

ISSN(p): 2394-2835

# **International Journal of Advances in Electronics and Computer Science (IAECS)**

**Volume No. 11  
Issue No. 1  
January - April 2024**



**ENRICHED PUBLICATIONS PVT. LTD**

**S-9, IInd FLOOR, MLU POCKET,  
MANISH ABHINAV PLAZA-II, ABOVE FEDERAL BANK,  
PLOT NO-5, SECTOR-5, DWARKA, NEW DELHI, INDIA-110075,  
PHONE: - + (91)-(11)-47026006**

# **International Journal of Advances in Electronics and Computer Science (IJAECs)**

## **Aim & Scope**

International Journal of Advances in Electronics and Computer Science (IJAECs) is a scholarly online, open access, peer-reviewed, interdisciplinary and fully refereed journal focusing on theories, methods and applications in Electronics Engineering and Computer Science . It is an international scientific journal that aims to contribute to the constant scientific research and training, so as to promote research in the field of Electronics and Computer Science Engineering

## Editor-in-Chief

**Dr. P. Suresh**  
Professor, KCE Coimbatore  
Life Member Of ISTE, ISNT, IIE

## Associate Editor

**J. K. Chen, Ph.D.**  
William And Nancy Thompson Professor, University Of Missouri  
Columbia, MO 65211, USA Life Member: Institute Of Research And Journals

## Other Editorial Board Members

**Dr Chi Hieu Le**  
MEng MSc DrEng, CEng, PGCHE, FHEA, MIMechE  
Pembroke, Central Avenue. Chatham Maritime  
Kent ME4 4TB. United Kingdom

**Prof. Dr. GOVINDARAJ THANGAVEL**  
Prof. & Head , EEE, Muthayammal Engineering College  
Rasipuram ,TamilNadu, India, Ph.D: Jadavpur University,  
India, Fellow Of Institution Of Engineers Of India  
(F.I.E.), Chartered Engineer (C Eng).  
Life Member Of Indian Society For Technical  
Education.(MISTE)

**Mohammad Samadi Gharajeh**  
Department Of Computer Engineering Tabriz Branch,  
Islamic Azad University Tabriz, Iran

**Prof. ANAND NAYYAR,**  
Assistant Professor, Department Of Computer  
Applications & IT,  
KCL Institute Of Management And Technology

**Prof (Dr) Manoj Wadhwa**  
Professor & Head , Dept. Of CSE & CSIT , Echelon Institute Of Technology, Faridabad, India  
Ph.D:Kurukshetra University, Kurukshetra,India, Member: IEEE USA (Delhi Region,ISTE,CSI



# International Journal of Advances in Electronics and Computer Science (IJAECs)

(Volume No. 11, Issue No. 1, January - April 2024)

## Contents

| Sr. No. | Article / Authors Name   | Pg. No. |
|---------|--|---------|
| 1       | Format-Preserving Encryption: Image Encryption under FF1 Scheme<br>- <i>Hadrian Rodriguez Cesar, Victor Gayoso Martinez, Luis Hernandez Encinas, Agustin Martin Munoz</i>        | 1 - 10  |
| 2       | Smart Irrigation using IOT and Virtuino<br>- <i>Sudarshan V, Prashanth E, Chirag A, Jananee V</i>  | 11- 16  |
| 3       | Data Breaches in IOT A Study and Solution by PUF Approach<br>- <i>Prakhar Sahu, S. K. Singh, Pankaj Kumar</i>  | 17 - 26 |
| 4       | A Recommendation Model for Cloud Service Selection using Fuzzy Logic<br>- <i>Parthasarathy. M, Aniruth. R, Gokulakrishnan. R, Krishnan. V. C, Sundaraganapathy. L, Subha. T.</i> | 27 - 36 |
| 5       | Use of Big Data and Data Mining to Enhance Productivity in the Business<br>- <i>M. G. P. M. Samarasinghe, K. A. Dilini. T. Kulawansa</i>   | 37 - 50 |



