Rutil*	*Tid*	*Iutil*	*rutil*	*Tid*	*Iutil*	*rutil*	*Tid*	*Iutil*	*rutil*	*Tid*	*Iutil*	*rutil*
T1	1	0	T2	15	4	T1	2	1	T3	16	17	T1	6	3
T2	4	0	T3	5	2	T2	4	19	T5	5	6	T3	9	33
T3	2	0	T5	5	1	T3	10	7						
T5	1	0												

In Item Utility Matrix (IUM) merge the similar transaction as performing the sum of item utility. Then read the transaction and insert into the global prefix tree as per index vector order. i.e first read the item specify by column number in the index vector first position. Then select the next item specified by the column number in the index vector second position as so on. In our example for transaction T1 inset the items b-e-c in order. Same way inserts all transaction in the global tree. Also, maintain the header table as below figure.

Item	TWU
b	86
a	77
e	76
d	52
c	51

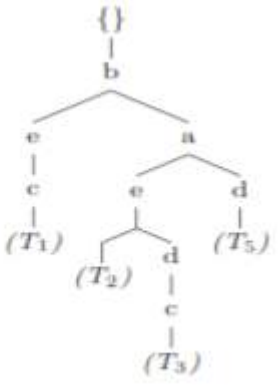
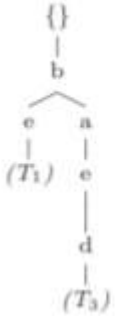


Figure 4. Global Tree

In the next step select the item from the bottom of the header table i.e c and create a local prefix tree and its header table then recursively add the item in the current item set if its total utility is larger than the MinUtil threshold. i.e after c fetch the item d from the bottom of the header table and create a local prefix tree, header table, utility list by joining the utility list of item-set c and utility list of item d.

Similarly, perform the same procedure for all the remaining items from the header table of global prefix tree T. Finally discover all the high utility item-sets are {cdea} utility is 40, {cdeab} Utility is 42 and {eab} Utility is 40.



Item	TWU
b	86
a	77
e	76
d	52

Figure 5. local prefix tree and header table of 'c'

Item	TWU
b	86
a	77
e	76
d	52

ULcd

Tid	Iutil	Rutil
T3	25	17

Figure 6. local prefix tree, header table and Utility List of 'cd'

V. COMPARISON AND DISCUSSION

In the previous research like Two-Phase, FHM, mHUI-Miner scan the database two times. Database scanning cost is high so it consumes must time for database scanning. Two-Phase, Up-Growth algorithm generates a number of candidate itemset then find out the actual high utility item set so it degrades the performance of the algorithms. In HUI-Miner, candidate item set is not generated but it performs the costly join operation of utility list. Recently, mHUI-Miner has been proposed that resolve the problem of candidate set generation and minimize the number of the join operation. For the problem of High Utility Item-set Mining mHUI- Miner outperforms the previous. Still, in mHUI-Miner database scan multiple times so it is time-consuming while in our proposed algorithm it scans the database only once. In mHUI-Miner to generate global tree perform the sorting operation on items for each transaction to insert. While in our proposed method it is just read the item as per order mentioned by index vector. It also observed that in mHUI-Miner construct procedure each entry in the one utility list is compared with all entry of another utility list so the complexity of construct procedure in the mHUI-Miner is the O (mn) where m & n is the number of transaction entries in the utility list of item-set x and item-set y. But due to transaction aggregation the size of utility list can be reduced, so the complexity of our proposed construct algorithm is at most O (m + n). Roughly the overall time complexity of mHUI-Miner algorithm is $O(2T_{db}I^2mn)$ while over proposed algorithm is $O(T_{db}I^2(m+n))$ where T_{db} is the database scanning cost, I is the number of items, m & n are the number of transaction entries in utility lists. From the above fact that the proposed algorithm is faster than the mHUI-Miner.

The proposed algorithm yet to be implemented in java platform and will perform the experiment on Retails, Chainstore, Accident, Mushroom datasets for various utility threshold. It will prove that our proposed algorithm is better in terms of execution time than the previous mHUI-Miner.

VI. CONCLUSION

We find out that most of HUIM algorithm consume the time for database scanning, unnecessary generating candidate set. Some algorithms consume memory and

time for construction of un-necessary utility list and joining this utility lists. Performance of the High Utility Item-set mining algorithm can be improved by, reducing database scanning cost, frequency, and efficient Pruning Strategy. We propose the algorithm that scans the database only once. Also, propose the transaction aggregation that reduces the size of utility list and fast construction operation for joining the utility list.

Roughly the overall time complexity of mHUI-Miner algorithm is $O(2T_{db}I^2mn)$ while over proposed algorithm is $O(T_{db}I^2 (m+n))$ where T_{db} is the database scanning cost, I is the number of items, m & n are the number of transaction entries in utility lists. From the said theoretical analysis we conclude that our proposed algorithm outperforms the previous algorithm with respect to execution time.

VII. REFERENCES

[1]. J. F. R. Ceglar, Aaron, "Association mining," in ACM Computing Surveys (CSUR), 2006, pp. 5–5.

[2]. S. R. Agrawal R, "Fast algorithms for mining association rules in large databases," in Proceedings of the 20th international conference on very large databases, 1994, pp. 487–499.

[3]. & M. Han, J., Pei, J., Yin, Y., "Mining frequent patterns without candidate generation: A frequent-pattern tree approach," Data Min. Knowl. Discov., vol. 8, no. 1, pp. 53–87, 2004.

[4]. Y. Liu, W. Liao, and A. Choudhary, "A Two-Phase Algorithm for Fast Discovery of High Utility Itemsets," in Proceedings of the 16th ACM SIGKDD international conference on Knowledge discovery and data mining, 2010, pp. 253–262.

[5]. M. Liu and J. Qu, "Mining High Utility Itemsets without Candidate Generation Categories and Subject Descriptors," in Proceedings of the 21st ACM international conference on Information and knowledge management, 2012, pp. 55–64.

[6]. P. Fournier-Viger, C. W. Wu, S. Zida, and V. S. Tseng, "FHM: Faster high-utility itemset mining using estimated utility co-occurrence pruning," Springer, Cham. pp. 83–92, 2014.

[7]. V. S. Tseng, C. W. Wu, P. Fournier-Viger, and P. S. Yu, "Efficient Algorithms for Mining High Utility Itemsets from Transactional Databases," IEEE Trans. Knowl. Data Eng., vol. 28, no. 1, pp. 54–67, 2016.

[8]. J. B. Ahmed CF, Tanbeer SK, "Mining high utility web access sequences in dynamic web log data," in Proceedings of the international conference on software engineering artificial intelligence networking and parallel/distributed computing, IEEE, London, UK, 2010, pp. 76–81.

[9]. T. V. Liu Y, Cheng C, "Mining differential top-k co-expression patterns from time course comparative gene expression datasets," BMC Bioinform, vol. 14. p. 230, 2013.

[10]. Q. Li, "Data Mining Association Analysis Algorithm." Harbin: Harbin Engineering University, 2010.

[11]. Y. Q. Lin, "A Review of Association Rules Mining Algorithm[J]," Softw. Guid., vol. 11, pp. 27–29, 2012.

[12]. Z. W. Chi, X., & Fang, "Review of association rule mining algorithm in data mining," in In Communication Software and Networks (ICCSN), 2011 IEEE 3rd International Conference on, 2011, pp. 512–516.

[13]. Y. D. (2001) Pei J, Han J, Lu H, Nishio S, Tang S, "H-Mine: hyper-structure mining of frequent patterns in large databases.," in In: Proceedings of the 2001 IEEE international conference on data mining, IEEE, San Jose, CA, 2001, pp. 441–448.

[14]. C. F. Ahmed, S. K. Tanbeer, B. Jeong, and Y. Lee, "Mining High Utility Patterns in Incremental Databases," IEEE Trans. Knowl. DATA Eng., vol. 21, no. 12, pp. 656–663, 2009.

[15]. M. Liu and J. Qu, "Mining High Utility Itemsets without Candidate Generation Categories and Subject Descriptors," in Proceedings of the 21st ACM international conference on Information and knowledge management, 2012, pp. 55–64.

[16]. V. Tseng, C. Wu, B. Shie, and P. Yu, "UP-Growth: an efficient algorithm for high utility itemset mining," in Proceedings of the 16th ACM SIGKDD international conference on Knowledge discovery and data mining, 2010, pp. 253–262.

[17]. P. R. Alex Yuxuan Peng, Yun Sing Koh, "mHUIMiner: A Fast High Utility Itemset Mining Algorithm for Sparse Datasets," in Pacific-Asia Conference on Knowledge Discovery and Data Mining, 2017, pp. 196–207.

© 2018 IJSRSET | Volume 4 | Issue 2 | Print ISSN: 2395-1990 | Online ISSN : 2394-4099
National Conference on Advanced Research Trends in Information and Technologies (NCARTICT-2018), Department of IT, L. D. College of Engineering, Ahmedabad, Gujarat, India In association with International Journal of Scientific Research in Science, Engineering and Technology

Machine Translation Using Deep Learning : A Survey

Janhavi R. Chaudhary[1], Ankit C. Patel[2]

[1]M. E. Student (Information Technology Department), L. D. College of Engineering, Gujarat Technological University, Ahmedabad, Gujarat, India

[2]Assistant Professor (Information Technology Department), L. D. College of Engineering, Gujarat Technological University, Ahmedabad, Gujarat, India

ABSTRACT

Machine Translation using Deep Learning (Neural Machine Translation) is a newly proposed approach to machine translation. The term Machine Translation is used in the sense of translation of one language to another, with no human improvement. It can also be referred to as automated translation. Unlike the traditional statistical machine translation, the neural machine translation aims at building a single neural network that can be jointly tuned to maximize the translation performance. This survey reveals the information about Deep Neural Network (DNN) and concept of deep learning in field of natural language processing i.e. machine translation. It is better to use Recurrent Neural Network(RNN) in Machine Translation. This paper studies various techniques used to train RNN for various language corpuses. RNN structure is very complicated and to train a large corpus is also a time-consuming task. Hence, a powerful hardware support (Graphics Processing Unit) is required. GPU improves the system performance by decreasing training time period.

Keywords: Machine Translation, Neural Networks, Neural Machine Translation, Deep Learning

I. INTRODUCTION

A. Deep Learning

Deep learning is part of machine learning methods based on learning data representations, as opposed to task-specific algorithms[9]. Learning can be supervised, partially supervised or unsupervised. Deep learning architectures such as deep neural networks, deep belief networks and recurrent neural networks[1] have been applied to fields including computer vision, speech recognition, natural language processing, audio recognition, social network filtering, machine translation and bioinformatics where they produced results comparable to and in some cases superior human experts.

B. Deep Neural Networks

A deep neural network (DNN) is an ANN with multiple hidden layers between the input and output layers[1]. DNNs are typically feed forward networks in which data flows from the input layer to the output layer without looping back. Recurrent neural networks (RNNs), in which data can flow in any direction, are used for applications such as language modeling. Long short-term memory is particularly effective for this use. Convolutional deep neural networks (CNNs) are used in computer vision. CNNs also have been applied to acoustic modeling for automatic speech recognition (ASR)[1].

C. Machine Translation

Machine Translation (MT) is a sub-field of computational linguistics that investigates the use of computer software to translate text or speech from one natural language to another. At its basic level, MT performs simple substitution of words in one natural language for words in another. Machine Translation system are needed to translate literary works which from any language into native languages. The literary work is fed to the MT system and translation is done. Such MT systems can break the language barriers by making available work rich sources of literature available to people across the world[11]. Figure 1[11] shows process of Machine Translation in the form of pyramid.

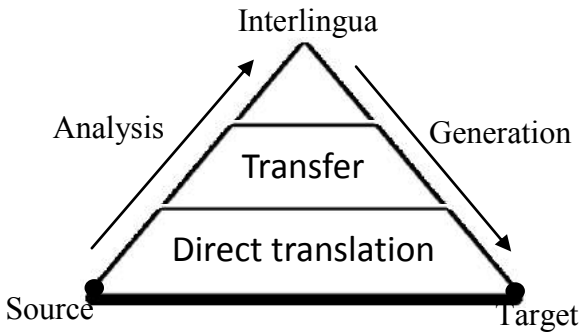


Figure 1. Machine Translation Pyramid

D. Neural Machine Translation

Neural Machine Translation is the approach of modeling the entire MT process via one big artificial neural network[18]. Figure 2[18] shows architecture or encoder and decoder.

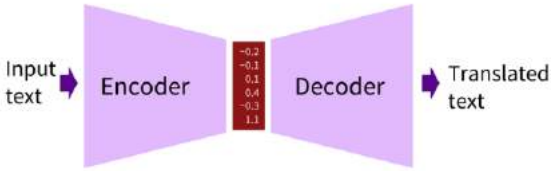


Figure 2. Neural encoder-decoder architecture

II. LITERATURE SURVEY

Jiajun Zhang and Chengqing Zong have designed many kinds of Deep Neural Networks. There are five popular neural networs introduced in their research work[1].

- ✓ Feed Forward neural network
- ✓ Recurrent neural network
- ✓ Recursive auto-encoder
- ✓ Recursive neural network
- ✓ Convolutional neural network

FNN(Feed Forward Network) :

The feed-forward neural network (FNN) is one of the simplest multilayer networks. Figure 3[1] shows an FNN architecture with hidden layers as well as input and output layers[1].

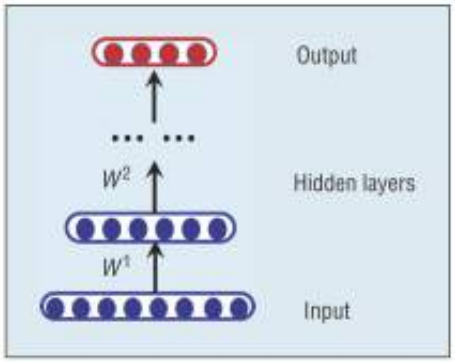


Figure 3. FNN Architecture

n-gram is collected from a text or speech corpus. The FNN attempts to predict the conditional probability of the next word given the fixed-window history words.

RNN (Recurrent Neural Network):

The recurrent neural network (RecurrentNN) is theoretically more powerful than FNN in language modeling due to its capability of representing all the history words rather than a fixed-length context as in FNN.

RAE (Recursive Auto Encoder):

The RAE provides a good way to embed a phrase or a sentence in continuous space with an unsupervised or semisupervised method.

Recursive Neural Network:

RecursiveNN differs from RAE in four points: RecursiveNN is optimized with supervised learning; the tree structure is usually fixed before training; RecursiveNN doesn't have to reconstruct the inputs; and different matrices can be used at different nodes.

Convolutional Neural Network:

The convolutional neural network (CNN) consists of the convolution and pooling layers and provides a standard architecture that maps variable-length sentences into fixed-size distributed vectors. The CNN model takes as input the sequence of word embeddings, summarizes the sentence meaning by convolving the sliding window and pooling the saliency through the sentence, and yields the fixed-length distributed vector with other layers, such as dropout and fully connected layers.

Kyunghyun Cho, Aaron Courville, and Yoshua Bengio describe Systems that learn to attend to different places in the input, for each element of the output, for a variety of tasks: machine translation, image caption generation,

video clip description, and speech recognition[2]. The attention-based neural machine translation uses a bidirectional recurrent neural network (BiRNN) as an encoder. The forward network reads the input sentence from the first word to the last, resulting in a sequence of state vectors

$$\{\overrightarrow{h_1}, \overrightarrow{h_2}, .., \overrightarrow{h_T}\}$$

The backward network, on the other hand, reads the input sentence in the reverse order, resulting in

$$\{\overrightarrow{h_T}, \overrightarrow{h_{T-1}}, .., \overrightarrow{h_1}\}$$

The use of the BiRNN is crucial if the content-based attention mechanism is used. The content-based attention mechanism relies solely on a so-called content-based scoring, and without the context information from the whole sentence, words that appear multiple times in a source sentence cannot be distinguished by the attention model[2].

Attention based Neural Network[2]

The content-based attention mechanism computes the relevance of each spatial, temporal or spatio-temporally localized region of the input,while the location-based one directly returns to which region the model needs to attend,often in the form of the coordinate such as the – coordinate of an input image or the offset from the current coordinate.

Sanjanaashree P and Anand Kumar M presented a new area of Machine Learning approach termed as a Deep Learning for improving the bilingual machine transliteration task for Tamil and English languages with limited corpus[3]. This technique precedes Artificial Intelligence. The system is built on Deep Belief Network (DBN), a generative graphical model, which has been proved to work well with other Machine Learning problem. They have obtained 79.46% accuracy for English to Tamil transliteration task and 78.4 % for Tamil to English transliteration. Bilingual Machine Transliteration task for Tamil and English languages is proceeded using DBN[3]. Figure 4[3] shows DBN architecture.

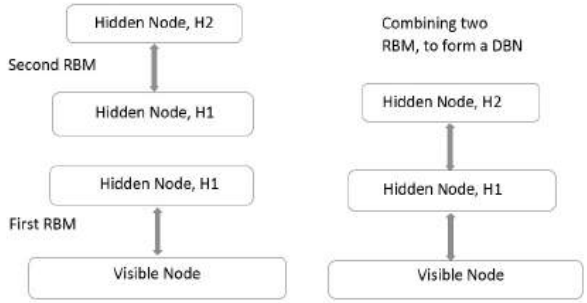


Figure 4. DBN Architecture

The system architecture is shown in Figure 5[3]. The two layers RBM on the right side is the encoders for source language and the left side is named as target language encoders. The uppermost layer is called Joint layer that concatenates the output of the top layer of source and target encoders. For a word to be transliterated, the source language word is passed through the source encoders and then reaches the joint layer and at last traverses downwards through the target encoders resulting with the transliterated word as the final output. In this architecture, each layer has 'n' number of neurons. Greedy layer-wise training is performed, the output of a layer becomes an input for the preceding layer.

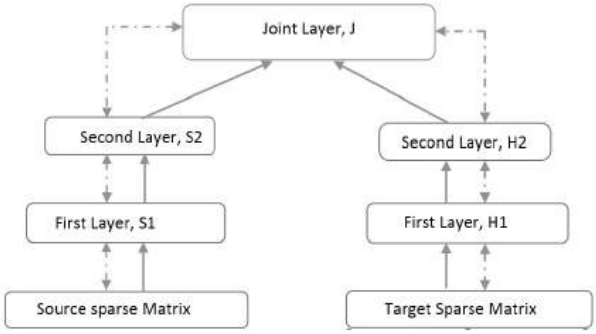


Figure 5. System Architecture for Transliterian

It concludes that a Deep Learning approach for bi-lingual transliteration is handled for English and Tamil languages using Deep Belief Networks. DBN is fully bi-directional, supports dimensionality reduction and it supports unsupervised way of learning feature and a supervised way for transliteration. with the small amount of corpus accounting to 3000 names, a decent result of accuracy around 79% is obtained in both ways, bi-lingual transliteration.

Zhen Yang, Wei Chen, Feng Wang and Bo Xu validate the hypothesis and propose a simple and flexible framework, which enables the NMT model to only focus on the relevant sense type of the input word in current context[4]. Firstly, this is the first effort to introduce the multi-sense representation, which represents each sense type of the word with a sense-specific embedding, into NMT. Secondly, propose a sense search module which can detect the sense type of the word automatically. Multi-Sense model is able to detect the sense type of the word exactly, and achieves remarkable improvements on

every test set over competitive baselines. The proposed sense search module enables the model to detect the right sense type of the input word automatically. Since the sense search module is task independent, it can be applied to any other semantic related NLP tasks with little modification[4].

XIAO-XUE WANG, CONG-HUI ZHU, SHENG LI, TIE-JUN ZHAO, DE-QUAN ZHENG have proposed trilingual NMT[5]. Based on the Encoder-Decoder and attentional mechanism, we translate source language to target language, meanwhile translate another parallel source language to target language. They provide two approaches called splicing-model and similarity-model. Both of the approaches are in order to enhance the semantic representation of input sequences.

Splicing model[5]

In this model, we get a new vector c which includes information of vector c' from source language1 and vector c'' from source language2. We think the new vector c is the semantic representation of parallel source language1 and source language2.

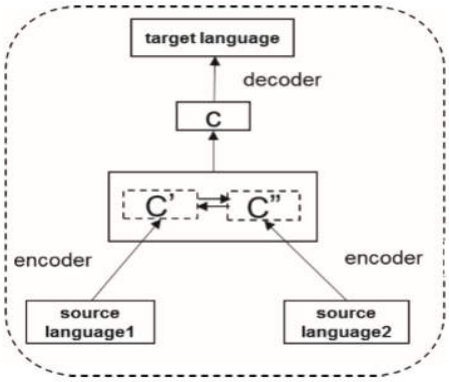


Figure 6[5]. The process of Splicing model

Similarity-model

Figure 7[5] shows the process. Because the whole process is based on similarity of vectors, we call the model similarity-model in this paper. And as you see in Figure 2, two systems are independent, so parameters are independent. Only in the training process we need to input source language1 and source language2 simultaneously. Once the model is trained, we could test the performance of the system from source language1 to target language by only inputting source language1 and test the performance of the system from source language2 to target language by only inputting source language2.

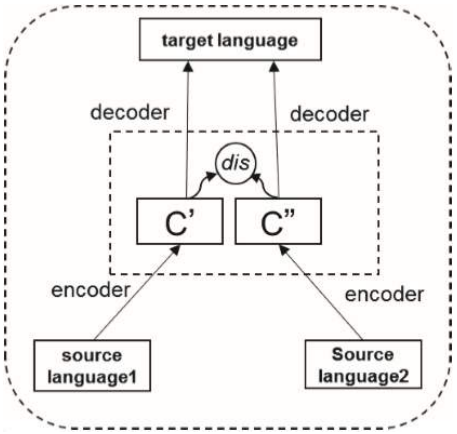


Figure 7. The process of similarity model

Feng Wang, Wei Chen, Zhen Yang, Xiaowei Zhang, Shuang xu and Bo Xu have proposed NMT model with class-specific copy network, which is referred to as CSCNMT[6]. With the network, the proposed NMT model is able to decide which class the words in the target belong to and which class in the source should be copied to. Experimental results on Chinese-English translation tasks show that the proposed model outperforms the traditional NMT model with a large margin especially for sentences containing the rare words. Copy mechanism has been proposed to deal with rare and unseen words for the neural network model using attention. But the negative point is that it's only able to decide whether to copy or not. It is unable to detect which class should the rare word be copied to, such as person, location, and organization. This paper proposes a new NMT model by novelly incorporating a class-specific copy network to overcome this issue[6].

Andi Hermanto, Teguh Bharata Adji, Noor Akhmad Setiawa have proposed RNN model for English-Indonesian machine translation[7]. In this research, a comparison between neural based network that adopts Recurrent Neural Network (RNN) and statistical based network with n-gram model for two-way English-Indonesian Machine Translation (MT) is conducted. The perplexity value evaluation of both models show that the use of RNN obtains a more excellent result. Meanwhile, Bilingual Evaluation Understudy (BLEU) and Rank-based Intuitive Bilingual Evaluation Score (RIBES) values increase by 1.1 and 1.6 higher than the results obtained using statistical based.

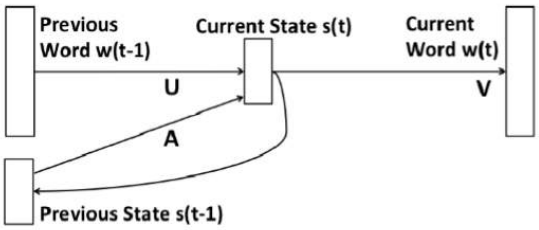


Figure 8. Recurrent Neural Language Model

Table 1. Literature Review

Review Paper	Title	Publication and year	Description
Paper-1 [1]	Deep neural networks in machine translation : An overview	IEEE, 2015	Various Neural Networks used in Different NLP methods are described
Paper-2 [2]	Describing multimedia content using attention based Encoder-Decoder network	IEEE Transactions on Multimedia, 2015	Describes Attention Based Neural Networks
Paper-3 [3]	Joint Layer based Deep Learning Framework for Bilingual Machine Transliteration	IEEE, 2015	Use of Deep Belief Network for Transliteration task for Tamil and English languages
Paper-4 [4]	Multi-Sense Based Neural Machine Translation	IEEE, 2017	Shows comparison between RNN and Sense-based model and introduces the multi-sense representation
Paper-5 [5]	Neural Machine Translation Research Based On The Semantic Vector Of The Tri-lingual Parallel Corpus	IEEE, 2016	Describes two approaches for trilingual translation with parallel corpus based on RNN and attention mechanism
Paper-6 [6]	A Class-specific Copy Network for Handling the Rare Word Problem in Neural Machine Translation	IEEE, 2017	Proposes a class-specific copy network model to overcome some issues faced in handling rare words
Paper-7 [7]	Recurrent Neural Network Language Model for English-Indonesian Machine Translation: Experimental Study	IEEE, 2015	Uses RNN language model for English-Indonesian Translation and shows comparison of this model with Statistical based language model

III. PERFORMANCE IMPROVEMENT USING GPUs

Deep learning application requires high computations because there exists large matrix multiplication, parallel processing and number of calculations during training phase. Graphics processing unit (GPU) is very good option for parallel processing and fast computation as compare to the CPU. GPU not only provides better energy efficiency but it also archives substantially higher performance over CPUs.

IV. CONCLUSION

In the present time, machine translation is a very hot research topic in natural language processing area. Deep learning helps to train a translation system like a human brain. RNN and RAE provides better result in text processing as compare to other neural networks. From above all papers about Machine Translation I have analyzed that deep learning is much better than any other methods for giving accurate result. But the processing time is high, also it needs more time to training a system.

V. REFERENCES

[1]. Jiajun Zhang and Chengqing Zong,"Deep neural networks in machine translation : An overview", in IEEE intelligent system, 2015.

[2]. Kyunghyun Cho, Aaron Courville, and Yoshua Bengio, "Describing multimedia content using attention based Encoder-Decoder network", IEEE TRANSACTIONS ON MULTIMEDIA, VOL. 17, NO. 11, NOVEMBER 2015.

[3]. Sanjanaashree P, and Anand Kumar M, "Joint Layer based Deep Learning Framework for Bilingual Machine Transliteration", IEEE, 2015.

[4]. Zhen Yang, Wei Chen, Feng Wang, Bo Xu, "Multi-Sense Based Neural Machine Translation",in IEEE, 2017.

[5]. XIAO-XUE WANG, CONG-HUI ZHU, SHENG LI, TIE-JUN ZHAO, DE-QUAN ZHENG, "Neural Machine Translation Research Based On The Semantic Vector Of The Tri-lingual Parallel Corpus", in IEEE, 2016.

[6]. Feng Wang, Wei Chen[1] ,Zhen Yang, Xiaowei Zhang, Shuang xu, Bo Xu, "A Class-specific Copy Network for Handling the Rare Word Problem in Neural Machine Translation", in IEEE, 2017.

[7]. Andi Hermanto, Teguh Bharata Adji, Noor Akhmad Setiawan, "Recurrent Neural Network Language Model for English-Indonesian Machine Translation: Experimental Study", in IEEE, 2015.

[8]. Biao Zhang, Deyi Xiong, Jinsong Su, and Hong Duan, "A Context-Aware Recurrent Encoder for Neural Machine Translation", IEEE/ACM TRANSACTIONS ON AUDIO, SPEECH, AND LANGUAGE PROCESSING, VOL. 25, NO. 12, DECEMBER 2017.

[9]. Deep Learning – Wikipedia, https://en.wikipedia.org/wiki/Deep_learning .

[10]. Xiao Sun, Xiaoqi Peng, Fuji Ren, Yun Xue, "Human-Machine Conversation Based on Hybrid Neural Network", in IEEE, 2017.

[11]. Sandeep Saini, Vineet Sahula, "A Survey of Machine Translation Techniques and Systems for Indian Languages", in IEEE, 2015.

[12]. Eric Greenstein, Daniel Penner, "Japanese-to-English Machine Translation Using Recurrent Neural Networks", available at https://github.com/lisagroundhog/GroundHog.

[13]. Bing Zhao, Yik-Cheung Tam, "BILINGUAL RECURREN T NEURAL NETWORKS FOR IMPROVED STATISTICAL MACHINE TRANSLATION", in IEEE, 2014.

[14]. Preeti Dubey, "Need for Hindi-Dogri Machine Translation System", in IEEE, 2014.

[15]. Heeyoul Choi, Kyunghyun Cho, Yoshua Bengio, "Context-dependent word representation for neural machine translation", in Elsevier, 2017.

[16]. Todd Law, Hidenori Itoh, Hirohisa Seki,"A neural network assisted Japanese-English Machine translation system", International Joint Conference on Neural Networks, 1993.

[17]. KyungHyun Cho, Yoshua Bengio,"NEURAL MACHINE TRANSLATION BY JOINTLY LEARNING TO ALIGN AND TRANSLATE ", in ICLR, 2015.

[18]. Neural Machine Translation Tutorial- https://sites.google.com/site/acl16nmt/home

© 2018 IJSRSET | Volume 4 | Issue 2 | Print ISSN: 2395-1990 | Online ISSN : 2394-4099
National Conference on Advanced Research Trends in Information and Computing Technologies (NCARTICT-2018), Department of IT, L. D. College of Engineering, Ahmedabad, Gujarat, India In association with International Journal of Scientific Research in Science, Engineering and Technology

Survey on Different Methods to Improve Accuracy of The Facial Expression Recognition Using Artificial Neural Networks

Chirag Ravat[1] , Shital A. Solanki[2]
[1]M.E., I.T. Department, L.D College Of Engineering, Ahmedabad, Gujarat, India
[2]Assist. Prof., I.T. Department, L.D College Of Engineering, Ahmedabad, Gujarat, India

ABSTRACT

Facial expression recognition by computer plays a key role in human computer interaction. FER has many applications such as Human-Robot interaction, surveillance, Driving-safety, Health-care, Intelligent tutorial system, music for mood, etc. Basically, Facial expression recognition can be done using Artificial Neural Network (ANN) and Support Vector Machine (SVM). So the accuracy of facial expression depends on these two phases, Feature extraction phase and classification phase. In this paper I'm going to survey different methods of FER and even face recognition methods.

Keywords: Facial Expression, Face recognition, CNN

I. INTRODUCTION

Artificial Neural Networks are computing systems inspired by the biological neural networks that constitute animal brains. Such systems learn to do tasks by considering examples, generally without task-specific programming. The human brain is composed of 86 billion nerve cells called neurons[1]**.** They are connected to other thousand cells by Axons**.** ANNs are composed of multiple nodes, which imitate biological neurons of human brain. The neurons are connected by links and they interact with each other.

The nodes can take input data and perform simple operations on the data. The result of these operations is passed to other neurons. The output at each node is called its activation or node value. Each link is associated with weight. ANNs are capable of learning, which takes place by altering weight values[1].
However Convolutional Neural Networks(CNNs) have an edge over conventional MLPs in terms of image recognition and classification. The brief introduction of CNN is given in chapter III.

Facial Expression Recognition System

Almost all prediction based systems consist of mainly two phases: Training phase and Testing phase. In either Supervised or Un-supervised learning method first feature vectors are trained based on given labelled data in case of Supervised learning or on given attributes in case of Un-supervised learning. These trained vectors are then used to test an unseen data and gives label accordingly.

In FER it is identified as Feature Extraction and Classification phases. In feature extraction phase certain features of an image is extracted based on which classifier, in classification phase, classifies image to one of the label of domain.

For fast and better training, first of all image is pre-processed to reduce noise. Pre-processing also includes converting image into grayscale and resizing it into predefined dimensions. For different FER methods different sizes can be considered.

Then next step is to detect face from the pre-processed image. There are lots of techniques available for detecting face from an image like Haar classifier, Ada

boost technique by Viola-jones, Adaptive skin colour, etc., which gives output as face image or non-face image[4].

After face is detected another step is to extract features from face image. It can be done directly from given image or also from video frames. Features such as eyes, nose, mouth, eyebrows, ears, etc. are detected. Feature extraction process works significantly well if face is already detected. There are two types of methods for feature extraction : Appearance based and Geometric based. Geometric based methods are more suitable for real time applications as Appearance based method consumes more power, time and memory but also highly discriminative. There are many feature extraction methods are there but Gabor feature and LBP[2] gives optimal results. There are also pre-defined Neural Net structures available such as AlexNet[6], which is a Convolutional Neural Net, useful in feature extraction.

In last classifier is used to train model using extracted features and classify them into label. Classifier learns the mapping between the given image and given label, so when we give a new image after sufficient training it can classify input image into certain label.

Also testing phase includes image pre-processing and feature extraction, based on these features classifier directly classifies that image into some label.

II. LITERATURE REVIEW

A. Facial Expression Recognition Using General Regression Neural Network[2]

In this paper authors proposed two phases for FER, namely Training phase and Testing phase. Both phases includes Image Pre-processing and Feature extraction. In training phase these extracted features are trained and mapping is learned between image and label for that image. And in testing phase extracted image can be directly classified into some label.

Here, in image pre-processing section they intent to reduce noise in image, converting image into grayscale and resizing image into various block sizes. They trained the model using various sized blocks : 256×256, 128×128, 64×64 and 32×32. Their experiment showed that 64×64 block size gives optimal result in terms in accuracy, Where 256×256 gives lowest one.

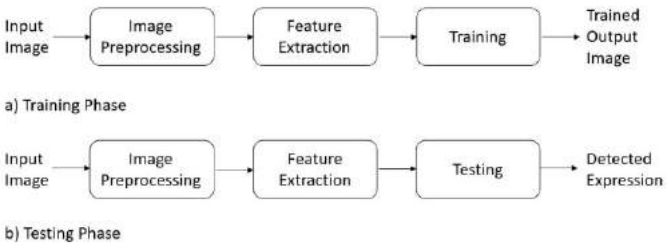


Figure 1. FER system[2]

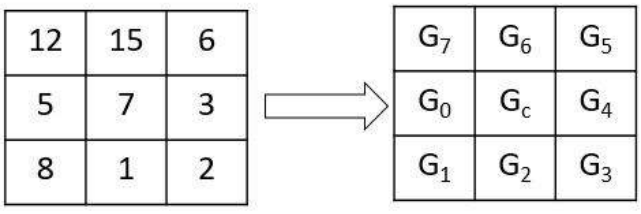
Then face detection module has been applied to pre-processed image. It targets the face and crops it. The cropped face is normalized using histogram equalization. Here Ada boost technique is used in Viola-Jones face detection algorithm.

Now comes the most important part of any image based classification method, Feature extraction. In this paper they used Local Binary Pattern(LBP) to extract feature from input image. LBP is used to determine feature vector which reduces the entire image data to a single feature vector ready for classification. Main purpose of feature extraction technique is to reduce large amount of pixel data into small most meaningful feature vector to ease classification problem.

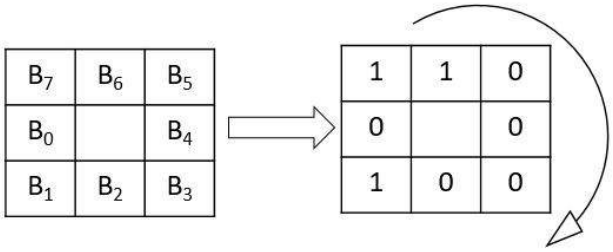
The last step is classification. Extracted features are classified into six already known expression classes : Happy, Unhappy, Surprise, Disgust, Fear and Angry. For classification of these feature Unsupervised approach General Regression Neural Net(GRNN) is used. It trains network faster and does not require iterative training procedure.

Local Binary Pattern

LBP was first introduced in 1996 by Ojala et al as a basic binary operator. It works as powerful texture classifier. LBP is simple tool for detection of feature. It is widely used because of its robustness and simplicity.



(a)



(b)

Figure 2. LBP Coding[2]

The pixels of the image are labelled by the binary operator by comparing the center pixel value with the 3×3 neighbourhood pixels. If value of pixel is greater than central pixel then it is assigned 1, if lower then assigned 0. Then these pixels traced circularly and formed LBP code.

This 8 bit code (Byte) is converted into decimal number. Advantage of LBP is, it calculates code relatively to the central pixel, so even if image is lighten or darken, relatively code will be same. So it is robust against illumination conditions.

General Regression Neural Network

GRNN consists of four layers: input layer, pattern layer, summation layer and output layer. Input nodes are fully connected to the pattern layer. GRNN estimates output by calculating the Euclidean distance between the train and test data.

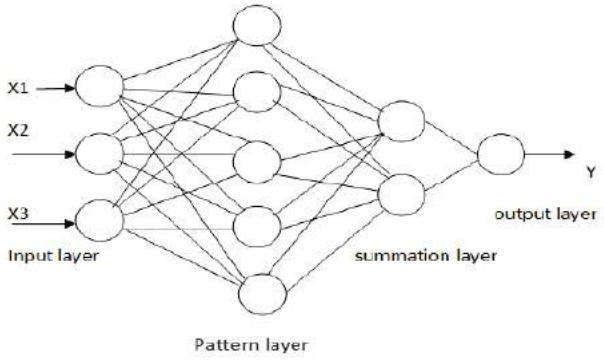


Figure 3. GRNN Architecture[2]

It uses free parameter named spread constant reduces Mean Square Error (MSE) and improve efficiency.
Overall accuracy achieved is 94.25-95.48%, Which is really good. But there are contradictions in various expression classes. Neutral, happy, surprise, disgust classes are detected almost for 100% of the time, though unhappy, fear and angry classes give high rate of confusion and have come down to 62.5% for some tests.

B. Spontaneous facial micro-expression detection based on deep learning[3]

In this paper, method for automatic detection of facial micro-expressions detection is proposed. Dataset used in this method is a video dataset. Temporal Interpolation Method is used to normalize the video length.

Here, Dlib machine learning toolkit is used to detect face in each frame. Rest of the process follows as below.
After detecting face, very first and crucial part is facial landmark localization. It is done by Deep CNNs model. CNNs can transform raw pixels of an image to facial landmark positions and other related attributes. Here CNN architecture consists of convolutional layer, pooling layers, fully connected layer and loss layers.

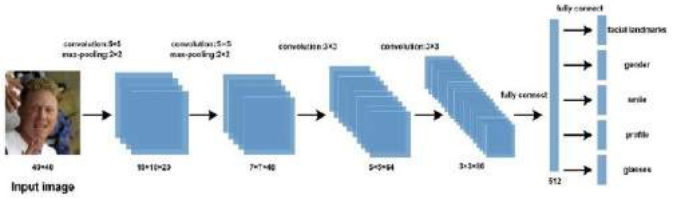


Figure 4. Architecture of CNN[3]

Convolutional layers filters the input image with 20 kernels of 5*5*3 with stride of 1 pixels. ReLU function is used as activation function. Then max pooling layers are applied. And in last fully connected layers are used to summarize all the learning. Here Euclidean loss function is used for facial landmarks and softmax loss function for other auxiliary tasks.
Concept of transfer learning is implemented here. Facial landmark localization on sparse facial landmarks, the trained model is then used to detect dense facial landmarks (68 points) , which really improves the training speed and capacity.

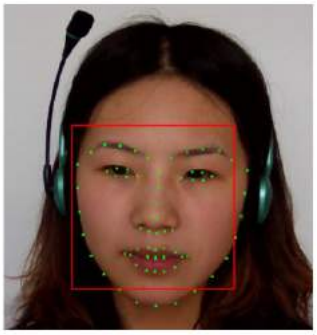


Figure 5. Detecting 68 landmarks[3]

As pipeline goes, facial region is split into 12 sub region according to 68 facial landmarks and Facial Action Coding System (FACS). These 12 regions are known as 12 ROIs (Region Of Interest).

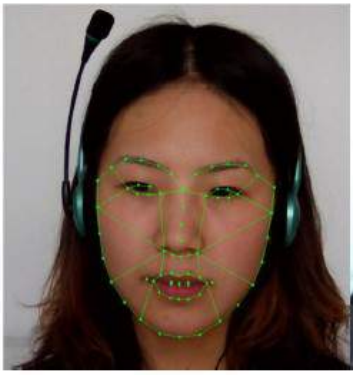


Figure 6. Detecting 12 ROIs[3]

Every ROI is then converted to HOOF (Histogram of Oriented Optical Flow) feature and normalized.
Finally Support Vector Machine (SVM) is used to classify micro-expressions.
Accuracy achieved is 80%, which is human level when it comes to detect micro-expression instead of just expressions.

C. A Robust Method for Face Recognition and Face Emotion Detection System using Support Vector Machine[4]

The proposed methodology of this paper is divided into two part : Face recognition and Emotion recognition.

Face Recognition

Author designed algorithm in such a way that if the person is recognizing for the first time then the system considers him as a new user and performs each step of operation, otherwise it is considered as Registered user and recognizes him.

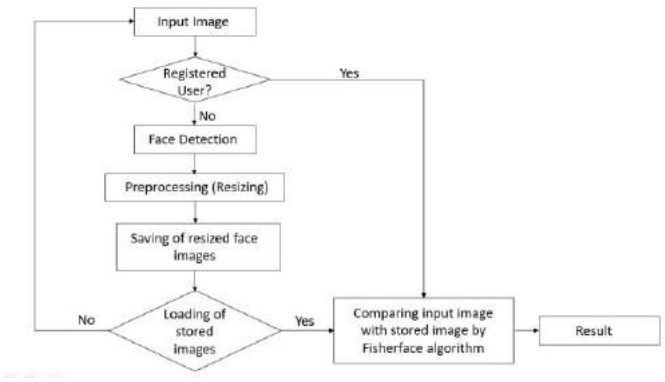


Figure 7. Face Recognition System[4]

Here OpenCV (Open Computer Vision) is used. It contains cascade classifiers in which Viola-jones algorithm is implemented. Haar-cascades is used to detect faces. It outputs image as positive or negative. Negative image is ignored for further process.

Here Fisherface algorithm is used to detect different users, so that false images can be discarded. Fisherface algorithm performs "Leave-one-out" cross validation for user identification.

Emotion Recognition

For detection of emotion, Face detection is must. So face detection part is done same as above. Then feature extraction is done using Dlib Machine Learning toolkit.

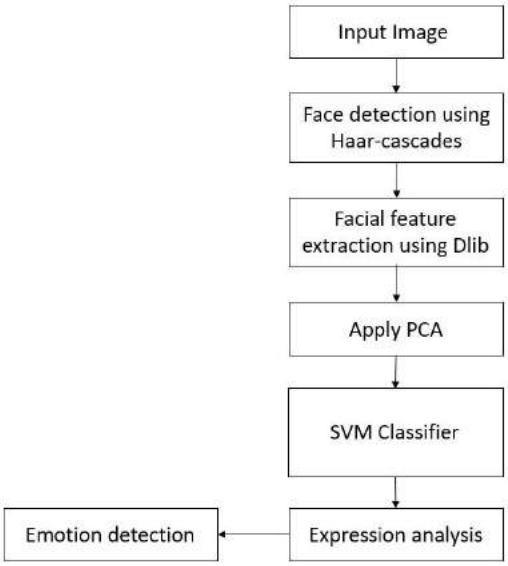


Figure 8. Emotion detection system[4]

Principal Component Analysis(PCA) is applied to training images to reduce the dimensionality. Higher dimensions images take more time to train, so considering the sufficient quality for feature extraction 273×273 size is used. Facial features such as nose, eyes, lips, face contour are considered as keys. These keys are applied to SVM classifier. It analyses the features and labels according to that. The trained model is then used to test new images.

The method is robust as seems, but training phase is not that strong, as it is trained on CK+ database, which has very small number of labelled images. Overall running time of system is significantly less. And achieves accuracy over 90%.

D. Smile Detection Using Pair-wise Distance Vector and Extreme Learning Machine[5]

In this paper, as a feature vector for smile detection pair-wise distance vector is used. It is extracted only from points around mouth, because they convey the most detail whether face is a smile face or not. Here 68 facial points an 7 face regions are considered. From MUG Facial Expression Database, which contains 401 images of 26 subjects, has each smile-face with its corresponding neutral image. These two image of same subject is treated as a pair of images. The movements are calculated and these variances are added together for each corresponding component and portions are shown in fig. Of course, the movements around the mouth area is highest.