# Format-Preserving Encryption: Image Encryption under FF1 Scheme

<sup>1</sup> Hadrian Rodriguez Cesar, <sup>2</sup> Victor Gayoso Martinez, <sup>3</sup> Luis Hernandez Encinas, <sup>4</sup> Agustin Martin Munoz

<sup>1,2,3,4</sup> Institute of Physical and Information Technologies (ITEFI), Spanish National Research Council (CSIC) C/ Serrano 144, 28006-Madrid, Spain

E-mail: <sup>1</sup>hadrian.rodriquez, <sup>2</sup>victor.gayoso, <sup>3</sup>luis, <sup>4</sup>agustin@iec.csic.es

## ABSTRACT

*In this work, we have analyzed the NIST (National Institute of Standards and Technology) special publication focused on the FF1 scheme for format-preserving encryption. Using that information, we have built an encryption model applied to image encryption, upon which we have performed several tests, namely histogram and analysis NPCR (Net Pixel Change Ratio) scores, and computational cost.*

**Keywords - Cryptography, Format-Preserving Encryption, FF1, FFX, Image Encryption.**

## I. INTRODUCTION

In the literature, Format-Preserving Encryption (FPE) refers to any cryptosystem that maintains the format of the plaintext once it is encrypted. More formally, the domain and range of the encryption (thus decryption) function are the same.

In practice, FPE is linked to any cryptosystem used for encrypting data whose format needs to be preserved, that is why despite that AES is technically a FPE (it encrypts a 128-bit string as another 128-bit string) we will focus on those domains whose format is interesting to be kept unchanged. Namely, document numbers, telephone numbers, geospatial coordinates, and potentially everything that can be stored in a database.

The original idea arose on 1997 under the name Datatype-Preserving encryption, but it did not become popular until 2002 [1]. Black and Rogaway analyzed three different approaches to encrypt a plaintext using a block cipher for a domain of the type  $[0, k-1]$  for any integer  $k$ : prefix cipher, cycle-walking cipher and generalized-Feistel cipher. The first method has major drawbacks and it only works on fairly small domains,  $k < 2^{30}$ . The second method works better if the message space is "dense" in the block cipher domain,  $k > 2^{60}$ , while the Feistel method is more efficient for  $> 2^{128}$ . Then, there exists a gap for  $k \in [2^{30}, 2^{60}]$  for which no efficient solution is known yet. The same happens for those domains that do not efficiently admit a bijection to a set of the form  $[0, k-1]$ .

Back in 2010, Miare, Rogaway and Spies published the first algorithm to effectively encrypt data using the FPE technique with an unbalanced Feistel structure and a tweak [2]. As they stated, FPE has applications like data sanitization and field encrypting in legacy databases. Originally named FFX (Format-Preserving, Feistel-Based), it produced a deterministic ciphertext on a  $|\text{Chars}|^n$  size alphabet where  $n$  is the size of the plaintext and Chars is the number of symbols of the used alphabet. In their paper they called FFX-A2 and FFX-A10 to two recommended instances for some parameter choices that achieved security against CPA and PRP-security against CCA2. It was proposed to NIST (National Institute of Standards and Technology) under the name FFX[Radix] and accepted after some fixes on parameter choices. In 2013 becoming FF1 on the SP 800-38G.

VAES3 was a set of parameters for FFX proposed to the NIST and accepted under the name FF2 on the same Special Publication as FF1 in 2013, but it was taken away in 2015 after a chosen-plaintext attack was found, due to this the 128 security bits were not achieved anymore. Currently VAES3 is under research.

The last FPE cryptosystem approved on the SP 800-38G is an instance of the so-called BPS with a 128-bit block cipher, named FF3 on that special publication. The idea was to create a framework called FFX where FF1, FF2 and FF3 were just specific instances for different parameter sets. For more information, it is recommended to review [3].

During the last years, few attacks to FPE schemes were discovered [4], [5] and a theoretical framework for an identity-based FPE system was drawn up [6]. In addition, some examples have been studied like [7], [8] with good results specially for IoT applications.