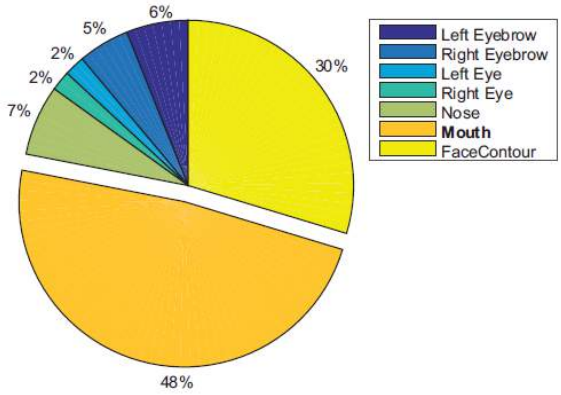


Figure 9. Impact factor of smiled face compared to neutral face[5]

Whole process is divided into 3 parts : Pre-processing, Feature extraction and Classification.

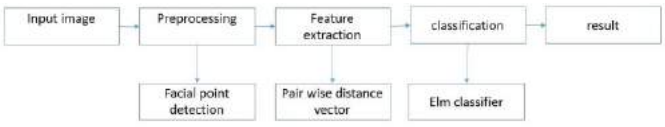


Figure 10. Smile detection system[5]

Pre-processing includes facial landmark detection, which should be accurate. State-of-the-art method named Coarse-to-Fine Auto-encoder Networks(CFAN), which is extremely fast and has achieved highest accuracy at present in terms of facial landmarking. It consists of a global Stacked Auto-encoder Network (SAN) and several local SANs. Global SAN can quickly reach the approximated location of facial landmarks and local SANs refine these locations step by step.

After pre-processing 68 points are obtained, but in training only 20 points are considered, which were around the mouth.

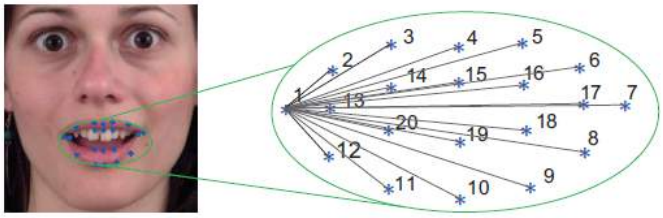


Figure 11. Detecting 20 points around the mouth[5]

Next step is feature extraction, in this part pair-wise distance vector is used. Pair-wise distance vector is an upgradation to the existing Euclidean distance matrix. This feature vector has important properties like Translation Invariance, Rotation Invariance and Scaling Invariance.

These extracted features are classified by Extreme Learning Machine, which is a single-hidden-layer feed-forward neural network. ELM assigns the parameters of the hidden nodes randomly without any iterative tuning, which makes it really faster than traditional net.

Accuracy achieved is above 90% for most of the time. Highest achieved is 94.96%, which is almost 95%. But algorithm running speed is relatively low and causes more sufficient hardware for real time implementation. Use of CFAN gives a boost to higher accuracy, CFAN is fast and most accurate facial landmark detection method as of now. By which this method achieves accurate landmarks and hence better post processing.

E. Gender and Age Classification of Human Faces for Automatic Detection of Anomalous Human Behaviour[6]

In this paper, authors aim to develop a system for automatic detection of anomalous Human behaviour. Main focus of paper is to detect Gender and Age of Human which will eventually support system to detect anomaly activity.

Convolutional Neural Nets achieves excellent performance in image feature extraction, but at the same time training a deep CNN from scratch requires vast amount of time, resources, datasets, high computational

power. In some cases it may go over days to train a deep CNN. So, the concept of Transfer Learning is used.
CNNs trained on some dataset can be tuned for a new task, even in a different domain.

As there are very few datasets available with labels of Gender and Age, Transfer learning approach is used.
Here, AlexNet – pre trained CNN is used, which has achieved excellent classification results on ImageNet competition. Input of ALexNet is 227×227 RGB image. Network has 5 convolutional layers and 3 fully connected layers.

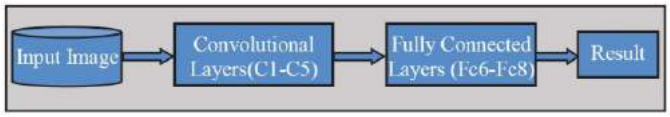


Figure 12. AlexNet architecture[6]

Feature extracted from AlexNet is classified using general classifier SVM.

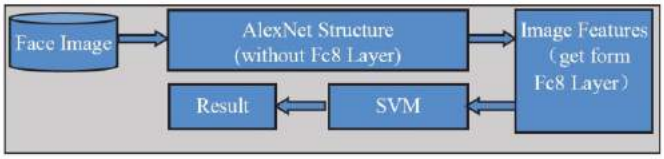


Figure 13. AlexNet with SVM[6]

Before entering into ALexNet images are resized into fixed size 227×227, so net can perform well.
Experiments also showed that if we use Haar-like features to extracts features from image and classify them using SVM, the accuracy tents to go really low, whereas AlexNet performs outstanding.

Gender classification gives 90.33% accuracy in average and Age class gives 80.17%. Use of transfer learning concept with pre-trained CNN architecture AlexNet makes it easy to train efficiently on datasets having less images.

III. CONVOLUTIONAL NEURAL NETWORKS

Convolutional Neural Network (CNN) architecture is inspired by mammalian visual cortex. Visual cortex processes images in hierarchical manner, first low level features and then high level features. CNN also works same as visual cortex, it first processes low level features of an image, such as curves, edges, then bit higher features like small part of an image and this hierarchy is continued layer by layer and in last whole image is processed.

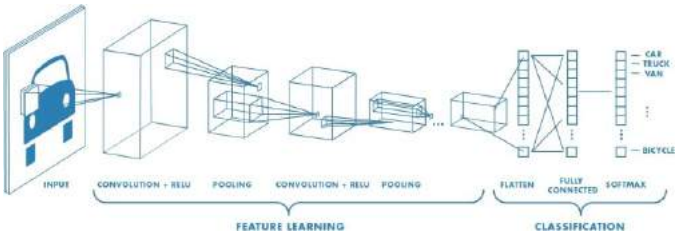


Figure 14. Basic CNN architecture

Basic layer of CNNs are convolutional layer, pooling layer, ReLU layer, Fully connected layer, loss layer, softmax layer, etc.

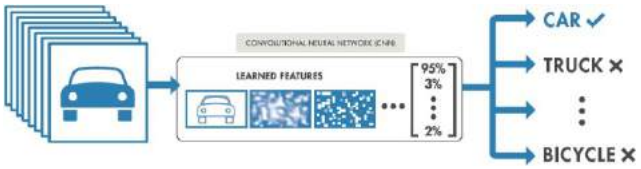


Figure 15. Result scheme of CNNs

Convolution layer takes an image as input and convolute it with feature vector or weight matrix and output more meaningful image. This image imported to other layer.
Pooling layer of CNN is used to extract most meaningful feature from each section of an image.

ReLU rectifies the image, it simply applied to check whether image gives some information or not. If image pixels are 0, the ReLU discards it.

Fully connected layer is used to sum up all learned features by connecting all neurons of previous layer to the next layer. It is usually used in later part of CNN architecture.

Loss layer in CNN, is used to apply different loss functions. For each attribute, a loss function can be applied. For example, softmax loss function is useful in multi-class classification and gives output a label with probability.

CNN is combination of these layers, not necessarily in same order.

IV. COMPARISON

Table 1

Method used	Advantages	Disadvantages	Accuracy
Adaboost by Viola-Jones+ LBP + GRNN [2]	Accuracy Especially in Neutral, Happy, Surprise, classes (mostly 100%).	Confusion rates are higher in Unhappy, fear and angry classes. 62.5 % is the lowest achieved in angry class. Gives optimal solution only on 64×64 size block.	94.25 - 95.48 %
Dlib toolkit+ CNN + Face split in 12 ROIs + HOOF Feature calculation + SVM classifier[3]	Uses transfer learning, which reduces the training time. Uses more effective features, so accuracy is good.	Eye blinking in clip is ignored as noise. Efficient only for short duration clips.	80 %
Face Recognition : Haar-cascades + Fisherface EmotiOn Detection : Haar-cascades + Dlib +PCA + SVM classifier[4]	Flow of method is good. Overall method takes significantly less time to classify image, Especially face detection part.	Trained under a dataset (CK+) having very less number of images. Noise removal is not efficient	≥ 90 %
CFAN + Pair-wise distance vector + ELM[5]	ELM outperforms SVM, Adaboost, which gives better accuracy comparatively.	Running speed of algorithm is relatively slow, so real time implementation is difficult and cause more efficient hardware.	93.42 ± 1.46 %
Pre-trained CNN – AlexNet + SVM [6]	Use of transfer learning through efficient pre-trained CNN – AlexNet gives very good results even dataset is small.	Even by using efficient feature extraction method, accuracy tents to low.	Gender : 90.33 % Age : 80.17 %

V. CONCLUSION

Based on above literature review, Comparison shows that highest accuracy for Facial Expression is achieved by GRNN classifier[2], which is an Un-supervised learning method. But it also gives higher confusion rates in some of emotion classes. 2nd highest is achieved in smile detection method using pair-wise distance vector and ELM[5], it also uses efficient facial landmark detection method CFAN[5]. On the basis of review, it can be seen that Feature extraction and classification plays major role for achieving higher accuracy. It can also be seen that parameters such as time complexity of algorithm for training, size of algorithm, etc are ignored and focused solely on accuracy rate.

VI. REFERENCES

[1] Nikhil Buduma, 2015. Fundamentals of Deep Learning.2th Edn., Sebastopol, CA., ISBN: 978-1-491-92561-4

[2] Kiran Talele, Archana Shirsat, Tejal Uplenchwar, Kushal Tuckley, "Facial Expression Recognition Using General Regression Neural Network", IEEE Bombay Section Symposium(IBSS), 2016.

[3] Xiaohong Li, Jun Yu, Shu Zhan, "Spontaneous facial micro-expression detection based on deep learning" in IEEE, 2016.

[4] Rajesh K M, Naveenkumar M, "A Robust Method for Face Recognition and Face Emotion Detection using Support Vector Machines", IEEE, 2016 International Conference on Electrical, Electronics, Communication, Computer and Optimization Techniques(ICEECCOT), 2016.
[5] Dongshun Cui, Guang-Bin Huang, Tianchi Liu, "Smile Detection Using Pair-wise Distance Vector and Extreme Learning Machine" in IEEE, 2016
[6] Xiaofeng Wang, Azliza Mohd Ali, Plamen Angelov, "Gender and Age Classification of Human Faces for Automatic Detection of Anomalous Human Behavior" in IEEE, 2017
[7] Jiaxing Li, Dexiang Zhang, Jingjing Zhang, Jun Zhang, "Facial Expression Recognition with Faster R-CNN", Science Direct, Procedia Computer Science 107 (2017) 135 – 140, 2017.
[8] Nazima kauser, Jitendra Sharma, "Automatic Facial Expression Recognition: A Survey Based on Feature Extraction and Classification Techniques", IEEE, 2016.
[9] Shubhada Deshmukh, Manasi Patwardhan, Anjali Mahajan, "Survey on real-time facial expression recognition techniques", IET (The Institute of Engineering and Technology) Journals, pp. 1-9, 2016.
[10] Kewen Yan, Shaohui Huang, Yaoxian Song, Wei Liu, Neng Fan, "Face Recognition Based on Convolutional Neural Network", IEEE, Proceedings of the 36th Chinese Control Conference, 2017.

© 2018 IJSRSET | Volume 4 | Issue 2 | Print ISSN: 2395-1990 | Online ISSN : 2394-4099
National Conference on Advanced Research Trends in Information and Technologies (NCARTICT-2018), Department of IT, L. D. College of Engineering, Ahmedabad, Gujarat, India. In association with International Journal of Scientific Research in Science, Engineering and Technology

An Adjacency Matrix Based Apriori Algorithm for Frequent Itemsets Mining

Mahendra N. Patel[1], Suresh B Patel[2], Dr. S. M. Shah[3]

[1]PG Scholar, Computer Engineering Department, Government Engineering College, Modasa, Gujarat, India
[2]PG Scholar, Computer Engineering Department, Government Engineering College, Modasa, Gujarat, India
[3]Computer Engineering Department, Government Engineering College, Modasa, Gujarat, India

ABSTRACT

Finding frequent itemsets is a most researched field in data mining. Currently, the finding of frequent itemsets problem's solution has been proposed by many researchers. The Apriori algorithm is the basic algorithm for frequent itemsets mining. In Apriori algorithm, there are main two issues: scanning the database multiple times and generating a large number of candidate sets. In recent years several improved apriori algorithms have been defined and evaluated to improve efficiency. Our main goal is to define a new optimized algorithm and to compare its performance with the existing algorithms. The main focus of our work is to propose a new optimized algorithm and to compare its performance with the state of the art methods. In proposed work, adjacency matrix will be employed in order to improve the operating efficiency and eliminate the candidate sets. In a proposed system not require the pruning step. Performance of the proposed method will be evaluated on existing datasets. A secondary data set is used to find frequent itemsets with using our proposed algorithm and existing algorithm. The effect of our proposed algorithm is presented.

Keywords: Apriori, Data Mining, Frequent Itemsets Mining (FIM), Adjacency Matrix, FI-generator

I. INTRODUCTION

Recent days have an explosive growth in generating data in all fields of business, science, medicine, defense etc. the same rate of growth in the processing power of evaluating and analyzing the data did not follow this growth. Due to this phenomenon, a large volume of data is still kept without being studied. Data mining is a research field that tries to overcome this problem, processes some methods for the extraction of significant and potentially useful patterns from these large collections of data.

Data mining technique is used to find valid, novel, useful and ultimately understandable patterns in data [1]. In general, there are different kinds of patterns that can be discovered from data. For example, association rules can be mined for market basket analysis, classification rules can be used for accurate classifiers, clustering can be used for customer relationship management.

Most current studies on frequent pattern mining adopt an Apriori-like approach, which is based on an "anti-monotone Apriori heuristic[1][2]: if any length k-pattern is not frequent in the database, its length (k+1) super pattern can never be frequent". The original Apriori algorithm requires k scans or passes over the data where k is the length of the longest frequent itemset.

In this paper, we define Adjacency matrix base Apriori algorithm and analyzed its performance. Paper is organized as follows. Section 2 introduces related work on apriori algorithm and existing algorithm based on adjacency matrix. Section 3 presents our proposed algorithm for frequent itemsets mining. Section 4 presents the time complexity of an existing algorithm

and our proposed algorithm. Section 5 conclusions and summarizes the paper.

II. RELATED WORK

Data mining is a process to discover hidden information or knowledge automatically from huge database [3]. First presented by Agrawal was to find frequent itemsets using mining of association rules[1]. In the following section, we try to show the concepts of frequent itemsets mining. Let TDB (Transactions databse) = {T1, T2, T3, …, Tm} where m is number of transactions. Each transaction Ti = {i1, i2, i3, …}contains a set of items from I = {i1, i2, i3, …, in} where n is number of items. An itemset X with k items from I is called a k-itemset. A transaction Ti contains an itemset X if and only if X ∈ Ti [4]. The support of a number of transactions in TDB containing X. An itemset is frequent if its support, supp(X) is greater than a support threshold called *minimum support*. For example, suppose we have a

TDB = {T1, T2, T3, T4, T5,T6} and
I = {A, B, C, D, E} where
T1 = {B, C, D, E},
T2 = {B, C, D},
T3 = {A, B, D},
T4 = {A, B, C, D, E},
T5 = {A, B, C}, and
T6 = {B, E}.

Thus, for instance, support({A}) = 3, can be achieved because A occurs only in T_3, T_4, and T_5 transactions. For k itemsets where k ≥ 2, for instance, support({A, E}) = 1, because it occurs only one time in all transactions. And support({A, B, D, E}) = 1 can be computed from all transactions using the same way for generating all frequent itemsets.

A. Apriori Algorithm

The Apriori algorithm for finding frequent itemsets was originally presented by Agrawal and Srikant [5]. It finds frequent itemsets according to a user-defined minimum support. In the first pass of the algorithm, it constructs the candidate 1-itemsets. The algorithm then generates the frequent 1-itemsets by pruning some candidate 1-itemsets if their support values are lower than the minimum support. After the algorithm finds all the frequent 1-itemsets, it joins the frequent 1-itemsets with each other to construct the candidate 2-itemsets and prune some infrequent itemsets from the candidate 2-itemsets to create the frequent 2-itemsets. This process is repeated until no more candidate itemsets can be created.

Consider the Apriori algorithm with the example shown in Figure 1 for the transaction database shown in Table 1 and also assume that minimum support is 2. In this C1 represent1-candidate itemset and after pruning it generate L1 which represents 1-frequent itemset. After L1 join operation applies and then C2 is generated. The same way other C3, L2, and L3 are generated. However, there are no 4-itemsets in pass 4 created because no candidate 4-itemsets can be created. Thus, the process stops.

Table 1. Transaction Dataset

TID	List Of Items IDs
T1	I1, I2, I5
T2	I2, I4
T3	I2, I3
T4	I1, I2, I4
T5	I1, I3
T6	I2, I3
T7	I1, I2
T8	I1, I2, I3, I5
T9	I1, I2, I3

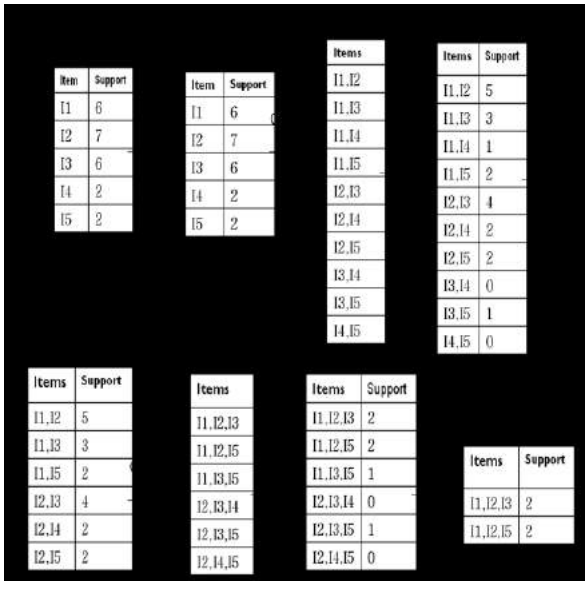


Figure 1. The apriori example

B. Improved Apriori Algorithm based on Adjacency matrix

Now in this section, we discuss FI-generator algorithm proposed by Archana Singh and Jyoti Agarwal [6]. In this method used Adjacency matrix as a data structure.

In this method [6] first scan the database and generate the frequency-wise sorted table of items in descending order. In first pruning step all the items, having frequency count less than minimum support is removed from the sorted table (TS). Again scan the database and create the adjacency matrix say R. In pruning step all the entries from adjacency matrix with frequency count less than minimum support are pruned. Now select the first item from the sorted table as 1-frequent itemset say L1. Then generate the candidate 2-itemset is the item set which has relation with L1 in adjacency matrix R. From the candidate 2-itemset select the item whose frequency is more and take union with L1, so we get 2-frequent itemset say L2. Then generate candidate 3-itemset is the item set which has relation to all items of L2 in adjacency matrix (R). Thus the size of Candidate set is reduced in the successive iteration.

Working of an adjajency matrix based [6] apriori algorithm

Let's take the example of given algorithm based on one sample database transactions shown in Table 1. Assume the minimum support threshold is 2. Example of given algorithm for frequent itemset generation shown in Figure 2 [7].

For example shows first the sorted table, second the adjacency matrix R and third pruned adjacency matrix. In the example, we select {I2} as a 1-frequent itemset from a sorted table whose frequency is higher than other. Then generate candidate 2-itemset which associated with {I2}. Next, we select {I1, I2} is a 2-frequent itemset whose frequency is higher than other. Same ways finally select {I1, I2, I3} as a 3-frequent itemset.

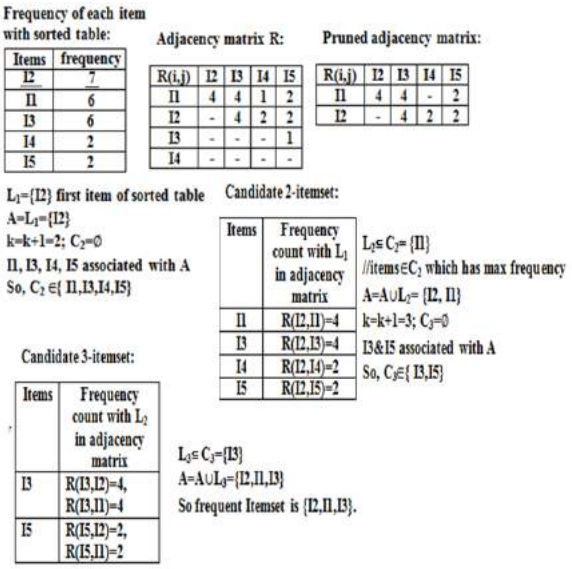


Figure 2. The FI-generator example

III. PROPOSED WORK

In this proposed method, we are improving the performance of an apriori like algorithm. In this proposed method we can enhance the efficiency of an enhanced apriori algorithm based on adjacency matrix [6].

In this proposed method, we take adjacency matrix as our data structure. In adjacency matrix R(i,j) = frequency count of item (i) and item(j) appears together in database. Below show the proposed algorithm:

Pseudo Code of Proposed Algorithm

Input: Transaction Database T, min_sup
Output: Frequent item set

Method:

Step 1: Create Transaction IDs table
Scan the transaction database and create T_ids table. T_ids first column consist item and second column consist transaction id's associated with that item.

Step 2: Creation of adjacency matrix (R)
Create an adjacency Matrix (R), where rows and column shows the different items from T_ids table.

Rows i=1 to n-1 //where n is number of items Columns j=2 to n In adjacency matrix R(i,j) =frequency count of item(i)and item(j) appears together in database.

Step 3: Frequent item set generation:

Initialize vector C & FI // empty

```
For i=1 to n-2
        Initialize a=0
        For j=i+1 to n
        If R(i,j) value ≥ min_support then
                C(a)=j
                Increment a
        End If
        End For
        Initialize b=0
        For k=b to a-1
                For l=k+1 to a
                If R(C(k),C(l)) value ≥
                min_support then

                    Insert into FI Frequent
                    Item Set (i,C(k),C(l))

                End If
                    Increment l

                End For

        End For

End For
```

Step 4: Display result and stop

Display FI contains frequent item set

Working of an proposed algorithm

Let's understand this step through an example, for that use Table 1 as sample transaction database. First, create Transaction_id table shown in Table 2. Then generate adjacency matrix (R) from that Transation_id table shown in Table 3. Finally, apply search technology on this adjacency matrix. In this example we take min_support value is 2.

Table 2. Transaction_id

Item	**T_id's**						
I1	1	4	5	7	8	9	
I2	1	2	3	4	6	8	9
I3	3	5	6	7	8	9	
I4	2	4					
I5	1	8					

Table 3. Adjacency matrix R

R(I,j)	**I2**	**I3**	**I4**	**I5**
I1	4	4	1	2
I2	-	4	2	2
I3	-	-	0	1
I4	-	-	-	0

Steps for search technique:

Step 1: fetch the row for i=1 to n-2where n = number of items(5).

Here first take I1 as a first-row item

Step 2: With this row find the associated value whose value greater or equal to min_support (User input=2).

So here associated with I1 are I2, I3, and I5 whose value is 4, 4, and 2 respectively in the adjacency matrix.

Step 3: Check the (I2, I3), (I2, I5) and (I3, I5) are frequent or not. Means here find R(2,3)= 4, R(2,5) = 2 and R(3,5) = 1 from adjacency matrix. So here we find (I2, I3), (I2, I5) are frequent.

Step 4: Display frequent item set

In step 3 find (I2, I3) and (I2, I5) which associate with row I1, so we display (I1, I2, I3)and (I1, I2, I5) as frequent item set.

Step 5: Repeat step 1 to step 4 for next row detail given below.

For the second row is I2.

Find I3, I4, and I5 associated with I2 and whose value is 4, 2, and 2 respectively.Here find R(3,4) = 0, R(3,5) = 1 and R(4,5) = 0, so here no one are frequent.

Now we check for the third row is I3

Find I4 and I5 associated with I3 and whose value is 0 and 1 respectively. Here all values are less than min_support, so here no one is frequent.

End of the procedure because search up to n-2 which is 3 was completed.

Finally, from the example shown above, we get the output of 3-frequent itemset is {(I1, I2, I3), (I1, I2, I5)}. This result is same as apriori algorithm.

IV. COMPARISONS AND DISCUSSION

We compare proposed system and existing system with respect to time complexity as shown below:

Time to execute steps in pseudo code, where n is a number of items, m is a number of transactions and T is transaction database.

First, in existing System [6] require time to find the frequency of each item ∈ T is O (nm), Time to create of the sorted table and pruning of sorted table is O (n). In proposed system require time to create transaction_id table is O (nm).

Second, in the existing system requires time to create adjacency matrix is O $(n)^2$ and Time to prune adjacency matrix is O $(n)^2$. So the overall time complexity is O $(2n^2)$. Whereas in the proposed system require time to create adjacency matrix is O $(n)^2$.

Third, in the existing system requires two times to generate frequent itemset is O (n-c), c is the number of items pruned in the candidate set. In proposed system require time to generate frequent itemset is O $(n)^2$.

So theoretically, we say that performance of the proposed system is better than the existing system. The proposed algorithm is yet to be implemented on Java platform and perform on dataset [6] and it will practically prove that proposed system is better than existing one.

V. CONCLUSION

Frequent itemsets mining is one of the most important areas of data mining. Existing implementations of the Apriori-based algorithms focus on the way candidate itemsets generated, the optimization of data structures for storing itemsets, and the implementation details. A key contribution of this research paper is to provide the user with a simple but yet powerful, adjacency matrix structure for efficient frequent pattern mining. There are various frequent itemset mining methods are available. These frequent itemset mining scans database multiple times and generates more candidate set. So currently available frequent itemset mining methods are lacking in speed. Current existing method based on adjacency matrix [6] does not generate all frequent itemset. The main objective of this research paper is to propose a new algorithm to find all frequent itemset without generating candidate sets by single time database scanning. So it reduces the execution time and memory requirement.

VI. REFERENCES

[1]. Agrawal and Rakesh,"Fast Algorithms for Mining Association Rules in Large Databases",Proceedings of the ACM SIGMOD International Conference Management of Data,Washington,Vol. 22,Issue 2,pp.207-216,1993.

[2]. Jiawei Han and Micheline Kamber,Book:"Data Mining,Concept and Techniques".

[3]. G. Piatetski-Shapiro,"Discovery,Analysis,and Presentation of Strong Rules. In Knowledge Discovery in Databases",Piatetski-Shapiro,G.,editor,AAAI/MIT Press,pp. 229-248,1991.

[4]. Ye,Yanbin,and Chia-Chu Chiang. "A parallel apriori algorithm for frequent itemsets mining." In Software Engineering Research,Management and Applications",Fourth International Conference on,pp. 87-94. IEEE,2006.

[5]. R. Agrawal and R. Srikant,"Fast Algorithms for Mining

[6]. Association Rules," Proceedings of the 20th International Conference on Very Large Data Bases,pp. 487-499,1994.

[7]. Archana Singh and Jyoti Agarwal,"Proposed algorithm for frequent itemset generation",Seventh IEEE International Conference on Contemporary Computing,pp. 160-165,2014.

[8]. DixitaTandel,Neha Soni,"Survey on Improvement of Apriori Algorithm",International Journal of Innovative Research in Computer and Communication Engineering Vol. 5,Issue 1,ISSN 2320-9801, January 2017.

[9]. X. Luo and W. Wang,"Improved AlgorithmsResearch for Association Rule Based on Matrix," International Conference on intelligent computing and Cognitive Informatics,pp. 415–419,Jun. 2010.

[10]. Jiaoling Du,Xiangli Zhang,Hongmei Zhang and Lei Chen,"Research and Improvement of Apriori Algorithm",IEEE Sixth International Conference on Science and Technology,pp.117-121,2016.

[11]. Jaishree Singh, Hari Ram, Dr. J.S. Sodhi," Improving Efficiency of Apriori Algorithm Using Transaction Reduction",International Journal of Scientific and Research Publications Volume 3,Issue 1,ISSN 2250-3153,January 2013

[12]. Jiao Y.,"Research of an Improved Apriori Algorithm in Data Mining Association Rules",International Journal of Computer & Communication Engineering,2(1)),pp.25-27,2013.

[13]. Singh,Archana,Jyoti Agarwal,and Ajay Rana,"Performance measure of similis and FP-growth algorithm",International Journal of Computer Applications,2013.

[14]. Han,Jiawei,Jian Pei,and Yiwen Yin,"Mining frequent patterns without candidate generation.",ACM sigmoid record,Vol. 29. No. 2,2000.

© 2018 IJSRSET | Volume 4 | Issue 2 | Print ISSN: 2395-1990 | Online ISSN : 2394-4099
National Conference on Advanced Research Trends in Information and
g Technologies (NCARTICT-2018), Department of IT, L. D. College of Engineering,
Ahmedabad, Gujarat, India In association with
International Journal of Scientific Research in Science, Engineering and Technology

A Comprehensive study of Geometric and Appearance based Facial Expression Recognition Methods

Nazeerah A. Sheth[*1], Mahesh M. Goyani[2]
[*1]Department of Computer Engineering, Government Engineering College, Modasa, Gujarat, India
[2]Department of Computer Engineering, Government Engineering College, Modasa, Gujarat, India

ABSTRACT

It is a well-known fact that facial expressions are one of the key reflectors of the emotional state of a person and the research on the same has been spanning for a long time. Being an essential requirement in Human Computer Interaction as well as other applications such as automobile safety, mental health detection, animations, etc. recognizing facial expressions with precision has become vital. This paper presents a survey on various important and effective techniques present in literature along with their variations used recently. Prominent techniques of each step and a detailed discussion on feature extraction methods have been provided along with a detailed comparison of few recent approaches.

Keywords: Facial Expressions, Survey, Emotion Recognition

I. INTRODUCTION

Facial Expressions are one of the most well-known types of non-verbal communication [1]. A single expression can perhaps convey more information than thousands of words combined as it reflects the emotional state of a person which a human brain can efficiently recognize and interpret. Automatic Facial Expression Recognition (AFER) intends to transfer a certain, if not full level of that ability to computers. With the advancements and increasing use of Human-Computer Interaction (HCI) in many applications, the need for effective AFER is all the more necessary.

Apart from socially sensitive Human-Computer Interaction [2], AFER can be utilized in many applications such as detection of mental disorders [3], safety against road rage [4], security [5], animations and video games [6], automate applications, etc. It can also be used to include emotion-related information in automatic image captioning systems [7].

Earlier, the facial expression related study was confined to psychology, medical, artistic or acting fields [8], but with the availability of high computing resources, new and smart technologies, and increasing research in image processing and machine learning, it has garnered a significant amount of interest from computer scientists as well [9]. The study of AFER can be traced back to Darwin, who in his popular book "The Expression of the Emotions in Man and Animals" [10] established general principles of expression and grouped them into various categories, for example, anger and fear combined in one category. He also classified the facial deformations that occur within each category. In 1971, Ekman and Friesen classified emotions into six main categories, namely angry, sad, happy, fear, disgust and surprise [11] as shown in Fig. 1 [12]. To this date, these six emotions are considered to be universal across different races, cultures, gender and ethnicities [13]. With the availability of high computational power and advances in fields of Computer Vision and Robotics, a lot of work has been done on AFER since the 1990's [9].

AFER can be looked at from various perspectives like real-time vs non-real time, geometric vs appearance based, message vs sign judgement, global vs local, static vs dynamic, etc. Some of the significant work done on AFER seems to be biased towards non-real time systems as there are various issues such occlusions, pose

variation, low-intensity input, identification of baseline frame, etc. that holds back the performance levels in real-time [14].

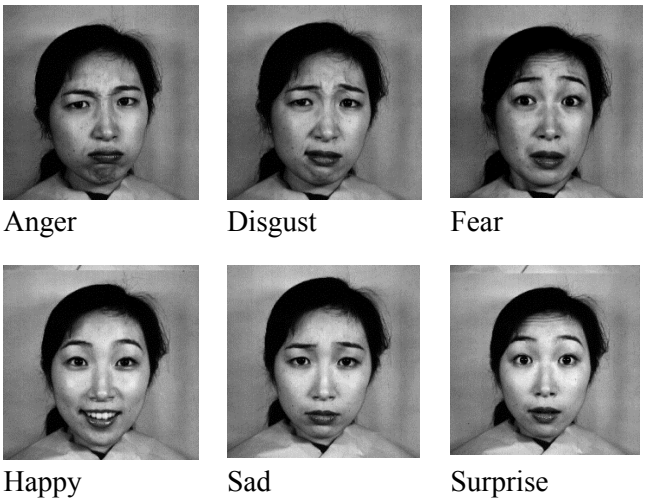


Figure 1. Basic six prototypical expressions

Geometry based approaches consider the position, curvature and deformation information of facial components to extract facial features whereas the appearance based approaches filter the intensity values. Also, both these approaches can be combined in order to combine their advantages. Both these methods will further be discussed in detail in section 3. Features extracted could either be holistic, i.e. generated from the whole face or part based, i.e. generated from specific region of interests like eyes, eyebrows, lips etc. [15].

Irrespective of the perspective followed, a basic AFER system comprises of mainly the stages as shown in Figure 2 [16]. Pre-processing includes enhancing the image in order to facilitate the further process and reduce any such possible conditions that might affect the recognition performance. It also includes Face detection and tracking which involves detecting the face region from the frame and tracking it across subsequent frames. The next stage deals with extracting the expression related features from the face followed by reducing them in order to filter out redundant features and decrease the complexity. The last step includes feeding the features to a classifier which then narrows down to one of the expression classes. The detailed survey on facial expression recognition can be found in [8], [17], [18].

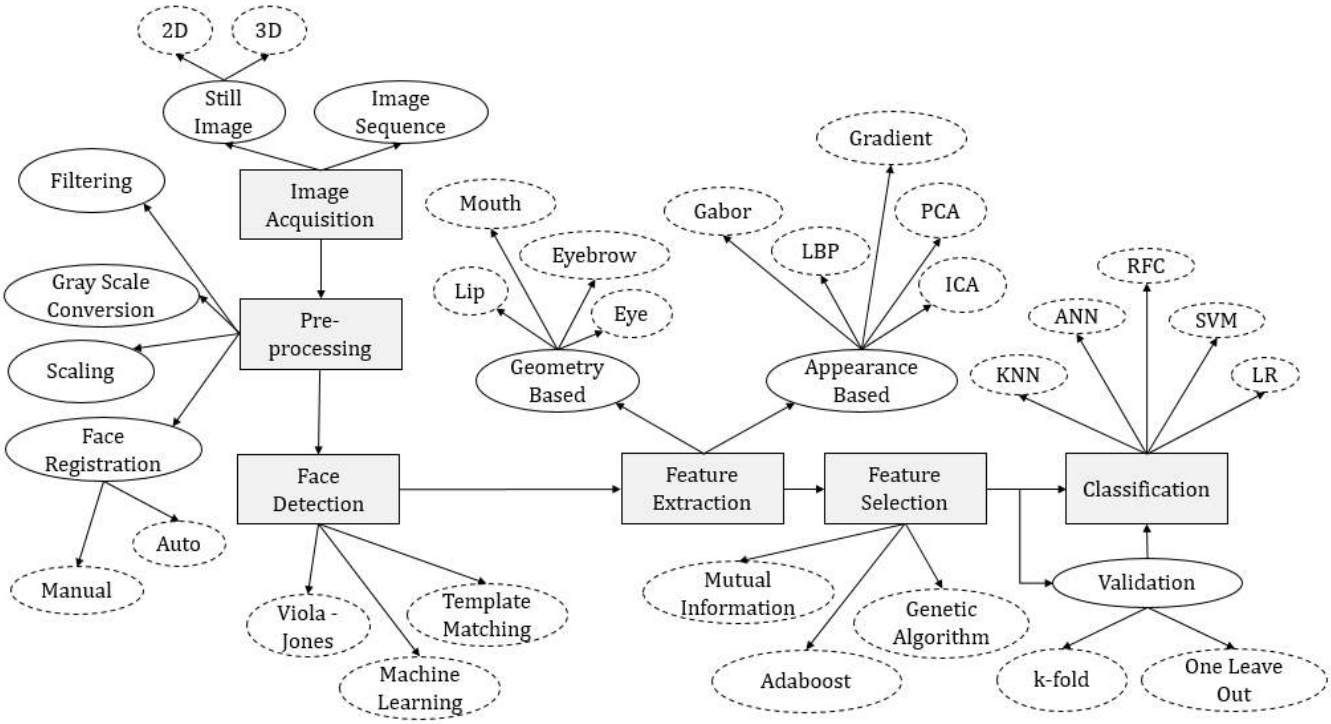


Figure 2. Basic AFER system

The rest of the paper is structured as follows. Section 2 (State Of The Art Methods) contains an in-depth review of various feature extraction techniques including both geometric and appearance based methods followed by Section 3 (Comparison and Analysis) which provides a detailed comparison of each step of various approaches.

Section 4 (**Conclusion and Future Scope**) concludes with various challenges and future scope of the system.

II. STATE OF THE ART METHODS

Feature extraction is probably the most important step in the FER process as it involves identifying and extracting discriminating features that contribute towards various facial expressions. Good features maximize between class dissimilarity and minimize within-class differences and are easily extracted with lower dimensions [19]. These features can be mostly found in the regions around eyes, mouth, nose, and face edges [20]. Methods that extract features relative to the whole face coordinates are known as global methods, while those that extract features relative to inner facial features or regions of interest are known as local methods [21], [15].

Most methods, such as [22], [23], [24], [25], [26] use global feature extraction methods so that a complete set of features are used for better classification purpose. However, approaches like [27], [20], [28], [29] used local techniques so that any redundant or unusable information is not taken into consideration. Another reason for choosing local methods is that they are not affected by face geometry, ageing, varying pose and face rotation [8]. Also, they are computationally less expensive as fewer number of features have to be processed. However, care has to be taken in choosing regions of interest so that vital discriminative features are not lost. Both these techniques can also be combined together as in [30], [31], [32] to balance the tradeoffs. Both local and global methods can generally be classified into geometric and appearance based techniques, an example of which is shown in Figure 3.

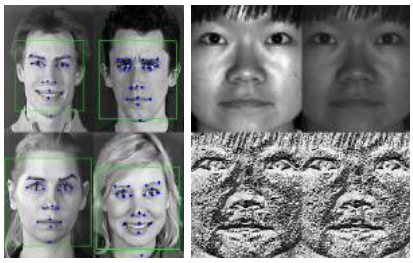


Figure 3. Geometry (left) and Appearance (right) based features

A. Geometric Based Feature Extraction

The Facial Action Coding System (FACS) [33] developed by Ekman and Friesen determined Action Units (AUs) based on landmark facial points whose movements were traced to define facial expressions. Action Unit is a measurement of smallest possible change in facial muscle. There are 44 Action Units defined consisting of various parts of a face, some of which are shown in Figure 4. Based on this approach, a lot of geometric methods have been used for the recognition of facial expressions.

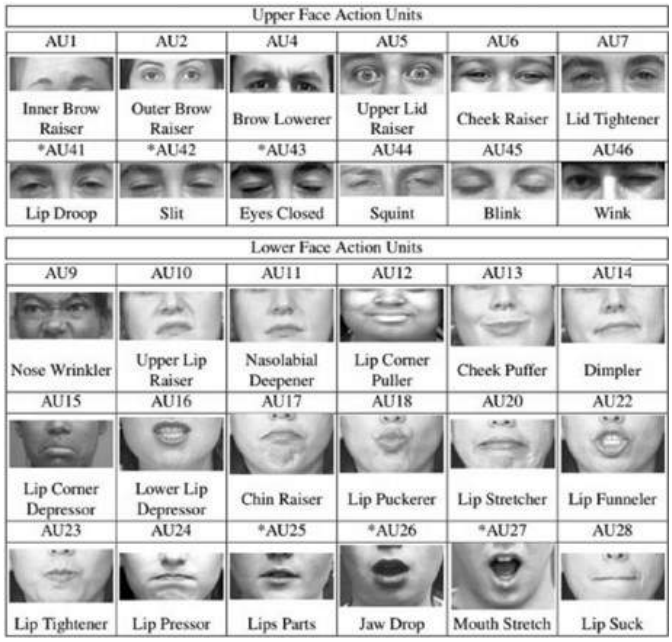


Figure 4. Action Units of upper and lower face parts

Geometric methods do not focus on face texture or intricate settings, rather they take feature indications from the geometry, deformation and tracking of fiducial points [16]. Usually, the methods require a reference to a baseline frame. Features are then extracted by tracking the landmarks through subsequent frames. For example, [34] uses triangles formed by 3 out of 68 landmarks, inspired by polygon mesh representation. The changes in angles and areas of these triangles are evaluated between consecutive frames to extract features. Ghimire et al. [35] used face graph normalization to bring all graphs to normal shape before extracting points, lines and triangles. These are determined by initializing and tracking facial points through a suitable technique.

Since few years, many researches based on 3D facial geometry have been proposed owing to the robustness of 3D image to scale, pose and illumination variations [36]. Zeng et al. [37] proposed a unique surface conformal representation in the form of Conformal Factor Image (CFI) and Mean Curvature Image (MCI). A conformal

map is computed and normalized which preserves the angles of facial geometry along with 3 main landmark points which are then used to extract features. Lemaire et al. [38] use 3D depth maps to generate Differential Mean Curvature Maps (DMCP) which improves multi-scale facial expression surface topology uniqueness. The DMCP are then normalized and Histogram of Gradients (HOG) [39] algorithm is used to generate features. Berretti et al. [36] identifies a set of facial key points and computes Scale Invariant Feature Transform (SIFT) [40] feature descriptors from depth images of the face around key points.

B. Appearance Based Feature Extraction

Geometric methods face many shortcomings such as dependency on face geometry and pose, failure to detect landmarks in case of occlusions, no tolerance against face representation errors, etc. [16] Owing to the various pitfalls of geometric methods, appearance based feature extraction methods have gained much popularity in terms of higher accuracy and lower error conditions [16]. Appearance based features take cue from the intensity levels of the face image to determine features contributing to facial expression. These features are extracted by generating a suitable filter which is the convolved around the face [21], [15], [16].

Among the pool of various appearance features, the ones with binarized local texture have shown promising results and effectiveness [21], [19]. In 1996, Ojala et al. [41] formulated the Local Binary Patterns (LBP) which is an efficient and easy non-parametric method to describe mild and intricate texture information and summarize the local structure of facial components. Earlier it was used for face description by Ahonen et al. [42] and texture description by Ojala et al. [43]. Later, many LBP based facial expression recognition systems have been [44], [45], [46], [47], [48]. Huang et al. [19] provides a comprehensive study on various LBP techniques on facial image analysis along with their variations. LBP is robust to monotonic illumination change and misalignment [21] and is computationally simple [49]. However, since the original LBP operates on a 3×3 neighbourhood, it may miss out vital structural information [50]. Therefore Improved LBP was proposed to include neighbourhood of all sizes to capture complete information [43].

Local Directional Pattern (LDP) is another technique proposed by Jabid et al. [51], to address the non–robustness problem of LBP towards random noise and non-monotonic illumination. It consists of directional information by comparing edge response values of each pixel in eight directions using Kirsch masks representing the impact of edge in each direction. The feature shows superior performance and can be represented in low-dimensional feature space with high accuracy for even low-resolution images. However, the LDP codes can be problematic in smooth regions as they focus more on edge response values and produces inconsistent patterns in uniform regions [52]. Jun et al. determined a new robust local gradient coding (LGC) which further implied that gradient differences among pixels improved accuracy [53]. Local Directional Pattern Variance (LDPv) was developed by Kabir et al. [54] to include contrast information along with edge response values by introducing variance as an adaptive weight to modify the LDP code.