In this work, we have analyzed the NIST proposal for format-preserving encryption. In particular, we have considered the FF1 scheme and based on it, we have built a novel encryption model to be applied to image encryption. The rest of this paper is organized as follows: In Section II we show the most important properties and characteristics of the FF1 scheme. In Section III, the proposal of how to apply this scheme for image encryption is presented. The results obtained when some properties of this proposal have been studied are considered in Section IV. Finally, Section V includes some conclusions.

## II. A QUICK GLANCE AT Ff1

Ff1 was standardized from his predecessor FFX with a specific set of parameters. Roughly it is an unbalanced Feistel-based cryptosystem that takes a numeral string  $X$  in a base radix, with length  $n \in [\text{minlen}, \text{maxlen}]$ , (that is  $X[i] \in [0, \text{radix}]$  for  $i \in [0, n]$ ) and it encrypts it by using a 128-bit block cipher



and a byte string T called "weak", being the result another numeral string with the same length and base. For this purpose, it uses a 128-bit key, K, and a tweak T of length t (the use of the tweak is optional as t = 0 is allowed. As described in [3] the parameters must meet the following conditions:

- radix  $\in [2, 2^{16}]$ ,
- radix<sup>minlen</sup>  $\geq 100$ ,
- $2 \leq \text{minlen} \leq \text{maxlen} < 2^{32}$ .

The FF1 encryption scheme is composed by following steps (Fig. 1 shows FF1 encryption and decryption schemes):

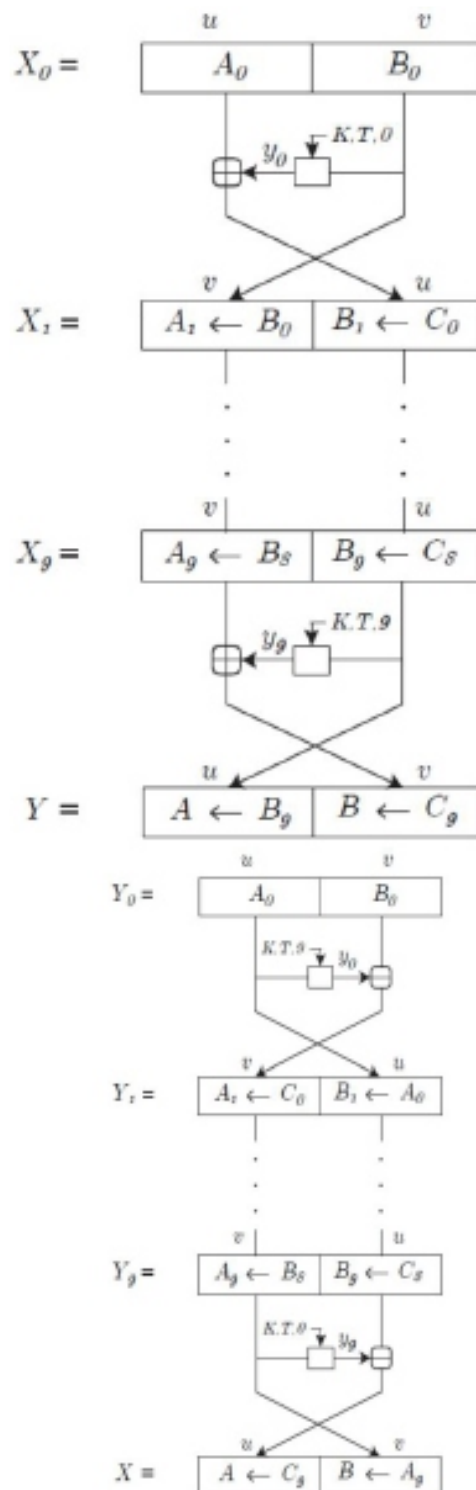
1. Let  $u = \lfloor n/2 \rfloor$ ,  $v = n - u$ .
2. Let split  $X = A || B$ , being  $A = X[1, u]$  and  $B = X[u+1, v]$ .
3. Let  $b = \lfloor \lfloor v \cdot \text{LOG}(\text{radix}) \rfloor / 8 \rfloor$ ,  $d = \lfloor 4b/4 \rfloor + 4$ .
4. Form  $P = [1]^1 || [2]^1 || [1]^1 || [\text{radix}]^3 || [10]^1 || [u \bmod 256]^1 || [n]^4 || [t]^4$ , where  $[a]b$  stands for the byte representation of a with length b.
5. For i from 0 to 9:
  - a) Let  $Q = T || [0]^{(-t-b-1) \bmod 16} || [i]^1 || [\text{NUMradix}(B)]^b$ .
  - b) Let  $R = \text{PRF}(P || Q)$ . The PRF function is detailed below.
  - c) Let S be the first d bytes of the string  $R || \text{EK}(R \oplus [1]^{16}) || \text{EK}(R \oplus [2]^{16}) || \dots || \text{EK}(R \oplus [\lfloor d/16 \rfloor - 1]^{16})$ .
  - d) Let  $y = \text{NUM}(S)$ , where  $\text{NUM}(\cdot)$  transforms a byte string into its integer representation.
  - e) Let  $m = u$  if i is even,  $m = v$  otherwise.
  - f)  $c = (\text{NUMradix}(A) + y) \bmod \text{radix}^m$ .
  - g)  $C = \text{STRmradix}(c)$ , where  $\text{STRmradix}(\cdot)$  takes an integer and transforms it into a string representation in base radix and length m.
  - h) Let  $A = B$  and  $B = C$ .
6. Return  $A || B$ .

The PRF function: It takes a bit string X whose length is multiple of 128 and performs the following steps:

1.  $m = \text{LEN}(X)/128$ ,  $X = X_1 || \dots || X_m$ .
2. Let  $Y_0 = [0]^{16}$ . For j from 1 to m:
  - a) Let  $Y_j = \text{EK}(Y_{j-1} \oplus X_j)$ .
3. Return  $Y_m$ .

The FF1 decryption scheme works similarly. The differences between them are marked in bold below:

1. Let  $u = \lfloor n/2 \rfloor$ ,  $v = n - u$ .



**Fig. 1. FF1 encryption and decryption schemes.**

2. Let  $X=A||B$ , being  $A=X[1, u]$  and  $B=X[u+1, v]$ .
3. Let  $b=\lfloor \lfloor v \cdot \text{LOG}(\text{radix}) \rfloor / 8 \rfloor$ ,  $d=\lfloor 4b/4 \rfloor + 4$ .
4. Form  $P = [1]^1 || [2]^1 || [1]^1 || [\text{radix}]^3 || [10]^1 || [u \bmod 256]^1 || [n]^4 || [t]^4$ , where  $[a]b$  stands for the byte representation of  $a$  with length  $b$ .

5. For  $i$  from 9 to 0:

a) Let  $Q = T \parallel [0]^{(-t-b-1) \bmod 16} \parallel [i]^1 \parallel [\text{NUMradix}(A)]^b$ .

b) Let  $R = \text{PRF}(P \parallel Q)$ . The PRF function is detailed below.

c) Let  $S$  be the first  $d$  bytes of the string  $R \parallel \text{EK}(R \oplus [1]^{16}) \parallel \text{EK}(R \oplus [2]^{16}) \parallel \dots \parallel \text{EK}(R \oplus [\lfloor d/16 \rfloor - 1]^{16})$ .

d) Let  $y = \text{NUM}(S)$ .

e)  $c = (\text{NUMradix}(B) + y) \bmod \text{radixm}$ .

f)  $C = \text{STRmradix}(c)$ .

g) Let  $B = A$  and  $A = C$ .

6. Return  $A \parallel B$ .

### III. FF1 APPLIED TO IMAGE ENCRYPTION

As we have seen, FPE has already been applied to different fields such as datetime strings, credit card numbers or geospatial coordinates. In this paper we describe the tests performed with a Java 11 implementation of FF1 for image encryption developed from scratch by the authors of this contribution.