Local Ternary Patterns [55] with an extra discrimination level and ternary codes were introduced to address limitations of LDP and also to tackle non-uniform noise. Ahmed et al. [56] proposed Gradient Local Ternary patterns (GLTP) which combines the advantages of LTP and LBP by encoding more robust gradient magnitude values in a three-level scheme to achieve consistent texture patterns in random noise and varying illumination. An improved version of GLTP with a better gradient operator and dimensionality reduction technique was used by Holder et al. [20] which is shown to give better results.

Recently, Ryu et al. [26] defined Local Directional Ternary Pattern (LDTP) which efficiently encodes edge directional information of emotion-related features in edge regions and neglects the less significant smooth regions based on magnitude encoded within the ternary pattern. It is shown to be a highly discriminable and robust pattern, but the recognition rates drop significantly with varying camera angles and ethnic variations [57]. Arshid et al. [31] proposed a Multi Stage Binary Pattern (MSBP) which aims to tackle real-world recognition issues such as varying features, intricate settings and complex backgrounds by retaining local texture variations along with gradient changes along edges such as eyes, wrinkles etc.. In MSBP, a multi-

stage binary code is generated for each comparison against neighborhood pixel.

Recently, Goyani and Patel presented two local appearance based FERS techniques: Local Mean Binary Operator (LMBP) [58] and Multi-Level Haar wavelet (MLH) based FERS [59]. MLBP computes the 256 bin histogram of the LBP code of the mean centred 3 × 3 patch, whereas MLH extracts approximation coefficients of the image at multiple scales using Haar wavelets. Haar are inherently good at noise suppression where as LMBP achieves the robustness to noise and non-monotonic change in illumination.

Gabor filter is a linear parametric filter which analyses whether there are any specific frequency content in the image in specific directions in a localized region around the region of analysis [60]. Gabor features suffer from identity bias [16] and hence used for identity recognition in many places [61], [62]. Lyons et al. [61] first used Gabor wavelets to generate facial expression features. It has proved to be an effective method and has since been used by many different approaches like in [63], [64], [65], [66], [67], for facial expression recognition. Local Phase Quantization (LPQ) was a technique used by [68] for blur insensitive texture classification which computes Short Term Fourier Transform (STFT) on local image window. It is then quantized using a scalar binary quantizer for calculating the phase information. Though it is computationally expensive than LBP, Dhall et al. [69] showed that it gives better recognition rates for facial expressions recognition. Histogram of Oriented Gradients (HOG) [70] is a descriptor that counts the occurrences of localized gradient orientation in images and is used for object recognition since it is efficient in encoding shape information. PHOG is an extension to HOG in which gradients in each grid are joined at pyramid level which is shown to give better performances. PHOG have been used for static facial expression analysis in [71] and [72]. To address the problems of high dimensionality and low variance between expression classes, another common technique employed for feature extraction is Fisherfaces Linear Discriminant Analysis (FLDA) [73] in which the discriminating features of an image are preserved with reduced dimensions. It intends to reduce intra class variance and increase inter class variance but fails in case of multimodal classes [73]. To overcome this, Local Fisher Discriminant Analysis (LFDA) was introduced [74], [75]. In this, local between class variance is increased while minimizing the local within-class variance. But it fails to determine essential mixed structure when face image space is highly nonlinear [25]. In order to solve the variance problem along with reduced dimensions, several more methods are present like Kernel Discriminant Analysis (KDA) [76], General Discriminant Analysis (GDA) [77] and Linear Discriminant Analysis (LDA) [76]. LDA is more widely used but is less flexible when comes to complex datasets [25]. Siddiqi et al. [25] used an enhanced version of LDA called Stepwise Linear Discriminant Analysis (SWLDA) which is computationally less expensive with higher predictive ability and gives better results. Another line of approach seen in literature is to cascade two or more effective methods one after the other. In this, one method is applied over the response of the other. One such pattern, LGBP (Local Gabor Binary Pattern) was proposed by Zhang et al. [78] in which Gabor features capture orientations and scales while LBP focuses subtler texture details. LGBP is robust to variations in lighting and expressions. Among others, LGBP was also used in [79] where expressions were classified into dominant and complementary emotions like happily surprised to cover a wider range of emotions instead of basic prototypes. In a recent approach, Sun et al. proposed a novel method in [23] in which instead of applying LBP over the features described by Gabor, both Gabor and LBP features are extracted from the input image and then those features are fused together using feature fusion to form the final feature vector. The approach has yielded superior performance since features described using different perspectives are used together. Similar method was proposed in [80] where Global features extracted by PCA (discussed under feature reduction) are fused with local features extracted from the mouth area using LBP. PCA is an effective method as it reduces dimensionality while selecting distinct features, but the features extracted by it are subject to environment changes [80]. This is combatted by fusing in LBP which focuses on texture details and hence the results show improved robustness. Since LBP is sensitive to random noise and non-monotonic illumination, this method was improved by Luo et al. in [32] where instead of LBP, LDP is used to extract local features from the eyes and mouth region. The method shows improved performance than PCA and LBP combined, since LDP has good stability against random noise. Rajesh et al. experimented various fusion

of different methods in [81] such as fusion of LBP with LGC, HOG with LDP, and HOG with wavelets. The results showed that fusion of HOG with wavelets outperformed all others and fusion of LDP with HOG improved LDP's performance. LMBP operator presented by Goyani and Patel [58] is computationally efficient compared to LDP as well it is robust to noise and illumination.

So far, the methods we have discussed work on static facial images with no temporal information. Image sequences can be used to incorporate temporal information which may represent expressions from onset to peak to further offset. Static images usually just include the peak expression image. Image sequences help to incorporate expression at different levels and also makes it easier to differentiate different expressions and have shown better descriptions as shown by [82]. The only drawback is that it takes a toll on the computation as more number of images have to be processed.

Few of the methods discussed earlier have been extended to include spatiotemporal information. LBP was extended to LBP-TOP [82] to include multiple frames. To avoid high computations, feature vectors from only three orthogonal planes namely, XY, XT and YT are concatenated to describe facial expressions [21]. The XY plane gives the static information while temporal information is provided by the other two [83]. Similarly, LPQ and LGBP are also extended to LPQ-TOP [83] and LGBP-TOP [84]. But as indicated in [27], these kind of methods lack: 1). temporal correspondences among different phases of an expression, and 2). semantic representation, since these low-level representations do not convey any semantic meaning of individual parts of an expression. For example, if the indicated expression is happy, then the low-level representation do not convey a smiling component. To overcome both these shortcomings, [27] have proposed a mid-level representation through expressionlets. Expressionlets aim to bridge the gap between low level representations and high level semantics by modelling each video clip as a Spatio Temporal Manifold (STM) composed of low level features. A Uniform Manifold Model (UMM) is built based on all STMs and each expression is then fitted as an instance of the UMM. Owing to the novel solution, this method has shown significant improvement over other video based recognition methods, but still lacks accuracy for challenging real-time datasets. Spatio Temporal Texture Map (STTM) [22] was developed to capture subtle spatial and temporal variations of expressions with low computational cost. It does so by extracting information from 3D Harris Corner function and representing the features in the form of histograms. The model is better suited for real-time expressions considering the lower computational cost. Compared to expressionlets it gives much better performance recognizing the disgust emotion, but the accuracy for others falls slightly behind. Another dynamic expression recognition approach was developed recently by Kamarol et al. [85] which couples intensity estimation along with expression recognition with low level computation. The model is based on the fact that each person expresses emotions at different intensities and thus a system that adapts to individual person would be useful in applications like pain detection from videos.

III. COMPARISON AND ANALYSIS

The previous section evaluated and analyzed various methods for feature extraction starting LBP and its variations developed over the years. These methods have shown great performance in representing mild texture details and have managed to overcome various challenges pertaining to effective FER. Parametric methods that use frequency and phase information have also shown to be effective. The effectiveness of methods that combine two or more techniques in different ways indicates that right combination of techniques can significantly improve recognition rate. FER systems based on image sequences are used to include time-based information and different levels of expressions and have shown better performance.

In this section we present a comparison of various FER approaches on various parameters and techniques used for each step. Table 1 provides the FER systems with geometry based feature extraction whereas Table 2 shows the systems with appearance based feature extraction.

Table 1. Comparision Of Geometry Based Facial Expression Recognition Methods

Ref.	PP	FE	FS	S/D	L/G	Classifier	NOE	Val	Dataset	P/S	Accuracy
[24]	-	ASM	-	S	G	Double SVM, sample selected SV	6	-	JAFFE	P	98.25
[34]	-	Angles, areas by landmark points	Exhaustive analysis of 68 landmarks	D S	G	CRF c-kNN (citation Nearest neighbour)	7	LOO	CK+	P	86.7 82.2
[35]	Face graph normalization, V&J, EBGM, KLT– point detection & track	Point, line, triangles	FS AdaBoost	D	G	SVM	6	10F CD	CK+ MMI MUG	P S P	97.80 77.22 95.50
[38]	Align model to frontal pose, cropping	DMCM HOG	DMCM Norm	D	G	SVM	7	10F	BU-3DFE	P	76.6
[36]	-	SIFT	mRMR	S	G	SVM	7	10F	BU-3DFE	P	78.43
[85]	Procrustes analysis	AAM	-	D	G	HMM	6	SI 10F LOO	CK+ BUFED	P P	82.4 62.86

Table 2. Comparision Of Appearance Based Facial Expression Recognition Methods

Ref	PP	FE	DR	FS	S/D	G/L	Classifier	NOE	Val	Dataset	P/S	Accuracy
[46]	Re-projected from 3D model to 5 images of different angles, V&J	LBP and variations	-	-	S	L G	SVM	6	10F	BU-3DFE MPIE	P	67.96 73.26
[86]	Regression Trees for FD	LBP NCM	ES	-	S	L	SVM	6 7	5F	CK+	P	94.83 97.25
[23]	Rotation, BG crop, Histogram equilization, AdaBoost	Gabor + LBP	PCA	LDA	S	G	SVM	6 7	10F	CK+	P	97.42(6) 95.45(7)
[27]	Normalization	Expressionlet	PCA	DL	D	L	Multi class linear SVM	6 7	PI 10F	CK+, MMI, CASIA, AFEW	P P P S	94.19 75.12 74.59 31.73
[22]	BG crop, V&J	STTM (3D harris corner)	-	-	D	G	Multi class SVM	6	2F	CK+, CASME II, AFEW	P S S	97.70 98.61 90.68
[25]	-	SWLDA	-	-	D	G	Hidden CRF	6	10F nF LOO	CK+ JAFFE B+ MMI	P P P P	96.37

[32]	Normalization, Adaboost detection	PCA + LDP	-	-	S	G L	SVM	7	LOO	-	P	91.61
[31]	Normalization, V & J	MSBP	Divide MSBP into 2 parts	-	S	G L	BF tree Simple-logistic kNN Bagging NB	7	10F	SFEW	S	96 (G) 60 (L)
[87]	Illumination enhancement, Face location, Normalization BG crop, V & J	LGBP	-	-	D	L	SVM	30+	Votin g	CK+ JAFFE	P	73.88
[26]	Normalization	LDTP	-	-	S	L	SVM	7	NF LOO	CK+ JAFFE MMI CMU-PIE GEMEP-F BU-3DFE	P	94.2 94.8 99.8 89.5 98.1 88.1
[20]	Crop face & components, Landmark locations and V&J	Improved GLTP	PCA	-	S	L	SVM	6 7	10F LOO PI	CK+ JAFFE	P P	95.17 85.05
[67]	Normalization, Integral projection of horizontal and vertical direction to detect eye lips	Gabor	-	FSDD	S	L	Local weighted AdaBoost	2	-	JAFFE CK	P P	96.17 95.10
[66]	Normalization, V & J	Gabor Template	-	M-FS, TM	S	G	Norm based SVM	7	10F	CK JAFFE	P P	95.1-90.8 80.3 78.4
[83]	Procrustes Transform, Automatic point detector	LPQ-TOP	-	GB	D	G	SVM MM	AUs	nF CD	MMI UNBCSAL GEMEP-FERA CK SEMAINE	P S S P P S	-
[84]	Normalization, Local Evidence Aggregated Regression (LEAR) detect points	LGBP-TOP	-	F1 based FS	D	G	SVM	AUs	10F SI CV	MMI CK	P P	-

PP- Pre Processing, FE- Feature Extraction, DR- Dimensionality Reduction, FS- Feature Selection, S-Static, D-Dynamic, L-Local, G-Global, NOE-Number of Expressions, P-Posed, S-Spontaneous, Val-Validation, BG-Background, nF- n-Fold Cross Validation, CD- Cross-Database, PI- Person Independent, LOO- Leave One Out

IV. CONCLUSION

This paper presented a review of the various state of the art techniques used for facial expression recognition. The survey starts with a quick description about the importance, applications and some background information about facial expression recognition systems, followed by a take on various perspectives an AFER system can be looked at. Various steps of a basic AFER system are then discussed including the assessment of numerous methods and techniques corresponding to each step. An in-depth analysis of various feature extraction methods has been provided including their motives and shortcomings.

The study has shown that despite a lot of research and achievements in the literature, there is still scope for improvement considering the dynamic nature of facial expressions and various challenges associated with it. For an AFER system to be effectively implemented, we need the system to be robust and accurate. This precision can be reached if all the challenges that reduce the efficiency are tackled. As it is observed that different techniques address different problems, it makes sense that a certain type of combination of some of them would achieve the desired goal. It is also deduced that instead of just the prototypical expressions, the range of expressions can be widened to cover more detailed and subtle emotions.

V. REFERENCES

[1]. A. Mehrabian, "Communication without words," Psychology Today, vol. 2. pp. 53-55, 1968.

[2]. F. Cid, J. A. Prado, P. Bustos, and P. Nunez, "A Real Time and Robust Facial Expression Recognition and Imitation approach for Affective Human-Robot Interaction Using Gabor filtering," IEEE International Conference on Intelligent Robots and Systems, pp. 2188-2193, 2013.

[3]. W. Gaebel and W. Wolwer, "Facial expression and emotional face recognition in schizophrenia and depression," European Archives of Psychiatry and Clinical Neuroscience, vol. 242, no. 1, pp. 46-52, 1992.

[4]. F. Nasoz, K. Alvarez, C. L. Lisetti, and N. Finkelstein, "Emotion recognition from physiological signals using wireless sensors for presence technologies," Cognition, Technology & Work, vol. 6, no. 1, pp. 4-14, 2004.

[5]. A. Butalia, M. Ingle, and P. Kulkarni, "Facial Expression Recognition for Security," International Journal of Modern Engineering Research, vol. 2, no. 4, pp. 1449-1453, 2012.

[6]. M. S. Bartlett, G. Littlewort, I. Fasel, and J. R. Movellan, "Real Time Face Detection and Facial Expression Recognition: Development and Applications to Human Computer Interaction," IEEE Conference on Computer Vision and Pattern Recognition Workshop, vol. 5, pp. 53-53, 2003.

[7]. R. Cowie et al., "Emotion recognition in human-computer interaction," Signal Processing Magazine, IEEE, vol. 18, no. 1, pp. 32-80, 2001.

[8]. B. Fasel and J. Luettin, "Automatic facial expression analysis: A survey," Pattern Recognition, vol. 36, no. 1, pp. 259-275, 2003.

[9]. V. Bettadapura, "Face Expression Recognition and Analysis : The State of the Art," arXiv preprint arXiv:1203.6722, pp. 1-27, 2012.

[10]. Charles Darwin, The expression of the emotions in man and animals. Oxford University Press, USA, 1998.

[11]. P. Ekman, "Universal and Cultural differences in facial expressions of emotions," Nebraska Symposium on Motivation, vol. 19. pp. 207-282, 1971.

[12]. M. J. Lyons, "Automatic classification of single facial images," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 21, no. 12, pp. 1357-1362, 1999.

[13]. P. Ekman, "Strong Evidence for universals in facial expressions," Psychol. Bull, vol. 115, no. 2. pp. 268-287, 1994.

[14]. S. Deshmukh, M. Patwardhan, and A. Mahajan, "Survey on real-time facial expression recognition techniques," IET Biometrics, vol. 5, no. 3, pp. 243-251, 2016.

[15]. C. A. Corneanu, M. O. Simón, J. F. Cohn, and S. E. Guerrero, "Survey on RGB, 3D, Thermal, and Multimodal Approaches for Facial Expression Recognition: History, Trends, and Affect-Related Applications," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 38, no. 8, pp. 1548-1568, 2016.

[16]. E. Sariyanidi, H. Gunes, and A. Cavallaro, "Automatic analysis of facial affect: A survey of registration, representation, and recognition,"

IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 37, no. 6, pp. 1-22, 2015.

[17]. M. Pantic and L. J. M. Rothkrantz, "Automatic analysis of facial expressions : the state of the art," IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 22, no. 12, pp. 1424-1445, 2000.

[18]. M. Goyani, "A Comprehensive Study on Texture Feature Based Facial Expression Recognition," 2017.

[19]. D. Huang, C. Shan, M. Ardebilian, and L. Chen, "Facial Image Analysis Based on Local Binary Patterns : A Survey," IEEE Transactions on Image Processing, 2011.

[20]. R. P. Holder and J. R. Tapamo, "Improved gradient local ternary patterns for facial expression recognition," EURASIP Journal on Image and Video Processing, vol. 2017, no. 1, pp. 42-57, 2017.

[21]. B. Martinez, M. F. Valstar, B. Jiang, and M. Pantic, "Automatic Analysis of Facial Actions : A Survey," IEEE Transactions on Affective Computing, 2017.

[22]. S. K. A. Kamarol, M. H. Jaward, J. Parkkinen, and R. Parthiban, "Spatiotemporal feature extraction for facial expression recognition," IET Image Processing, vol. 10, no. 7, pp. 534-541, 2016.

[23]. Y. Sun and J. Yu, "Facial Expression Recognition by Fusing Gabor and Local Binary Pattern Features," International Conference on Multimedia modeling. Springer, pp. 209-220, 2017.

[24]. T. Yu and X. Gu, "Facial Expression Recognition Using Double-Stage Sample-Selected SVM," International Conference on Intelligent Computing, Springer, pp. 304-315, 2017.

[25]. M. H. Siddiqi, R. Ali, A. M. Khan, Y. T. Park, and S. Lee, "Human Facial Expression Recognition Using Stepwise Linear Discriminant Analysis and Hidden Conditional Random Fields," IEEE Transactions on Image Processing, vol. 24, no. 4, pp. 1386-1398, 2015.

[26]. B. Ryu, A. Ramirez Rivera, J. Kim, and O. Chae, "Local Directional Ternary Pattern for Facial Expression Recognition," IEEE Transactions on Image Processing, vol. 26, no. 12, pp. 6006-6018, 2017.

[27]. M. Liu, S. Shan, R. Wang, and X. Chen, "Learning expressionlets on spatio-temporal manifold for dynamic facial expression recognition," in IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2014, pp. 1749-1756.

[28]. D. Ghimire, S. Jeong, J. Lee, and S. H. Park, "Facial expression recognition based on local region specific features and support vector machines," Multimedia Tools and Applications, vol. 76, no. 6, pp. 1-19, 2016.

[29]. A. Uçar, Y. Demir, and C. Güzeliş, "A new facial expression recognition based on curvelet transform and online sequential extreme learning machine initialized with spherical clustering," Neural Computing and Applications, vol. 27, no. 1, pp. 131-142, 2016.

[30]. A. Mollahosseini, D. Chan, and M. H. Mahoor, "Going deeper in facial expression recognition using deep neural networks," IEEE Winter Conference on Applications of Computer Vision, pp. 1-10, 2016.

[31]. S. Arshid, A. Hussain, A. Munir, A. Nawaz, and S. Aziz, "Multi-stage binary patterns for facial expression recognition in real world," Cluster Computing, pp. 1-9, 2017.

[32]. Y. Luo, T. Zhang, and Y. Zhang, "A novel fusion method of PCA and LDP for facial expression feature extraction," International Journal for Light and Electron Optics, vol. 127, no. 2, pp. 718-721, 2016.

[33]. P. Ekman and W. Friesen, Manual for the Facial Action Coding System. Consulting Psychologist Press, Palo Alto, 1977.

[34]. D. Acevedo, P. Negri, M. E. Buemi, F. G. Fernandez, and M. Mejail, "A Simple Geometric-Based Descriptor for Facial Expression Recognition," 12th IEEE International Conference on Automatic Face & Gesture Recognition, pp. 802-808, 2017.

[35]. D. Ghimire, J. Lee, Z. N. Li, and S. Jeong, "Recognition of facial expressions based on salient geometric features and support vector machines," Multimedia Tools and Applications, vol. 76, no. 6, pp. 7921-7946, 2016.

[36]. S. Berretti, B. Ben Amor, M. Daoudi, and A. Del Bimbo, "3D facial expression recognition using SIFT descriptors of automatically detected keypoints," Visual Computer, vol. 27, no. 11, pp. 1021-1036, 2011.

[37]. W. Zeng, H. Li, L. Chen, J.-M. Morvan, and X. D. Gu, "An automatic 3D expression recognition framework based on sparse representation of conformal images," 10th IEEE International Conference and Workshops on Automatic Face and Gesture Recognition, pp. 1-8, 2013.

[38]. P. Lemaire, M. Ardabilian, L. Chen, and M. Daoudi, "Fully automatic 3D facial expression recognition using differential mean curvature maps and histograms of oriented gradients," 10th IEEE International Conference and Workshops on Automatic Face and Gesture Recognition, pp. 1-7, 2013.

[39]. Dahmane, Mohamed, and J. Meunier, "Emotion recognition using dynamic gridbased HoG features," in IEEE International Conference on Automatic Face and Gesture Recognition and Workshops, 2011, pp. 884-888.

[40]. Lowe and G. David, "Distinctive image features from scale-invariant key points," International Journal of Computer Vision, vol. 60, no. 2, pp. 91-110, 2004.

[41]. T. Ojala, M. Pietikäinen, and D. Harwood, "A comparative study of texture measures with classification based on featured distributions," Pattern Recognition, vol. 29, no. 1, pp. 51-59, 1996.

[42]. T. Ahonen, A. Hadid, and M. Pietikäinen, "Face description with local binary patterns: application to face recognition," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 28, no. 12, pp. 2037-2041, 2006.

[43]. T. Ojala, M. Pietikäinen, and T. Mäenpää, "A generalized local binary pattern operator for multiresolution gray scale and rotation invariant texture classification," Advances in Pattern Recognition, pp. 399-408, 2001.

[44]. C. Shan, S. Gong, and P. W. Mcowan, "Facial expression recognition based on Local Binary Patterns : A comprehensive study," Image and Vision Computing, vol. 27, no. 6, pp. 803-816, 2009.

[45]. C. Shan, S. Gong, and P. W. McOwan, "Robust facial expression recognition using local binary patterns," IEEE International Conference on Image Processing, vol. 2, pp. 170-173, 2005.

[46]. S. Moore and R. Bowden, "Local binary patterns for multi-view facial expression recognition," Computer Vision and Image Understanding, vol. 115, no. 4, pp. 541-558, 2011.

[47]. R. Hablani, N. Chaudhari, and S. Tanwani, "Recognition of facial expressions using Local Binary Patterns of important facial parts," International Journal of Image Processing, vol. 7, no. 2, pp. 163-170, 2013.

[48]. X. Zhao and S. Zhang, "Facial expression recognition using local binary patterns and discriminant kernel locally linear embedding," EURASIP Journal on Advances in Signal Processing, vol. 2012, no. 1, p. 20, 2012.

[49]. D. Huang, C. Shan, M. Ardabilian, Y. Wang, and L. Chen, "Local Binary Patterns and Its Application to Facial Image Analysis: A Survey," IEEE Transactions on Systems Man and Cybernetics Part C-Applications and Reviews, vol. 41, no. 6, pp. 765-781, 2011.

[50]. H. Jin, Q. Liu, H. Lu, and X. Tong, "Face detection using improved LBP under Bayesian framework," in International Conference on Image and Graphics, 2004, pp. 306-309.

[51]. T. Jabid, M. H. Kabir, and O. Chae, "Facial expression recognition using Local Directional Pattern (LDP)," 17th IEEE International Conference on Image Processing, pp. 1605-1608, 2010.

[52]. F. Ahmed and M. H. Kabir, "Directional ternary pattern (dtp) for facial expression recognition," in IEEE International Conference on Consumer Electronics, 2012, pp. 265-266.

[53]. B. Jun and D. Kim, "Robust face detection using local gradient patterns and evidence accumulation," Pattern Recognition, vol. 45, no. 9, pp. 3304-3316, 2012.

[54]. H. Kabir, T. Jabid, and O. Chae, "Local Directional Pattern Variance (LDPv):Robust Feature Descriptor for Facial Expression Recognition," The International Arab Journal of Information Technology, vol. 9, no. 4, pp. 382-392, 2012.

[55]. X. Tan and B. Triggs, "Enhanced local texture feature sets for face recognition under difficult lighting conditions," in Analysis and Modeling of Faces and Gestures, 2007, pp. 168-182.

[56]. F. Ahmed and E. Hossain, "Automated facial expression recognition using gradient-based ternary texture patterns," Chinese Journal of Engineering, pp. 1-8, 2013.

[57]. R. Ahmed, A. Meyer, H. Konik, and S. Bouakaz, "Framework for reliable , real-time facial expression recognition for low resolution images," Pattern Recognition Letters, vol. 34, no. 10, pp. 1159-1168, 2013.

[58]. M. Goyani and N. Patel, "Robust Facial Expression Recognition using Local Mean Binary Pattern," Electronic Letters on Computer Vision and Image Analysis, vol. 16, no. 1, pp. 54-67, 2017.

[59]. M. Goyani and N. Patel, "Multi-Level Haar Wavelet based Facial Expression Recognition using Logistic Regression," Indian Journal of Science and Technology, vol. 10, no. 9, pp. 1-9, 2017.

[60]. G. Donato, M. S. Bartlett, J. C. Hager, P. Ekman, and T. J. Sejnowski, "Classifying facial actions," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 21, no. 10, pp. 974-989, 1999.

[61]. M. Lyons and S. Akamatsu, "Coding Facial Expressions with Gabor Wavelets," 3rd IEEE Conference on Automatic Face and Gesture Recognition, pp. 200-205, 1998.

[62]. M. Lades et al., "Distortion Invariant Object Recognition in the Dynamic Link Architecture," IEEE Transactions on Computers, vol. 42, pp. 300-311, 1993.

[63]. W. Gu, C. Xiang, Y. V. Venkatesh, D. Huang, and H. Lin, "Facial expression recognition using radial encoding of local Gabor features and classifier synthesis," Pattern Recognition, vol. 45, no. 1, pp. 80-91, 2012.

[64]. F. Chen and K. Kotani, "Facial Expression Recognition by SVM-based Two-stage Classifier on Gabor Features," IAPR Conference on Machine Vision Applications, pp. 453-456, 2007.

[65]. M. J. Lyons, J. Budynek, A. Plante, and S. Akamatsu, "Classifying Facial Attributes using a 2-D GaborWavelet Representation and Discriminant Analysis," in 4th International Conference on Automatic Face and Gesture Recognition, 2000, pp. 202-207.

[66]. L. Zhang, D. Tjondronegoro, and V. Chandran, "Random Gabor based templates for facial expression recognition in images with facial occlusion," Neurocomputing, vol. 145, pp. 451-464, 2014.

[67]. Y. Xing and W. Luo, "Facial Expression Recognition Using Local Gabor Features and Adaboost Classifiers," IEEE Internation al Conference on Progress in Informatics and Computing, pp. 1-5, 2016.

[68]. V. Ojansivu and J. Heikkil, "Blur insensitive texture classification using local phase quantization," Image and Signal Processing, pp. 236-243, 2008.

[69]. A. Dhall, A. Asthana, R. Goecke, and T. Gedeon, "Emotion recognition using PHOG and LPQ features," IEEE International Conference on Automatic Face and Gesture Recognition and Workshops, pp. 878-883, 2011.

[70]. N. Dalal and B. Triggs, "Histograms of Oriented Gradients for Human Detection," International Conference on Computer Vision and Pattern Recognition, vol. 1, pp. 886-893, 2005.

[71]. Z. Li, J. Imai, and M. Kaneko, "Facial-component-based Bag of Words and PHOG Descriptor for Facial Expression Recognition," in International Conference on Systems, Man, and Cybernetics San, 2009, pp. 1353-1358.

[72]. Y. Bai, L. Guo, L. Jin, and Q. Huang, "A novel feature extraction method using pyramid histogram of orientation gradients for smile recognition," in 16th IEEE International Conference on Image Processing, 2009, pp. 3305-3308.

[73]. P. N. Belhumeur, J. P. Hespanha, and D. J. Kriegman, "Eigenfaces vs . Fisherfaces : Recognition Using Class Specific Linear Projection," vol. 19, no. 7, pp. 711-720, 1997.

[74]. M. Sugiyama, "Dimensionality Reduction of Multimodal Labeled Data by Local Fisher Discriminant Analysis," Journal of Machine Learning Research, vol. 8, pp. 1027-1067, 2007.

[75]. Y. Rahulamathavan, R. C. Phan, J. A. Chambers, and D. J. Parish, "Facial Expression Recognition in the Encrypted Domain Based on Local Fisher Discriminant Analysis," vol. 4, no. 1, pp. 83-92, 2013.

[76]. S. Mika, G. Ratsch, J. Weston, B. Scholkopf, and K. Muller, "Fisher Discriminant Analysis with Kernels," in IEEE Signal Processing Soceity Workshop Neural Network Signal Process IX, 1999, pp. 41-48.

[77]. G. Baudat and F. Anouar, "Generalized Discriminant Analysis using a Kernel Approach,"

Neural Computing, vol. 12, no. 10, pp. 2385-2404, 2000.

[78]. W. Zhang, S. Shan, W. Gao, and X. Chen, "Local gabor binary pattern histogram sequence (lgbphs): a novel non-statistical model for face representation and recognition," IEEE 10th International Conference on Computer Vision, vol. 1, pp. 786-791, 2005.

[79]. C. Loob et al., "Dominant and Complementary Multi-Emotional Facial Expression Recognition Using C-Support Vector Classification," IEEE 12th International Conference on Automatic Face & Gesture Recogition, pp. 833-838, 2017.

[80]. Y. Luo, C. Wu, and Y. Zhang, "Facial expression recognition based on fusion feature of PCA and LBP with SVM," International Journal for Light and Electron Optics, vol. 124, no. 17, pp. 2767-2770, 2013.

[81]. J. Kumari, R. Rajesh, and A. Kumar, "Fusion of features for the effective facial expression recognition," International Conference on Communication and Signal Processing, pp. 457-461, 2016.

[82]. G. Y. Zhao and M. Pietikainen, "Dynamic Texture Recognition using Local Binary Pattern with an Application to Facial Expressions," Transactions on Pattern Analysis and Machine Intelligence, vol. 2, no. 6, pp. 915-928, 2007.

[83]. B. Jiang, M. Valstar, B. Martinez, and M. Pantic, "A dynamic appearance descriptor approach to facial actions temporal modeling," IEEE Transactions on Cybernetics, vol. 44, no. 2, pp. 161-174, 2014.

[84]. T. R. Almaev and M. F. Valstar, "Local gabor binary patterns from three orthogonal planes for automatic facial expression recognition," in Humaine Association Conference on Affective Computing and Intelligent Interaction, 2013, pp. 356-361.

[85]. S. K. A. Kamarol, M. H. Jaward, H. Kalviaineninen, J. Parkkinen, and R. Parthiban, "Joint facial expression recognition and intensity estimation based on weighted votes of image sequences," Pattern Recognition Letters, vol. 92, pp. 25-32, 2017.

[86]. D. Ghimire, S. Jeong, J. Lee, and S. H. Park, "Facial expression recognition based on local region specific features and support vector machines," Multimedia Tools and Applications, vol. 76, no. 6, pp. 7803-7821, 2017.

[87]. C. Loob et al., "Dominant and Complementary Multi-Emotional Facial Expression Recognition Using C-Support Vector Classification," IEEE 12th International Conference on Automatic Face & Gesture Recogition, pp. 833-838, 2017.

[88]. H. Naveen Kumar, S. Jagadeesha, and A. K. Jain, "Human Facial Expression Recognition from static images using shape and appearance feature," 2nd International Conference on Applied and Theoretical Computing and Communication Technology, pp. 598-603, 2016.

[89]. F. Bashar, A. Khan, F. Ahmed, and M. H. Kabir, "Robust facial expression recognition based on median ternary pattern (MTP)," International Conference on Electrical Information and Communication Technology, pp. 1-5, 2013.

© 2018 IJSRSET | Volume 4 | Issue 2 | Print ISSN: 2395-1990 | Online ISSN : 2394-4099
National Conference on Advanced Research Trends in Information and
g Technologies (NCARTICT-2018), Department of IT, L. D. College of Engineering,
Ahmedabad, Gujarat, India In association with
International Journal of Scientific Research in Science, Engineering and Technology

Image Deblurring Techniques – A Detail Review

Mariya M. Sada[*1], Mahesh M. Goyani[2]

[*1]Computer Engineering Department, Government Engineering College, Modasa, Gujarat, India
[2]Computer Engineering Department, Government Engineering College, Modasa, Gujarat, India

ABSTRACT

Images are nowadays an integral part of our lives, whether in scientific applications or social networking and where there is an image, the concept of blurring might occur. Blurring is a major cause of image degradation and decreases the quality of an image. Blur occur due to the atmospheric commotion as well as the improper setting of a camera. Along with blur effects, noise also corrupts the captured image. Deblurring is the process of removing blurs and restoring the high-quality latent image. Blur can be various types like Motion blur, Gaussian blur, Average blur, Defocus blur etc. There are many methods present in literature, and we examine different methods and technologies with their advantages and disadvantages.

Keywords: Blur Types, Survey, Deblurring, Blur Detection, Blur Classification

I. INTRODUCTION

Images are blurred due to many reasons such as imperfections in capturing pictures, atmospheric problems and low-intensity level during camera exposure. The blurring of an image is a major cause of image degradation. Due to blurring, we cannot get exact details of the original image. Deblurring is a process to remove the blur and restore the image with high quality. Noise also corrupts the image so we need to perform de-noising on image [1], [2]. Image de-noising is also a part of image Deblurring. Applications of Deblurring include Iris recognition [3], Image segmentation [4], [5], Information retrieval [6], Astronomy [7], Microscopy [8], Space observation [9], Video object extraction [10], etc. There are many types of blurs like an Average blur, Motion blur, Defocus blur, Gaussian blur, etc.

Average blur: Average blur can be scattered in a vertical and horizontal direction [11]. The Average filter used to remove this type of blur and it is useful when noise present and affect the whole image.

Motion Blur: Motion blurs [12], [13], [14], [15] can be caused by relative motion between camera and scene during the exposure time.

Defocus Blur: Defocus blur [14], [16], [17], [18] is caused by an optical imaging system. Defocus blur is employed to blur a background and "pop out" the main object using large aperture lenses.

Gaussian Blur: Gaussian blur is simulated by Gaussian function. Effect of this blur is produced through a Gaussian filter that follows a bell-shaped curve. The blurring is dense in the center and fluff at the edge side [11], [19], [20].

(a) (b)) (d)

Figure 1. (a) Average Blur (b) Motion Blur (c) Defocus Blur (d) Gaussian Blur

Surroundings and outside conditions can affect the image quality during the process of the image acquiring. The blurring effect is caused by many reasons; one of the most important reasons include out of focus or camera shaken image. Image blurring is based on the degradation model. Blurring process can be formulated as the convolution of a clear or original image with point-spread-function (blur kernel) plus noise, given by

$$G(x,y)=f(x,y)\otimes h(x,y)+\eta(x,y)$$

Where $\otimes$ denotes a convolution operation. $G(x,y)$ is degraded image or blur and noisy image. $f(x,y)$ is a clear and original image. $h(x,y)$ is point-spread-function or blur kernel, and $\eta(x,y)$ is noise.

Photographical defocusing is another common type of blurring, known as defocus blur, mainly due to the finite size of the camera aperture [16]. In addition, camera shake is a common degradation in images. A small movement of the camera and object can degrade the image when the image is poorly lighted and requires a long exposure time.

Image Deblurring mainly includes two techniques: Blind Image Deconvolution and Non-Blind Image Deconvolution. In [21], many other techniques like subspace analysis [22], deblurring with noisy image pairs [23], deblurring with Richardson-Lucy algorithm [11], deblurring using Wiener Filtering [11], [12] are used. In Blind deconvolution, the PSF and clear image are being estimated but it is an ill-posed problem. In Non-Blind deconvolution, the clear image is estimated using known PSF. It includes Weiner filtering, deconvolution using RL method and deconvolution using a Regularized filter.

The organization of this document is as follows. In Section 2 (Review of State of Art Method), many deblurring methods and blurred region detection techniques have been discussed followed by conclusions and discussion in Section 3(Conclusion).

II. REVIEW OF STATE OF ART METHOD

Noise present in the image is a random variation of brightness which is caused due to taking an image in low-intensity level or due to many atmospheric problems [24]. An image contains different types of noise like Gaussian noise, white noise and salt and pepper noise, uniform noise etc. Even a small amount of noise can degrade the quality of image and image de-blurring techniques are sensitive to image noise. Literature [1], [2], [25] have applied a de-noising package as a pre-processing step.

A. Blur Region Detection and Classification

Nowadays many types of research are done on how to identify blur region and blur type. Subsequently, in the literature [14], [15], [26], blur identification and classification performed based on features. Kim et al. [26] introduced a 3-way blur identification method using features like Magnitude of gradient, Directional coherence, which classified the image into defocus blur, motion blur, and non-blur regions. The linear SVM classifier is employed to estimate each pixel's label into one of the blur regions. Super-pixel segmentation technique refines the rough blur region identification.

Singular value feature and alpha channel feature is used for blur region detection and blur classification [15]. Singular value feature differentiates the blurred region from a non-blurred region and estimates the blur degree by taking a ratio between the first few significant singular values (Eigen values) and all singular values that are computed over a local image patch surrounding each image pixel. Few most significant Eigen-images of a blurred image patch usually have higher weights or singular values than an image patch with no blur. The alpha value feature is used for classifying the blur based on the gradient distribution pattern, either it may be motion blur or defocus blur. The motion blurred image regions will have much larger values of variation of distance array compared with defocus blurred image regions.

Askari et al. [27] proposed a novel blur metric, which can significantly distinguish blur and non-blur regions and generate a blur map to encode the amount of blurriness for individual pixels. Estimated blur map is segmented into blur and non-blur regions by applying a pixon based technique. Pixon is a region that is made up of a set of connected pixels with associated properties such as intensity, texture, color. The original image is constructed by the union of the pixons. The Fuzzy C-Means algorithm is used for efficient segmentation. The estimation of blurriness at each pixel is performed by considering three different size blocks around each pixel. The purpose of taking these three blocks is to increase accuracy in estimating the blurriness value. The average value of these three blocks is defined as the blurriness value of the corresponding pixel.

Blur region detection and blur classification of partial blur image are challenging but some features are used

which is based on spectral, gradient, and color information [14]. Blur features modeled by gradient histogram span [28], [29], maximum saturation, power spectrum slope is used to identified blur image regions from the partially blurred image. The whole image is segmented into patches and all operations are performed on these patches.

Gradient Magnitudes of natural images usually follow a heavy-tailed distribution. A blurred region rarely contains sharp edges which result in small gradient magnitude. The distributions of the log gradient magnitude for blurred regions should have shorter tails.

The maximum value of saturation in blurred regions is consequently expected to be smaller than the un-blurred regions and un-blurred regions have more vivid colors than blur regions. Power spectrum slope of a blurred region tends to be steeper than that of an un-blurred region. If the power spectrum of the whole image is less than the patch, then this patch is considered as a blurred patch.

Local autocorrelation congruency is used for blur classification whether it is motion blur or defocus blur. This function is measured how well a signal matches a time-shifted version of itself.

Shi et al. [28] present a method for constructing a blur feature representation directly from an input image. Image Gradient distribution, Spectra in Frequency Domain and Local Filters are used to differentiate blurred region from un-blurred. The image is divided into patches and the peakedness of a gradient distribution is measured by Kurtosis from these patches. Peakedness of blur region is less than the un-blurred region. Average Power spectrum for the blurred patch is smaller than the sharp patch. Groups of linearly independent filters are used to best separate blurred and un-blurred patch sets. Local filters are like as a Gabor filter and spatial filters that used for blurred detection task. They capture local band-pass or high-pass information that supplements frequency and gradient domain features.

Dong Yang and Shiyin Qin represented the restoration algorithm for a partial blurred image which is based on blur detection and blur classification [29]. Blur detection can be implemented using image segmentation. Each region is compared with its neighbor regions. For each region of a partially blurred image, we count the number of the blurred points and the number of the points in the region. If this ratio is greater than a threshold then this region is treated as a blurred region.

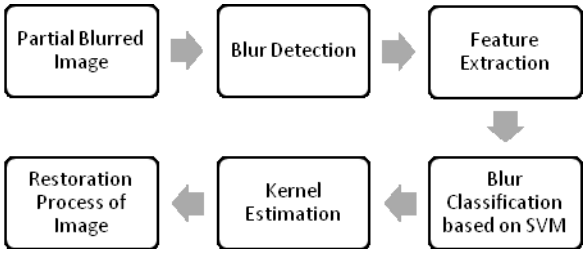


Figure 2. Restoration of Partial Blurred Image [29]

Gradient feature, Radon transform feature and an edge feature of the Fourier transform are used to classify the blur. Image gradient feature and radon transformation differentiates blur into non-defocus and defocus blur. After taking logarithmic Fourier transformation of a non-defocus blur, we classify the blur into motion blur and blend blur. The central area of the Fourier Transformation image of the defocus blur image is in the form of disk while motion blur have a strip of lines and blend blur have a combination of both. The non- blind de-blurring algorithm is used for restoration of blurred regions in the partially blurred image.

B. Kernel Estimation for Image Deblurring

In literature [1], [17] , [18], [30] various methods are available for kernel estimation. We can derive the kernel from the blur map and blur map can be created based on neighbor's behavior. Zhang et al. [17] proposed a method to remove the spatially defocus blur based on the estimated blur map and blur map is estimated by utilizing the KNN interpolation and edge information. By segmenting the blur map according to the blur amount of local regions and image contours, we get the local kernels. For Deconvolution, they used the local kernel and BM3D based non-blind deconvolution method.

Tang et al. [18], creates a blur map based on two assumptions: 1). If two neighbor image regions share the same color and belong to the same object then they should share similar blurriness, and 2). If two neighbor image regions share same gradient distribution then it also shares similar blurriness. This blur map is being used for blur detection refinement. The kernel can be estimated by parameters like blur angle, blur length, etc.

Hough transform method [2], [31] is used to detect the blur angle and blur length. Using this parameter and augmented Lagrangian method, the kernel is estimated.

Hough transform was one of the methods for motion blur angle estimation in which dark lines are observed in the Fourier spectrum of the blurred image after taking logarithmic transformation. Blur angle is between the corresponding x-axis and these dark lines. Once the locations of dark lines are found the blur angle can be calculated. Some central lobes are found by rotating the blurred image by blur angle and this blurred image has maximum values of pixels are found in all the column of an image and stored in a one-dimensional matrix. Blur length is calculated by; the mean lobe width is divided by the width of the image.

Cepstrum [29], [31] and radon transform [29], [32] are available for estimating these parameters but they are not accurate and contain many drawbacks especially when the number of columns and rows are not equal which leads distortion.