In order to encrypt an image, we have treated each pixel individually as a byte string with fixed length 3. NIST parameter specifications are respected as  $\text{radix} = 256$ ,  $\text{minlen} = \text{maxlen} = 3$  and  $\text{radixminlen}$  is way greater than 100.

Moreover, the tweak length must verify  $t = \text{LEN}(T) \geq \lceil \text{LOG}(N)/4 \rceil$ , where  $N = \text{width} \cdot \text{height}$ . This length will ensure the use of one different tweak for each pixel, so encryption is not just a palette change. This scheme starts by using a given tweak  $T$  and it is increased for each encryption/decryption. It starts at the left upper corner, which corresponds to coordinates  $(0,0)$ , and finishes at the right bottom corner, with coordinates  $(w-1, h-1)$ .

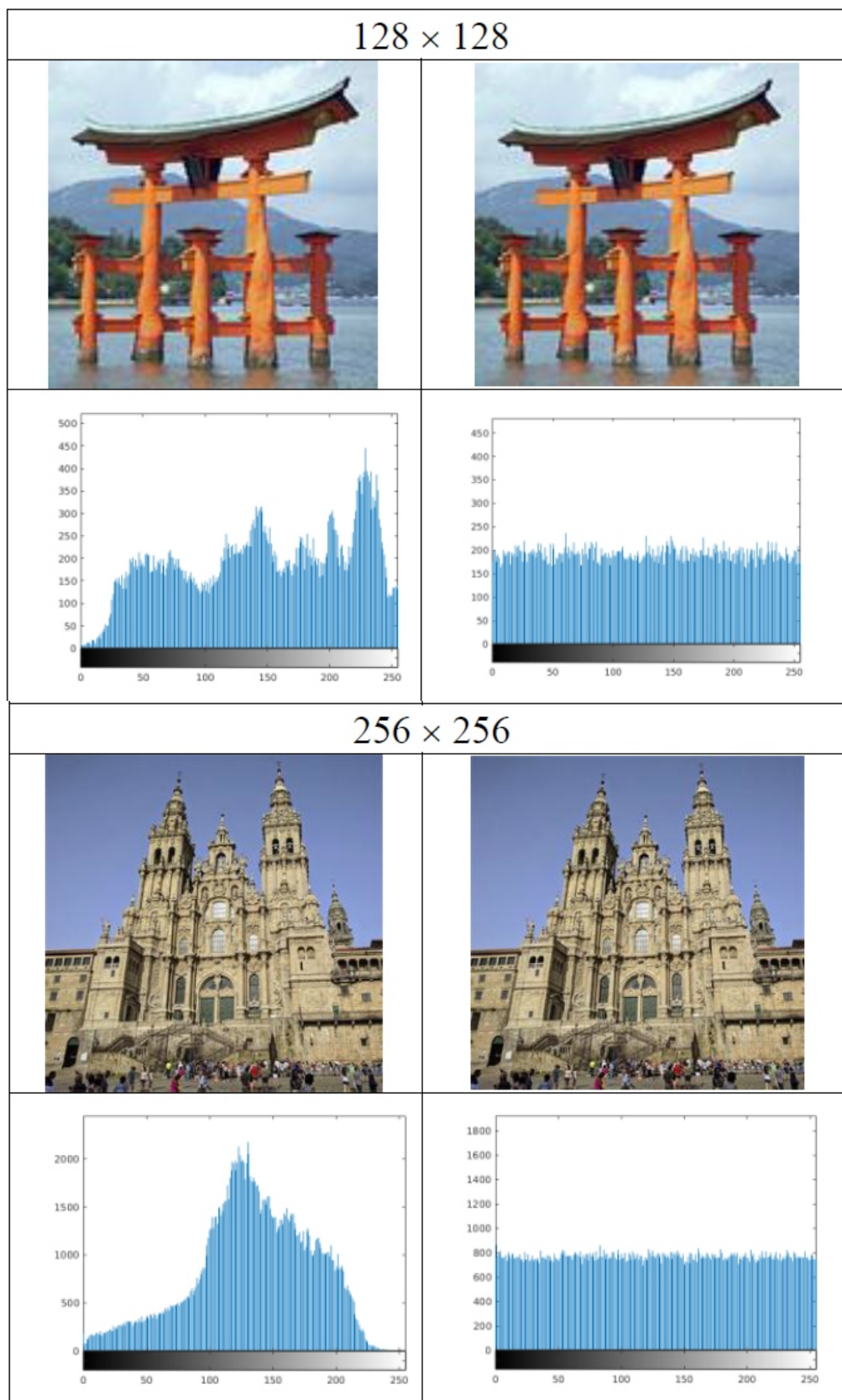
The process for encrypting the pixel  $(i, j)$  is as follows: Let be  $I_{i,j}$  for  $(i, j) \in [0, w-1] \times [0, h-1]$  the numeral string of a given colour pixel and  $T$  the initial tweak for the image  $I$ . Then:

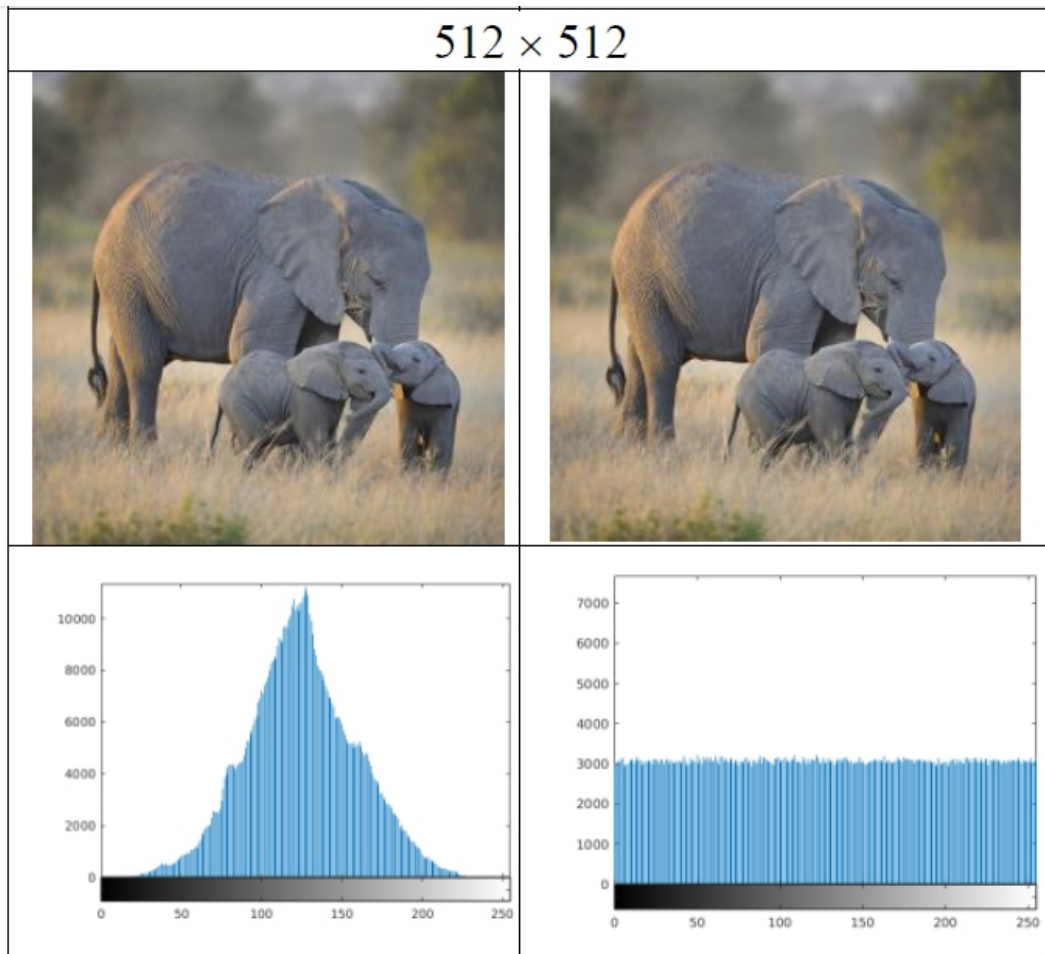
$$\text{Enc}_{K,T}(I_{i,j}) = \text{FF1.Encrypt}(K, T+i+h \cdot j, I_{i,j})$$

For this scheme, we propose to use a hash of the image (namely SHA3-256) to use as the initial tweak. Thus slightly different images will provide completely different encrypted outputs. We will denote by  $E_{\text{FF1},K}(I)$  the encryption with this technique of the image  $I$ .

#### IV. RESULTS

We have chosen three coloured pictures of sizes  $128 \times 128$ ,  $256 \times 256$ , and  $512 \times 512$ , respectively, to test this scheme (see Fig. 2) but, as commented before, it is generalized for any given size.





**Fig. 2. Histograms of several encrypted and decrypted images with the FF1 scheme.**

### A. Security and performance analysis

We have followed the approach made on [9] and performed these analyses: histogram, NPCR (Net Pixel Change Ratio), key space and a computation complexity analysis.

#### 1) Histogram

For these examples, we have chosen three images of three different sizes. Once encrypted, the image histogram approaches that of a noise image as we see in Fig. 2.

#### 2) Net Pixel Change Ratio

NPCR measures the changes detected in the encrypted image when one pixel in the plain image is changed. Let be  $C = E_{FF1,K}(I)$  the encryption with  $K$  of the original image and  $C' = E_{FF1,K}(I')$  where  $I'$  is the image with only one random pixel changed. Then we define

$$NPCR = 100\% \times (\sum_{i=1}^H \sum_{j=1}^W D(I, j)) / (W \times H),$$

where 0, if  $C_{ij} = C'_{ij}$ ,  $D(i, j) = 1$ , otherwise.

By performing these calculations 1000 times on the first image, we have obtained the results shown in Table I.

| Channel | NPCR         |
|---------|--------------|
| R       | 99.999961121 |
| G       | 99.999961304 |
| B       | 99.999960693 |

**Table I: NPCR Results.**

### 3) Computational cost

Let be I an image with height h and width w. The cost of the encryption/decryption of a single pixel is identical because they perform the same operations:

- Since  $n = 3$ ,  $t = 32$ , we will keep  $b = 2$ ,  $d = 4$ , and P unchanged throughout the whole execution with no cost.
- The loop cost:
- $NUM_{256}(\cdot)$  is calculated twice,  $NUM(\cdot)$  once and  $STR_{m256}(\cdot)$  once too.
- The PRF function requires four rounds since  $P \parallel Q$  is 64 bytes long. So each loop performs 4 AES executions (recall that since  $d = 4$  no AES executions are needed in step 5b).
- One addition modulo either 256 or 2562 if i is even or odd, respectively.

We can optimize the conversion functions NUM and STR being the most expensive operation the AES encryptions that are performed 40 times per pixel. That is,  $40 \cdot w \cdot h$  times per image. We have presented in the Table II few image dimensions. Additionally one SHA3-256 is needed for the image encryption.

| Image size         | AES executions |
|--------------------|----------------|
| $128 \times 128$   | 655360         |
| $256 \times 256$   | 2621440        |
| $512 \times 512$   | 10485760       |
| $2560 \times 1440$ | 147456000      |

**Table II: Number of AES executions by size image.**

### B. Drawbacks

As we have seen in section IV.A, our scheme achieves good scores except for the running time since, for one encryption, we need to compute a huge number of AES executions. We have discarded other approaches like taking several pixels at once because that would create a bigger array of length n. Then, when computing the modular exponentiation the values grow exponentially.



## V. CONCLUSION

We have built an image encryption scheme different from usual approaches by using the tweak option given by the FF1 scheme. This could be applied to a biometric database together with an Identity-Based model as defined in [6], since the size and number of images might be small enough to perform all the needed encryptions in a reasonable time.