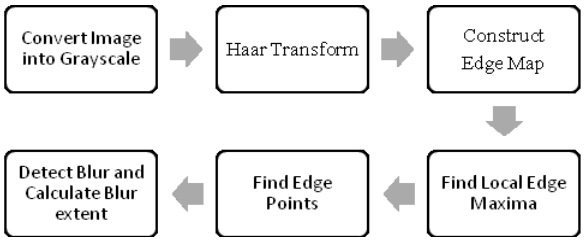


Figure 3. Blur Detection using Haar Transform [33]

Literature [33], [34] used the Haar wavelet transform for blur detection in the image. In [33], firstly it converts the image into gray scale and applies HWT for the edge analysis. Based on the analysis the edge map is constructed and then finds the local edge maxima for each local area. The process of finding local edge maxima is repeated by three times. If the ratio of image edges and edge points is less than a threshold, then it is a blurred region. An edge point is calculated by the local edge maxima and it is stored in edge vector. HWT is used due to its faster performance and better effectiveness. In literature [34], For de-blurring, alternating direction method of multiplier (ADMM) technique is used.

Pan et al. [35] proposed an algorithm by exploiting reliable edges and removing outliers in the intermediate latent image for robust kernel estimation. Initial kernel estimation can take place using the MAP framework and this framework also deals with the outlier. The Intermediate latent image can be obtained by the initialized kernel and the hyper-laplacian prior. The intermediate salient edges are selected from blurred image and outlier is detected from these selected. After detecting these outliers, the outliers are removed and the estimated kernel is updated. IRLS (Iteratively Re-weighted Least Squares) method used this estimated kernel to restore the image. IRLS is used to solve certain optimization problems with objective functions. It is performed using Maximum Likelihood approach.

An approach of kernel fusion stated that each kernel contributes to the final kernel and combining multiple kernels from different methods can lead to better kernel [36]. In baseline method, combination strategy takes the average of all individual estimated kernels but a disadvantage is that they compromise the good kernel. Another method is Gaussian Conditional Random Fields (GCRF) which predicts the kernel value at each kernel element individually and captures the relationship between the values of adjacent kernel elements. Performance of GCRF is better than baseline method.

Since a small amount of noise can degrade the quality of blur kernel estimation [1], for removing the noise, series of directional low pass filtering are applied on a blurred input image. The directional filter applies on a different orientation of the image. After filtering, blur kernel from each filtered image is estimated using inverse Radon transform. It also introduces a noise tolerant non-blind deconvolution technique that generates a high-quality final image. When we apply the directional filter, it damages the kernel.

C. Blind Image Deconvolution

Image de-blurring includes Blind Image deconvolution which jointly estimates the clear image and blur kernel [37]. It is ill-posed problem due to the loss of information on both images and blurring process [38]. Blind Deconvolution as the name indicates, works blindly where there is no information about point spread function. Point spread function is a point input, represented as a single pixel in the “ideal” image, which will be reproduced as something other than a single pixel in the “real” image [11], [39]. There are two approaches to blind deconvolution; projection based and maximum likelihood restoration based.

In former approach ([11], [40], [41]) it restores the true image and PSF by making the initial estimation. The technique is cylindrical in nature. This cyclic process is repeated until a predefined convergence criterion is met. The advantage of this method is, it is insensitive to noise. This approach involves the simultaneous evaluation of the recovered image and PSF that leads to a more sophisticated computational algorithm.

In latter approach, the maximum likelihood [42] estimates the parameter like PSF and covariance matrices. The PSF estimation is not unique in factors like symmetry, size etc. The definition of PSF separately from the restored image means the use of this information later by applying one of the known classical methods of restoration. Estimation of the image restoring and PSF are a separate procedure. The algorithm applied to implement such methods is computationally simple. It has low computational complexity and also helps to obtain blur, noise and power spectra of the true image. Maximum likelihood estimation can be seen as a special case of maximum a posteriori estimation (MAP). Literature [12], [16], [43], [44] uses blind deconvolution for de-blurring the image.

Shan et al. [12] proposed the method to de-blur the image using a unified probabilistic model of both blur kernel estimation and blurred image restoration. These terms include a model of spatial randomness of noise in the blurred image as well as new local smoothness prior that reduces ringing artifacts. Using this method input kernel is assumed to be quite inaccurate so the optimization is performed by MAP approach.

Literature review of many de-blurring techniques is given in [16], [45], [46], [47]. Removing motion blur from images is a typical blind problem. Blind image restoration is iteratively performed as linear and nonlinear processing. The nonlinear processing technique is used that is based on the compact representation of the image edge by means of Local Radon Transform (LRT) [32]. Radon transform is used to estimate the blur kernel and the overall restoration produces a sharp and focused image. The blind image deconvolution considers estimating the blur kernel in the gray domain. As a result, Xu et al. [43] present an estimation of blur kernel which is performed through RGB channel and it also stated that the blur effects for each color channel are usually different. The cost function is required for kernel estimation in this literature.

Blind image de-blurring only makes weak assumptions about the blurring filter. To overcome the ill-posed problem, the method includes a technique which initially focuses on the main edges of the image. The initialization of kernel estimation is performed using this edge information [48].

For estimating the kernel, different methods are used like kernel estimation using MAP approach, Kernel estimation through TV (Total Variation) regularization approach. Blur kernel estimation and image restoration can be done using MAP framework [35], [49] and this task can be solved by the minimization problem. For kernel estimation, it is assumed that the set of disjoint segmentation mask is given and fixed. In [50], kernel estimation is performed using a TV regularization method. TV regularization is widely used since it has good edge-preserving property. This approach makes the restoration algorithm achieved much better-recovered quality. TV Regularization and MAP approaches are closely related to the method of Maximum Likelihood (ML) estimation. The blind deconvolution methods have the inability to obtain preliminary information about the scene.

Deblurring using Motion Density Function (MDF): In Literature [11], [51] was proposed only for a single blurred image to remove camera shake. MDF represents the motion of camera and MDF is used for recording that time fraction which was spent in each discretised part of the space of all the possible camera poses. This method is useful for directly estimating spatially varying PSF and cannot work for multiple images. At any location in the image, MDF can be used to generate the kernel. This approach is used to model the spatial variation of blur kernels to run a blind deconvolution method at each pixel.

D. Non-Blind Deconvolution

Non-Blind deconvolution estimates only the clear image using known kernel. The prior knowledge about the parameters of blur kernel is required (point spread function length and angle). Non-blind deconvolution is used in the literature [1], [30]. Non-Blind deconvolution is performed by an already estimated kernel. Kernel estimation can be performed using radon transform [1]

and it also states that small amount of noise can affect the kernel estimation task. Yang et al. [30], performed non-blind deconvolution using the kernel which is estimated and calculated by dark channel and minimization can be performed using a MAP approach and the regularized term is used for minimization computational term. The Dark channel is the smallest value in a local neighborhood and it also proves that the dark channel of a blurred image is less sparse than the original image.

Non-Blind Deconvolution Technique includes Wiener filtering [17], [52] Lucy-Richardson method [31], [53], Regularization approach [54], [55] etc.

Deconvolution using Wiener Filter: Some knowledge about point spread function is required in this approach. An estimate of a target random process can be made by Wiener filter using Linear Time Invariant (LTI) of an observed noisy image. Wiener deconvolution can be used effectively when adaptive noise and the frequency characteristics of the image are known. Wiener filtering is optimal in terms of the mean square error. It minimizes the overall mean square error in the process of inverse filtering and noise smoothing and it is a linear estimation of the original image. In the absence of noise, the wiener filter reduces to the ideal inverse filter [31]. Wiener filtering is used to store latent image in [17], [46], [56].
Wiener filtering also performs de-blurring using estimated kernel. Zhang et al. [17] presented a method in which kernel can be estimated by segmenting the blur map which is constructing using KNN interpolating method and edge information. After estimating kernel deconvolution can be performed using Wiener filtering. Wiener filter minimizes the MSE between the desired process and estimated processes [57]. Three conditions must be executed for guaranteeing the MSE: 1). The undistorted image and noise do not correlate with each other. 2). The undistorted image or noise must have zero average value. 3). The estimation linearly depends on the distorted image. Wiener filter has relatively higher noise immunity.

Deconvolution using Lucy-Richardson Method: William Richardson and Leon Lucy [58], [59] invented the LR which is an iterative procedure for recovering a latent image that has been blurred by a known PSF. In this case, PSF is identified but no information is available for the noise. One main problem with basic LR method is how many times process should repeat. If the numbers of iterations are very large then it will slow down the computational process and also introduce ringing artifacts. The equation of Richardson-Lucy algorithm is [45], [60]

$$f^{n+1} = f^{n}\, H^{*}\left(\frac{g}{Hf^{n}}\right)$$

Where f^{n+1} the new estimate from the previous one is f^{n}, (g) is the blurred image, (n) is the number of the step in the iteration, (H) is the blur filter (PSF) and (H*) is the Adjoin of (H).

The LR algorithm reduces the effect of noise amplification. It is time consuming because more iteration is required. In literature [60], [61], [62], RL algorithm is being used to restore the final and latent image. Shaked et al. [60] represent an extended version of RL algorithm for estimation of medical images for their blurred measurements corrupted by Poisson noise and it's particularly simple algorithmic structure which implies straightforward implementation. Blind deconvolution is performed using RL algorithm in the literature [63]. Each iteration of RL method is used to guess the PSF and using this PSF image is restored. RL was developed from Bayes's theorem. A blurred image can be reconstructed using a combination of RL algorithm and Pyramid structure [62]. By using pyramid structure, the images with different frequency bands are generated. The combination of these two reduces the computational complexity and avoids the ringing effect.

The reason for the popularity of the RL algorithm is its implementation of maximum likelihood and its apparent ability to produce a reconstructed image of good quality in the presence of high noise levels [64]. RL provides the good estimation of the blurring function and gives the better PSNR within limited iterations.

Deconvolution using Regularized Filter: This technique is useful when limited knowledge about noise is present. The blurred and noisy image is restored by a constrained least square restoration algorithm that uses regularized filter [47]. The regularization filter is often chosen to be a discrete Laplacian. This filter can be understood as an approximation of a Wiener filter. Although the wiener filtering is the optimal tradeoff of

inverse filtering and noise smoothing, in this case when the blurring filter is singular, the wiener filtering actually amplifies the noise. The implementation of the regularized inverse filter involves the power spectrum estimation spectrum of the original image in the spatial domain.

Wang et al. [55] developed a Linear Time Invariant Regularized Backward Heat Diffusion (LTI RBHD) method that estimates the blur kernel with low and high width, better results are achieved with right kernel width and it is better than wiener filtering. The prediction of the original image is obtained from the watermark image by using the regularized filter [54]. By subtracting the predicted image from the watermarked image we can recover embedded watermark image blindly and this is the combination of blind and a regularized filter. Wiener filtering has a better result than a regularized filter.

Deblurring using Blurred/noisy pair images [23] can also be done. Both the blurred and noisy images are used to estimate the exact blur kernel, which otherwise is very difficult to get the blur kernel from a single image. The initial kernel can be assumed by simple constrained Least Square Optimization technique. Using both images, deconvolution is performed and this significantly reduces the ringing artifacts. The remaining ringing artifacts are further suppressed by gain controlled deconvolution process. Non-blind deconvolution can be divided into two parts: non-blind kernel estimation and non-blind image deconvolution.

In kernel estimation, the very accurate initial kernel can be recovered from the blurry image by exploiting large-scale, sharp image structures in the noisy image. The main advantage of this approach is that it takes both the blurred and noisy image, as a result, it produces high-quality latent image [40]. The disadvantage of this approach is that the point spread function is invariant.

Deblurring using subspace analysis [40], [65] can be done by constructing a feature space which includes blurred images which are degraded by same point spread function. A statistical model that represent prior knowledge of predefined point spread function sets in feature space is learned which estimates the blur kernel of a query image and compare with each model in feature space and selects the closest one for PSF inference [22]. The given query image is de-blurred using the blur kernel corresponding to that particular model. The inferred PSFs were used to sharpen both target and query images. This approach has not yet been proven for images blurred with multi unknown factors like camera shake.

Liao et al. [66] present a novel algorithm for hyperspectral image de-blurring with Principal Component of Analysis (PCA) and total variation. The first step is to de-correlate the hyperspectral image and separate the information content from the noise using PCA. The first k PCA channel contains most information about the HS image and remaining B-k channels contains the information about noise. If de-blurring is performed on these high-dimensional B-k PCs and noise, then it will amplify the noise of the data cube and cause high cost of computation during data processing, which is not desirable. Therefore fast TV denoising and de-blurring methods are applied for first k PCs and a soft-thresholding denoising scheme is applied to remove the noise from remaining B-k PCs.

Kumar et al. [67] presented a learning-based image de-blurring technique using an artificial neural network. Using this approach blur PSF is assumed as uniform. The image is divided into patches and deconvolution algorithm is applied on different patches. The network is then trained using back propagation algorithm. When we give the blurred patch to the neural network as an input, the center pixel of this patch is considered as the output of original patch.

Various blind and non-blind deconvolution based methods are compared in Table 1 and Table 2, respectively.

Table 1. Survey On Blind Deconvolution Based Image Deblurring Methods

Ref.	Blur Type	Kernel Estimation	Type of Kernel	Features used	Dataset	Classifier
[2]	Motion Blur	Estimate blur angle and length by Hough transform, TV regularization	Regularized filter		In-house	
[14]	Motion Blur, Defocus Blur		Gaussian filter	LPSS, GHS, MS, LAC	Flickr.com and PBase.com,	Naive Bayes
[15]	Motion Blur, Defocus Blur		Low pass filtering	SVF, ACF	In-house	
[18]	Defocus Blur, Motion Blur	Blur map can be created by blur metric which is based on assumption		Log spectrum	Public blurred image dataset	
[24]	Motion Blur	Using TV regularization approach	Low pass filtering		In-house	
[43]	Motion Blur	By predicting image salient edges and using Gaussian prior and cost function is required	Gaussian filtering, shock filter		In-house	
[32]	Motion Blur	Using Radon transform	Wiener filter		In-house	
[49]	Motion blur	Jointly solve blur segmentation and PSF estimation, MAP framework, two-phase kernel estimation,			In-house	Naive Bayes
[35]	Motion Blur	Removing the outliers, using MAP approach	Shock filter		Benchmark dataset	
[51]	Motion Blur	Motion Density Function used to generate PSF at any location			In-house	
[33]	Motion Blur, Gaussian blur	Using Haar Wavelet Transform blur detection performed			Caltech-256 Object Category, INRIA Holidays and LabelMe, ARCHIVES	SVM
[34]	Motion Blur	Estimate the PSF and LI using ADMM	Edge Filter		In-House	
[26]	Motion blur, Defocus blur			MG,DC	Google Image, Flicker, and DPChallenge	SVM
[27]	Motion blur, Defocus blur	Estimating Blurriness using Pixon based technique, different size of blocks is used for accuracy in estimation of blurriness	Low pass filtering		In-house	Naive Bayes
[28]	Motion Blur, defocus Blur		Gaussian filter	IGD, SFD, LF	In-house	Naive Bayes

LPSS = Local Power Spectrum Slope, GHS = Gradient Histogram Span, MS= Maximum Saturation, LAC = Local Autocorrelation Congruency, SVF = Singular Value Feature, AVF = Alpha Value Feature, IGD = Image Gradient Distribution, SFD = Spectra in Frequency Domain, LF = Local Filters, SVM = Support Vector Machine, LI = Latent Image, MG = Magnitude of Gradient, DC = Directional Coherence

Table 2. Survey On Non-Blind Deconvolution Based Image Deblurring Methods

Ref.	Blur Type	Kernel Estimation	Type of Kernel	Features used	Dataset	Classifier
[1]	Camera Shake	from each filer image and using radon transform	Low pass filter		In-house	
[12]	Motion Blur	Kernel estimate blindly and use non-blind technique for de-blurring	Wiener filter		In-house	
[17]	Defocus Blur	Kernels derived from segmenting the blur map	BM3D Filtering, Wiener filtering		In-house	
[22]	Defocus Blur, Motion Blur	Using a feature vector the closest one selected			In-house	NN
[23]	Camera Shake	A least squares optimization and Landweber algorithm are used for iteratively update	Bilateral filtering		In-house	
[44]	Motion Blur, Gaussian Blur, Average Blur, Defocus Blur	Using MAP estimation approach	Wiener filter, Laplacian filter	Moment Invariant, HOG, Zernike Moment	In-house	
[30]	Motion Blur	Using Dark channel and minimization perform by MAP approach	Bilateral filtering		Bench-mark dataset	
[31]	Motion Blur	Estimate blur length and angle and then kernel estimated from these parameters	Gaussian filter		In-house	
[56]	Motion Blur	Kernel estimated by parameters like blur angle and blur length	Wiener filtering		In-house	
[66]	HS image	Using PCA and TV regularization			In-house	
[36]		Kernel estimate by kernel fusion in which different kernel estimate by different method	Low pass filtering		Flickr, facebook, Google plus websites for dataset	
[29]	Motion blur, defocus Blur, Blend blur	Radon transform and Cepstrum is used for kernel estimation	Wiener filter	GF, RTF,EF	In-house	SVM

GF= Gradient Feature, RTF = Radon Transform Feature, EF = Edge Feature, SVM = Support Vector Machine, NN=Nearest Neighbour.

III. CONCLUSIONS

This paper presents the review of state of art methods of image deblurring. Blind deconvolution jointly estimates the point spread function and clear image and we have no prior knowledge about PSF, types of blur, the presence of noise and the resultant image is better than any other technique but it is time consuming. We discuss non-blind deconvolution which includes three methods: Wiener filtering, Richardson-Lucy algorithm, Regularized filter. In the presence of Gaussian noise, Wiener filter gives the best result and it is optimal in terms of mean square error. Richardson-Lucy gives better results with higher PSNR but it gives ringing

artifacts with an increase in a number of iterations. The regularized filter is one of the best techniques to deblurring when there is no noise in the image but when noise is present with blur, the Richardson-Lucy technique gives better performance. Many approaches use blur detection and blur classification as the blind estimation of PSF is difficult. Therefore, by classifying the blur we get the structure model of PSF and using this structure model of blur kernel we can easily estimate the true PSF. In general, blind deconvolution techniques show better results in comparison with non-blind deconvolution techniques.

IV. REFERENCES

[1]. L. Zhong, S. Cho, D. Metaxas, S. Paris, and J. Wang, "Handling Noise in Single Image Deblurring using Directional Filters," In proceeding of IEEE conference on computer vision and pattern recognition, pp. 612-619, 2013.

[2]. M. S. Shakeel and W. Kang, "Efficient blind image deblurring method for palm print images," IEEE International Conference on Identity, Security and Behavior Analysis, pp. 1-7, 2015.

[3]. N. Liu, J. Liu, Z. Sun, and T. Tan, "A Code-level Approach to Heterogeneous Iris," IEEE Transactions on Information Forensics and Security, vol. 6, no. 1, 2007.

[4]. W. Ren, J. Pan, X. Cao, and M. Yang, "Video Deblurring via Semantic Segmentation and Pixel-Wise Non-Linear Kernel," arXiv preprint arXiv:1708.03423 (2017).

[5]. T. Askari, J. Hamid, and H. Vahid, "Local motion deblurring using an effective image prior based on both the first- and second-order gradients," Machine Vision and Applications, vol. 28, no. 3-4, pp. 431-444, 2017.

[6]. A. Husni, M. Shapri, and M. Z. Abdullah, "Accurate retrieval of region of interest for estimating point spread function and image deblurring," The Imaging Science Journal, vol. 65, no. 6, pp. 327-348, 2017.

[7]. A. Fiandrotti, S. M. Fosson, C. Ravazzi, E. Magli, and P. Torino, "GPU-Accelerated Algorithms for Compressed Signals Recovery with Application to Astronomical Imagery Deblurring," International Journal of Remote Sensing, pp. 1-22, 2017.

[8]. L. Han and Z. Y. B, "Refocusing Phase Contrast Microscopy Images," International Conference on Medical Image Computing and Computer-Assisted Intervention, pp. 65-74, 2017.

[9]. X. Zhihai, Y. Pengzhao, C. Guangmang, F. Huajun, L. Qi, and C. Yueting, "Image restoration for large-motion blurred lunar remote sensing image," Confernce of Chinese Society for Optical Engineering, pp. 1-8, 2017.

[10]. H. Jeon, S. Member, J. Lee, and Y. Han, "Multi-Image Deblurring using Complementary Sets of Fluttering Patterns," IEEE Transactions on Image Processing, vol. 26, no. 5, pp. 1-16, 2017.

[11]. S. Yadav, C. Jain, and C. Aarti, "Evaluation of Image Deblurring Techniques," International Journal of Computer Applications, vol. 139, no. 12, pp. 32-36, 2016.

[12]. Q. Shan, J. Jia, and A. Agarwala, "High-quality Motion Deblurring from a Single Image," in ACM Transactions on Graphics, 2008, vol. 27, no. 3.

[13]. M. Tico, M. Trimeche, and M. Vehvilainen, "Motion Blur Identification Based on Differently Exposed Images," IEEE International Conference on Image Processing, pp. 2021-2024, 2006.

[14]. R. Liu, Z. Li, and J. Jia, "Image Partial Blur Detection and Classification," IEEE international conference on computer vision and pattern recognition, pp. 1-8, 2008.

[15]. B. Su, S. Lu, and C. L. Tan, "Blurred Image Region Detection and Classification," in In proceeding of 19th ACM International Conference on Multimedia, 2011, pp. 1397-1400.

[16]. G. Air, M. Indaco, D. Rolfo, L. O. Russo, P. Trotta, and P. Torino, "Evaluation of image deblurring algorithms for real-time applications," IEEE 9th conference on Design & Technology of Integrated System, pp. 1-6, 2014.

[17]. R. Wang and W. Wang, "Spatially Variant Defocus Blur Map Estimation and Deblurring from a Single Image," Journal of Visual Communication and Image Representation, vol. 35, pp. 257-264, 2016.

[18]. C. Tang, J. Wu, Y. Hou, P. Wang, and W. Li, "A Spectral and Spatial Approach of Coarse-to-Fine Blurred Image Region Detection," IEEE Signal Proceeding Letters, vol. 23, no. 11, pp. 1652-1656, 2016.

[19]. Z. Al-ameen, G. Sulong, and G. Johar, "A Comprehensive Study on Fast image Deblurring Techniques," International Journal of Advanced Science and Technology, vol. 44, pp. 1-10, 2012.

[20]. S. Jain, A. Dubey, D. S. Chundawat, and P. K. Singh, "Image Deblurring from Blurred Images," International Journal of Advanced Research in Computer Science & Technology, vol. 2, no. 3, pp. 2-6, 2014.
[21]. M. Poulose, "Literature Survey on Image Deblurring Techniques," International Journal of Computer Applications Technology and Research, vol. 2, no. 3, pp. 286-288, 2013.
[22]. H. Takeshima, N. Masashi, and H. Abdenour, "Facial Deblur Inference Using Subspace Analysis for Recognition of Blurred Faces," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 33, no. 4, pp. 838-845, 2011.
[23]. L. Yuan, J. Sun, L. Quan, and H.-Y. Shum, "Image Deblurring with Blurred and Noisy Image Pairs," ACM Transactions on Graphics, vol. 26, no. 3, pp. 1-10, 2007.
[24]. Y. T. S. Lin, "Motion-aware noise filtering for deblurring of noisy and blurry images," IEEE Conference on Computer Vision and Pattern Recognition, pp. 17-24, 2012.
[25]. N. Kamarudin, F. Layth, K. Asem, and R. A. Ramlee, "Natural image noise removal using nonlocal means and hidden Markov models in transform domain," The Visual Computer, pp. 1-15, 2017.
[26]. H. Lee and C. Kim, "Blurred Image Region Detection and Segmentation," IEEE International Conference on Image Processing, pp. 4427-4431, 2014.
[27]. T. Askari, J. Hamid, and H. Vahid, "Automatic estimation and segmentation of partial blur in natural images," The Visual Computer, 2015.
[28]. J. Shi and J. Jia, "Discriminative Blur Detection Features," IEEE international conference on computer vision and pattern recognition, pp. 2965-2972, 2014.
[29]. D. Yang and S. Qin, "Restoration of Partial Blurred Image Based on Blur Detection and Classification," Journal of Electrical and Computer Engineering, 2016.
[30]. J. Pan, D. Sun, and M. Yang, "Blind Image Deblurring Using Dark Channel Prior," IEEE Conference on Computer Vision and Pattern Recognition, pp. 1628-1636, 2016.
[31]. S. Tiwari, V. P. Shukla, A. K. Singh, and B. S.R, "Review of Motion Blur Estimation Techniques," Journal of Image and Graphics, vol. 1, no. 4, pp. 176-184, 2013.
[32]. S. Colonnese, P. Campisi, G. Panci, and S. Gaetano, "Blind Image Deblurring Driven by Nonlinear Processing in the Edge Domain," The European Association for Signal Processing Journal on Advances in Signal Processing, no. 16, pp. 2462-2475, 2004.
[33]. G. S. Trani, T. P. Nghieml, N. Quang, A. Drogoul, and L. C. Mai, "Fast Parallel Blur Detection of Digital Images," IEEE Research, Innovation and vision for the future International Conference on Computing & Communication Technologies, pp. 147-152, 2016.
[34]. S. Pendyala, P. Ramesha, A. V. Bns, and D. Arora, "Blur Detection and Fast Blind Image Deblurring," IEEE Annual Conference in India, pp. 1-4, 2015.
[35]. J. Pan, Z. Lin, Z. Su, and M. Yang, "Robust Kernel Estimation with Outliers Handling for Image Deblurring," Proceeding of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 2800-2808, 2016.
[36]. L. Mai and F. Liu, "Kernel Fusion for Better Image Deblurring," IEEE Conference on Computer Vision and Pattern Recognition, pp. 371-380, 2015.
[37]. A. Levin, Y. Weiss, F. Durand, and W. T. Freeman, "Understanding Blind Deconvolution Algorithms," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 33, no. 12, pp. 2354-2367, 2011.
[38]. P. Campisi and K. Egiazarian, Blind image deconvolution: theory and applications. 2016.
[39]. A. P. Abhilasha, S. Vasudha, N. Reddy, V. Maik, and K. Karibassappa, "Point Spread Function Estimation and Deblurring Using Code V," IEEE international Conference on Electronics, Information and Communications, pp. 1-4, 2016.
[40]. M. Poulose, "Literature Survey on Image Deblurring Techniques," International Journal of Computer Applications Technology and Research, vol. 2, no. 3, pp. 286-288, 2013.
[41]. D. Singh and R. K. Sahu, "A Survey on Various Image Deblurring Techniques," International Journal of Advanced Research in Computer and Communicationa Engineering, vol. 2, no. 12, pp. 4736-4739, 2013.

[42]. A.K. Katsaggelos and K. T. Lay, "Maximum likelihood blur identification and image restoration using the EM algorithm," IEEE Transactions on Signal Processing, vol. 39, no. 3, pp. 729-733, 1991.
[43]. X. Xu, H. Liu, Y. Li, and Z. Yi, "Image Deblurring with Blur Kernel Estimation in RGB Channels," IEEE International Conference on Digital Signal Processing, pp. 681-684, 2016.
[44]. R. Chokshi, D. Israni, and N. Chavda, "An Efficient Deconvolution Technique by Identification and Estimation of Blur," IEEE international Conference on Electronics Information and Communications Technology, pp. 17-23, 2016.
[45]. Z. Al-Ameen, G. Sulong, and M. G. M. Johar, "A Comprehensive Study on Fast image Deblurring Techniques," International Journal of Advanced Science and Technology, vol. 44, pp. 1-10, 2012.
[46]. T. Singh and S. B.M, "Comparative Analysis of Image Deblurring Techniques," International Journal of Computer Applications, vol. 153, no. 5, pp. 39-44, 2016.
[47]. A. Kaur and C. Vinay, "A Comparative Study and Analysis of Image Restoration Techniques Using Different Images Formats," International Journal of Science and Emerging Technologies with Latest Trends, vol. 2, no. 1, pp. 7-14, 2012.
[48]. M. S. C. Almeida and A. L. B, "Blind and Semi-Blind Deblurring of Natural Images," IEEE Transactions on Image Processing, vol. 19, no. 1, pp. 36-52, 2010.
[49]. T.-L. Wang, L. Kuan-Yun, and Y.-C. F. Wang, "Partial Image Blur Detection and Segmentation from a Single Snapshot," IEEE international Conference on Acoustic, Speech and Signal Processing, pp. 1907-1911, 2017.
[50]. L. Ma and T. Zeng, "Image Deblurring Via Total Variation Based Structured Sparse Model Selection," Journal of Scientific Computing, vol. 67, no. 1, pp. 1-19, 2016.
[51]. A. Gupta, N. Joshi, C. L. Zitnick, M. Cohen, and B. Curless, "Single image deblurring using motion density functions," European Conference on Computer Vision, pp. 171-184, 2010.
[52]. X. Zhang, R. Wang, X. Jiang, W. Wang, and W. Gao, "Spatially variant defocus blur map estimation and deblurring from a single image," Journal of Visual Communication and Image Representation, vol. 35, no. April, pp. 257-264, 2016.
[53]. M. Tico, M. Trimeche, and M. Vehvilainen, "Motion Blur Identification based on Differntly Exposed Images," IEEE International Conference on Image Processing, pp. 2021-2024, 2006.
[54]. K. Thongkor, P. Supasirisun, and T. Amornraksa, "Digital Image Watermarking based on Regularized Filter," 14th International Association of Pattern Recognition International Conference on Machine Vision Applications, pp. 493-496, 2015.
[55]. L. Wang, S. Luo, and Z. Wang, "Image Deblur with Regularized Backward Heat Diffusion," 17th IEEE International Conference on Image Processing, pp. 1141-1144, 2010.
[56]. R. Dash and B. Majhi, "Optik Motion blur parameters estimation for image restoration," Optik - International Journal for Light and Electron Optics, vol. 125, no. 5, pp. 1634-1640, 2014.
[57]. A. Bogoslovsky, I. Zhigulina, and E. Bogoslovsky, Image Deblurring Based on Physical Processes of Blur Impacts. 2018.
[58]. W. H. Richardson, "Bayesian-Based Iterative Method of Image Restoration," Journal of Optical Society of America, vol. 62, no. 1, pp. 55-59, 1972.
[59]. Lucy and L. B, "An iterative technique for the rectification of observed distributions," The Astronomical Journal, vol. 79, 1974.
[60]. E. Shaked, S. Dolui, and O. V Michailovich, "Regularized Richardson-Lucy Algorithm for Reconstruction of Poissonian Medical Images," IEEE International Symposium on Biomedical Imaging: From Nano to Macro, pp. 1754-1757, 2011.
[61]. D. a. Fish, a. M. Brinicombe, E. R. Pike, and J. G. Walker, "Blind deconvolution by means of the Richardson-Lucy algorithm," Journal of the Optical Society of America A, vol. 12, no. 1, pp. 58-65, 1995.
[62]. J. Ding, W. Chang, Y. Chen, and S. Fu, "Image Deblurring Using a Pyramid-Based Richardson - Lucy Algorithm," 19th IEEE International Conference on Digital Signal Processing, pp. 204-209, 2014.
[63]. D. A. Fish, A. M. Brinicombe, E. R. Pike, J. G. Walker, and R. L. Algorithm, "Blind

deconvolution by means of the Richardson - Lucy algorithm," vol. 12, no. 1, 1995.
[64]. M. Thakur and S. Datar, "Image Restoration Based On Deconvolution by Richardson Lucy Algorithm," International Journal of Engineering Trends and Technology, pp. 161-165, 2014.
[65]. M. Nishiyama, H. Takeshima, J. Shotton, T. Kozakaya, and O. Yamaguchi, "Facial Deblur Inference to Improve Recognition of Blurred Faces," IEEE Conference on Computer Vision and Pattern Recognition, pp. 1115-1122, 2009.
[66]. W. Liao et al., "Hyperspectral Image Deblurring with PCA and Total Variation," 5th IEEE Workshop on Hyperspectral Image and Signal Processing: Evolution in Remote Sensing, pp. 1-4, 2013.
[67]. N. Kumar, R. Nallamothu, and A. Sethi, "Neural Network Based Image Deblurring," 11th IEEE Symposium on Neural Network Applications in Electrical Engineering, pp. 219-222, 2012.

© 2018 IJSRSET | Volume 4 | Issue 2 | Print ISSN: 2395-1990 | Online ISSN : 2394-4099
National Conference on Advanced Research Trends in Information and
g Technologies (NCARTICT-2018), Department of IT, L. D. College of Engineering,
Ahmedabad, Gujarat, India In association with
International Journal of Scientific Research in Science, Engineering and Technology

A Survey on State of the Art Methods of Fingerprint Recognition

Urmi A. Soni[*1], Mahesh M. Goyani[2]

[*1]Computer Engineering Department, Government Engineering College, Modasa, Gujarat, India
[2]Computer Engineering Department, Government Engineering College, Modasa, Gujarat, India

ABSTRACT

Fingerprint recognition is one of the most well-known and publicized biometrics for personal identification and authentication. With the advancement of technology and for security purposes in various civilian, defence and commercial applications, fingerprint recognition is studied since a long time. Since second millennium BCE when fingerprints were used as signature, researchers have done studies on different aspects and features of fingerprint. This paper specifies a literature survey of the widely used techniques in fingerprint recognition system. Some of the renowned techniques of feature extraction and matching modules are discussed in deep.

Keywords: Fingerprint recognition, Minutiae points, Fingerprint classes

I. INTRODUCTION

Biometrics technologies are automated methods which uses physiological or behavioural characteristic of an individual for verifying or recognizing the identity of a living person [1], [2], [3], [4]. Fingerprint [5], face, voice, iris [2], retina geometry [6], palm print [7], gait, hand vein [8], signature etc. are popular biometric characteristics [9] widely used in automatic identification of individuals [10]. Among them, fingerprint recognition is most reliable biometric due to its uniqueness among the people, stability, originality and fairly good performance in terms of accuracy, speed and robustness [11].

Fingerprint identification is one of very few techniques employed in forensic science for criminal investigation as FBI uses fingerprints to store data and records of criminals [12] . Also it is used in legal affairs as it can be used as evidence on legal documents in place of signature [13]. In safeguarding work [14] such as providing protection to health, wealth and human rights, fingerprint recognition system has been used in doors and lockers. National IDs' such as unique ID, passports use fingerprint to identify an individual. Real time Automated Personal IDentification System (RAPIDS) at DoD (Department of Defence in USA) [15] uses fingerprint for personal identification.

Fingerprint recognition is used in numerous applications that include civilian and commercial applications like military, law enforcement, medicine, education, payment using ATM [16], civil service, forensics [14], driving license registration, cellular phone access, computer log-in [16].

As discussed in [17], fingerprint identification began in the late 19th century with the development of identification bureaus for storage and verification of criminal records using any physiological characteristic instead of names. Slowly, the focus shifted from verification to identification of a single person from a large database [18].

Fingerprint impression of a person consists of ridges (black lines) and valleys (white region) [19]. A ridge can spread into two ways: ridge ending and ridge bifurcation. Combination of ridge ending and ridge bifurcation is called basic Minutiae points [19]. Figure 1 shows a sample fingerprint consisting of minutiae points.

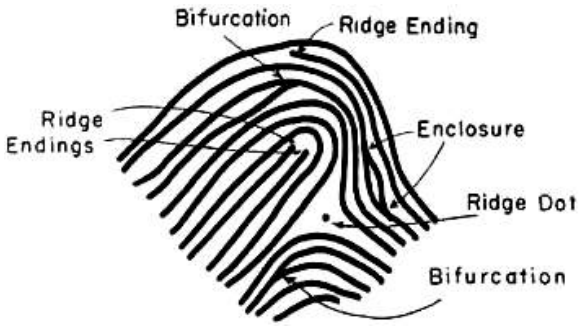


Figure 1. Minutiae points in sample fingerprint

Fingerprint of a person contains numerous distinctive features which makes them unique from a person to person. Figure 2 shows that fingerprint features can be classified mainly as three levels [4], [20] :

1) Level 1 (Global): refers ridge flow patterns (orientation) and singular points such as core and delta.

2) Level 2 (Local): introduces minutiae details such as ridge ending and ridge bifurcation.

3) Level 3 (Fine details): includes intra-ridge details such as width, shape, curvature, edge contours and even sweat pores.

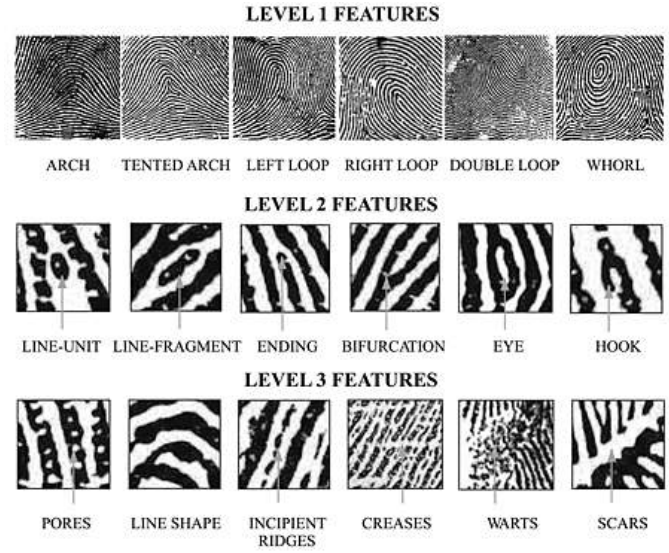


Figure 2. Feature classification

Among these different levels of features, level 1 feature can be used for fingerprint classification as they are defined from global charac2teristics. Level 2 and 3 features are commonly used for fingerprint matching [21] as they allow to claim for the individuality of a fingerprint. Once features are classified, the fingerprint pattern can be classified in six common classes [19]:

1) Arch: ridges enter from one side, rise to form a small bump, and then go down and to the opposite side. No loops or delta points are present.

2) Tented Arch: similar to the arch except that at least one ridge has high curvature, thus one core and one delta points.

3) Left loop: one or more ridges enter from one side, curve back, and go out the same side they entered. Core and delta are present.

4) Right loop: same as the left loop, but different direction.

5) Whorl: contains at least one ridge that makes a complete 360 degree path around the centre of the fingerprint.

6) Unclassified: Fingerprints in databases are non-uniformly distributed in these classes. Fig. 3 shows above 6 classes of fingerprint.

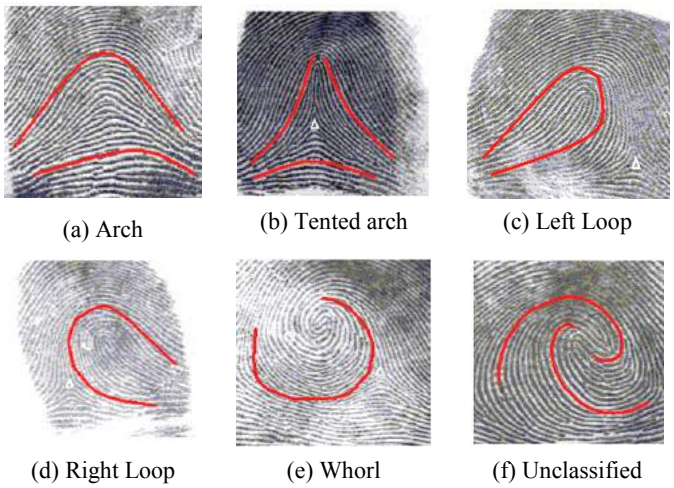


Figure 3. Fingerprint classes

Conventional fingerprint recognition system consists of 3 main sub-domains [22]: Image acquisition and pre-processing; Feature extraction and Matching/ Classification. Fingerprint images are acquired by the sensors adopted by the system or mobile device. These acquired images are pre-processed to improve the quality of image. Pre-processing such as image enhancement and morphological operations such as thinning, thickening etc. is done on the images. Popular enhancement techniques are directional median filter [23], directional weighted median filter [24] and STFT analysis [25].

Minutiae features such as ridge ending and bifurcations are extracted after refining of the thinned image and detecting the minutiae points. Additional features such

as core, delta, island, pores, crossovers etc. can also be detected to improve the matching process [26]. The most popular technique of minutiae detection is through the use of the crossing numbers approach [22]. Gabor filter is also used for feature extraction [27]. Once the features are extracted from fingerprint image, matching is performed with already registered template or by machine learning techniques. Matching score is computed using score computation metrics and performance evaluation is done using performance metrics. Main modules of fingerprint recognition system are shown in Figure 4.

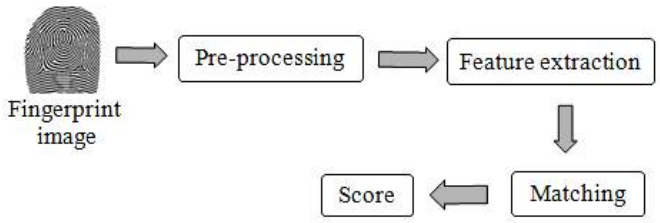


Figure 4. Main modules of fingerprint recognition system

Fingerprint matching includes numerous techniques according to extracted features used for matching and in terms of processing. Among them there are mainly two types: minutiae-based approach and non-minutiae based approach. Non-minutiae based approach [28], [29] includes Image-based method, Ridge-feature based method, Level3 feature based method and feature-point-based method. Minutiae-based method [20] includes local minutiae matching and global minutiae matching. Techniques also use hybrid approach [29], [30] in which combination of minutiae features and non-minutiae features are combined to improve accuracy since only minutiae features may not be sufficient for matching [31].

Image-based method [30], [32], [33], [34] is a non-minutiae based technique that directly compares the entire fingerprint patterns by finding the correlation between two images but this method is vulnerable to the alignment error caused by non-linear deformation. Texture features such as Gabor response [30], [32], Local Binary Pattern (LBP) [33], Histogram of Oriented Gradient (HoG) [34] are also used which is also sensitive to noise, skin condition or nonlinear deformation.

This problem of nonlinear deformation can be solved by using Ridge-feature-based approaches that utilize the topological information of the ridge patterns which includes ridge orientation [35] and frequency [36] information but the problem with this method is only several ridges of minutiae are used in extraction and ridge information cannot be used sufficiently for matching. For validating the usefulness of ridge features, Level3 feature based method is used which incorporates ridge details such as pores [28], dots and incipient ridges [37] and ridge contours along with minutiae features to improve performance. However these features can only be detected in high-resolution fingerprint images of 1000 dpi and over. Several researches proposed partial fingerprint matching based on Scale Invariant Feature Transform (SIFT) [38] or accelerated KAZE (AKAZE) [39], [40] features which are typically used for object recognition and image matching these approaches are relatively sensitive to the large textural variations caused by noise or skin condition. Table I show non-minutiae based fingerprint recognition methods and features used in each of them.

Table 1. Non-Minutiae Based Fingerprint Recognition Methods

Method	**Features**
Image-based method	Pixel intensities [30]; Texture features such as Gabor response [30] [32], LBP [33], [41], HoG [34]
Ridge feature-based method	Ridge orientation [35], Ridge frequency [36], Ridge count [42]
Level 3 feature-based method	Pores [28], Dots and incipient ridges [37], Ridge contour
Feature-point-based method	Key-points such as SIFT [38], A-KAZE [39], [40]

The organization of this document is as follows. In Section 2 (Survey of existing methods) discussion of various state of the art methods of pre-processing and feature extraction in fingerprint recognition system is given. Finally, in Section 3 (Conclusion and discussion) the paper is concluded and suggests the future work.

II. SURVEY OF EXISTING METHODS

Survey of various techniques of pre-processing and feature extraction modules of fingerprint recognition is discussed in this section.

A. Pre-processing

Images captured by different sensors may contain various noises, false traces, blurred ridges and indistinct boundaries [43] which may result in poor quality images. Hence before proceeding for matching images need to be pre-processed which includes noise reduction, contrast enhancement, improving contrast among ridges and valleys of grayscale fingerprint pictures and morphological operations [44]. Fingerprint enhancement can be conducted on either binary ridge images or gray-scale images thus Input image is normalized so that it has a pre-specified mean and variance. In [45] two approaches are described for fingerprint image enhancement: Spatial Domain Method, which operates directly on pixels and Frequency Domain Method, which operates on Fourier transfer of an image.

Histogram equalization (HE) is most extensively utilized global contrast enhancement technique due to its simplicity and ease of implementation [46] which distributes the pixel values uniformly such that the enhanced image has linear cumulative histogram. This technique is used to make the intensity distribution uniform by re-assigning the intensity values of pixels and effectively spreading out the most frequent intensity values across the entire spectrum of pixels (0-255) [47]. Through this adjustment, the intensities can be better distributed on the histogram and lower local contrast can gain a higher contrast.

Directional Median Filter (DMF) [23] can be used for reducing Gaussian distributed noises (by anisotropic filter) and impulse noises along the direction of ridge flow (by DMF). This algorithm may fail when image regions are contaminated with heavy noises and orientation field in these regions can hardly be estimated. An improved method was proposed in [24] over [23] that is Directional Weighted Median Filter (DWMF). A new impulse detector [24], which is based on the differences between the current pixel and its neighbors aligned with four main directions. It is appropriate to use Gabor filters as band pass filters to remove the noise and preserve true ridge/valley structures as Gabor filters have both frequency-selective and orientation-selective properties and have optimal joint resolution in both spatial and frequency domains [48]. In [49], it is stated that the previous fingerprint image enhancement methods based on Traditional Gabor filter (TGF) banks have some drawbacks in their image-dependent parameter selection strategy, which leads to artifacts in some cases. To overcome this issue, modified Gabor filter (MGF) have been developed as an improved version of the TGF. Its parameter selection scheme is image-independent.

Fourier transform [3] is an important mathematical tool used to decompose an image into sine and cosine components that divide the fingerprint image into small processing blocks and enhance each block independently. Wavelet transform is an effective tool in reducing noise in which images are analyzed at multiple scales [43], [50]. Directional Wavelet Transform (DWT) [50], [51] can be used for image enhancement and noise removal. DWT along with Gabor filter is also recommended for noise removal. Gabor wavelet filter bank (GWT) and Directional Median Filter (DMF) together used as robust approach to fingerprint image enhancement. Gaussian-distributed noises are reduced effectively by Gabor Filters and impulse noises by DMF. Short Time Fourier Transform (STFT) is a new approach for fingerprint enhancement which is a well known technique in signal processing to analyze non-stationary signals [25]. This algorithm estimates all the intrinsic properties of the fingerprints such as the foreground region mask, local ridge orientation and local ridge frequency.

Directional filter bank [52] reduces the influence of noise on the ridges and valleys, enhances the ridges' moving shape and preserves the spatial characteristics at minutiae and singular points. In [53], minutiae can be directly extracted from gray-level fingerprint images. Their algorithm is based on a gray-level ridge tracing [54] which extracts ridges by sequentially following each gray-level ridge until it ends or bifurcates. Their algorithm does not binarize the gray-level fingerprint image directly when conducting minutiae extraction, but binarization is still conducted implicitly by the gray-level ridge tracer. The problem of binarization of gray-level images [55] by Blayvas I et al. by acquiring images under non-uniform illumination and using different method to determine an adaptive threshold surface which gives faster binarization and better performance..

Captured fingerprint contains two components, original area captured by contacting the sensor called foreground and noisy are at the borders called background. Fingerprint segmentation [20] is used to separate two areas of fingerprint called foreground and background from each other and also for removing false features. A segmentation algorithm in [56], based on pixels features, using the criterion of Rosenblatt's perceptron to classify the pixels. The disadvantage of this algorithm is based on pixels features and moderate performance which gives low speed. N. Ratha, et al. [57] used morphological operators to detect and remove spikes.

B. Feature Extraction

Different levels of features can be extracted using feature extraction techniques according to the matching technique used. Singular points are extracted by its symmetry properties using symmetry point extraction as proposed in [58]. Complex orientation field is used to in which complex filters are conducted in multiple resolution scales. The symmetry is detected by its strong response to complex filters [59].

PoinCare Index (PI) [40], [60], [61] is a popular method to detect singular points which is computed as the sum of the difference between the orientation of point and each of its neighbor [62]. Depending upon value of PI singular points can be detected whether it is core or delta. However a challenging task is to improve reliability of orientation field as wrong orientation of points due to noise or poor quality of image can result in spurious detection. Chikkerur and Ratha [60] presented significant improvements in singular point detection algorithm based on complex filtering principles originally proposed by Nilsson and Bigun in [63]. J. Zhou et al. [61] proposed to use the DORIC (Differences of the ORIentation values along a Circle) feature for singular point verification, which can remove spurious detections and provide more discriminative information. The method needs more heuristics and sophisticated filters to detect singular points [62].

Convolution Neural Network (CNNs) [31], [62] are biologically inspired variants of multilayer perceptrons (MLP) [64], used to automatically detect singular points with lack dependence of any prior knowledge and human effort. In [62], a two layer CNN feature extraction and non-linear classifier is used for singular point detection by learning Multi-Stage (MS) features using CNN architecture provided by D. Sermanet et al. in [65] which provides richer complementary information such as local textures and fine details. A deep CNN is employed in [31] to learn global and minutiae features followed by K-Nearest Neighbor policy for proper triplet selection. S. Kim et al. [13] proposed a fingerprint liveness detection method based on a deep belief network (DBN) in which features are learned by Convolution network with random filters followed by the Principle Component Analysis (PCA). In [66], a convolution network is used to extract features, whose dimension is reduced by the principal component analysis (PCA), to be used by a SVM for fingerprint liveness detection.

Crossing numbers (CN) [22], [54], [67] are the most commonly employed method of minutiae extraction concept. Rutovitz's definition for crossing number involves the use of the skeleton image where the ridge flow pattern is eight-connected which extracts both true and false minutiae [54] where H-point elimination method to remove several types of spurious minutiae. In [22], an advancement of crossing number is used from which ridge ending and bifurcation can be detected using threshold value of CN and from applying it, 24 possible bifurcation templates are generated for ease in extraction. Applying same to the algorithm after image enhancement step, M. M. Min et al. [67] proposed 8 different possible termination and bifurcation features.

Gabor filter are used in the region around the fingerprint core point in a well-known image-based feature extraction technique, FingerCode [68] to extract local and global features which computes the standard deviation from mean of grey levels to define feature vector. The main problem of this approach is detection of a reliable core. In [69], [70] two FingerCode variants are proposed, where fingerprints are first zaligned using minutiae before extraction In [69] a set of 8 Gabor filters are used for hybrid approach of minutiae and ridge flow information to capture ridge strength at equally spaced orientation whereas 2D wavelet decomposition and convolution with 16 Gabor filter are used for extracting fingerprint pattern and minutiae in [70]. These hybrid approaches [68], [69], [70] tend to give better results than conventional minutiae-based fingerprint matching. L. Nanni et al. [33] proposed a hybrid approach where fingerprints are pre-aligned using minutiae, and then image-based features are extracted by invariant local

binary patterns (LBP) from the fingerprint image convolved with Gabor filters which together called GLBP (combination of Gabor filter and LBP). Here, Gabor filter is applied on different wavelet sub-bands [70] and invariantly LBP [33] to show that combining descriptors may improve the discrimination power.

Local Binary Pattern (LBP) operator is used for feature extraction in [41] to get LBP histogram on extracted pores and calculate the chi-square distance between two average LBP histograms is performed for matching that gives the best match. [41] Combines the pore-LBP based level-3 matching score and minutiae-based level-2 matching score in decision level. Local Phase Quantization (LPQ) [71] is statistics of labels computed in the local pixel neighborhoods through filtering and quantization. These methods describe each pixels neighborhood by a binary code which is obtained by first convolving the image with a manually predefined set of linear filters and then binarizing the filter responses. The bits in the code string correspond to binarized responses of different filters. As proposed in [71], BSIF is inspired by LBP and LPQ to automatically learn a fixed set of filters from a small set of natural images, instead of using handcrafted filters such as in LBP and LPQ.

Minutiae Cylindrical Code (MCC) [21] uses the Minutia Cylinders to code the fingerprint into fixed length using binary representation. The algorithm proposed in [72] uses binary minutia cylinder code (MCC) as minutia descriptor which encodes locations and directions of neighbor minutiae around each minutia into a fixed-length bit vector. MCC is more accurate and robust than minutia triplets, but its dimensionality is higher.

Locality Sensitive Hashing (LSH) is effective and efficient in finding approximate nearest neighbor of high-dimensional vector [73]. Locality Sensitive Hashing (LSH) is used to efficiently find out hypothesis correspondences of MCC instead of traditional quantization scheme. An improved Locality Sensitive Hashing algorithm for MCC descriptor is proposed in [74] combining with pose estimation algorithm.

Minutiae Vicinity (MV) was proposed by Guoqiang Li et al. in [75]. A minutiae vicinity is a fundamental unit which is framed by four details including an inside minutia and its three nearest neighboring details sorted by climbing request in view of their Euclidean distance with inside minutia. Minutiae Vicinity (MV) and Minutiae Cylinder Code (MCC) both are used in score level fusion method [75], [76] for fingerprint indexing approach where matching is quite complex for a large database. In [75] new designed indexing method extracts a feature vector including 9 components from minutiae details and a triplet which is contained in a Minutiae Vicinity. In [76], it is proposed that MV- features and MCC-Features are extracted from the sample. After extraction these features are fused and a single feature vector or a template is obtained. Principal component analysis (PCA) [35], [66] is a powerful algorithm used for dimensionality reduction of the data to make the algorithm faster and more efficient.

M. Yamazaki et al. [38] adopts scale invariant feature transform (SIFT) for partial fingerprint matching. Gaussian blurring causes ridges diffusion, which reduces the distinctiveness of fingerprint texture. Mathur at el. [39] applies A-KAZE feature to address this problem. A-KAZE features use non-linear diffusion, which preserves fingerprint ridges. Histogram of Gradients (HoG) [34] is an image-based matcher proposed by N. Dalal et al. [77] where the histogram of gradients features are used to describe the fingerprint which represent an image is represented by a set of local histograms which count occurrences of gradient orientation in a local cell of the image gives best results for stand-alone descriptors as it outperforms Gabor filters and LBP [34]. Inspired by histograms of oriented gradients (HOG) and the scale invariant feature transform (SIFT), a new approach of Histogram Invariant Gradient (HIG) is proposed in [78] which achieves average accuracy for liveness detection.

As various techniques are used for fingerprint recognition depending on the extracted features, each technique has different aspects and metrics. A comparative study of widely used existing methods is shown in Table 2.

Table 2. Survey On Various Fingerprint Recognition Methods

Ref.	Technique	L/G	Features	No of thumb Fingerprints used	Database Used	Minutiae extraction	Classifier	Accuracy
[4]	Multi-feature based score fusion method	L	Minutiae + texture information (Ridge count)	48 people with 10 impressions of each finger	FVC 2000 DB1 FPC DB	Traditional minutiae-based matching method	Hausdorff distance	Matching score: 4.57% 2.04% 97.05% boost EER: Lower than 50% for all
[29]	Partial fingerprint-matching method incorporating new RSFs with minutiae	L	Minutiae + edge shapes (concave and convex) are	-	FVC20 02, FVC20 04 and BERC	Local neighborhood of minutiae (Ridge shape features RSF) Proposed algorithm	Euclidean distance	-
[74]	Minutiae – based indexing with pose constraint	L	Ridge orientation information	-	FVC2000 DB1 FVC2000 DB2 FVC2000 DB3 NIST Special Database 4 NIST Special Database 14	local 1-D normalized histogram of ridge orientations Pose estimation Algo + Improved LSH for MCC	SVM	-
[79]	Fingerprint recognition by euclidean distance	G	Core & bifurcation points	172 400	ST-BIO Card Reader Model: BCR100T V3.0 VeriFinger Sample DB database [23].	Core: ASDF – Average Square Directional Field	Shortest Euclidian distance core & bifurcation points	Overall precision : 85% ST-BIO Card Reader : 95% VeriFinger Sample DB : 75%
[80]	Minutiae Extraction	L	Minutiae features	-	-	Iterative, parallel thinning algorithm Enhanced thinning algorithm Minutiae Marking: Crossing Number	False Minutia Removal Minutiae Match: alignment-based match algorithm (elastic match algo)	EER: 25
[51]	Texture Feature based Fingerprint Recognition	L	Core	300 fp images	In-house (256*256) FVC 2000	Feature extraction: DWT Core point:	Matching: Euclidian distance	98.98% boost EER: 2.80%

						PoinCare Index		
[81]	Pattern Recognition System	L	Minutiae	900 fp images (each has 500dpi 256 gray level) 269 fingers 450 imgTrain 450 imgTest	IBM HURSLEY database	Learning Vector Quantizer (LVQ)	Feature refinement rule-based classifier LVQ based classifier	95% training 87% testing
[62]	Convolution Neural Networks (CNNs)	G	Singular points	-	FVC 2002 DB (core 2738 Delta 731)	CNN	Two-layer nonlinear CNN classifier	False detection rate: Core 7.5% Delta 6%
[31]	Deep CNN (train minutiae) + K-NN based triplet selection(select proper triplet)	G	Minutiae + Global Pattern	-	In-house DB PHF300 FVC2000 FVC2006	Deep CNN	-	-

FDR: False Detection Rate, CNN: Convolution Neural Network, MCC: Minutiae Cylinder Code, PR: Precision Rate, ASDF: Average Square Directional Field, SVM: Support Vector Machine, LSH: Locality Sensitive Hashing, EER: Equal Error Rate, A-KAZE: Accelerated KAZE, L: Local, G: Global

III. CONCLUSION AND DISCUSSION

This paper presented a review of state of the art methods of fingerprint recognition system. The paper starts with the introduction of biometric technologies which are widely used in forensic applications like criminal investigations, terrorist identification and other security issues in applications such as defence, civilian, commercial era. Fingerprint applications and its history is briefly discussed which is followed by the features of fingerprint. The structure of a conventional fingerprint recognition system is shown which includes various modules. Fingerprint pre-processing techniques such as enhancement techniques, segmentation, binarization, morphological operations etc. used to improve the quality of fingerprint images before further stages of extraction and matching are discussed. Various minutiae and non-minutiae based features extraction techniques are analysed. Combination of these techniques along with advantages and limitations are specified. Template matching and machine learning techniques are presented with a brief discussion.

It is observed that, non-minutiae features are not sufficient for matching purpose because of their distinctiveness. Also according to survey, minutiae points solely do not provide satisfying accuracy and results as these points does not give any details about surrounding pattern whether it is ending or bifurcation. Thus extracting local structure of these minutiae points provides better results.

IV. REFERENCES

[1]. W. James, A. Jain, D. Maltoni, and D. Maio, An Introduction to Biometric Authentication Systems. Springer, London, 2005.

[2]. F. Besbes, H. Trichili, and B. Solaiman, "Multimodal biometric system based on fingerprint identification and iris recognition," 3rd International Conference on Information and Communication Technologies: From Theory to Applications, pp. 1–5, 2008.

[3]. P. T. Selvi and N.Radha, "Multimodal Biometrics Based Authentication Against Dictionary Attacks," International Journal on Computer Science and Engineering, vol. 02, no. 08, pp. 2652–2658, 2010.

[4]. Q. Li, C. Jin, W. Kim, J. Kim, S. Li, and H. Kim, "Multi-feature based score fusion method for fingerprint recognition accuracy boosting," Annual Summit and Conference on Asia-Pacific

Signal and Information Processing Association, pp. 1–4, 2017.
[5]. M. Ephin, S. Mohan, and N. A. Vasanthi, "Survey on Multimodal Biometric using Palm print and Fingerprint," International Journal of Computer Applications, pp. 36–41, 2013.
[6]. C. Marino, M. G. Penedo, M. Penas, M. J. Carreira, and F. Gonzalez, "Personal authentication using digital retinal images," Pattern Analysis and Applications, vol. 9, no. 1, pp. 1–21, 2006.
[7]. D. Zhang, W.-K. Kong, J. You, and M. Wong, "Online palmprint identification," IEEE Transactions on pattern analysis and machine intelligence, vol. 25, no. 9, pp. 1041–1050, 2003.
[8]. S.-K. Im et al., "An biometric identification system by extracting hand vein patterns," Journal-Korean Physical Society, vol. 38, no. 3, pp. 268–272, 2001.
[9]. D. Bhattacharyya, R. Ranjan, F. A. A, and M. Choi, "Biometric Authentication : A Review," International Journal of u-and e-Service, Science and Technology, vol. 2, no. 3, pp. 13–28, 2009.
[10]. A. Jain, A. and Ross, and S. Prabhakar, "An Introduction to Biometric Recognition," IEEE Transactions on Circuits and Systems for Video Technology, vol. 14, no. 1, pp. 4–20, 2004.
[11]. G. S. F. Daniel Souza, Aquiles Burlamaqui, "Improving biometrics authentication with a multi-factor approach based on optical interference and chaotic maps," Multimedia Tools and Applications, pp. 1–20, 2017.
[12]. N. Yager and A. Amin, "Fingerprint verification based on minutiae features: A review," Pattern Analysis and Applications, vol. 7, no. 1, pp. 94–113, 2004.
[13]. S. Kim, B. Park, B. S. Song, and S. Yang, "Deep belief network based statistical feature learning for fingerprint liveness detection," Pattern Recognition Letters, vol. 77, pp. 58–65, 2016.
[14]. A. C. K. Sheng Li, "Fingerprint Combination for Privacy Protection," IEEE Transactions on Information Forensics and Security, vol. 8, no. 2, 2013.
[15]. H. Jang, H. Choi, D. Kim, J. Son, and H. L. B, "Fingerprint Spoof Detection Using Contrast Enhancement and Convolutional Neural Networks," Information Science and Applications, vol. 424, pp. 331–338, 2017.
[16]. S.Padma Priya, "Biometrics and Fingerprint Payment Technology," International Journal of Advanced Research in Computer Science & Technology, vol. 5, no. 1, 2017.
[17]. R. B. Nalini Ratha, Automatic fingerprint recognition systems. 2007.
[18]. N. Ratha and R. Bolle, Automatic fingerprint recognition systems. 2004.
[19]. K. Karu and A. K. Jain, "Fingerprint classification," Pattern Recognition, vol. 29, no. 3, pp. 389–404, 1996.
[20]. D. Peralta et al., "A survey on fingerprint minutiae-based local matching for verification and identification: Taxonomy and experimental evaluation," Information Sciences, vol. 315, pp. 67–87, 2015.
[21]. R. Cappelli, M. Ferrara, and D. Maltoni, "Minutia Cylinder-Code: A New Representation and Matching Technique for Fingerprint Recognition," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 32, no. 12, pp. 2128–2141, 2010.
[22]. R. B. Anil Jain, Lin Hong, "Online fingerprint verification," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 19, no. 4, pp. 302–314, 1997.
[23]. C. Wu, Z. Shi, and V. Govindaraju, "Fingerprint image enhancement method using directional median filter," Biometric Technology for Human Identification, vol. 5404, pp. 66–75, 2004.
[24]. Y. Dong and S. Xu, "A new directional weighted median filter for removal of random-valued impulse noise," IEEE Signal Processing Letters, vol. 14, no. 3, pp. 193–196, 2007.
[25]. S. Chikkerur, A. N. Cartwright, and V. Govindaraju, "Fingerprint enhancement using STFT analysis," Pattern Recognition, vol. 40, no. 1, pp. 198–211, 2007.
[26]. Q. Li, C. Jin, W. Kim, J. Kim, S. Li, and H. Kim, "Multi-feature based score fusion method for fingerprint recognition accuracy boosting," Annual Summit and Conference on Asia-Pacific Signal and Information Processing Association, 2017.
[27]. M. U. Munir and D. M. Y. Javed, "Fingerprint matching using gabor filters," National Conference on Emerging Technologies., vol. 3, pp. 147–151, 2004.

[28]. M. . P. Jain, A., Chen, Y., Demirkus, "Pores and Ridges: High-Resolution Fingerprint Matching Using Level 3 Features," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 29, no. 1.

[29]. W. Lee, S. Cho, H. Choi, and J. Kim, "Partial fingerprint matching using minutiae and ridge shape features for small fingerprint scanners," Expert Systems with Applications, vol. 87, pp. 183–198, 2017.

[30]. A. Jain, A. Ross, and S. Prabhakar, "Fingerprint matching using minutiae and texture features," Proceedings 2001 International Conference on Image Processing (Cat. No.01CH37205), vol. 3, pp. 282–285, 2001.

[31]. F. Zhang, S. Xin, and J. Feng, "Combining global and minutia deep features for partial high-resolution fingerprint matching," Pattern Recognition Letters, 2017.

[32]. F. Benhammadi, M. N. Amirouche, H. Hentous, K. Bey Beghdad, and M. Aissani, "Fingerprint matching from minutiae texture maps," Pattern Recognition, vol. 40, no. 1, pp. 189–197, 2007.

[33]. L. Nanni and A. Lumini, "Local binary patterns for a hybrid fingerprint matcher," Pattern Recognition, vol. 41, no. 11, pp. 3461–3466, 2008.

[34]. L. Nanni and A. Lumini, "Descriptors for image-based fingerprint matchers," Expert Systems with Applications, vol. 36, no. 10, pp. 12414–12422, 2009.

[35]. X. Si, J. Feng, J. Zhou, and Y. Luo, "Detection and rectification of distorted fingerprints," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 37, no. 3, pp. 555–568, 2015.

[36]. J. Feng, "Combining minutiae descriptors for fingerprint matching," Pattern Recognition, vol. 41, no. 1, pp. 342–352, 2008.

[37]. Y. Chen and A. K. Jain, "Dots and incipients: Extended features for partial fingerprint matching," Biometrics Symposium, 2007.

[38]. M. Yamazaki, D. Li, T. Isshiki, and H. Kunieda, "SIFT-based algorithm for fingerprint authentication on smartphone," 6th International Conference on Information and Communication Technology for Embedded Systems, pp. 4–8, 2015.

[39]. S. Mathur, A. Vjay, J. Shah, S. Das, and A. Malla, "Methodology for partial fingerprint enrollment and authentication on mobile devices," International Conference on Biometrics, no. c, 2016.

[40]. S. Mathur, A. Vjay, J. Shah, S. Das, and A. Malla, ""Localization of corresponding points in fingerprints by complex filtering," 2016 International Conference on Biometrics, ICB 2016, no. c, 2016.

[41]. P. Agrawal, R. Kapoor, and S. Agrawal, "Partial fingerprint matching: Fusion of level 2 and level 3 features," Proceedings of the 5th IEEE International Conference on Confluence The Next Generation Information Technology Summit, pp. 504–508, 2014.

[42]. H. Choi, K. Choi, and J. Kim, "Fingerprint matching incorporating ridge features with minutiae," IEEE Transactions on Information Forensics and Security, vol. 6, no. 2, pp. 338–345, 2011.

[43]. P. S. P. W. Jiajia Lei, Qinmu Peng, Xinge You, Hiyam Hatem Jabbar, "Fingerprint Enhancement Based on Wavelet And Anisotropic Filtering," International Journal of Pattern Recognition and Artificial Intelligence, vol. 26, no. 01, 2012.

[44]. A. V Telore, "Study of Distortion Detection and Enhancement Methods for Fingerprint Images," 2016.

[45]. D. B. S. S. Raju Sonavane, "Noisy Fingerprint Image Enhancement Technique for Image Analysis: A Structure Similarity Measure Approach," International Journal of Computer Science and Network Security, vol. 07, no. 09.

[46]. K. Singh and R. Kapoor, "Image enhancement using Exposure based Sub Image Histogram Equalization," Pattern Recognition Letters, vol. 36, no. 1, pp. 10–14, 2014.

[47]. Y. Zhu and C. Huang, "An Adaptive Histogram Equalization Algorithm on the Image Gray Level Mapping," Physics Procedia, vol. 25, pp. 601–608, 2012.

[48]. A. K. Jain and F. Farrokhnia, "Unsupervised Texture Segmentation Using Gabor Filters," Pattern Recognition, vol. 24, no. 12, pp. 1167–1186, 1991.

[49]. F. Y. Yang J, Liu L, Jiang T, "A modified gabor filter design method for fingerprint image enhancement," National Laboratory of Pattern Recognition, Institute of Automation, Chinese

Academy of Sciences, Pattern Recognition, pp. 1805–1817, 2003.
[50]. M. Ch, "Image enhancement using Wavelet transforms and SVD," International Journal of Engineeering Science and Technology, vol. 4, no. 03, pp. 1080–1087, 2012.
[51]. Z. M. Win and M. M. Sein, "Texture feature based fingerprint recognition for low quality images," International Symposium on Micro-NanoMechatronics and Human Science, pp. 333–338, 2011.
[52]. M. K. K. and M. A. K. Muhammad Talal Ibrahim, Imtiaz A. Taj, "Fingerprint Image Enhancement Using Decimation Free Directional Adaptive Mean Filtering," Indian Conference on Computer Vision, Graphics and Image processing, pp. 950–961, 2006.
[53]. M. D. Maio D, "Quantitative-qualitative firiction ridge analysis: an introduction to basic and advanced Ridgeology," CRC Press, Boca Raton, 1990.
[54]. F. Zhao and X. Tang, "Preprocessing and postprocessing for skeleton-based fingerprint minutiae extraction," Pattern Recognition, vol. 40, no. 4, pp. 1270–1281, 2007.
[55]. A. B. and R. K. Ilya Blayvas, "Efficient Computation of Adaptive Threshold Surfaces for Image Binarization," Computer Science Department, Technion Institute of Technology Haifa, Israel.
[56]. A. M. Bazen and S. H. Gerez, "Segmentation of fingerprint images," Workshop on Circuits, Systems and Signal Processing, Veldhoven, The Netherlands, 2001.
[57]. A. K. J. Ratha N, Chen S, "Adaptive flow orientation based feature extraction in fingerprint images," Pattern Recognition1, vol. 28, pp. 1657–1672, 1995.
[58]. K. Nilsson and J. Bigun, "Localization of corresponding points in fingerprints by complex filtering," Pattern Recognition Letters, vol. 24, no. 13, pp. 2135–2144, 2003.
[59]. J. Bigun, "Pattern recognition in images by symmetries and coordinate transformations.," Understanding, vol. 68, no. 3, pp. 290–307, 1997.
[60]. S. Chikkerur and N. Ratha, "Impact of singular point detection on fingerprint matching performance," Proceedings - Fourth IEEE Workshop on Automatic Identification Advanced Technologies, AUTO ID 2005, vol. 2005, pp. 207–212, 2005.
[61]. J. Zhou, S. Member, F. Chen, J. Gu, and S. Member, "A Novel Algorithm for Detecting Singular Points from Fingerprint Images," Analysis, vol. 31, no. 7, pp. 1239–1250, 2009.
[62]. M. Zomorodian et al., "Automatic Detection of Singular Points in Fingerprint Images Using Convolution Neural Networks," Asian Conference on Intelligent Information and Database Systems. Springer, Cham, 2017., vol. 7197, pp. 11–21, 2012.
[63]. K. Nilsson and J. Bigun, "Complex filters applied to fingerprint images detecting prominent symmetry points used for alignment," Biometric authentication, pp. 39–47, 2002.
[64]. K. Fukushima, "Neocognitron: A self-organizing neural network model for a mechanism of pattern recognition unaffected by shift in position," Biological Cybernetics, vol. 36, no. 4, pp. 193–202, 1980.
[65]. P. Sermanet, S. Chintala, and Y. LeCun, "Convolutional Neural Networks Applied to House Numbers Digit Classification," 2012.
[66]. R. F. Nogueira, R. D. A. Lotufo, and R. C. Machado, "Evaluating software-based fingerprint liveness detection using Convolutional Networks and Local Binary Patterns," Proceedings of the IEEE Workshop on Biometric Measurements and Systems for Security and Medical Applications (BIOMS), pp. 22–29, 2014.
[67]. M. M. Min, "Intelligent Fingerprint Recognition System by Using Geometry Approach," IEEE International Conference on Current Trends in Information Technology, pp. 1–5, 2009.
[68]. A. K. Jain, S. Prabhakar, L. Hong, and S. Pankanti, "Filterbank-based fingerprint matching," IEEE Transactions on Image Processing, vol. 9, no. 5, pp. 846–859, 2000.
[69]. A. Ross, A. Jain, and J. Reisman, "A hybrid fingerprint matcher," Pattern Recognition, vol. 36, no. 7, pp. 1661–1673, 2003.
[70]. L. Nanni and A. Lumini, "A hybrid wavelet-based fingerprint matcher," Pattern Recognition, vol. 40, no. 11, pp. 3146–3151, 2007.
[71]. L. Ghiani, A. Hadid, G. L. Marcialis, and F. Roli, "Fingerprint Liveness Detection using Binarized Statistical Image Features," IEEE 6th International

Conference on Biometrics: Theory, Applications and Systems, BTAS 2013, 2013.
[72]. R. Cappelli, M. Ferrara, and D. Maltoni, "Fingerprint indexing based on minutia cylinder-code," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 33, no. 5, pp. 1051–1057, 2011.
[73]. S. Har-Peled, P. Indyk, and R. Motwani, "Approximate Nearest Neighbor: Towards Removing the Curse of Dimensionality Sariel," Theory of Computing, vol. 8, no. 1, pp. 321–350, 2012.
[74]. Y. Su, J. Feng, and J. Zhou, "Fingerprint indexing with pose constraint," Pattern Recognition, vol. 54, pp. 1–13, 2015.
[75]. G. Li, B. Yang, and C. Busch, "A score-level fusion fingerprint indexing approach based on minutiae vicinity and minutia cylinder-code," 2nd International Workshop on Biometrics and Forensics, IWBF 2014, pp. 0–5, 2014.
[76]. P. A. Parmar and S. D. Degadwala, "Approach Based on MV and MCC using SVM Classifier," International Conference on Communication and Signal Processing, pp. 1024–1028, 2016.
[77]. N. Dalal and B. Triggs, "Histograms of Oriented Gradients for Human Detection," 2010.
[78]. C. Gottschlich, E. Marasco, A. Y. Yang, and B. Cukic, "Fingerprint liveness detection based on histograms of invariant gradients," IJCB 2014 - 2014 IEEE/IAPR International Joint Conference on Biometrics, 2014.
[79]. C. Pornpanomchai and A. Phaisitkulwiwat, "Fingerprint Recognition by Euclidean Distance," Computer and Network Technology (ICCNT), 2010 Second International Conference on, pp. 437–441, 2010.
[80]. K. Manvjeet, S. Mukhwinder, G. Akshay, and S. S. Parvinder, "Fingerprint Verification System using Minutiae Extraction Technique," World Academy of Science, Engineering and Technology International Journal of Computer, Electrical, Automation, Control and Information Engineering, vol. Vol.2, no. No.10, pp. 3405–3410, 2008.
[81]. S. Prabhakar, A. K. Jain, S. Pankanti, and R. M. Bolle, "Minutia verification and classification for fingerprint matching," Proceedings 15th International Conference on Pattern Recognition. ICPR-2000, vol. 1, pp. 25–29, 2000.

© 2018 IJSRSET | Volume 4 | Issue 2 | Print ISSN: 2395-1990 | Online ISSN : 2394-4099
National Conference on Advanced Research Trends in Information and
g Technologies (NCARTICT-2018), Department of IT, L. D. College of Engineering,
Ahmedabad, Gujarat, India In association with
International Journal of Scientific Research in Science, Engineering and Technology

Review Paper on Throttled Load Balancing Algorithm in cloud computing environment

Smaranika Parida*1, Bakul Panchal*2
*[1]ME student, IT Department, L.D. College of engineering, Ahmedabad, Gujarat, India
[2]Assistant Professor, IT Department, L.D. College of engineering, Ahmedabad, Gujarat, India

ABSTRACT

The popularity of cloud computing is increasing day by day. It is a platform for providing dynamic pool of resources and virtualization to the cloud users. To properly handle user requests and utilize cloud resources, load balancing plays a vital role. For load balancing many algorithms have been designed to carry out the client's request towards the resources in order to maintain balance. Load balancing plays a vital role which can improve the Quality of Service (QoS) metrics, including resource utilization, performance, response time, cost and throughput. There are various algorithms are there to balance the load in cloud computing environment, such static and dynamic types. In this survey paper the focus is given to various throttled algorithms used for load balancing in cloud computing environment.

Keywords: Cloud Computing, Load Balancing, Throttled load balancing Algorithm.

I. INTRODUCTION

Cloud computing technology is an emerging technology, the computing power and storage technology sharing through the Internet. Cloud computing, fulfils user needs in pay as per the use basis, aiming in maximum resource utilization to prevent the waste of resources, to achieve maximum efficiency. Its features include on-demand service, extensive network access, resource pooling, brisk elasticity and measurement service.

However, the main issue in cloud computing is load balancing means assigning resources to the incoming requests efficiently with less response time. For executing the client's request with a minimal response time many algorithms are there like RR, Throttled, Stochastic Hill Climb etc [1]. In this paper focus is given on various throttled algorithm proposed by several researchers.

The organization of this document is as follows. In Section II,load balancing types are discussed, in section III' various proposed throttled algorithms are discussed. In section IV concludes the paper.

II. LOAD BALANCING TYPES

Load balancing algorithm has been designed to balance the whole load, and two types of load balancing algorithms are introduced:

A. Static Load Balancing :

Some characteristics of static load balancing algorithm are as follows: the load distribution decision does not depend on the current state of the system; they need prior knowledge about the system [2]. This algorithm is suitable for homogenous system environment [3]. They are not flexible. The performance of processors is explained at the starting of the execution and which does not change the executing process at run time for making changes in the system load [4]. Some of the static load balancing algorithms are: FCFS,RR, Max-min, Min-Min etc

B. Dynamic Load Balancing :

In dynamic Load balancing algorithm the load distribution decision is based on the current state of the

system, they do not require any prior knowledge about system resources [2]. This is suitable for heterogeneous system. Load balancing decision is dynamic in nature, means load balancing decision can be at run time [3]. This algorithm provides outstanding improvement in performance than static algorithm [2]. Some of the dynamic algorithm examples are ant colony, throttled, GA, Honey Bee Algorithm etc.

III. VARIOUS PROPOSED THROTTLED ALGORITHMS:

All paragraphs must be indented. All paragraphs must be justified, i.e. both left-justified and right-justified.

A. Throttled Load Balancing Algorithm:

Throttled Load Balancing Algorithm [5][6] is a dynamic algorithm for distribution of load that completely deploy on VMs. In VMs are grouped according to the requests they can handle. When the client sends the request, the load balancer immediately gets alert and searches for the group which can manage easily and allocates that request. But the issue is that the load balancer has to search for the suitable VM group, which results in the delay in operation. Working process of throttled load balancing is shown in figure 1.

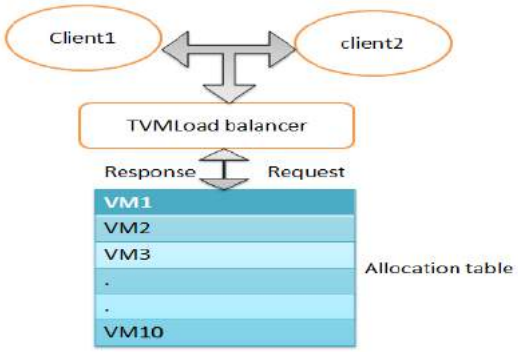


Figure 1. Throttled scheduling process

B. Modified Throttled Load Balancing Algorithm:

Modified Throttled algorithm [6] is modified version of Throttled Load Balancing Algorithm. It maintains a set of *VM* named as VM index table and that states the position of the VMs (i.e. Available/busy).when user requests come, VM at first index is initially selected depending upon the state. If VM is available then that request is assigned to that VM, and it returns (-1) to the DCC(Data Centre Controller), if VM is not found. Then the VM next to the already allocated VM is choose, when the next request arrives and this process is repeated continuously up to the index table size is reached, which is shown in Figure 2 below.

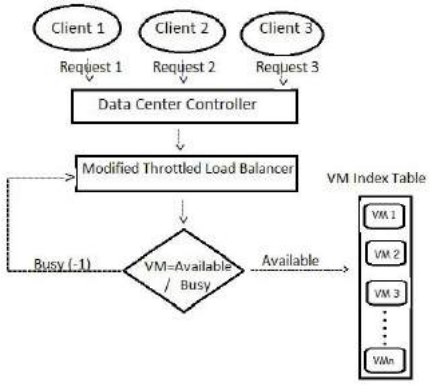


Figure 2. Flowchart of Modified Throttled Load Balancing Algorithm

Throttled algorithm has with better response time than Round-Robin but it failed to distribute load uniformly, it only overloads initial VMs , other VM are under loaded or get idle. In modified throttled, it distributes load uniform among VMs, and also improved response time as compared to existing throttled algorithms.

C. Hybrid Approach Using Throttled And ESCE Load Balancing[7]:

This is a hybrid load balancing algorithm, which is formed by Throttled and Equally Spread Current Execution (ESCE) algorithms.

Hybrid algorithm contains the excellent features of both ESCE and Throttled algorithms. In ESCE algorithm, load balancer maintain equal load to all the VMs connected to the data center. Using this technique the load balancer gets number of VMs by maintaining an index table and queue for the number of requests that are currently assigned to the VM. The load balancer scans the queue and the index list of VMs continuously. When a request comes, it is allocated to the available VM, that can handle the incoming request. If on the other hand there are two VMs, one is free and another one needs to be freed of the load, then the load balancer distributes some of the tasks of that VM to the free one so as to reduce the overhead of the former VM. Similarly the Throttled algorithm maintains an index table that lists all VMs as well as their states (i.e. Available or Busy).

Whenever a request comes to the data center the throttled load balancer scans the index table list from top

until the first available VM is found or the index list is scanned fully. If a match is found on the basis of size and availability of the VM, then the load balancer allocates the task to that VM. If no VM is available then the load balancer returns -1 and the request is get queued. This hybrid approach utilizes best features of both the algorithms to get more efficient load balancing algorithm. The hybrid algorithm maintains two things, index of list of VM allocation status and list to count the allocated request. The allocated request list count is compared with the index list of VMs. If VMs index list if greater than allocated request list implies that there are more VMs available to handle the request else request has been queued in order to wait for the availability of VM. Hybrid throttled load balancing shown in fig-3 below.

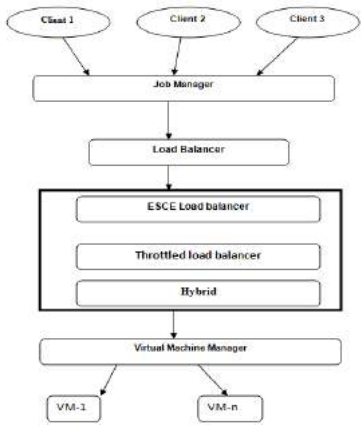


Figure 3. :Hybrid throttled load balancing Algorithms

D. Priority Based Modified Throttled Algorithm:

The Priority based Modified Throttled Algorithm (PMTA)[9] improves average execution time as compare to other existing load balancing algorithms such as Throttled and Round Robin. The working procedure of PMTA is shown in fig.4. In Cloudsim, DCC uses a VmLoadBalancer that determines for processing which VM should be assigned the next Cloudlet. VmLoadBalancers implementing '3' load balancing policies - RR Load Balancer VMs, Active Monitoring Load Balancer and Throttled Load Balancer, ensuring at any given time only a predefined number of Internet Cloudlets are allocated to a single VM. When a client request comes, the DCC (Data Centre Controller) queries for allocation table. Now if there is free VM available, then that request would directly be allocated to the suitable VM, if there is no such free VM available, then the VM's Load Balancer check for the job, executing in the V.M. which have the lowest priority then stops the execution and switch it to the Switching Queue. Switching Queue is used to hold the requests which have been removed temporarily from the VM due to the arrival of higher priority request .The high priority service request will be allocated to the vacant VM. The switched job have to be wait until availability of suitable VM but in this situation a problem may arise, If higher priority requests are continuously serve and the lower priority job would never get a chance to be executed and suffer from starvation. To overcome this problem, priority of the each waiting requests has increased. If any of the VM gets free, then it will check for new request and the waiting request, then selects the request having highest priority and allocate VM to it.

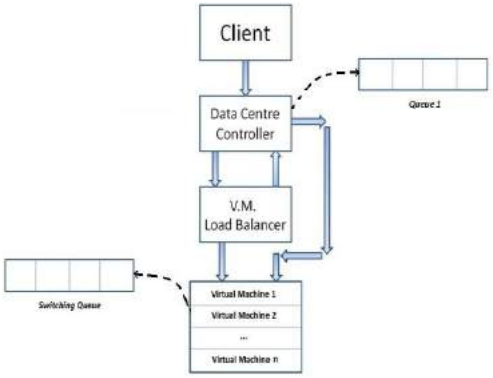


Figure 4. Priority Based Throttled Load Balancing Algorithm.

E. Weighted throttled load balancing approach

Weighted throttled load balancing algorithm [10]; is modification of the throttled load balancer, which assigns a weight to each of the VM in order to achieve better response time and processing time. In this weighted throttled load balancing approach it uses the concept of weights in throttled load balancing, the VMs are assigned different amount of the available processing power of server host to the individual application services. Tasks are assigned to the most powerful available VM between these VMs of different processing powers. In this algorithm a request is not queued and instead is assigned a busy VM that has a higher weight in comparison to others. This algorithm leads to better response time and processing time. The proposed weighted throttled VM load balancing algorithm is illustrates in figure 5 by use of a flowchart.

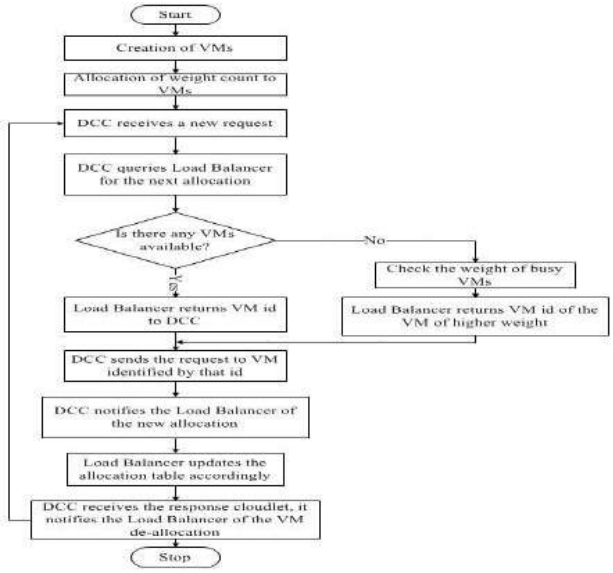


Figure 5. Weighted throttled load balancing algorithm

IV.CONCLUSION

The greatest challenge in cloud computing is load balancing, in which the user requests are allocated ensuring no VM get overloaded or under loaded. Throttled load balancing algorithm is a dynamic load balancing algorithm that can minimize the response time of service request. This survey paper focuses on different algorithms which are modified version of the throttled algorithm. These algorithms are designed to get better result and to overcome problems arise in previous one. In future more efficient throttled load balancing algorithm can be obtained which can give better result.

V. REFERENCES

[1]. Klaithem Al Nuaimi, Nader Mohamed, Mariam Al Nuaimi and Jameela Al-Jaroodi, "A Survey of Load Balancing in Cloud Computing: Challenges and Algorithms", in Second Symposium on Network Cloud Computing and Applications, 978-0-7695-4943-9/12, IEEE 2012

[2]. Desai, Tushar, and J. Prajapati. "A survey of various load balancing techniques and challenges in cloud computing." International Journal of Scientific & Technology Research 2.11 (2013): 158-161.

[3]. Raghava, N. S., and Deepti Singh. "Comparative study on load balancing techniques in cloud computing." Open Journal of Mobile Computing and Cloud Computing 1.1 (2014).

[4]. Ren, X., R. Lin and H. Zou, "A dynamic load balancing strategy for cloud computing platform based on exponential smoothing forecast" in proc. International Conference on. Cloud Computing and Intelligent Systems (CCIS), IEEE, pp: 220-224, September 2011

[5]. S. Mohapatra, K. Smruti Rekha, S. Mohanty, "A comparison of Four Popular Heuristics for Load Balancing of Virtual Machines in Cloud Computing," IJCA Journal, vol. 68(6), pp. 34-38, April 2013.

[6]. G. Domanal, G. R. Mohana Reddy, "Load Balancing in cloud computing Using Modified Throttled Algorithm," IEEE International Conference on Cloud Computing in Emerging Markets (CCEM), pp.1-7, October 2013.

[7]. Vishwas Bagwaiya, Sandeep k. Raghuwanshi, "Hybrid Approach Using Throttled And Esce Load Balancing Algorithms In Cloud Computing"

[8]. Shikha Garg, Rakesh Kumar Dwivediy, Himanshu Chauhanz, "Efficient Utilization of Virtual Machines in Cloud Computing using Synchronized Throttled Load Balancing", International Conference on Next Generation Computing Technologies (NGCT-2015) Dehradun, India, 4-5 September 2015.

[9]. Soumi Ghosh1, Chandan Banerjee1, "Priority Based Modified Throttled Algorithm in Cloud Computing"

[10]. Walugembe Hussein, Tao Peng and Guojun Wang," A weighted throttled load balancing approach for virtual machines in cloud environment", Int. J. Computational Science and Engineering, Vol. 11, No. 4, 2015

© 2018 IJSRSET | Volume 4 | Issue 2 | Print ISSN: 2395-1990 | Online ISSN : 2394-4099
National Conference on Advanced Research Trends in Information and
g Technologies (NCARTICT-2018), Department of IT, L. D. College of Engineering,
Ahmedabad, Gujarat, India In association with
International Journal of Scientific Research in Science, Engineering and Technology

A Review Paper on Multiplier Algorithms for VLSI Technology

Kajal Agrawal, Milind Shah, Gaurav Asari
Vishwakarma Government Engineering College, Chandkheda, Gujarat, India

ABSTRACT

In the era of digitalization, it is required to increase the speed of digital circuits while reducing area and power consumption. In any digital system, multiplication is a key element. One of the important parameter which affects the performance of entire system is performance of multiplier unit. Therefore, it is required to design efficient multiplier unit. To improve the efficiency of multiplier unit, it's needed to optimize various parameters such as speed and area. There are different multiplier algorithms discussed and compared in this paper for performance optimization.

Keywords: Array multiplier, Wallace tree multiplier, Booth algorithm, Karatsuba algorithm, Vedic multiplier.

I. INTRODUCTION

Multipliers play an important role in today's digital signal processing and various other applications. With advances in technology, many researchers have tried and are trying to design multipliers which offer either of the following design targets – high speed, low power consumption, regularity of layout and hence less area or even combination of them in one multiplier thus making them suitable for various high speed, low power and compact VLSI implementation.

The general multiplication method is performed by addition, subtraction and shifting operations. After each step of calculation partial product is generated, and this is the main factor that determines the performance of the multiplier. This repetitive addition requires more time to perform the operation [1].

To reduce the number of partial products to be added, Modified Booth algorithm is one of the most popular algorithms. To achieve speed improvements Wallace Tree algorithm can be used to reduce the number of sequential adding stages.

There are many types of multiplication algorithm such as Array multiplier, Wallace tree multiplier, Booth algorithm, Karatsuba algorithm, Vedic multiplier, etc.

The functionality of any algorithm is greatly dependent on functional parameters of multipliers. The parameters include delay, memory and power. For efficient working of any algorithm which includes multiplication the selection of multiplier plays a key role. The selection of multiplier is concerned by observing the delay and area of the multiplier [2].

The organization of this document is as follows. In Section 2(Multiplier Algorithms), detail of different algorithms for multiplication is discussed. In Section 3(Comparison and Discussion), comparison of different algorithms of multiplication is given. In Section 4(Conclusion)a conclusion is given based on comparison of different algorithms for FPGA implementation.

II. MULTIPLIER ALGORITHMS

A. Array Multiplier

The traditional method for multiplication is done by using array multiplier. Array multiplier is popular due to its regular structure. It is based on add and shift algorithm. In parallel multiplication operation, number of partial products to be added is the main parameter that determines the performance of the multiplier. Each partial product is generated by the multiplication

of the multiplicand with one multiplier bit. The partial products are then shifted according to their bit order and then added. Addition can be performed with normal carry propagate adder[3].

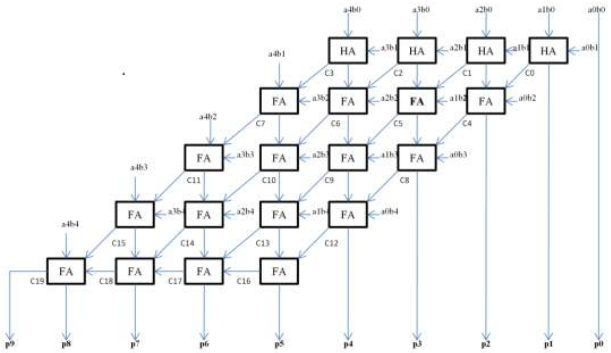


Figure 1. Structure of array multiplier

For n x n array multiplier, number of adders and gates needed are:

1. n(n-2) full adders
2. n half adders
3. n^2 AND gates

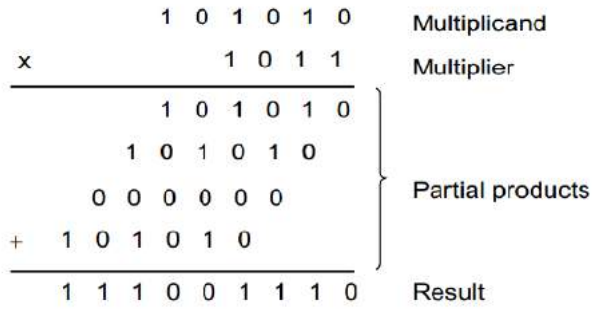


Figure 2. Example of array multiplier

The advantage of array multiplier is that it has minimum complexity and regular structure. And Disadvantages are large number of logic gates, so more chip area. It has high power consumption and it is limited to 16-bits.

B. Wallace-Tree Multiplier

A Wallace tree is an efficient hardware implementation of a digital circuit that multiplies two integers, devised by Australian Computer Scientist Chris Wallace in 1964.

The Wallace tree has three steps:

1. Multiply (that is – AND) each bit of one of the arguments, by each bit of the other, yielding n^2 results. Depending on position of the multiplied bits, the wires carry different weights.
2. Reduce the number of partial products to two by layers of full and half adders.
3. Group the wires in two numbers, and add them with a conventional adder.

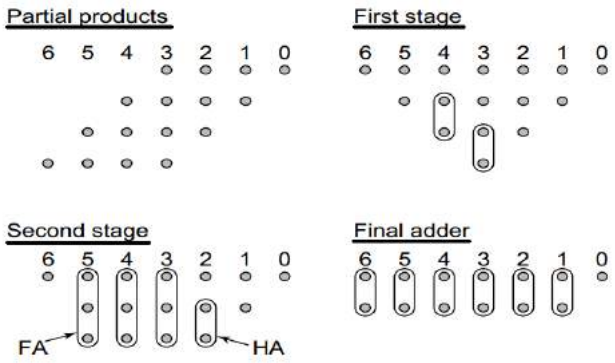


Figure 3. Example of Wallace tree multiplier

The advantage of Wallace-tree multiplier is that it becomes more pronounced for more than 16-bits. And Disadvantage is that a logarithmic depth reduction tree based CSA's has an irregular structure, therefore its design and layout is difficult.

C. Booth Algorithm

Booth's multiplication algorithm is a multiplication algorithm that multiplies two binary numbers in two's complement notation. The algorithm was invented by Andrew Donald Booth in 1950 while doing research on crystallographyatBirkbeck CollegeinBloomsbury, London. Booth's algorithm is of interest in the study of computer architecture. Booth multiplier is best for signed numbers. Booth used desk calculators that were faster at shifting than adding & created the algorithm to increase their speed.

Hence to reduce the iterations Booth's Algorithm isdeveloped for multiplying signed as well as unsignednumbers. It initiate with the ability to both add and subtract there are multiple ways to compute a product. Booth's algorithm is a multiplication algorithm that utilizes two's complement notation of signed binary numbers for multiplication. This multiplier can scan the three bits at a time hence the delay decreases. But the power consumption of this multiplier is more hence the efficiency of the system reduces [1].

Example: 3 x (-4).

m = 3 = 0011, r = (-4) = 1100.

Multiplication can be implemented by repeatedly adding one of two predetermined values A and S to a

product P, then performing a rightward arithmetic shift on P.

Step:1 Let x and y represent the number of bits in m and r.
Step:2 Determine the values of A and S , and the initial value of P. All of these numbers should have a length equal to $(x + y + 1)$.
A = Fill the most significant (leftmost) bits with the value of m. Fill the remaining $(y + 1)$ bits with zeros.
A= 001100000
S= Fill the most significant bits with the value of (−m) in two's complement notation. Fill the remaining $(y + 1)$ bits with zeros.
S = 110100000
P = Fill the most significant x bits with zeros. To the right of this, append the value of r. Fill the least significant (right most) bit with a zero.
P = 000011000

Step:3 Determine the two least significant (rightmost) bits of P.
If they are 01, find the value of $P + A$. Ignore any overflow.
If they are 10, find the value of $P + S$. Ignore any overflow.
If they are 00, do nothing. Use P directly in the next step.
If they are 11, do nothing. Use P directly in the next step.
A = 0011 0000 0, S = 1101 0000 0, P = 0000 1100 0

Step:4 Perform the loop four times:
P = 0000 1100 0. The last two bits are 00, Arithmetic right shift.
P = 0000 0110 0. The last two bits are 00, Arithmetic right shift.
P = 0000 0011 0. The last two bits are 10, So $P = P + S$.
P = 1101 0011 0. Then Arithmetic right shift.
P = 1110 1001 1. The last two bits are 11, Arithmetic right shift.
P = 1111 0100 1. Arithmetic right shift.
Result: The product is 1111 0100, which is −12.

The advantages ofbooth algorithm are that it used to reduce number of stages in multiplication and it performs two bits of multiplication at once, so it requires half number of stages. And its Disadvantage is that its each stage is more complex than simple multiplier.

D. Karatsuba Algorithm

The Karatsuba algorithm is a fast multiplication algorithm. It was discovered by Anatoly Karatsuba in 1960 and published in 1962. Karatsuba algorithm uses a divide and conquers approach. Where it breaks down the inputs into Most Significant half and Least Significant half.

Recursive application of Karatsuba Algorithm

If n are four or more, the three multiplications in Karatsuba's basic step involve operands with fewer than n digits. Therefore, those products can be computed by recursive calls of the Karatsuba algorithm. The recursion can be applied until the numbers are so small that they can (or must) be computed directly.

In a computer with a full 32-bit by 32-bit multiplier, for example, one could choose $B = 2^{31} = 2{,}147{,}483{,}648$, and store each digit as a separate 32-bit binary word. Then the sums $x_1 + x_0$ and $y_1 + y_0$ will not need an extra binary word for storing the carry-over digit (as in carry save adder), and the Karatsuba recursion can be applied until the numbers to multiply are only one digit long.

Karatsuba algorithm uses divide and conquer approach where it breaks down the inputs into Most significant half and Least significant half. Karatsuba algorithm is best suited for operands of higher bit length [4].
For multiplication, break down the input into two such as x_H and x_L.

Then apply following equations:

$$a = X_H Y_H$$

$$d = X_L Y_L$$

$$e = (X_H + X_L)(Y_H + Y_L) - a - d$$

$XY = a\, r^n + e\, r^{n/2} + d$

Where, r = 10

n = bit size

Example: 1234 x 4321

Ans:

a = 12 x 43

d = 34 x 21

e = (12+34) (43+21) – a – d

= (46 x 64) – a – d

Sub-problem: 1 a = 12 x 43

$a_1 = 1 \times 4 = 4$

$d_1 = 2 \times 3 = 6$

$e_1 = (1+2)(4+3) - 4 - 6$
$= 11$
$X_1Y_1 = (4 \times 10^2) + (11 \times 10) + 6$
$= 516$

Sub-problem: 2 d = 34 x 21
$a_2 = 3 \times 2 = 6$
$d_2 = 4 \times 1 = 4$
$e_2 = (3+4)(2+1) - 6 - 4$
$= 11$
$X_2Y_2 = (6 \times 10^2) + (11 \times 10) + 4$
$= 714$

Sub-problem: 3 46 x 64
$a_3 = 4 \times 6 = 24$
$d_3 = 6 \times 4 = 24$
$e_3 = (4+6)(6+4) - 24 - 24$
$= 52$
$X_3Y_3 = (24 \times 10^2) + (52 \times 10) + 24$
$= 1714$

Ans:
a = 12 x 43 = 516
d = 34 x 21 = 714
e = (12+34) (43+21) – a – d
= (46 x 64) – a – d
=1714 – 516 – 714
= 484

So,1234 x 4321 is,
$XY = (516 \times 10^4) + (484 \times 10^2) + 714$
$= 5332114$

The advantage of Karatsuba Algorithm is that it reduces number of multipliers by replacing them with adders.And its Disadvantage is that it is optimal if width of input is less than 16-bit.

E. Vedic Multiplier

Vedic Mathematics is a book written by the Indian monk Swami Bharati Krishna Tirtha and first published in 1965. It contains a list of mental calculation techniques claimed to be based on the Vedas. The mental calculation system mentioned in the book is also known by the same name or as "Vedic Maths". Its characterization as "Vedic" mathematics has been criticized by academics, who have also opposed its inclusion in the Indian school curriculum. Ancient mathematics has 16 different sutras, which are taken from Atharva Ved [3]. For multiplication, there are two sutras. First is Urdhva Tiryakbhyam sutra. Urdhva-Tiryagbhyam is one of the sutra from 16-Vedic sutras which performs the multiplication operation of two decimal numbers. Urdhva-Tiryagbhyam is the general formula applicable to all cases of multiplication of a large number by another large number. “Urdhva” means vertically and “Tiryagbhyam” means crosswise therefore it is also called as vertically and Crosswise Algorithm [3, 5]. It means “Vertical & Cross-wise”.

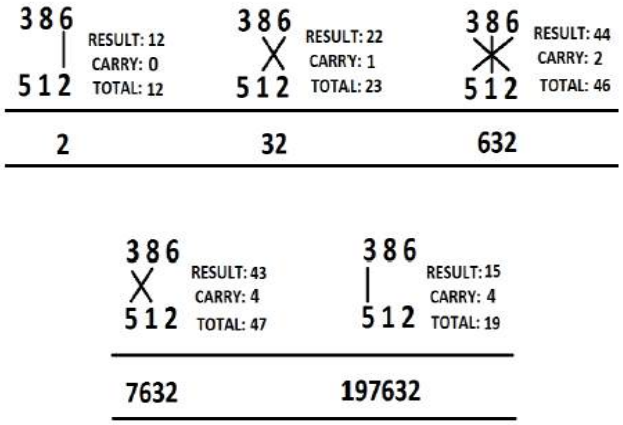


Figure 4. Example of Vedic multiplier [5]

The other one is Nikhilam Navatashcaramam Dashatah sutra. It means “All from 9 and last from 10”.The sutra basically means start from the left most digit and begin subtracting 9 from each of the digits; but subtract 10 from the last digit [1].

Its advantage is that it has Minimum Delay. And its Disadvantage is that as number of bits increases, multiplication process becomes tedious.

III. COMPARISON AND DISCUSSION

Delay is very high in booth algorithm compared to other algorithms. Vedic multiplier has low delay in comparison with Array, Wallace and Karatsuba algorithm. Total Power consumption is very low in Vedic multiplier and in Booth algorithm total power is very high. Memory requirement is very high in booth algorithm and Wallace multiplier and low in Vedic multiplier. Vedic multiplier consumes very less number of LUTs for FPGA implementation. Wallace multiplier consumes higher number of LUTs. Thus, Vedic multiplier gives efficient performance for FPGA implementation.

Comparison between above different algorithm is presented Table 1.

Table 1. Comparison Of Multiplier Algorithms For 16-Bit Multiplication

	Array	**Wallace**	**booth**	**Karatsuba**	**Vedic**
Delay (ns)	High	Moderate	Very High	Moderate	Very Low
Total Power	High	High	Very High	High	Low
Memory (kb)	Moderate	Very High	Very High	High	Very Low
Number of LUTs	Moderate	Very High	High	Moderate	Very Low

IV. CONCLUSION

This Paper presents different multiplier algorithms. From the comparison, Vedic multiplier is efficient among all the multiplication algorithms for 16-bit.Vedic multiplier shows the improved speed among all multipliers and it also reduces the memory requirement of the system. Therefore, Vedic multiplier is efficient and can be used in the DSP applications and Generic processors for faster computations.

V. REFERENCES

[1]. Hemangi P.Patil, S.D.Sawant. "FPGA Implementation of Conventional and Vedic Algorithm for Energy Efficient Multiplier" IEEE International Conference on Energy Systems and Applications (ICESA) Dr. D. Y. Patil Institute of Engineering and Technology, Pune, India 30 Oct-01 Nov, 2015.

[2]. G.Challa Ram, D.Sudha Rani, R.Balasaikesava, K.Bala Sindhuri. "Design of Delay Efficient Modified 16 bit Wallace Multiplier" IEEE International Conference On Recent Trends In Electronics Information Communication Technology, May 20-21, 2016, India.

[3]. S.P.Pohokar, R.S.Sisal, K.M.Gaikwad, M.M.Patil, Rushikesh Borse. "Design and Implementation of 16 x 16 Multiplier using Vedic Mathematics" International Conference on Industrial Instrumentation and Control (ICIC) College of Engineering Pune, India. May 28-30, 2015.

[4]. Arish S, R.K.Sharma. "An efficient floating point multiplier design for high speed applications using Karatsuba algorithm and Urdhva-Tiryagbhyam algorithm" IEEE International Conference on Signal Processing and Communication, March 16-18, 2015.

[5]. Paras Gulati, Harsh Yadav and Manoj Kumar Taleja. "Implementation of an Efficient Multiplier Using the Vedic Multiplication Algorithm"International Conference on Computing, Communication and Automation (ICCCA2016).

© 2018 IJSRSET | Volume 4 | Issue 2 | Print ISSN: 2395-1990 | Online ISSN : 2394-4099
National Conference on Advanced Research Trends in Information and g Technologies (NCARTICT-2018), Department of IT, L. D. College of Engineering, Ahmedabad, Gujarat, India In association with
International Journal of Scientific Research in Science, Engineering and Technology

Security and Privacy for Group Data Sharing Mobile Computing Environment

SK. Ahmad Shah[1], Dr. Penmesta V Krishna Raja[2]

[1]Assistant Professor, Department of IT, Aditya College of Engineering & Technology, Surampalem, Kakinada, East Godavari, Andhra Pradesh, India

[2]Principal & Professor, Kakinada institute of engineering and technology, Kakinada, East Godavari, Andhra Pradesh, India

ABSTRACT

Time and Trend has its own significance to build the technology smarter, better and easier to the end user. To the Better stretch of the Information Technology, the Innovation and renovation has changed computing approach to the next level. In this paper, we try to give the glimpse of the contextual virtual cloud storage in the public data distribution in terms of the Database in the virtualized Data Storage. These days cloud storage become common, but having the constraint towards the technical advancement is the Security. If we consider behavioral aspect of the cloud storage, we will come across much aspect. Hence, In this one we have overcome the public protection in terms of the privacy towards the authorization of public audit, which we call it as the best to the trend of the acknowledgement based identification with the cryptographic model where ever the node to the parallel cloud distributed elastic stretchable environment with the high end cloud data center marinating the graphics of the flow triggering the security in the public Domain. Considering all the aspects ; the classical Database and these days database has its own advantage to make overcome the hindrance to make the scalability, performance , efficiency and at the end the security is the major which the no-sqlite like Cassandra can overcome.

Keywords: Data storage, privacy preserving, Cloud database, Cassandra, confidentiality, encryption, adaptively, cost model

I. INTRODUCTION

Cloud Cover trusts the cloud platform. Clouds can attain suitable trustworthiness through trust management, replication, virtualization, and a variety of other technologies not typically available to mobile devices and other, stand-alone, cloud-assisted machines. Privacy preservation of computation results is beyond our scope. For such protection, we refer the reader to numerous related works on that subject, including Anonymous-Cloud; secure multiparty computation, and differential privacy. Proof validation through checkpoint chaining engenders a natural trade-o between assurance and computational expense through spot-checking. A spot-checking validated recomputed and checks each segment in the checkpoint chain with probability p. This reduces the total computation cost to a fraction p of the total, and detects erroneous computation results with probability p. This is necessary for the checkpoint equivalence. Cloud Cover therefore extends Java Continuation class with an equal's method that compares two suspended program states for semantic equivalence.

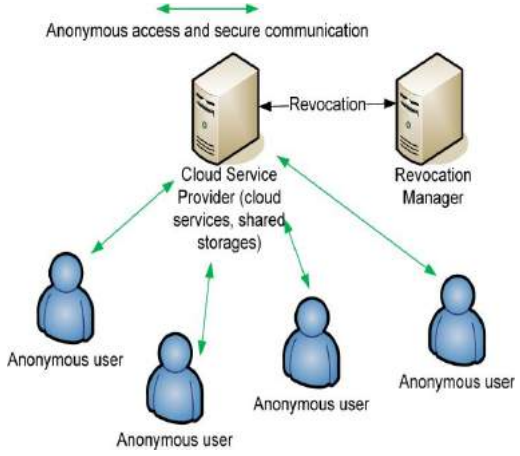


Figure 1. Illustration of the cloud server

Thus, clients may tune parameter p in accordance with their desired level of assurance and the expense of cloud computing time. Although Java supports suspension and resumption of computations via continuations, it does not support continuation equivalence-checking.

II. RELATED WORK

Cloud Cover proofs have the advantageous quality that the task of verifying them can be parallelized almost arbitrarily even when the original computation is not parallelizable. Thus, they derive maximal benefit from massively parallel architectures, like clouds. To demonstrate, we implement Cloud Cover for Hadoop Map Reduce , and use it validate non-parallelizable Java computations for message digest generation using SHA-1 (National Institute of Standards and Technology, 1995) and MD5 cryptographic hash functions. Experimental results indicate that Cloud Cover scales extremely well, with the only practical limit to parallelization stemming from the fixed overhead of dispatching new mappers and reducers. The checker is deployed on a Hadoop (Apache, 2013) cluster consisting of 6 Data Nodes and 1 Name Node. Node hardware is comprised of Intel Pentium IV 2.40, 3.00GHz processors with 2{4GB of memory each, running unix operating systems.

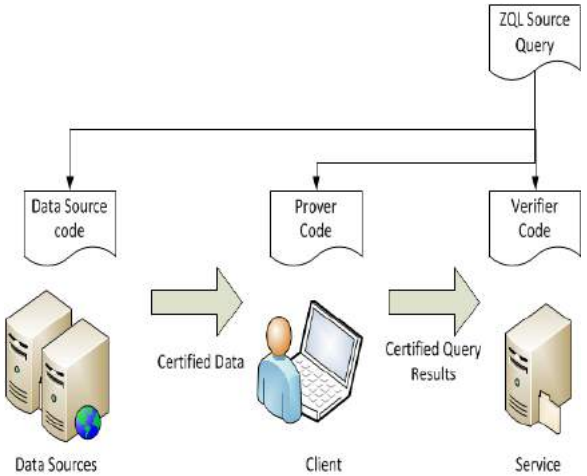


Figure 2. Data Center to Monitor the Data Flow

Java was installed and configured on each Data Node in the Hadoop distributed environment, making it available to distributed jobs. We implemented a mechanism for reading and writing checkpoints for map in Hadoop in an appropriate file format for equality-checking with Java. LZO compression was applied to all Hadoop file transfers to minimize transfer and storage costs. For trusted and trusted components of Cloud Cover, we use standard desktop computers with configurations similar to the individual cloud nodes above. For experiments, we select two non-parallelizable cryptographic.

III. METHODOLOGY

Technology has its own significance at the time when people having the extension for the more and more research and it's from Abacus to today's' cloud Computing. In the context of revolution of technology and its great advantage to its social, behavioral and other technical aspect where we come across the best of the cloud to province the virtual global village as the global world. In order for an attack against Anonymous Cloud to succeed, the manager or master node (or both) must be malicious. Managers are the only principals that receive decrypt able access tokens or credentials, and all other communications involving pseudonyms and data are conducted via Tor circuits having the master node as the only un-trusted endpoint. Managers are separate from CPs and have a much smaller attack surface because they do not process customer-submitted computations. Our experiments therefore assume that managers are trusted, but that master nodes are always malicious. In addition, we assume that a percentage p of slave nodes are also malicious and collude with the malicious master node in an effort to violate privacy. Aside from verifying checkpoint chain segments in parallel, we additionally parallelized the checkpoint equality checking procedure in our implementation. Continuations are stacks that can be partitioned arbitrarily into sub-stacks that can all be checked in parallel for equivalence. We implemented this for Java by introducing a continuation compare method. During comparison, instead of equality-checking each pair of objects inside the checkpoints, a map can redirect them to other map by submitting new jobs in Hadoop. The advantage is that if any individual checkpoint-pair is extremely large (e.g., very large stacks), then the checkpoint equality-checking job can be parallelized to compensate. In our experiments, the stacks are not that large, so this feature went unexercised.

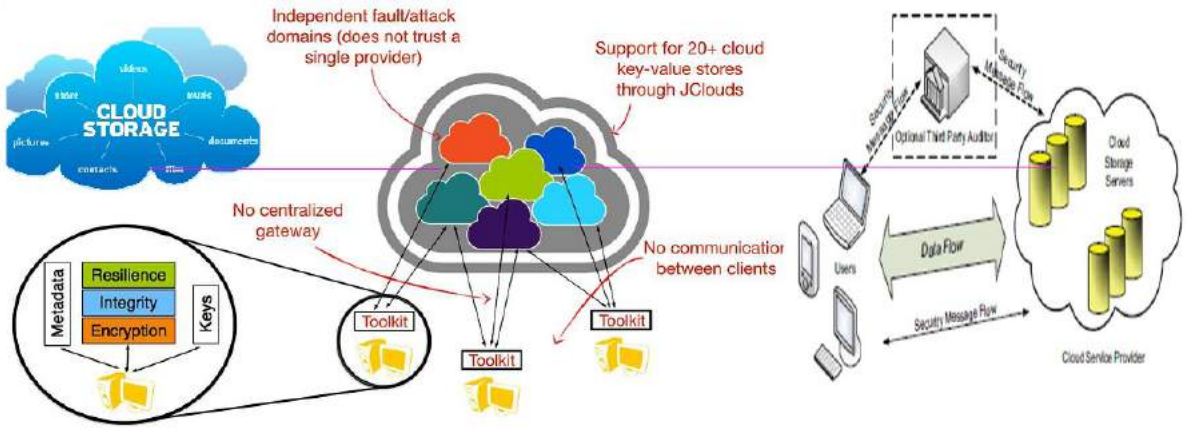


Figure 3. Architecture Design of the Secured Public Data in the Database of Cassandra

Therefore, using this approach not only the security and privacy concerns of cloud consumers can be addressed more effectively, but also the burden of managing end-users' identities and fine-granular access control will be reduced from cloud service providers. Unfortunately, all of these approaches require a significant redesign of most software. For example, typical Android apps are not easily modified to contain inextricable, secret computations or cryptographically verifiable compositions. Two states are equivalent if they consist of equal-length stacks whose corresponding slots contain equivalent values and objects. Deciding such semantic equivalence is non-trivial in general; for example, the states may contain objects with private fields to which the continuation object lacks access, or they may include fields whose values are semantically equivalent but non-identical. Fortunately, all Java objects have their own equal's methods, which encode an object-specific notion of semantic equivalence.

3.1 Evaluation and Analysis

As a result, few mainstream mobile computing devices have adopted these technologies. Moreover, many of these solutions rely on software obfuscation, which does not provide rigorous guarantees, since clever attackers can potentially reverse the obfuscation the sharp increase in communications overhead potentially invites denial-of-service attacks by customers who request unreasonably long circuits. Recall that master nodes can report computational expense information associated with anonymous jobs to managers by labeling it with the encrypted ownership data they received during authentication.

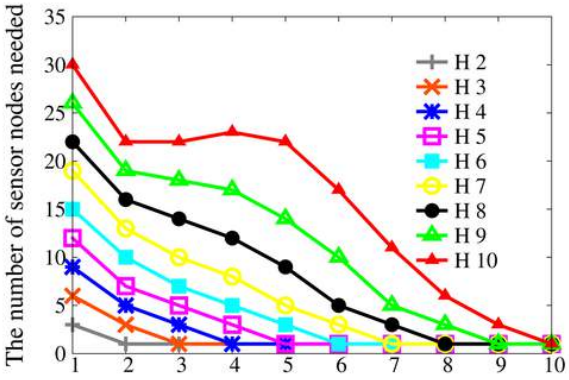


Figure 4. Node with the Peak Value

We therefore recommend incentivizing reasonable values of k by charging customers proportionally to the communications overhead incurred by their demanded level of privacy. This allows the master node to report the expense without knowing the identity of the customer. Managers may also want to impose a mandatory upper limit on k during authentication to further control congestion.

IV. CONCLUSION AND FUTURE WORK

Cloud computing is an emerging paradigm which its cost-effectiveness and edibility have given it a tremendous momentum. In this paper, we try to put forward the concept of the cloud in the aspect of the privacy preserving towards the public shared node data. It may lead to the extent of the cloud with the variant of the most suitable technological advancement of the recent solution. In the Public shared cloud computing where the data passed through the network which needs to be robust, secure and highly preserved in the sense no can cal replicate the data while reaching to the next node

of the cloud server? Hence, we can make the sense of the cloud for the further research oriented making the global world as the data can be se3cured in the cloud architectural designs model of the Data center. However, there are many security challenges that, if not addressed well, may impede its fast adoption and growth. This dissertation primarily addresses the problem of sharing, managing and controlling access to sensitive resources and services in an integrated cloud environment. The primary conclusion of our research is that adoption of user-centric security models and shifting certain parts of communication and computation to the client side allows us to provide the cloud consumers with more visibility and control over their resources.

V. REFERENCES

[1]. B. Wang, B. Li, and H. Li, "Oruta: Privacy-Preserving Public Auditing for Shared Data in the Cloud," Proc. IEEE Fifth Int'l Conf.Cloud Computing, pp. 295-302, 2012.

[2]. M. Armbrust, A. Fox, R. Griffith, A.D. Joseph, R.H. Katz, A. Konwinski, G. Lee, D.A. Patterson, A. Rabkin, I. Stoica, and M. Zaharia, "A View of Cloud Computing," Comm. ACM, vol. 53, no. 4, pp. 50-58, Apr. 2010.

[3]. K. Ren, C. Wang, and Q. Wang, "Security Challenges for the Public Cloud," IEEE Internet Computing, vol. 16, no. 1, pp. 69-73, 2012.

[4]. D. Song, E. Shi, I. Fischer, and U. Shankar, "Cloud Data Protection for the Masses," Computer, vol. 45, no. 1, pp. 39-45, 2012.

[5]. C. Wang, Q. Wang, K. Ren, and W. Lou, "Privacy-Preserving Public Auditing for Data Storage Security in Cloud Computing," Proc. IEEE INFOCOM, pp. 525-533, 2010.

[6]. B. Wang, M. Li, S.S. Chow, and H. Li, "Computing Encrypted Cloud Data Efficiently under Multiple Keys," Proc. IEEE Conf. Comm. and Network Security (CNS '13), pp. 90-99, 2013.

[7]. R. Rivest, A. Shamir, and L. Adleman, "A Method for Obtaining Digital Signatures and Public Key Cryptosystems," Comm. ACM, vol. 21, no. 2, pp. 120-126, 1978.

[8]. The MD5 Message-Digest Algorithm (RFC1321). https://tools.ietf.org/html/rfc1321, 2014.

[9]. G. Ateniese, R. Burns, R. Curtmola, J. Herring, L. Kissner, Z. Peterson, and D. Song, "Provable Data Possession at Untrusted Stores," Proc. 14th ACM Conf. Computer and Comm. Security (CCS '07), pp. 598-610, 2007.

[10]. H. Shacham and B. Waters, "Compact Proofs of Retrievability," Proc. 14th Int'l Conf. Theory and Application of Cryptology and Information Security: Advances in Cryptology (ASIACRYPT '08), pp. 90-107, 2008.

[11]. C. Erway, A. Kupcu, C. Papamanthou, and R. Tamassia, "Dynamic Provable Data Possession," Proc. 16th ACM Conf. Computer and Comm. Security (CCS'09), pp. 213-222, 2009.

[12]. Q. Wang, C. Wang, J. Li, K. Ren, and W. Lou, "Enabling Public Verifiability and Data Dynamic for Storage Security in Cloud Computing," Proc. 14th European Conf. Research in Computer Security (ESORICS'09), pp. 355-370, 2009.

[13]. C. Wang, Q. Wang, K. Ren, and W. Lou, "Ensuring Data Storage Security in Cloud Computing," Proc. 17th Int'l Workshop Quality of Service (IWQoS'09), pp. 1-9, 2009.

[14]. B. Chen, R. Curtmola, G. Ateniese, and R. Burns, "Remote Data Checking for Network Coding-Based Distributed Storage Systems," Proc. ACM Workshop Cloud Computing Security Workshop (CCSW'10), pp. 31-42, 2010.

[15]. Y. Zhu, H. Wang, Z. Hu, G.-J. Ahn, H. Hu, and S.S Yau, "Dynamic Audit Services for Integrity Verification of Outsourced Storages in Clouds," Proc. ACM Symp. Applied Computing (SAC'11), pp. 1550-1557, 2011.

[16]. N. Cao, S. Yu, Z. Yang, W. Lou, and Y.T. Hou, "LT Codes-Based Secure and Reliable Cloud Storage Service," Proc. IEEE INFOCOM, 2012.

[17]. B. Wang, B. Li, and H. Li, "Certificateless Public Auditing for Data Integrity in the Cloud," Proc. IEEE Conf. Comm. and Network Security (CNS'13), pp. 276-284, 2013.

[18]. C. Wang, S.S. Chow, Q. Wang, K. Ren, and W. Lou, "Privacy-Preserving Public Auditing for Secure Cloud Storage," IEEE Trans. Computers, vol. 62, no. 2, pp. 362-375, Feb. 2013.

© 2018 IJSRSET | Volume 4 | Issue 2 | Print ISSN: 2395-1990 | Online ISSN : 2394-4099
National Conference on Advanced Research Trends in Information and g Technologies (NCARTICT-2018), Department of IT, L. D. College of Engineering, Ahmedabad, Gujarat, India In association with
International Journal of Scientific Research in Science, Engineering and Technology

An Overview of GANs

Ketakee Nimavat[1], Tushar Champaneria[2]
[1]UG student, Computer Engineering Department ,L. D. C. E, Ahmedabad, Gujarat, India
[2]Assistant Professor, Computer Engineering Department ,L. D. C. E, Ahmedabad, Gujarat, India

ABSTRACT

GANs are the most recent and exciting development in the field of generative models. Inspired by game theory, the system consists of two models which learn on its own. In this paper, the architecture of the models is explored along with its ideology. Research on application in GANs have shown improvement in various possible areas over the traditional models. Areas pertaining to image such as image filling, text to image are explored. Finally, the paper also discusses various open ended research areas that would help make GANs a more stable and optimal system.
Keywords: GANs, Image Processing, Generative models

I. INTRODUCTION

Over the past few years, innumerable strides have been made in machine intelligence right from ELIZA to conquering jeopardy to beating the world's best player at chess and the most recent ones such as winning a game of GO and winning at DOTA. While we have excelled at various tasks of narrow domain, true human intelligence in terms of human characteristics such as empathy, creativity, spontaneity, context awareness and other such traits remain to be displayed yet. One of these characteristics, creativity, consists of generation of information, be it pictures or text or music. What makes us human is not the ability to calculate in our heads but the ability to understand context and respond. Generative algorithms have often been seen as the measure of 'understanding' that a machine has of the training data. Since they generate after learning from a set of information, if they generate correctly, they must've understood the information correctly as well. Hence, generation is seen as a measure of understanding of the system. They are used in generating texts, images, speech and to generate system behaviours or system states. It must be noted that GANs cannot be used for discrete data and hence applications with continuous data will be able to use GANs.

Generative algorithms span a broad spectrum, from the Naïve Bayes to deep belief networks. The most recent addition to the existing toolset of Gaussian Mixture model, Hidden Markov models, Latent Dirichlet Allocation and Restricted Boltzmann machines are Generative Adversarial Networks or GANs. In 2014, Ian Goodfellow et al proposed a new model called Generative Adversarial Networks (GANs) in the paper presented at NIPS.[1] Since then, they have been hailed as the next step in machine learning.

GANs have often been favoured for unlike previous networks they fit to complex real life data very well. For example, the quality of the images they produce and the closeness of the image to the real image is superior to that achieved using existing systems. What makes GANs different is that the system is based on an approach which is a combination of game theory and supervised learning. On the other hand, Deep Belief network, a popular neural network based method for generation, is a complex one and is found to be difficult to scale due to its reliance on energy functions that involve large amounts of computations. GANs hence give us a new perspective to approach machine learning with. GANs also are a new concept because of their architecture which enables them to grasp the major characteristics of the system. The architecture is based on Game theory and is discussed in the next section.

In this paper, the working of GANs is discussed in section 2. Section 3 explores various research work

where in GANs are applied in different tasks and in Section 4, the areas of GANs that require more work are discussed.

II. THE WORKING OF GANS

The architecture consists of two networks: one generator network and one discriminator network. The generator generates data that the discriminator guesses as either real or fake. The goal of the generator is to fool the discriminator into thinking the item it produced is the real one. The goal of the discriminator is to not be fooled by fake generations. An analogy of a forger and an original painting can be used here, the forger first uses random colours and tries to create a piece similar to the painting. Then it keeps on editing the painting until it looks like the original painting but it isn't exactly the same. The goal of the forger is to fool the buyer into thinking that it is indeed an original painting. The buyer's job is to reject the painting if it doesn't look like the original one, the buyer is a trained expert and knows fake paintings from real ones. The generator is the forger and the buyer is the discriminator network.

Both the generator and the discriminator can be multilayer perceptron or neural networks. Both the networks have back propagation enabled to learn better. Since the discriminator is trained to differentiate between real and fake. The discriminator maps from input space to [0,1] and is a binary classifier.

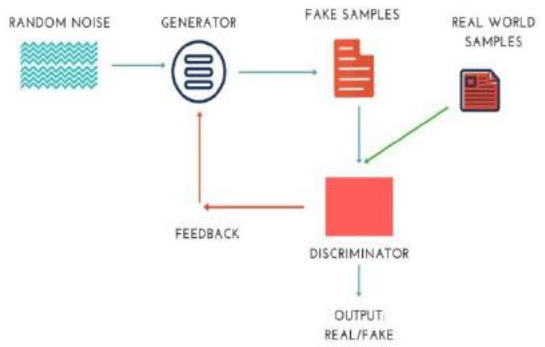


Figure 1. GAN components and working

Generator generates from random set with a pre-decided seed and from a predefined latent space. The key here is that the generator is trained after generating a sample not before. This way, the generator gets feedback on its generated input and can learn from it without wasting any chances of learning. The generated items are sent to a discriminator along with real samples and the discriminator either labels each image as real or fake. The feedback is then sent back to the generator network which learns via back propagation.

This particular nature of two competing networks have made GANs to be powerful systems, various research approaches have portrayed their efficiency. In the coming section, we discuss various possible uses of GANs and the research papers that have portrayed their use.

III. APPLICATION AND ARCHITECTURE OF GANS

Since they were first proposed in 2014, they have been experimented with and used for various real world scenarios. Variations of GANs involve using different types of networks and different combinations of networks. These variations are discussed below:

Image to image translation[2] (Conditional GAN(cGAN))
This approach uses a conditional GAN for solving image translation problems such as translation of an image from a Black and white space to a RGB space. CGANs prove to be a successful approach as compared to other existing approaches and also successfully maintain the details in the picture.

Another approach using a similar cGAN finds that cGANs are able to generate faces without overfitting the training dataset. It also proves that providing a conditional matrix to GAN does indeed improve performance.

Image generation[3] (Recurrent GAN)
This approach uses a recurrent GAN to show that an iterative convolutional neural net when trained to update features on a single canvas can generate images in general and not just images based on training images.

Enhancing resolution in images[4] (SRGAN(GAN for image Super Resolution))
When producing high resolution images from low resolution ones, GANs have seen to surpass traditional methods. In particular they have been successful at maintaining the details even at scale which other methods would blur out.

Text to image synthesis[5], [6][7]: (DCGAN, STACKGAN)
One approach uses a DCGAN which is conditioned on word embedding to generate images but finds that the images are recognizable and possess parts of the description when GANs are used with interpolation.

In another approach called StackGAN, two stacked GANs have been used to generate images. The first layer consists of generation of basic features and components based on the description and the second layer of GAN then edits the image for resolution and details to make it more accurate. The approach has seen to generate more photo realistic details at a higher resolution with more accuracy and more diversity than at lower resolutions.
Image filling[8]

Adobe created a recent application using GANs where in the user draws a shape and the system fills in the possible textures and colours. Such applications would help introduce more functions and enhance current functions in image editing softwares.

The applications of GANs are impressive but not without drawbacks. The drawbacks however present opportunities for making GANs more flexible and adaptable. In the next section, they are discussed with an aim to provide researchers a direction to work in and for implementers to be aware of the shortcomings of the system.

IV. OPEN ENDED RESEARCH AREAS IN GANS

In the end however, it mustn't be forgotten that these are mathematical approaches to generation that hope to arrive to a solution. Owing to this nature, there are some fundamental flaws such as drawing four eyes instead of two, lack of perspective and lack of 3D understanding.

This happens mainly because the system doesn't know the boundaries within which it can experiment. The system attempts to find which change is best suitable and closest to the original but in doing so it has no constraints other than the measures of similarity. It doesn't know the properties of the very object it is changing. This can lead to discrepancies such as a cat with two faces or a cow with two bodies. In image related applications, as the dimensionality of the picture increases, so does the scope of error. For such a scenario, perhaps finding a way to store information about boundaries of experimentation will help train the system faster. What this tells us is that the system does indeed learn the general representation of an object but fails to capture the characteristics of the object.

Mode Collapse: An open ended research area is the tendency of the system to collapse, also called mode collapse. The system is a combination of two interacting entities. Feedback and input from each is pivotal to the survival of the other. In this situation, they might go on forever or failure of one might cause overall failure.

Equilibrium: From the system's perspective, the system is designed to minimize the cost function of each the generator and the discriminator but it isn't designed to reach Nash Equilibrium which causes the system to keep going without reaching anywhere. Hence, convergence of the system is not guaranteed. Hence another open ended research area is algorithms that help a system reach equilibrium.

Repetition: GANs often fall into a trap where in when one image is accepted by the discriminator i.e. manages to fool the discriminator, they either create the same image again and again or give minor variations of the image as output. This limits the diversity in the output and introduces inflexibility in the system. One proposed solution to this issue to training the data in minibatches. This has shown to reduce redundancy.[9]

Evaluation metrics: As of now, there is a lack of proper evaluation metrics for GANs. The ones being used are nearest neighbour algorithm for measuring distance between the produced image/sample and the training set. Another proposed approach is battling two GANs to see which would win.[3]

Hence, GANs are in their nascent stage and hold a lot of potential after implementation. Various improvements such as training in mini batches have been proposed in [9]. Solving the above mentioned problems will help make GANs more accessible for various other fields of implementation such as speech generation and real time image generation applications such as image editors and video editors.

V. CONCLUSION

GANs produce quality outputs and provide a novel way to create generations similar to real world scenarios. Research is needed to make GANs more controllable and moldable. Apart from that, various problems mentioned above are areas where work can be done to provide a more usable result. GANs will be useful in image applications, spontaneous and realistic speech generation, cross field applications such as image to text, image description, image captioning etc. Image editing softwares could benefit largely from GANs. Yoshua Bengio has called GANs to be the most revolutionary technology in machine learning in the past decade. It is hoped that future work will bring a step closer in our attempt to embed human qualities such as creativity in the intelligence we build.

VI. REFERENCES

[1]. I. Goodfellow et al., "Generative Adversarial Nets," Adv. Neural Inf. Process. Syst. 27, pp. 2672-2680, 2014.

[2]. P. Isola, J.-Y. Zhu, T. Zhou, and A. A. Efros, "Image-to-Image Translation with Conditional Adversarial Networks," 2016.

[3]. D. J. Im, C. D. Kim, H. Jiang, and R. Memisevic, "Generating images with recurrent adversarial networks," 2016.

[4]. C. Ledig et al., "Photo-Realistic Single Image Super-Resolution Using a Generative Adversarial Network," 2016.

[5]. S. Reed, Z. Akata, X. Yan, L. Logeswaran, B. Schiele, and H. Lee, "Generative Adversarial Text to Image Synthesis," Icml, pp. 1060-1069, 2016.

[6]. H. Zhang et al., "StackGAN: Text to Photo-realistic Image Synthesis with Stacked Generative Adversarial Networks," 2016.

[7]. A. Radford, L. Metz, and S. Chintala, "Unsupervised Representation Learning with Deep Convolutional Generative Adversarial Networks," pp. 1-16, 2015.

[8]. J. Y. Zhu, P. Krähenbühl, E. Shechtman, and A. A. Efros, "Generative visual manipulation on the natural image manifold," Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics), vol. 9909 LNCS, pp. 597-613, 2016.

[9]. T. Salimans, I. Goodfellow, W. Zaremba, V. Cheung, A. Radford, and X. Chen, "Improved Techniques for Training GANs," pp. 1-10, 2016.

© 2018 IJSRSET | Volume 4 | Issue 2 | Print ISSN: 2395-1990 | Online ISSN : 2394-4099
National Conference on Advanced Research Trends in Information and ing Technologies (NCARTICT-2018), Department of IT, L. D. College of Engineering, Ahmedabad, Gujarat, India In association with International Journal of Scientific Research in Science, Engineering and Technology

Application of Data Mining Techniques in IoT: A Short Review

Akshat Savaliya*[1], Aakash Bhatia[2], Jitendra Bhatia[3]
*[1]Computer Department, LD college of Engineering, Ahmedabad, Gujarat, India
[2]Computer Department, DA-IICT, Gandhinagar, Gujarat, India
[3]Computer Department, Vishwakarma Government Engineering College, Ahmedabad, Gujarat, India

ABSTRACT

Internet of Things (IoT) has been growing rapidly due to recent advancements in communications and sensor technologies. Interfacing an every object together through internet looks very difficult, but within a frame of time Internet of Things will drastically change our life. The enormous data captured by the Internet of Things (IoT) are considered of high business as well as social values and extracting hidden information from raw data, various data mining algorithm can be applied to IoT data. In this paper, We survey systematic review of various data mining models as well as its application in Internet of Thing (IoT) field along with its merits and demerits. At last, we discussed challenges in IoT.

Keywords: Internet of Things, Data Mining, Machine Learning, Application of Data Mining

I. INTRODUCTION

The Internet of Things(IoT) refers to the type of the network which connect anything i.e. physical objects-devices, buildings, vehicles and other items embedded with software, sensors and network connectivity based on stipulated protocols that enables these objects to collect and exchange data. In our daily lives, we have become more reliant on IoT with our wearable tech, appliances, our cars, how we receive health care. Due to Seamless integration of classical networks with IoT, it enables a great vision that all things can be easily monitored and controlled which results in to voluminous data. So, in order to make IoT more smarter, lots of data analysis is needed for which one of the most solution is data mining. Much research in recent years has focused on data mining in Internet of Things (IoT) which connects physical objects, person to person, person to machine or machine to machine via internet and manages information [11].

Data mining process refers to the process of semiautomatically analyzing large databases for pattern mining which are innovative, legitimate, useful and understandable which is also known as Knowledge Discovery in Databases (KDD). Data mining or KDD process includes problem formulation, data collection, data cleaning i.e. preprocessing, transformation, choosing mining task/method and result evaluation/visualization. Knowledge discovery is an iterative process.

Data mining overlaps with other fields like statistics, machine learning, artificial intelligence, databases but mainly it focuses on automation of handling large heterogeneous data, algorithm and scalability of number of features and instances.

As of late, an expanding number of rising applications manage countless sensor information in Internet of Things (IoT) due to a wide assortment of sensor gadgets on detecting layer. The extensive scaling of heterogeneous sensor creates a problem of data

handling which is one of key issue for the IoT framework application. Sensors in IoT applications sense the complicated environment and generates an enormous data that must be filtered and cleaned so that it can be interpreted and user will be provided with insights of the data collected in form of patterns [13]. Across various network infrastructures, IoT allows sensing of the objects and remotely access which in turn enables opportunities for a better integration between real and computerized world. It results into an improved effectiveness, exactness and better economic outcomes. Each query can be recognized exceptionally by the utilization of its installed registering framework, However these objects can interoperate inside the current infrastructure of the Internet. Assessments recommend that IoT will be an accumulation involving roughly 50 billion objects before the finish of 2020 [4]. The way toward finding and investigating helpful patterns in a large amount of information is what we refer it as Data Mining. Data mining can likewise be characterized as a sensible procedure that is used to investigate and look through extensive measure of enormous information so as to discover more valuable information in it. Till date the pattern finding procedures were not full fledge utilized and the information gathered was only an static accumulation of databases. But with the method of finding patterns in the information, more usage of the information is being acquired which settles on better choices for the advancement of the business or social aspect.

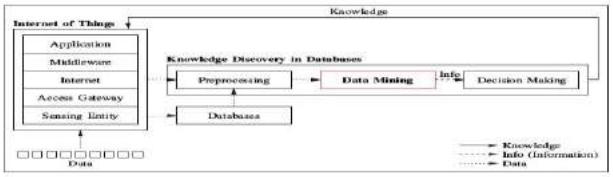


Figure 1. Data Mining Integrated IoT Architecture [13]

Figure 1 depicts clearly that, IoT gathers information from various sources, which may contain information for the IoT itself. KDD, when connected to IoT, will change over the information gathered by IoT into valuable data that would then be able to be changed over into learning. The information digging step is in charge of extracting designs from the yield of the information preparing step and after that encouraging them into the basic leadership step, which deals with changing its contribution to meaning information. It is essential to take note of that, every means of the KDD procedure may have a solid affect on the previous stage of mining. For instance, not all the traits of the information are valuable for mining; in this way, highlight determination is normally used to choose the key qualities from each record in the database for mining. The result is that information mining calculations may experience serious difficulties to discover valuable data (e.g., placing designs into suitable gatherings) if the chosen properties can't completely speak to the qualities of the information. It is likewise vital to take note of that the information combination, substantial scale information, information transmission, and decentralized processing issues may strongerly affect the framework execution and benefit the nature of IoT than KDD or information mining calculations alone may have on the customary applications. The key contribution of this paper includes:

- ✓ We explore the basic architecture of data mining assisted IoT.
- ✓ We discuss variants of data mining models for the IoTs.
- ✓ We survey various data mining application techniques used in IoT.

The rest of the paper includes the following. Section II discusses variants of work done in IoT using data mining techniques. Section III includes framework design of various data mining model. Section IV highlights various IoT application domain while section V discusses key issues involved in data mining of IoT. At the last, we conclude the paper.

II. RELATED WORKS

Since Internet of Things is a completely new concept, researches are still at the initial stage. Right now, there are few works regarding data mining in the IoT. Following are some of the work trending in this domain. Masciari [6] investigated mining in RFID data stream. Which tracks moving data made by different gadgets of IoT i.e. RFID sensor network, GPS gadgets, satellites and so on. Hector Gonzalez [5] proposed a model

utilizing which RFID data can be collected, which thus protect changes in it along with compression and path-dependent aggregation. Xiaolei Li [7] come up with a new system known as ROAM, which identify inconsistency in moving objects. SpatioTemporal Sensor Graphs (STSG) proposed by Betsy George [10] is used to model and mine sensor data. It can discover inconsistent patterns, incorporated areas at each time interim, and even hubs eligible to be future hotspot. Jae-Gil Lee [8] gave a new classification to track path followed by an object named TraClass using trajectory-based clustering and hierarchical region. Discovery of a knowledge from sensor data. Joydeep Ghosh [9] put forward a universal probabilistic system that permits supervised learning under computational/power/memory limitations. In the domain of data mining, a few broad organizations like Yahoo, Facebook, and Twitter pick up and supply attempts to open source ventures said by author in [10].

In [13], author proposed a design for high-performance data mining module of KDD for IoT with the three key considerations i.e. choosing objective, characteristics of data, and mining algorithm. Objective: The relevant mining techniques needs to be decided for the issue to be settled by the KDD. The suppositions, restrictions, and estimations of the issue should be determined first in order to accurately characterize the issue to be comprehended. With this data, the goal of the issue can be influenced precious stone to clear. Data: Another imperative worry of data mining is the characteristics of data, for example, size, distribution, and representation. Distinctive data typically should be processed in a different way. In spite of the fact that data originating from various issues might be alike, they may must be investigated distinctively if the implications of them are unique. Mining algorithm: Having above two parameters decided accurately, determining and selecting a data mining algorithm that suits to accomplish users task is very much easier task. In [13], author discussed three parameters which are very important to decide whether to develop new data mining algorithm or to use already designed algorithms. For an example, considering a scenario if we come to a derivation, that size and complexity of data that is needed to be processed is very high that are beyond available system capabilities to process and no other options or techniques are available to reduce size and complexity of data then it is supposed to be solved using novel mining algorithm.

III. VARIANTS OF DATA MINING MODEL FOR THE IoT

A. Multi-Layer Data Mining Model

As shown in Figure 2, model is partitioned into four layers namely information gathering layer, information administration layer, event processing layer and information mining service layer. Among them, information accumulation layer embraces devices , e.g. RFID sinks/readers and so forth., to gather intelligent information from different objects such as RFID stream information, GPS information, satellite information, positional information and sensor information and so on. Extraordinary sort of information requires distinctive information storage methodology. In the procedure of information collection, a progression of issues, e.g., vitality output, fault tolerance, data preprocessing, communications and so forth., ought to be very much explained [6]. Information i.e. data management layer focuses in centralized or disseminated database or information stockroom i.e. data warehouse to oversee gathered information.

An Event is a combination that joins information, time and different variables, so it gives an abnormal state component to information handling of IoT. An Event handling layer is utilized to investigate an events in IoT viably. Hence, it enables querying or investigation based on an event at this layer [10].

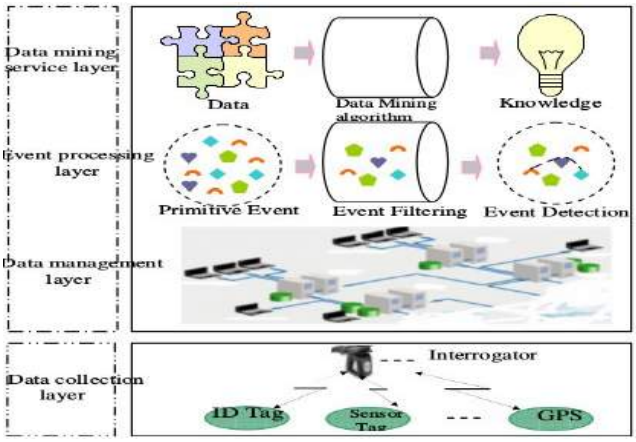


Figure 2. Multi Layer Model [1].

At that point, aggregation, sorting out and break down of information as per event can be done. Information mining administration layer is constructed in light of information administration and event handling. Different protest based or on the other hand management of event based information mining i.e. clustering, grouping, classification, forecasting, noise detection and mining of patterns, are provided for applications, e.g., SCM, inventory management and an optimization etc. The design of this layer is service-oriented.

B. Distributed data mining model

Comparing to traditional information which is raw in nature, Information in IoT has its own attributes. For instance, the information in IoT is dependably mass, appropriated, timerelated and position-related. At the same time, the information wellsprings of IoT are heterogeneous, and the assets of nodes are restricted. These attributes bring a few issues to unified information mining design. At initially, mass information of IoT is put away in distinctive locales. Consequently, it is troublesome for us to mine conveyed information by concentrated engineering. Furthermore, information in IoT is mass and needs preprocessing continuously.

For the thought of information security, information protection, adaptation to internal failure, business rivalry, legitimate requirements and different elements, the technique of assembling every pertinent datum is regularly not doable. In addition, the assets of nodes are restricted. The technique of transfering all information to central nodes does not enhance the utilization of vitality expensive transmissions. In most cases, the central node needn't bother with all information, yet a few evaluations of parameters. Hence, we can preprocess the raw information in the appropriated distributed node, and afterward send the fundamental information to the recipient.

Appropriated information digging model for IoT isn't just capable to take care of the issues brought by distributed capacity of nodes, but also decomposes problem complexity. In this way, the necessity of elite performance, high storage computing and processing power is decreased. In this model, the global control node is the center of the entire information mining framework. It picks the information mining calculation and the informational collections for mining, and afterward explores to the sub-nodes containing these informational collections. The sub-nodes get the raw information from different savvy objects. These raw information is supplied as an input to data filteration for preprocessing and then data abstraction and data compression, and finally, it get stored in the local data warehouse. Event separation, recognition and data mining at local nodes results into local models. Global models are the result of aggregation of local models are aggregated Subnodess trade protest information, process information and learning with each other. The entire procedure is restricted by the multi-agent based collaborative management module which is depicted in figure 3.

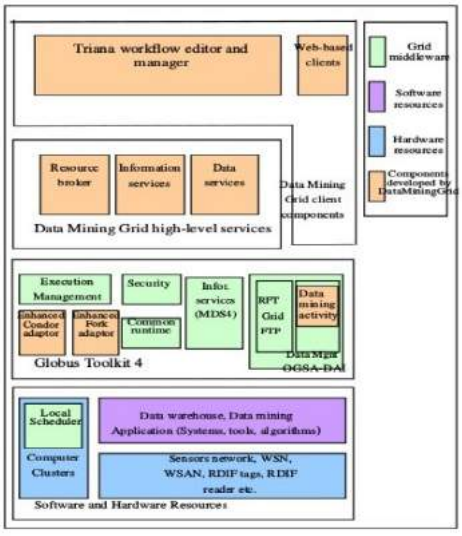


Figure 3. Distributed data mining model [1].

IV. APPLICATIONS

There are wide variety of application of data mining in Internet of Things. In [13] author suggested predicting user's preferences, nature and reaction to some situation, Object identification using different already available images of that object [13]. Video based classification where different scenarios and objects are identified, read facial expression of any person using already available devices like camera, microphone, etc, can also be done as suggested by author in [13]. Three-dimensional emotion model is used to identify human's emotions where a machine will have a very large data about a person different feeling and emotions that person posses while being in different situation and then deriving some patterns and conclude that persons emotions in particular given situation, Tracking movement of things sensing sound effects like human steps sound, door clapping, phone ringing, glass breaking [13].

Temperature, weather, wind speed, humidity prediction from previous data which may be very helpful to users like farmers or tourists before deciding their plans [2][16]. Also, Agriculture based on IoT, Cloud computing considered to be a great agricultural transformation [15]. Healthcare is booming domain for application of data mining using IoT devices and one can detect many deadly diseases in very early stage where getting rid of such disease is possible. Growth of disease in certain areas can be predicted using these techniques [13]. In academia domain, determining hot demanded areas according to student and market is easily achievable using data mining along with IoT devices [12]. Suggestions to deploy more public transport services in certain locations also can be evaluated using the data from IoT devices. Making prediction of usage of milk, grains, fruits, etc edible things in upcoming week or month or year can be derived from smart home IoT devices [3].

Utilization of IoT to gather information, which will be examined to get data valuable for basic leadership to enhance the web programs in Higher Education Institutions (HEIs) [5]. Many applications related to IoT essentially and implicitly consider occurrence of episodes (and events) with spatio-temporal constraints so as to initiate any further processing actions [9].

V. KEY ISSUES IN DATA MINING OF IOT

There are various issues involved in data mining in Internet of Things:

A. Efficiency in data gathering

Energy efficiency, scalability and fault tolerance should be taken into consideration when data is to be collected from distributed sensor networks [14]

B. Data abstraction and aggregation

Managing massive data generated from IoT is a challenging task. Efficient mechanism should be adopted for data deduplication.

C. Distributed data processing and mining

Due to nodes' constraints, paradigm shift is needed for prior level preprocessing of the data at each distributed nodes and an aggregated information is to be sent to sink node in order to optimize energy usage instead of sending all distributed data to server for processing.

D. Data mining towards the next age of Internet

In an upcoming generations of Internet, latest trends and technologies like ubiquitous computing, semantic web, IPv6 technologies are going to be integrated with IoT. This will give rise to challenges for Data Mining due to heterogeneous unstructure data [8].

VI. CONCLUSION

As a vital improvement of the next age of Internet, the Internet of Things pulls in numerous considerations by industry world and scholarly circles. IoT information has numerous qualities, for example, distributed storage, mass temporal and spatial related information, and constrained assets of nodes and so forth. These makes the issue of information mining in IoT turn into a test assignment.

VII. REFERENCES

[1] Shen Bin, Liu Yuan, and Wang Xiaoyi. Research on data mining models for the internet of things. In Image Analysis and Signal Processing (IASP), 2010 International Conference on, pages 127–132. IEEE, 2010.

[2] Olaiya Folorunsho and Adesesan Adeyemo. Application of data mining techniques in weather prediction and climate change studies. 4, 02 2012.

[3] Jeu Young Kim, Hark-Jin Lee, Ji-Yeon Son, and Jun-Hee Park. Smart home web of objects-based iot management model and methods for home data mining. In Network Operations and Management Symposium (APNOMS), 2015 17th Asia-Pacific, pages 327–331. IEEE, 2015.

[4] Saral Nigam, Shikha Asthana, and Punit Gupta. Iot based intelligent billboard using data mining. In Innovation and Challenges in Cyber Security (ICICCS-INBUSH), 2016 International Conference on, pages 107–110. IEEE, 2016.

[5] Alexander Muriuki Njeru, Mwana Said Omar, Sun Yi, Samiullah Paracha, and Muhammad Wannous. Using iot technology to improve online education through data mining. In Applied System Innovation (ICASI), 2017 International Conference on, pages 515–518. IEEE, 2017.

[6] Sebastian Scholze Claudio Cenedese Oliviu Matei, Carmen Anton. Multi-layered data mining architecture in the context of internet of things. In IEEE. IEEE, 2017.

[7] Rob Petersen. 20 companies do data mining and make their business better. 2016.

[8] Brijesh Iyer Prachi Deshpande. Research directions in the internet of every things(ioet). In International Conference on Computing, Communication and Automation (ICCCA2017), pages 1353–1357. IEEE, 2017.

[9] Vangipuram Radhakrishna, Puligadda Veereswara Kumar, Vinjamuri Janaki, and Shadi Aljawarneh. A computationally efficient approach for temporal pattern mining in iot. In Engineering & MIS (ICEMIS), International Conference on, pages 1–4. IEEE, 2016.

[10] Aashi Singh and Shilpi Sharma. Analysis on data mining models for internet of things. In I-SMAC (IoT in Social, Mobile, Analytics and Cloud)(I-SMAC), 2017 International Conference on, pages 94–100. IEEE, 2017.

[11] Xin Song, Cuirong Wang, and Jing Gao. An integrated framework for analysis and mining of the massive sensor data using feature preserving strategy on cloud computing. In Computational Intelligence and Design (ISCID), 2014 Seventh International Symposium on, volume 2, pages 337–340. IEEE, 2014.

[12] Fatima Merchant Suhem Parack, Zain Zahid. Application of data mining in educational databases for predicting academic trends and patterns. Technology Enhanced Education (ICTEE), 2012 IEEE International Conference, 2012.

[13] Chun-Wei Tsai, Chin-Feng Lai, Ming-Chao Chiang, Laurence T Yang, et al. Data mining for internet of things: A survey. IEEE Communications Surveys and Tutorials, 16(1):77–97, 2014.

[14] Dr. Shalini Bhaskar Bajaj Tushar Taneja, Aman Jatain. Predictive analytics on iot. In IEEE, pages 1312–1317. IEEE, 2009.

[15] Qiulan Wu, Yong Liang, Ying Li, and Yusheng Liang. Research on intelligent acquisition of smart agricultural big data. In Geoinformatics, 2017 25th International Conference on, pages 1–7. IEEE, 2017.

[16] Hai xiang Zhao and Fred´ eric´ Magoules.` A review on the prediction of building energy consumption. Renewable and Sustainable Energy Reviews, 16(6):3586 – 3592, 2012.

© 2018 IJSRSET | Volume 4 | Issue 2 | Print ISSN: 2395-1990 | Online ISSN : 2394-4099
National Conference on Advanced Research Trends in Information and Computing Technologies (NCARTICT-2018), Department of IT, L. D. College of Engineering, Ahmedabad, Gujarat, India In association with International Journal of Scientific Research in Science, Engineering and Technology

Proposed Cryptographic Approach for Securing IOT Device

D. M. Trivedi[1], Prof. T. J. Raval[2]

[1]Computer Department, L.D.College of Engineering, Gujarat Technological University, Gujarat, India
[2]Computer Department, L.D.College of Engineering, Gujarat Technological University, Gujarat, India

ABSTRACT

Internet of Things (IoT) connects sensing devices to the Internet for the purpose of exchanging information.The Internet of Things (IoT) shall be able to incorpo-rate transparently and seamlessly a large number of different and heterogeneous end systems,hence the security is most important factor of this system as the system will be useful in healthcare, industry etc. in this paper we are going to provide information about security and cryptographic algorithm that are best suited for IOT.

Keywords:

I. INTRODUCTION

The Internet of Things(IOT) is a interconnected physical devices through internet. physical devices include car,machine,person smartphone and other items embedded with electronics , sensor, actuators,software and network connectivity on it. signals are collected by the sensor and transmitted to the network.

1.1 Architecture of IOT

From the perspective of architecture, the Internet of things can be generally divided into three layers, namely perception layer, network layer, and application layer. The perception layer transforms the information of things to the readable digital signals via RFID, sensors, etc.

On the other hand, the network layer transmits these digital signals to corresponding platforms via a connected network. In the end, the application layer unscrambles and applies digital signals through corresponding software.

Application layer provide end- to end communication between node.as shown in the figure 1.

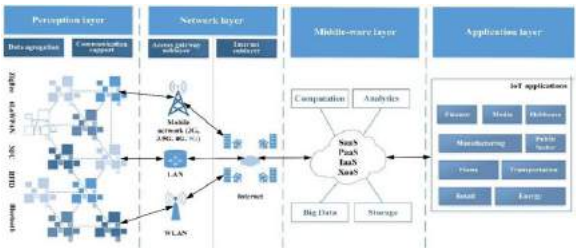


Figure 1. Architecture of IOT [16,17]

1.2 Need Of Security in IOT

Among these three layers, security is a problem to ensure signal is corrected collected, transmitted, and interpreted by the applications. Both the ordinary nodes and sink nodes are vulnerable to a variety of security attacks, such as denial of service attacks, or illegal control and failure. These attacks could compromise the sensitive information and result in malfunctions.

1.3 Types Of Attack on IOT

- ✓ Physical Attacks
- ✓ Side Channel attacks
- ✓ Cryptanalysis attacks
- ✓ Software Attacks
- ✓ Network Attacks

II. CRYPTOGRAPHIC ALGORITHM

There are various cryptographic algorithm available based on key distribution conventional cryptography is referred to as Conventional cryptography is referred to as symmetric encryption or single key encryption. Same key is used for encryption and decryption. This means that the encryption key is equal to the decryption key. Figure represents the simplified model for conversional encryption technique. In general, there are two types of the symmetric ciphers, namely, stream ciphers and block ciphers.

Figure 2. Symmetric Encryption Technique [1]

III. PROPOSED METHODS

AES-GCM

In the AES-GCM, only the AES encryption is utilized with the input and the output blocks of 128 bits. However, based on the security requirements, the key size could be determined as AES-128 (with 10 rounds), AES-192 (with 12 rounds), or AES-256 (with 14 rounds) [1]. In the AES encryption, all the rounds except for the last round have four transformations of SubBytes, ShiftRows, MixColumns, and AddRoundKey. For the last round, MixColumns is eliminated and only three transformations of SubBytes, ShiftRows, and AddRoundKey are used.

The transformation SubBytes (S-boxes) is implemented by 16 S-boxes. In the S-box, each byte of the input state is substituted by a new byte. In ShiftRows, the first row of the state remains intact and the four bytes of the last three rows of the input state are cyclically shifted. In the MixColumns transformation, each column is modified individually and in the final transformation, AddRoundKey, modulo-2 addition of the input state and the key of the corresponding round is performed [1].

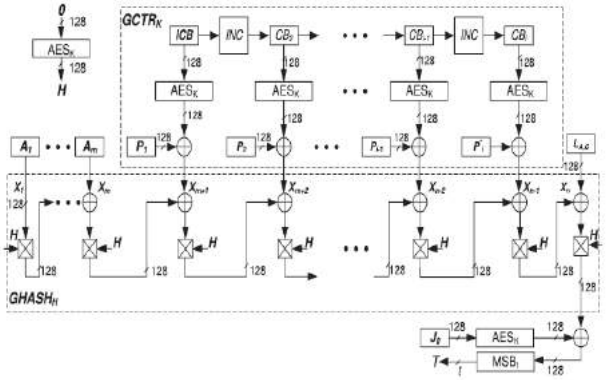


Figure 3. The GCM Authenticated Encryption Data Flow

NTRU :

NTRU (N-th degree Truncated polynomial Ring Unit) is an open source and patented public-key cryptosystem which uses lattice-based cryptography for encryption and decryption of files. The two keys used in this algorithm are: public key and private key. The key is used for the encryption is Public Key or to verify the digital signature but private key is used for decryption or to create digital signature

The Combination of AES-GCM, and NTRU will be used. it provides nonce misuse resistance, security over 256 bits and parallelizability. Stream ciphers take the plaintext as streams of characters with size of 1 bit or n-bit word. In cipher, the plaintext is encrypted (and decrypted) one character at a time. According to Alfred et al. [18], stream ciphers are used in real-time applications such as pay TV and communications. This is because they are able to run in high speed.

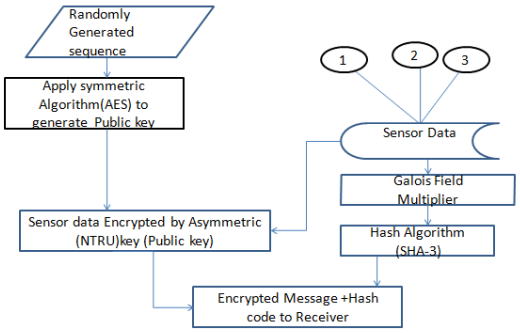


Figure 4. Proposed System Encryption Process.

Randomly generated Sequence has been applied to Symmetric algorithm AES and it will generate the public key, so the benefit of symmetric encryption has been achieved,the limitations of symmetric encryption has been overcome by encrypting the message with public key with NTRU algorithm.

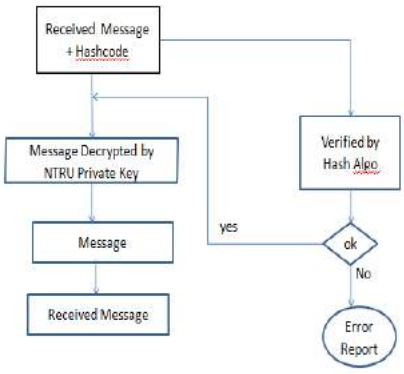


Figure 5. Proposed System for Decryption Process ecryption Process.

Encrypted message has been sent along with hash code to the receiver, the receiver will receive the message and first verify it by hash function output of hash function is the hash code in my proposed system hash code will be generated using sha-256, which is a non reversible function,if the verification function ouput is matched with the received code then decryption will be performed ,but if the verification output does not match withh received code then notification has been sent to the receiver.Decryption process is done with NTRU public key cryptographic algorithm and with private key of the user.

Table 1. Speed Time Of Han In Comparision With Two Other Encryption Algorithms

The Total Speed time of HAN in comparison with two other Encryption Algorithms			
Algorithm	HAN	AES	RSA
Period of time for implementing Whole algorithm per second	0,321081	2.718182	2.35072

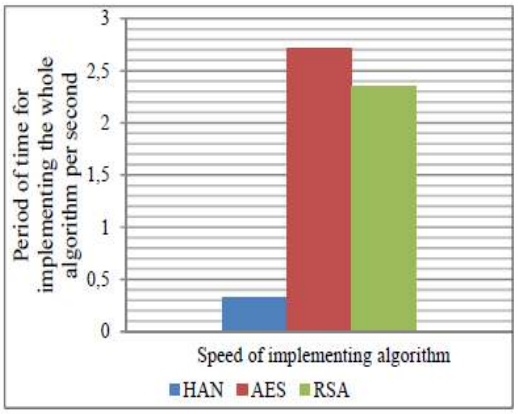
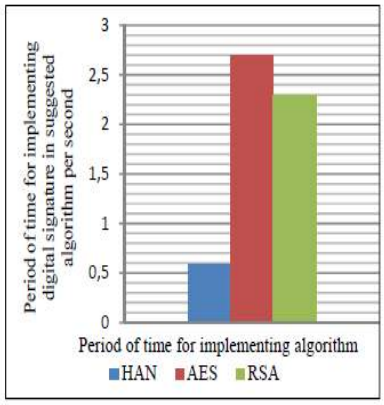


TABLE II
IMPLEMENTATION TIME OF DIGITAL SIGNATURE IN SUGGESTED ALGORITHM

ALGORITHM	Total time algorithm implementation (sec)
HAN by digital sign	0.58
AES without digital sign	2.718182
RSA without digital sign	2.350752



IV. CONCLUSION

The Table 1 represents the speed of various algorithm according to that hybrid encryption algorithm will provide better and faster result. The proposed system will be capable enough to provide parallelizability, efficiency in security with authentication and confidentiality.

Table 2 represents the implementation time of various algorithm including hybrid encryption algorithm.which shows with digital signature also the proposed system generates good result.

Advantages of Proposed System:

- ✓ More efficient encryption and decryption in both hardware and software implementations;-
- ✓ much faster key generation allowing the use of “disposable" keys (because keys are computationally “cheap" to create).low memory use allows it to use in applications such as IOT Device.
- ✓ Here the combination of symmetric AES GCM and NTRU asymmetric algorithm is used, so the benefit of security and faster performance id achieved.
- ✓ AES-GCM is Authenticated Encryption algorithm,by using this we can reduce the time to create digital signature separately.

V. REFERENCES

[1]. markups.kdanmobile.com/sharings/

[2]. Joe Ruether, Cryptography Primer, http://jrruethe.github.io/blog/2014/10/25/cryptography-primer/

[3]. William Stallings, "Cryptography and Network Security: Principles & Practices", 4th edition, Prentice Hall.
[4]. Atul Kahate ,"Cryptography and Network Security:
[5]. Mingyuan Xin,2016 , "A Mixed Encryption Algorithm Used in Internet of Things Security Transmission" International Conference on Cyber-Enabled Distributed Computing and Knowledge Discovery
[6]. T K Goyal," Lightweight Security Algorithm for Low Power IoT Devices, (ICACCI), Sept. 21-24, 2016
[7]. S. Singh, 2017,"Advanced lightweight encryption algorithms for IoT devices:survey, challenges and solutions",Springer-Verlag Berlin Heidelberg
[8]. A. Safi, 2017," Improving the Security of Internet of Things Using Encryption Algorithms", International Scholarly and Scientific Research & Innovation [5]2017 S Koteshwar,2016 "Comparative study of Authenticated Encryption" targeting lightweight IoT applications, 2168-2356 (c) 2016 IEEE,
[9]. Dr. S. S. Manikandasaran, 2016,"Security Attacks and Cryptography Solutions for Data Stored in Public Cloud Storage" (IJCSITS),
[10]. M. Katagi," Lightweight Cryptography for the Internet of Things",
[11]. J Choi "An improved LEA block encryption algorithm to prevent side-channel attack in the IoT system"
[12]. A. SAJID, 2016,"Cloud-Assisted IoT-Based SCADA Systems", A Review of the State of the Art and Future Challenges, Digital Object Identifier
[13]. MIKAEL ASPLUND, 2016,"Attitudes and Perceptions of IoT Security in Critical Societal Services"Digital (IEEE) [11] Dieter Uckelmann,2011,"An Architectural Approach Towards the Future Internet of Things", DOI 10.1007/978-3-642-19157-2_1,Springer-Verlag Berlin Heidelberg 2011
[14]. A Tiwari, 2017,Challenges and Ongoing Researches for IOT (Internet of Things):
[15]. R Khan,2012,Future Internet: The Internet of Things Architecture,Possible Applications and Key Challenges,
[16]. Dr. S. Baskaran,Implementation of Enhanced Honey Encryption for IoT Security, International Journal of New Technology and Research (IJNTR)
[17]. I.Cvitic,2016, CLASSIFICATION OF SECURITY RISKS IN THE IOT ENVIRONMENT DAAAM International,
[18]. M. Farooq, M. Waseem, A. Khairi, and S. Mazhar, A Critical Analysis on the Security Concerns of Internet of Things (IoT), International Jurnal of Computer Applications,
[19]. L. Zheng, H. Zhang, W. Han, and X. Zhou, Technologies, Applications, and Governance in the Internet of Things, in: O. Vermesan and P. Friess (Eds.), Internet of Things - Global Technological and Societal Trends From Smart Environments and Spaces to Green ICT, River Publishers, 2011, pp. 141–175.
[20]. S. Babar,2011,Proposed Embedded Security Framework for Internet of Things (IoT), 978-1-4577-0787-2/11/$26.00 2011 IEEE

© 2018 IJSRSET | Volume 4 | Issue 2 | Print ISSN: 2395-1990 | Online ISSN : 2394-4099
National Conference on Advanced Research Trends in Information and Technologies (NCARTICT-2018), Department of IT, L. D. College of Engineering, Ahmedabad, Gujarat, India. In association with
International Journal of Scientific Research in Science, Engineering and Technology

Ranking Model : Study & Algorithmic Approach towards Common Ranking for Higher Education Institutes

Abhishek Kumar*[1], Beena K. S.[2], Kalgi Shah[3]
*[1]Scientist (Computer Science), INFLIBNET Centre, Gandhinagar, Gujarat, India
[2]Department of Computer Science, Gujarat University, Ahmedabad, Gujarat, India
[3]Department of Computer Science, Gujarat University, Ahmedabad, Gujarat, India

ABSTRACT

Indian education system is one of the largest education systems in the world which comprise of Universities, Colleges, Technical Institutions, Institutions of National importance etc. Choosing the right Institution for the academic activity is always a challenging job. Competition between Universities, emerging of new Institutions with adequate infrastructure etc. makes difficulty to select the desired Institution in county. As many ranking agencies are available in other countries to rank various Institutions, India Rankings is an initiative of MHRD (Government of India) to rank the educational institutions in the country. This article is exploring the different agencies of the ranking model and their parameters. It also covers that assessment / evaluation of decisive parameters using algorithm (decision theory approach) because weightage of same parameters in different ranking model are different. It tends towards global ranking using common weightage across all countries and their ranking agencies.

Keywords: Ranking System, University Ranking, Ranking Parameters, Higher Education Ranking, India Ranking, National Institutional Ranking Framework (NIRF), Times Higher Education (THE), Quacquarelli Symonds (QS), Academic Ranking of World Universities (ARWU)

I. INTRODUCTION

Albert Einstein told "Not Everything that counts can be measured. Not everything that can be measured counts". Why Rank Institutes? To serve as a guide to students for the selection of universities based on set of criteria. To help universities to improve their research performance and its quality (major criteria is used in all rankings).To help the universities to identify areas of improvement and to become better at education students and conducting research. To help industry and employers to target specific students programmes and projects for hiring and research. To provide a stable formula or frame work for ranking Indian Universities. To achieve the higher ranks in global ranking.

A. World Ranking of Universities

The three longest established World Rankings are those produced by Shanghai ranking consultancy (the Academic Ranking of World Universities- ARWU), Times Higher Education (THE) and Quacquarelli Symonds (QS). All of these, along with other world rankings, primarily measure the Research performance of universities rather than their teaching.

Figure 1: World Ranking Systems of Universities

B. Parameters and Weightage of Global Ranking Systems

1) *Academic Ranking of World Universities (ARWU): ARWU was first published in June 2003 by the Centre for World Class Universities (CWCU), Institute of Higher Education of Shanghai Jiao Tong University, China, and updated on an annual basis. Since 2009 ARWU has been published and copyrighted by Shanghai Ranking Consultancy. ARWU also known as Shanghai Ranking. The ranking considers all the institution with the fields' medallist, highly citied researches, and papers published and paper indexed.*

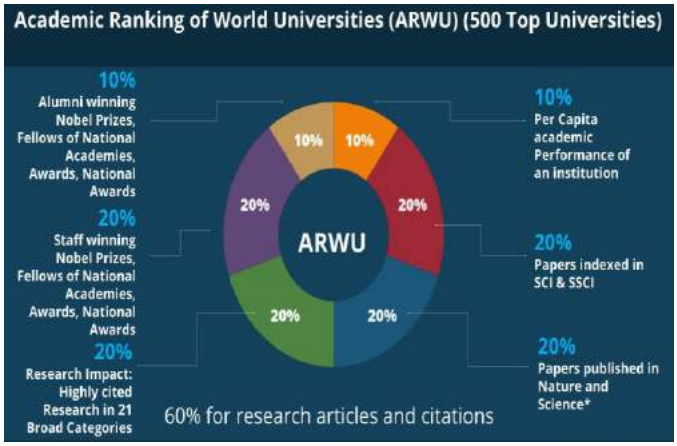


Figure 2 : ARWU Parameters with their Weightage

2) *Times Higher Education (THE): THE world university ranking lists the top universities in the world. Their core missions are teaching, knowledge transfer and international outlook. The Times Higher Education world university rankings, founded in 2004, provide the definitive list of the world's best universities. THE splits from its original partner Quacquarelli Symonds (QS) in 2009.*

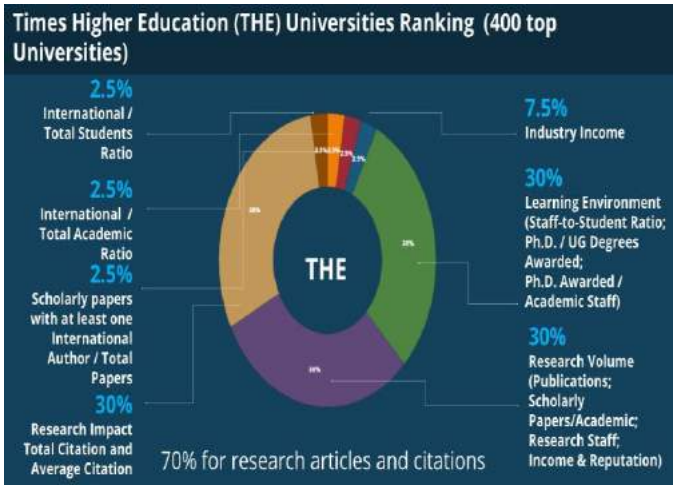


Figure 3: THE Parameters with their Weightage

3) *Quacquarelli Symonds (QS): QS world university ranking is a ranking of the world's top 500 universities by QS since 2004. QS Rankings were originally published in collaboration with Times Higher Education, and were known as the THE-QS World University Rankings. QS assumed sole publication of existing methodology and Times Higher Education split in order to create a new ranking methodology in 2010, which became the THE World University Rankings.*

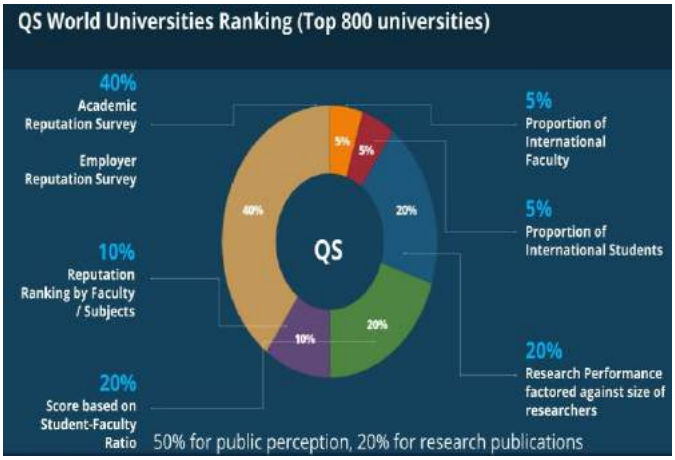


Figure 4: QS Parameters with their Weightage

C. Parameters and Weightage of India Rankings

The NIRF provides for ranking of institutions in five broad generic parameters, namely: a) Teaching, Learning and Resources; b) Research, Consulting and Collaborative Performance; c) Graduation Outcomes; d) Outreach and Inclusivity; and e) Perception.

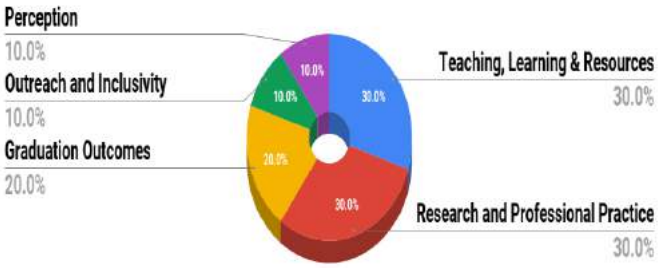
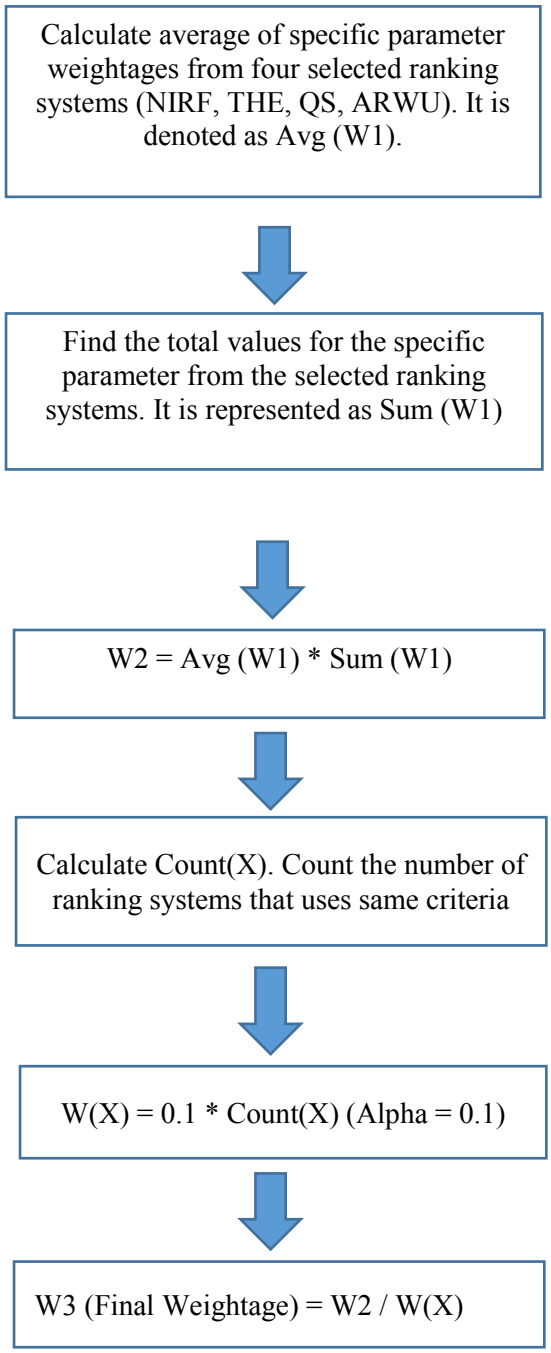


Figure 5: NIRF Parameters with their Weightage

II. METHODOLOGY

The global and national ranking systems follow different parameters for ranking universities. These different parameters are assigned with some weightages. The

differences in weightage for each and every parameters of different ranking systems produce different outputs from each other. To find out balance between distinct weightages we have applied the average approach (without using probabilities) of decision making theory. Here, we have considering three parameters a.) Teaching-Learning b.) Research Practice and c.) International Outlook and use the existing global and national ranking systems weightages. To find out balance weightage for parameters, we have applied following algorithm by using average approach.



Here, W3 is the Parameters Weightages.

III. RESULTS

Following Figure shows the weightage comparisons of selected parameters of various ranking systems with our calculated weightage outputs. It shows the balanced weightage among all the ranking model.

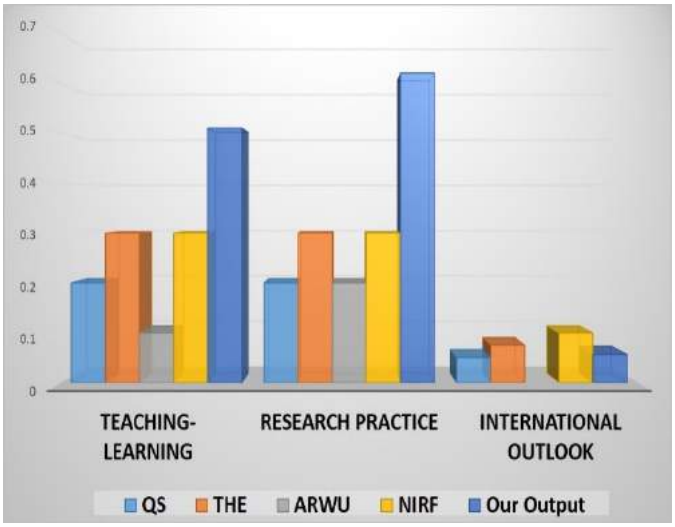


Figure 6: Correlation Chart for weightage of parameters from various ranking systems and our calculated outcomes

IV. CONCLUSION

The higher education in India is need of infusion of quality and clarity in its approach towards building world class universities in the Indian context and the environment. New benchmarks of quality need to be defined and put in place to help overall system to move up on the quality spectrum. Research assessment and national ranking of Indian universities can play an important role in improving performance and quality of academic institutions. One more aspect which is very much important in global scenario i.e. common ranking across globe need common weightage. Moreover, web tool need to be evolved to display ranking on the basis of different ranking agencies and their parameters.

V. REFERENCES

[1] "MHRD, National Institute Ranking Framework (NIRF)", *Nirfindia.org*, 2017. [Online]. Available: https://www.nirfindia.org/Home. [Accessed: 01-Sep- 2017].

[2] "Rankings", *Times Higher Education (THE)*, 2017. [Online]. Available:

https://www.timeshighereducation.com/world-university-rankings. [Accessed: 08- Sep- 2017].
[3] "ARWU World University Rankings 2017 | Academic Ranking of World Universities", *Shanghairanking.com*, 2017. [Online]. Available: http://www.shanghairanking.com/. [Accessed: 15- Sep- 2017].
[4] "University Rankings", *Top Universities*, 2017. [Online]. Available: https://www.topuniversities.com/university-rankings. [Accessed: 25- Sep- 2017].
[5] P. G. Altbach, "The State of the Rankings | Inside Higher Ed", *Insidehighered.com*, 2017. [Online]. Available: https://www.insidehighered.com/views/2010/11/11/state-rankings. [Accessed: 29- Sep- 2017].
[6] A. Marszal, "University rankings: which world university rankings should we trust?" *The Daily Telegraph*, 2012. [Online]. Available: http://www.telegraph.co.uk/education/universityeducation/9584155/University-rankings-which-world-university-rankings-should-we-trust.html. [Accessed: 06- Oct- 2017].
[7] A. Zirulnick, "New world university ranking puts Harvard back on top", *The Christian Science Monitor*, 2010. [Online]. Available: https://www.csmonitor.com/World/2010/0916/New-world-university-ranking-puts-Harvard-back-on-top. [Accessed: 10- Nov- 2017].
[8] Q. Zha, "University Rankings in Perspective | Inside Higher Ed", *Insidehighered.com*, 2016. [Online]. Available: https://www.insidehighered.com/blogs/world-view/university-rankings-perspective. [Accessed: 17- Nov- 2017].
[9] B. Bekhradnia, "International university rankings: For good or ill? - HEPI", *HEPI*, 2016. [Online]. Available: http://www.hepi.ac.uk/2016/12/15/3734/. [Accessed: 24- Nov- 2017].
[10] "Decision Making Without Probabilities: Optimistic, Conservative & Minimax Approaches - Video & Lesson Transcript | Study.com", Study.com. [Online]. Available: https://study.com/academy/lesson/decision-making-without-probabilities-optimistic-conservative-minimax-approaches.html. [Accessed: 10- Jan- 2018].

© 2018 IJSRSET | Volume 4 | Issue 1 | Print ISSN: 2395-1990 | Online ISSN : 2394-4099
National Conference on Advanced Research Trends in Information and echnologies(NCARTICT-2018), Department of CE&IT,L. D. College of Engineering, Ahmedabad, Gujarat, IndiaIn association with
International Journal of Scientific Research in Science, Engineering and Technology

Impact of Demographic Factors on Payment Method Chosen by Indian Consumer using ANOVA

Thomson Christain*[1], Lalit Sengunthar[2], Oneel Christain[3], Shanti Verma[4]
[1,2,3]PG Scholar , L.J. Institute of Computer Applications, Ahmedabad, Gujarat, India
[4]Assistant Professor, L.J. Institute of Computer Applications, Ahmedabad, Gujarat, India

ABSTRACT

Online shopping is a current wonder in the field of E-Business and is unquestionably going to be the eventual fate of shopping in the world. The vast majority of the organizations are running their online portals to offer their product/services online. In spite of the fact that internet shopping is exceptionally normal outside India, its development in Indian Market, which is a vast and vital buyer showcase, is still not in accordance with the worldwide market. The potential development of online shopping has set off directing a study on online shopping in India. The primary objective of the current study is to understand the behavior of Indian online consumer and find out opinion about online shopping to identifying what many factors Indian online consumer takes into consideration when purchasing products online. The data was collected through Questionnaires. Sample sizes of 1014 respondents are taken for the collection of the data. In this paper authors try to find the impact of demographic factor Gender and occupation on various payment method pick by Indian consumer. Authors apply Analysis of Variance (ANOVA) method to test hypothesis. Tool used to do analysis was SPSS. From the results of analysis authors conclude that the demographic factor does impact on payment method pick by consumer in India at 99% significance.

Keywords: Online Shopping, Survey, Analysis, Customer Behavior

I. INTRODUCTION

Online shopping is the process whereby consumers directly buy goods or services from a seller in real-time, without an intermediary service, over the Internet. It is a form of electronic commerce. Nothing is predictable in India except change. Currently the retail industry in India is accelerating. India is excited to grow to be a most important player in the retail market. Also India with a high on cross culture factor, it allows different companies bringing in variety of products Targeting different consumer segments [4]. According to the Global Retail Development Index 2016, India ranks second among the top 30 emerging markets for retail [2]. The implementation and introduction of internet technologies has created new market for manufacturers and service providers it also has provided new arena for innovative marketing strategies by the professionals. There are different reasons of moving the customers purchasing pattern towards online retail shops. The other prominent names for web based shopping are e-shop, web shop, and web-store and so on. These days' Mobile commerce is also one of the mainstream methods for shopping. The facilities of various discount scheme and coupon are also fascinating the customers in online shopping. From the past few years, online shopping is the prevalent way of doing dealings in the field of E-Business and is unquestionably going to be the future of shopping in the human race [5]. Because of wide communication network e-commerce has become the new mediator between the companies/manufacturers and their customers. The boost in use of internet by the customers in younger age bracket in India has given a promising viewpoint to online retailers. Earlier the

Indian customers were more attracted towards electronic devices and cell phones for online shopping but now the products like lifestyle, watches, apparels, perfumes, and beauty products are also in high demand through the medium of online shopping. The market is also developing for books, jewellery, home appliances and kitchen appliances etc.

II. LITERATURE REVIEW

Sultan and Henrichs (2012) [6] in his study concluded that the consumer's willingness to and preference for adopting the Internet as his or her shopping medium was also positively related to income, household size, and innovativeness. Vijay, Sai. T. &Balaji, M. S. (May 2012) [7], revealed that Consumers, all over the world, are increasingly shifting from the crowded stores to the one-click online shopping format. However, in spite of the convenience offered, online shopping is far from being the most preferred form of shopping in India. A survey among 150 internet users, including both users and non-users of online shopping, was carried out to understand why some purchase online while others do not. The results suggested that convenience and saving of time drive Indian consumers to shop online; while security and privacy concerns dissuade them from doing so.

The work of Kim and Park (2005) [8] using U.S. samples suggests that their positive attitudes as well as willingness to search for pre-purchase information leads to a strong likelihood that they will buy online. Online shoppers are required to have computer skills in order to use the Internet for shopping. Hence, those who are not comfortable with using the computer, will likely do their shopping at the traditional store, modern shop, or discount store because it will be faster shopping there than in the Internet shop[9][10].

Maignan and Lukas's research (1997) and Rowley (2000) studied that the financial risks had been cited as a main reason to stop internet shopping and security had become a major concern both in online transaction relationships.

Researcher examines the relationship among demographics, personal characteristics, and attitudes towards online shopping. These authors find that people who have a more wired lifestyle. And who are more time constrained tends to buy online more frequently, i.e., those who use the Internet as a routine tool and/or those who are more time starved prefer shopping on the Internet [1][3]. There are various statistical methods like Chi-square test, ANOVA (one way and two way) and correlation, used to find the association or impact of one parameter on other [11][12].

III. RESEARCH HYPTHESIS

H1: Demographic Factors of Indian consumer Significantly impact the payment method picked by consumer in online shopping

H1.1 Gender of Indian consumer significantly impact the * payment method.

H1.2 Occupation of Indian consumer significantly impact * payment method

*- credit card, debit card, net banking, cash on delivery, E-wallets.

IV. DISCREPTION OF SURVEY

A. Method of Data collection

For primary data collection author used online data collection method; Google forms.

A. Sample size

Sample size used for this study is 1014. As the scope of study is limited to only one city of Gujarat, India, This sample size is appropriate for study.

B. Nature of questionnaire

To fulfil the objectives of study various multiple choice questions was used in the questionnaire. The complexity of questions is defined as per targeted users education and willingness to submit response properly.

C. Tools & technique used for study

In data analysis and interpretation method of "Descriptive Statistics" is used to analyse the data with the help SPSS (Statistical Package for the Social Science) Tool. The technique used to analyse data is one way analysis of variance (ANOVA).

V. FINDINGS OF SURVEY

The following tables has broadly analysed for the survey questions. The data collected has given some interesting findings. It has provided the information about the buying behaviour of the consumers through online mode. Questionnaire designed in manner to collect data of customer perspective for online shopping. The findings

can be useful for taking strategic marketing decisions to capture the huge Indian retail market.

TABLE 1. CROSS TABULATION OF GENDER AND AGE

Gender * Age Cross tabulation							
Count							
		Age					Total
		Below 18	18-25	26-30	31-50	Above 50	
Gender	Male	16	85	227	174	5	507
	Female	4	71	197	228	7	507
Total		20	156	424	402	12	1014

1) Cross Tabulation is representing that in male there are 5 age groups in which below 18 there are 16 males, 18-25 age group there are 85 males, 26-30 age group there are 227 males, 31-50 age group there are 174 males and above 50 there are 5 males so total numbers of male is 507.

2) Cross Tabulation is representing that in female there are 5 age groups in which below 18 there are 4 females, 18-25 age group there are 71 Cross Tabulation is representing that in female there are 5 age groups in which below 18 there are 4 females, 18-25 age group there are 71.

TABLE 3. CROSS TABULATION OF OCCUPATION AND FREQUENCY OF PURCHASE

Occupation * How frequently do you purchase products online? Cross tabulation					
Count					
		How frequently do you purchase products online?			Total
		Weekly	Monthly	Yearly	
Occupation	Unemployed	31	39	49	119
	Employed	46	673	32	751
	Self-Employed	24	96	24	144
Total		101	808	105	1014

1) Cross Tabulation of occupation and frequency of purchase indicates that the 31 weekly purchases, 39 monthly purchases and 49 yearly purchases were totally 119 purchase frequencies by unemployed people.

2) Cross Tabulation of occupation and frequency of purchase is indicating that the 46 weekly purchases, 673 monthly purchases and 32 yearly purchases were totally 751 purchase frequencies by employed people.

3) Cross Tabulation of occupation and frequency of purchase is indicating that the 24 weekly purchases, 96 monthly purchases and 24 yearly purchases where totally 114 purchase frequency by self-employed people.

TABLE 3. SHOPPING ON INTERNET SAVES TIME

I think shopping on internet saves time				
		Frequency	Percent	Cumulative Percent
Valid	Strongly Disagree	13	1.3	1.3
	Disagree	18	1.8	3.1
	Neither Agree Nor Disagree	38	3.7	6.8
	Agree	114	11.2	18.0
	Strongly Agree	831	82.0	100.0
	Total	1014	100.0	

Shopping on internet save time was strongly agreed by 831 members out of 1014, it is equivalent to 82.0% of sample size.

TABLE 4. MORE DIFFICULT TO SHOP ON THE INTERNET

It is more difficult to shop on the internet				
		Frequency	Percent	Cumulative Percent
Valid	Strongly Disagree	48	4.7	4.7
	Disagree	752	74.2	78.9
	Neither Agree Nor Disagree	102	10.1	89.0
	Agree	52	5.1	94.1
	Strongly Agree	60	5.9	100.0
	Total	1014	100.0	

Shopping is more difficult on internet were disagree by 752 members out of 1014; it is equivalent to 74.2% of sample size.

TABLE 5. PRICES ARE LOWER THAN ACTUAL PRICE

I will prefer online shopping only if online prices are lower than actual price				
		Frequency	Percent	Cumulative Percent
Valid	Strongly Disagree	28	2.8	2.8
	Disagree	353	34.8	37.6
	Neither Agree Nor Disagree	52	5.1	42.7
	Agree	422	41.6	84.3
	Strongly Agree	159	15.7	100.0
	Total	1014	100.0	

Online prices are lower than actual price were agreed by 422 members out of 1014, it is equivalent to 41.6% of sample size.

TABLE 6. PAYING IN ADVANCE

I do not mind paying in advance for the products on the internet				
		Frequency	Percent	Cumulative Percent
Valid	Strongly Disagree	37	3.6	3.6
	Disagree	79	7.8	11.4
	Neither Agree Nor Disagree	67	6.6	18.0
	Agree	72	7.1	25.1
	Strongly Agree	759	74.9	100.0
	Total	1014	100.0	

I do not mind paying in advance for the products were strongly agreed by 759 members out of 1014; it is equivalent to 74.9% of sample size.

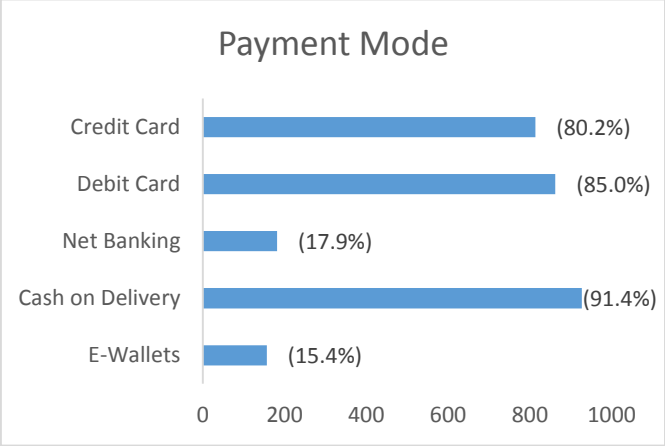


Figure 1. Payment Mode

From Figure 2 we highlight the following points:

1) 80.2% of sample size on online shopping out of 1014 members, 813 members were using credit card for payment.
2) 85.0% of sample size on online shopping out of 1014 members, 862 members were using debit Card for payment.
3) 17.9% of sample size on online shopping out of 1014 members, 181 members were using net banking for payment.
4) 91.4% of sample size on online shopping out of 1014 members, 927 members were using cash on delivery for payment.
5) 15.4% of sample size on online shopping out of 1014 members, 156 members were using e-wallets for payment.

VI. HYPOTHESIS TESTING

H1.1 Gender of Indian consumer significantly impact the * payment method.

TABLE 7. CONSUMER RESPONSE VARIATIONS FOR PAYMENT MODE*ACROSS GENDER

		Sum of Squares	df	Mean Square	F	Sig.
Credit Card	Between Groups	.009	1	.009	.056	.813
	Within Groups	161.148	1012	.159		
	Total	161.157	1013			
Debit Card	Between Groups	.773	1	.773	6.092	.014
	Within Groups	128.442	1012	.127		
	Total	129.215	1013			
Net Banking	Between Groups	1.823	1	1.823	12.565	.000
	Within Groups	146.868	1012	.145		
	Total	148.691	1013			
Cash on delivery	Between Groups	.829	1	.829	10.664	.001
	Within Groups	78.706	1012	.078		
	Total	79.536	1013			
E-Wallets	Between Groups	0.000	1	0.000	0.000	1.000
	Within Groups	132.000	1012	.130		
	Total	132.000	1013			

We here summarize the results of Table 7 in Table 8 to conclude the results.

TABLE 8. SUMMARIZE RESULTS OF ANOVA TABLE FOR PAYMENT MODE*ACROSS GENDER

ANOVA GENDER		
Dependent Variable	F	Sig.
Credit Card	0.056	0.813
Debit Card	6.092	0.014
Net Banking	12.565	0
Cash on delivery	10.664	0.001
E-Wallets	0	1

The results of table 8 shows that the Gender as one of the demographic factor of consumer does impact on payment method Credit card, Debit card, Net Banking, Cash on delivery. But Gender does not impact payment method E-wallets.

H1.2 Occupation of Indian consumer significantly impact * payment method

This hypothesis is used to show impact of consumer occupation on various payment method used by Indian consumer.

TABLE 9. CONSUMER RESPONSE VARIATIONS FOR PAYMENT MODE*ACROSS OCCUPATION

		Sum of Squares	df	Mean Square	F	Sig.
Credit Card	Between Groups	18.287	2	9.144	64.705	.000
	Within Groups	142.869	1011	.141		
	Total	161.157	1013			
Debit Card	Between Groups	21.652	2	10.826	101.757	.000
	Within Groups	107.563	1011	.106		
	Total	129.215	1013			
Net Banking	Between Groups	8.243	2	4.121	29.667	.000
	Within Groups	140.449	1011	.139		
	Total	148.691	1013			
Cash on delivery	Between Groups	.425	2	.213	2.716	.067
	Within Groups	79.110	1011	.078		
	Total	79.536	1013			
E-Wallets	Between Groups	.825	2	.413	3.180	.042
	Within Groups	131.175	1011	.130		
	Total	132.000	1013			

Here we summarize the results of Table 9 in Table 10 to conclude the ANOVA results.

TABLE 10. SUMMARIZED RESULTS OF ANOVA TABLE FOR PAYMENT MODE*ACROSS OCCUPATION

ANOVA OCCUPATION		
Dependent Variable	F	Sig.
Credit Card	64.705	0
Debit Card	101.757	0
Net Banking	29.667	0
Cash on delivery	2.716	0.067
E-Wallets	3.18	0.042

The results of table 10 shows that the occupation as one of the demographic factor of consumer does impact on payment method Credit card, Debit card, Net Banking, Cash on delivery and E-wallets.

VII. CONCLUSION

The overall results prove that the respondents have perceived online shopping in a positive manner. This clearly justifies the project growth of online shopping. The following
major outcomes are dervived from the survey based on questionnaire.

1) Males and females are equally participated in the survey. Where age group of major respondents is between 31-50.

2) Employed respondents shop very frequently and the frequency of doing shopping is on monthly basis.
3) Cloth and electronics are the highest demanding products by customers with 88% and 65% demand respectively.

Major respondents agree on the follwing points

1) Online shopping saves time.
2) Pay in advance is not an issue.
3) Arrangement of product in wedsite is very important.

Online shopping organizations can apply the relevant variables and factors, identified from the research, to create their strategies and tactics. The organizations can prioritize the Consumer's requirements in online shopping environment. The results can also be used by various organizations to identify their target customer segments.

From the results of ANOVA table we conclude that the demographic factor Gender and Occupation does impact on payment method opted by consumer. it clearly shoes that there is a effect of occupation and gender on payment method. In future we try to find out which occupation more effects various payment method like Credit card, debit card, net banking and e-wallets.

VIII. REFERENCES

[1] Barkhordari, Maryam, et al. "Factors influencing adoption of e-payment systems: an empirical study on Iranian customers." Information Systems and e-Business Management 15.1 (2017): 89-116.

[2] Nagra, Gagandeep, and R. Gopal. "An study of Factors Affecting on Online Shopping Behavior of Consumers." *International Journal of Scientific and Research Publications* 3.6 (2013): 1-4.

[3] Kumar, Naveen, and Upasana Gupta. "Factors Influencing Online Purchase Behaviour of Customers in Tier III Cities of India-A Factor Analysis Approach." International Journal of Marketing & Business Communication 6.1 (2017).

[4] Sharma, Renuka, Kiran Mehta, and Shashank Sharma. "Understanding online shopping behaviour of Indian shoppers." International Journal of Management & Business Studies 4.3 (2014): 9-18.

[5] Kanchan, Upasana, and Naveen Kumar. "A Study about Impact of Customer Characteristics on Online Purchase Behavior in Indian Context." American Journal of Economics and Business Administration 7.3 (2015): 130.

[6] Richa, Dahiya. "Impact of demographic factors of consumers on online shopping behaviour: A study of consumers in India." International Journal of Engineering and Management Sciences 3.1 (2012): 43-52.

[7] Moshrefjavadi, Mohammad Hossein, et al. "An analysis of factors affecting on online shopping behavior of consumers." International Journal of Marketing Studies 4.5 (2012): 81.

[8] Kalia, Prateek, Penny Law, and Richa Arora. "Determining Impact of Demographics on Perceived Service Quality in Online Retail." Encyclopedia of Information Science and Technology, Fourth Edition. IGI Global, 2018. 2882-2896.

[9] Verma, Shanti, and Kalyani Patel. "Impact of Consumer Gender on Expenditure Done in Mobile Shopping Using Test of Independence." Information and Communication Technology for Sustainable Development. Springer, Singapore, 2018. 13-19.

[10] Verma, Shanti, and Kalyani Patel. "Dependency between type of product purchased and payment method opted by Indian consumer in Mobile shopping." Cloud Computing, Data Science & Engineering-Confluence, 2017 7th International Conference on. IEEE, 2017.

[11] Verma, Shanti. "Deciding Admission Criteria For Master of Computer Applications Program in India using Chi-Square Test." Proceedings of the Second International Conference on Information and Communication Technology for Competitive Strategies. ACM, 2016.

[12] Verma, Shanti, and Jignesh Doshi. "Correlation Between Text Book Usage and Academic Performance of Student in Higher Education Using 'R'." Proceedings of International Conference on Communication and Networks. Springer, Singapore, 2017.

© 2018 IJSRSET | Volume 4 | Issue 1 | Print ISSN: 2395-1990 | Online ISSN : 2394-4099
National Conference on Advanced Research Trends in Information and
ng Technologies(NCARTICT-2018), Department of IT,L. D. College of Engineering,
Ahmedabad, Gujarat, IndiaIn association with
International Journal of Scientific Research in Science, Engineering and Technology

An Improved K-Means Clustering Algorithm

Ekta Joshi[1], Dr. D. A. Parikh[2]
[1]Computer Engineering, L.D. College of Engineering, Ahmedabad, India
[2]HOD Computer Engineering, L.D. College of Engineering, Ahmedabad, India

ABSTRACT

This Vast spread of computing technologies has led to abundance of large data sets. Thus, there is a need to find similarities and define groupings among the elements of these big data sets. One of the ways to find these similarities is data clustering. Currently, there exist several data clustering algorithms which differ by their application area and efficiency. Increase in computational power and algorithmic improvements have reduced the time for clustering of big data sets. But it usually happens that big data sets can't be processed whole due to hardware and computational restrictions. Clustering techniques, like K-Means are useful in analyzing data in a parallel fashion. K-Means largely depends upon a proper initialization to produce optimal results.

Keywords: K means, Clustering, Data Mining, Big Data.

I. INTRODUCTION

There has been a tremendous growth in the volume of data in the recent times. Data, whether it be structured or unstructured contribute to this enormous collection. To draw meaningful insights from this mountain of data we need algorithms which can perform analysis on this data. Clustering is the process of grouping data into groups called clusters, so that the objects in the same cluster are more similar to each other and more different from the objects in the other group [1]. It is one of these various important analysis techniques that is employed to large datasets and finds its application in the fields like search engines, recommendation systems, data mining, knowledge discovery, bioinformatics and documentation to name a few.

Nowadays, the data being generated is not only huge in volume, but is also stored across various machines all around the world. We need to process this data in parallel to reduce the cost of processing. K-Means is one of the most famous algorithms in the field of data mining [6]. Its scalability to large datasets and simplicity can be considered as one of the major reasons for its popularity. It is simple in data analysis and provides good performance. But it has a great dependence on the initial cluster center. The selection of initial cluster centers determines the quality of clustering. Therefore, it is an important step to select a reasonable set of initial cluster centers in K-means algorithm.

II. THE TRADITIONAL K - MEAN ALGORITHM [6]

K-means algorithm is a clustering algorithm based on partition, proposed by McQUeen in 1976. The aim of Kmeans algorithm is to divide M points in N dimensions into K clusters so that the precision rate and the recall rate are maximum. It is not practical to require that the solution has maximum against all partitions, except when M, N are small and K=2. The algorithm seeks instead of "local" optima solution, such that no movement of an object from one cluster to another will reduce the within-cluster sum of squares.

The basic principle of the traditional K-means algorithm is: firstly, each data object in the data set is regarded as a single cluster, randomly select K data objects as the initial clustering centers; secondly, successively calculate the distance of the rest data objects to each of the K cluster center, each data object will be categorized into the nearest cluster, and then recalculate the centroid

of each cluster; repeat iteratively until the cluster partition is no longer changed. The process of K-means algorithm is as follows:

Input: data set contained n data objects, k(the number of clusters) ;
Output: k clusters;
Step1: Randomly select K data objects as the initial cluster centers;
Step2: Calculate the distances from the remaining data objects to initial cluster centers, assigned the remaining n-k data objects to the nearest cluster;
Step3: Recalculate the cluster centers of each cluster;
Step4: repeat step2 and step3 until convergence;

K-means algorithm is a simple and efficient clustering algorithm [6]. Its time complexity is close to O(n*k). When the differences between categories are small or the scale of data set is large, K-means algorithm will perform more efficient, and get better clustering results. It has two major drawbacks- (1) A priori fixation of the number of clusters (2) Random selection of initial centers. So there are different methods to improve the algorithm while maintaining its simplicity and efficiency.

III. TYPE OF METHODS

A. K-mean+:a developed clustering algorithm for big data [2]

In this paper, it proposed a new approach for fast clustering. It divides first instances into blocks applying block operation.
Block operation: let dataset D has M attributes and N instances for each attribute; range is divided by *f* equal width. The feature space of D is separated to blocks of a size. N instances are assigned to these blocks and processed as one instance but weighted by number of instances in single block.
Distance calculation: location of block is decided by the center of a block and then Manhattan Distance as function of distance measure instead of Euclidean distance, causes much floating point arithmetic.
Iteration: Cluster center and cluster labels of block update is iterated. When cluster center do not change, iteration stops. The result is rounded to integer to reduce complexity as cluster centers are weighted average of blocks.

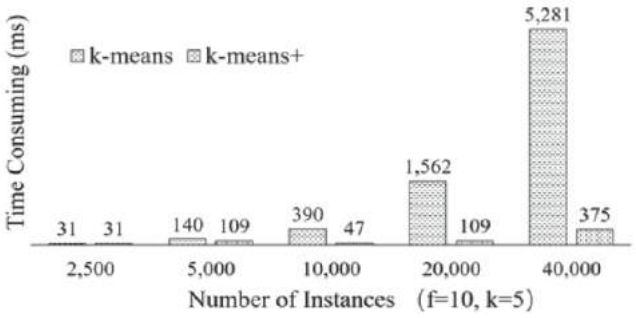


Figure 1. Time complexity [2]

B. Batch Clustering Algorithm for Big Data Sets [4]

In this paper, it is proposed to cluster a given large data set in batches by using k-means algorithm. That is take some portions of data elements from the given data set and process it. Then take next portion and process it and so on until all the elements of data set are processed. After that the whole given data set is also clustered by k-means algorithm for efficiency and quality comparison. Later qualitative indicators for both of these approaches are measured. For qualitative indicators the below are considered:

- Time T required for calculation of centroids and assigning data elements to centroids (clusters);
- The value of the objective function J (squared error function):

$$j = \sum_{j=1}^{k} \sum_{i=1}^{n} \left\| x_i^{(j)} - c_j \right\|^2$$

Where $\left\| x_i^{(j)} - c_j \right\|^2$ is a given distance (Euclidean) measure between a data element $x_i^{(j)}$ and the cluster center c_j , i.e. j is an indicator of the Euclidean distance of the n data elements from their corresponding cluster centers.
Assuming that we have a data set consisting of n data elements, m is a number of elements in a small subset of data set, k is a number of centroids, the proposed batch clustering algorithm can be defined as follows:

1. We take m number of data elements from the given data set of n elements (m < n);
2. m number of data elements are processed in RAM of computer by the k-means algorithm to find k number of centroids (k < m);
3. Then take the k centroids along with m − k number of data elements from the remaining data set;

4. Repeat steps 2 – 3 until all the elements of the initial data set are processed;
5. Assign all elements to the k centroids calculated in the last step 4.

The result of calculations is below:

Classic k-means
• Time spent calculating centroids: T = 225 sec
• Time spent calculating objective function: T = 979 sec
• Objective function: J = 33997

Batch clustering
• Number of elements in each portion: m =100,000 elements.
• Time spent calculating centroids: T = 223 sec
• Time spent calculating objective function: T = 932 sec
• Objective function: J = 33996

As seen from the above calculation results, batch clustering algorithm produces better results over classic k-means algorithm. Besides, we have a big gain in using computational power with restricted resources.

C. New Approach for Clustering of Big data: DisK-means[1]

The proposed new algorithm called DisK-Means. The traditional K-Means algorithm produces varying results over several runs on large datasets. It is also very time consuming. Our algorithm reduces the time of execution along with improving the quality of clusters.

Our algorithm divides the complete dataset into *m* parts, where each part must have more than minimum the sample size to be representative of *Y*. It is not necessary that it should be equal to the computational cores available as in CK-Means. In step 2, for each part obtained from step 1, a distance matrix [15] is made which contains the shortest distance from each point to every other point in the set. *D(i,j)* represents the shortest distance from point *i* to *j* and *M(i,j)* represents the matrix that has *i* rows and *j* columns. Now for each subset, choose the point that has the minimum sum of distances. In step 5, K-Means++ is performed on each subset obtained in the initial stage. Here, the initial points for each subset are not chosen randomly but from each subset in the previous step. Next, fitness measure of each cluster is calculated and stored in the array *Ti*. We use WSSQ method as it is easy to calculate and does not add further complexity. The cluster that has the minimum value of WSSQ is the best fit and its centers are put in the set *C* which contains the initial centers for the next step. The re-clustering is further performed using K-Means to finally obtain *k* centers.

Algorithm 3 DisK-Means [1]

1. Partition *y* into *y1, y2,..., ym* ;
2. For each $i \in \{1,2,.....m\}$ do
3. Calculate the distance *D(i,j),* from each point to every other point and form the distance matrix *M(i,j).*
4. Find the node with minimum sum of distances and put it into *ICi*;
5. Run K-Means++ on *yi* using *ICi* as the initial node to get *k* centroids *Ci'* and clusters (*clyi*);
6. *Ti* ←*f*(*clyi*);
7. *C*=*Ci'* where i←Best-fit;
8. Re-clustering using K-Means with *C* as set of initial centers.

TABLE 1. COMPARISION BASED ON RUNNING TIME [1]

S.no.	Running Time (in sec)		
	Values of k	K-Means	*Disk-Means*
1	50	300	220
2	50	340	223
3	100	200	127
4	100	213	130

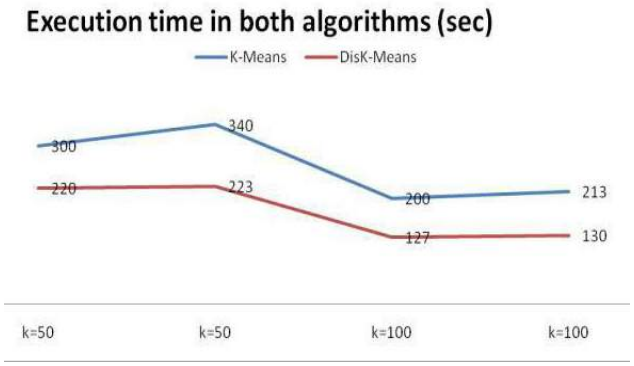


Figure 2. Execution time for K-Means and DisK-Means [1]

It can be observed that in both cases, when k=50 and k=100, our new DisK-Means takes less time to execute with less deviation when compared to K-Means. A major drop is observed when the value of k is changed

dramatically from 50 to 100. It is expected that as the number of clusters increase, the time for execution would decrease.

D. A hybrid clustering algorithm : the FastDBSCAN [7]

Proposed FastDBSCAN including two steps:

(1) To Partition the data set by K-means and then use Min-Max method to sample data. The main principle of min-Max is: first a starting point y1 is randomly chosen from the dataset D. Then all the other points in Y are chosen among the points of dataset X that maximize their minimal distance from the point already in Y. Thus when t points already belong to y, the process that selects the point y_{t+1} from X can be formalized as shown in equation below:

where d(.) denotes the distance defined in the space of the objects. In each iteration methods selects the point that exhibits the largest label uncertainty according to the previous answers of the user.

Algorithm 1: Min-Max method;
Input: data set D, the number of samples k;
Output: set of points selected by Min-Max Y.

1. Take any reference point r;
2. Insert r in Y
3. Temp = 1;
4. while |temp|$\leqslant$ k+1
5. Find the point x that maximize their minimal distance from the points already in Y
6. Insert x in Y
7. temp = temp +1
8. endwhile
9. remove r from Y
10. return Y

(2) Clustering sampled data by FastDBSCAN. K-means in first step guarantee that the data chosen for step 2 will cover the whole data set. Then it extracts t percent of points by Min-Max method. This new set is used by DBSCAN.

Algorithm 2: FastDBSCAN
Input: A data set D, the number of clusters for K-Means k, the proportion of data t;
Output: Clusters and noises.

1. Initialize k centers
2. Partition data by K-Means,
3. Take a proportion t of points (Min-max algorithm) from clusters to form a new data set E; build a correspondence list to associate each selected point with its cluster.
4. Perform DBSCAN clustering on the set E,
5. Recover the clusters detected by DBSCAN to form final clusters.

Comparison of FastDBSCAN and DBSCAN is with two aspects: clustering accuracy and time calculation. The accuracy of the algorithm is better as k-means is used for clustering in the step one. The time calculation of FastDBSCAN is less compared to DBSCAN.

E. An Improved K-means text clustering algorithm By Optimizing initial cluster centres[5]

The basic idea of the improved K-means algorithm is: at first, calculate the density parameter of all data objects in the data set, and determine data objects which are isolated points.

If a data object is isolated, it will be removed from the data collection. After deleting the isolated points, we will get a data set with high density parameter. Then, select a data object with the highest density parameter in the set as the first initial cluster centre; Next, select a data object from the rest of high-density data collection as the second initial cluster centre, which is the furthest from the first initial cluster centre; And so on, until find k initial cluster centres. Based on this k initial cluster centres, use the traditional K-means algorithm to do clustering.

The process of the algorithm is described as follows:
Input: text set D = {d1, d2, ⋯ , dn} containing n data objects, and k (the number of clusters);
Output: k clusters;

Step1: Calculate distances between any two data objects in data set D and the average distance, using formula (1) and (2) respectively;

Step2: Calculate density parameters of all data objects in the data set D and the average density parameter of the set D, using formula (3) and (4);

Step3: According to the formula (5), determine isolated data objects, and delete them from the set D, thus obtain a collection A with high density parameter;

Step4: Select a data object with the highest density parameter from collection A as the first initial clustering center, and add it to the collection B, and remove it from collection A;

Step5: From collection A, select a data object which is furthest from collection B as the next initial cluster center, and add it to collection B, and remove it from collection A;

Step6: Repeat Step5, until the number of data objects in collection B is k;

Step7: Based on the k cluster center, use the traditional Kmeans to do clustering;

This algorithm is divided into two stages, one is to determine the initial cluster centers, and the other is to use the traditional K-means to do clustering based on the initial cluster centers. To determine the initial cluster centers needs to calculate the distances between all data pairs, and its time complexity is O(n2). Because of the extra computation, this algorithm's time complexity is higher than that of the original algorithm, but it can get a better clustering result.

TABLE 2. Precision Evaluation [5]

Clustering algorithm	**Evaluation index**	**Art**	***Economy***	***Environment***	***Political***	***Sports***
K-MEANS	P (%)	80.33	78.84	79.53	80.95	78.62
	R (%)	81.14	82.45	78.97	81.83	79.69
Improved k-mean	P (%)	81.74	80.42	81.21	82.71	81.58
	R (%)	82.23	83.59	82.36	83.75	82.64

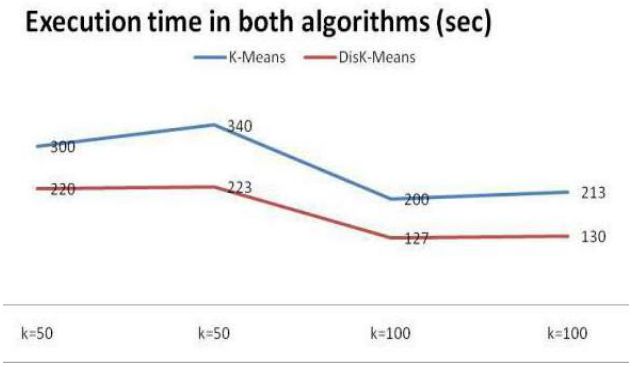


Figure 3. Execution time for K-Means and DisK-Means [5]

be controlled by noise or isolated data, leading to the inaccurate or even wrong clustering results.

IV. ANALYSIS

In this paper, a study of different types of algorithms of k-means modified been represented each have improved the one or the other shortcomings of the k-means algorithm and hence provided a better and efficient algorithm to learn a large data set. A verity of methods are been introduced want to tackle a large amount of dataset and provide a fast clustering while keeping the spark of k-mean simplicity and efficiency.

V. CONCLUSION

The problem of k-mean clustering algorithm has a lot of shortcomings: (1)It requires the user to specify the number of clusters in advance. However, in the beginning, the user does not know how many clusters should be divided into. (2) It has a great dependence on the initial cluster center, and it is easy to produce the local optimal solution. Since K-means' clustering criterion function is a non-convex squared error evaluation function, which leads to there is only one global minimum, but there are a number of local minimum. The randomly selected initial clustering centers tend to fall into the non-convexity, causing the algorithm deviates from the searching range of global optimal solution. So, when the initial clustering centers are selected improperly, the clustering results will be unstable and inaccurate. (3)It is sensitive to isolated points and noise data. K-means algorithm takes average point as cluster's center, and adds it to the next round of the algorithm, resulting in the cluster's center may be away from the dense regions of data set, and the cluster's center may be a noise point or an isolated point. Therefore, if the data set contains a lot of isolated points or noise data, to a great extent, the clustering results will

Different aspects of k-mean algorithm are deal with and different approaches are introduced working on different issues of k-means clustering algorithm. One has reduced time to response, other to reduce the complexity of algorithm using k-mean. One of papers is dealing with a large amount of data dividing the work into batches. All

these methods enhanced traditional k-mean algorithm to an efficient algorithm.

VI. REFERENCES

[1] Anu Saini, G. B. Pant ,Jaypriya Ubriani "New Approach for Clustering of Big Data: DisK-Means", 2016 IEEE ,International Conference on Computing, Communication and Automation ,pp 122-126;

[2] Kun niu, zhipeng gao,haizhen jaog ,haijie deng "K-mean+:a developed clustering algorithm for big data", 2016 IEEE , Proceedings of CCIS2016,pp 141-144;

[3] Vadlana Baby,Dr. N. Subhash Chandra "Distributed threshold k-means clustering for privacy preserving data mining",2016 IEEE,Conference on Advances in Computing, Communications and Informatics (ICACCI);

[4] Rasim Alguliyev , Ramiz Aliguliyev , Adil Bagirov , Rafael Karimov "Batch Clustering Algorithm for Big Data Sets";

[5] Caiquan Xiong, Zhen Hua, Ke Lv, Wuhan Hubei ,"An Improved K-means text clustering algorithm By Optimizing initial cluster centers", 2016 IEEE, International Conference on Cloud Computing and Big Data,pp 265-268;

[6] Jiawei Han, Jian Pei, Micheline Kamber "Data Mining: Concepts and Techniques" 3rd edition;

[7] Vu Viet Thang, D.V. Pantiukhin, A.I. Galushkin "A hybrid clustering algorithm : the FastDBSCAN" 2015 International Conference on Engineering and Telecommunication,pp 69-74;

[8] Tahereh Kamali, Daniel Stashuk "A Density-Based Clustering Approach to Motor Unit Potential Characterizations to Support Diagnosis of Neuromuscular Disorders" 2016 IEEE Transactions on Neural Systems and Rehabilitation Engineering ;

[9] Bin Jiang, Jian Pei, Yufei Tao and Xuemin Lin, Member, IEEE "Clustering Uncertain Data Based on Probability Distribution Similarity" IEEE TRANSACTIONS ON KNOWLEDGE AND DATA ENGINEERING, VOL. 25, NO. 4, APRIL 2013;

[10] Chang Lu , Yueting Shi, Yueyang Chen, Shiqi Bao, Lixing Tang "Data Mining Applied to Oil Well Using K-means and DBSCAN" 2016 7th International Conference on Cloud Computing and Big Data;

[11] Jianbing Shen, Xiaopeng Hao, Zhiyuan Liang, Yu Liu, Wenguan Wang,and Ling Shao, Member, IEEE "Real-time Superpixel Segmentation by DBSCAN Clustering Algorithm" 2016 IEEE TRANSACTIONS ON IMAGE PROCESSING;

[12] Dongming Tang.Affinity propagation clustering for bid data based on Hadoop. Computer Engineering and Applications, 2015, 51(4):29-34;

[13] Joshua M.Dudik a, AtsukoKurosu b, JamesL.Coyle b, ErvinSejdić a,n "A comparative analysis of DBSCAN, K-means, and quadratic variation algorithms for automatic identification of swallows from swallowing accelerometry signals", Computers in Biology and Medicine 59 (2015);

[14] Jesal Shethna "Data Mining Techniques available from https://www.educba.com/7-data-mining-techniques-for-best-results/" November 7, 2016;

[15] Martin Brown "Key techniques from https://www.ibm.com/developerworks/library/ba-data-mining-techniques/" Published on December 11, 2012;

[16] Data Mining tutorials "Data Mining Techniques from http://www.zentut.com/data-mining/data-mining-techniques/"

[17] Saurabh Arora, Inderveer Chana "A Survey of Clustering Techniques for Big Data Analysis" 2014 5th International Conference- Confluence The Next Generation Information Technology Summit (Confluence),pp 59-65.

[18] Martin Ester, Hans Peter Kriegel, Jorg Sander, Xiaowei Xu, "A Density-Based Algorithm for Discovering Clusters in Large Spatial Databases with Noise", Published in Proceedings of 2nd International Conference on Knowledge Discovery and Data Mining (KDD-96)

Organised by

Information Technology Department,
L.D. College of Engineering (LDCE), Ahmedabad, Gujarat, India

Publisher

Technoscience Academy

Website : www.technoscienceacademy.com
Email: info@technoscienceacademy.com