Given that the Draft upon which this research is based is being currently discussed by NIST [10], the validity of this proposal is subject to changes if NIST withdraws or modifies the publication.

## REFERENCE

- [1] J. Black and P. Rogaway, "Ciphers with Arbitrary Finite Domains," in *Topics in Cryptology - CT-RSA 2002* (B. Preneel, ed.), pp. 114-130, Springer Berlin, 2002.
- [2] M. Bellare, P. Rogaway, and T. Spies, "The FFX Mode of Operation for Format-Preserving Encryption draft 1.1," 2010.
- [3] M. Dworkin, "SP 800-38G: Recommendation for Block Cipher Modes of Operation: Methods for Format-Preserving Encryption," March 2016.
- [4] V. T. Hoang, S. Tessaro, and N. Trieu, "The Curse of Small Domains: New Attacks on Format-Preserving Encryption," in *Advances in Cryptology - CRYPTO 2018*, pp. 221-251, Springer International Publishing, 2018.
- [5] M. Bellare, V. T. Hoang, and S. Tessaro, "Message-Recovery Attacks on Feistel-Based Format-Preserving Encryption," in *Proceedings of the 2016 ACM SIGSAC Conference on Computer and Communications Security, CCS'16*, (New York, NY, USA), pp. 444-455, ACM, 2016.
- [6] M. Bellare and V. T. Hoang, "Identity-Based Format-Preserving Encryption." *Cryptology ePrint Archive*, Report 2017/877, 2017, <https://eprint.iacr.org/2017/877>.
- [7] J. L. Zheli Liu, Chunfu Jia and X. Cheng, "Format-Preserving Encryption for DateTime," in *2010 IEEE International Conference on Intelligent Computing and Intelligent Systems*, vol. 2, pp. 201-205, Oct 2010.
- [8] A. Lenk, P. Marcus, and I. Pova, "Geo-FPE: Format Preserving Encryption of Geospatial Data for the Internet of Things," pp. 172-175, July 2018.
- [9] M. Kumari, S. Gupta, and P. Sardana, "A survey of image encryption algorithms," *3D Research*, vol. 8, p. 37, Nov 2017.
- [10] M. Dworkin, "Sp 800-38G rev. 1 (DRAFT): Recommendation for Block Cipher Modes of Operation: Methods for Format-Preserving Encryption," February 2019..





# Smart Irrigation using IOT and Virtuino

<sup>1</sup> Sudarshan V, <sup>2</sup> Prashanth E, <sup>3</sup> Chirag A, <sup>4</sup> Jananee V

<sup>1,2,3</sup>UG Student, Department of Computer Science and Engineering, Rajalakshmi Engineering College, Chennai, India

<sup>4</sup>Assistant Professor, Department of Computer Science and Engineering, Rajalakshmi Engineering College, Chennai, India

E-mail: <sup>3</sup>chirag.attivaram@gmail.com

## **ABSTRACT**

*India is a country that depends on agriculture and the climatic conditions that depends on the rainfall and if not leads to the lack of water resources. Water scarcity has become a serious problem due to the increase in the population in our country and that the farmers are facing today as well. The main objective of this paper is to provide a solution to automatic transformation and to reduce manual labor as technology needs to be upgraded and the farmers need to be aware of this situation. The apparatus consists of an audio microcontroller and a few sensors that takes in the input that has been programmed and then the output is seen through the pump.*

**Keywords - Virtuino, Audruino IDE, ESP8266 (Wi-Fi Module), Microcontroller and Motor**

## **I. INTRODUCTION**

Smart Irrigation is a technology which is used to water the plants by scheduling the time or used to water the plants automatically, when the level of water in plants goes down or the moisture gets dried. As Water is the most basic component for all the living organisms, it is very important to reserve them and thus must not be wasted. In this project we have implemented sensors which detects the water level of the plant and moisture of the soil. The sensor is used to water the plants at regular intervals and at the same time it intimates a message to the farmer regarding the content and water level of the plants. It also shows a graphical representation on the increase and decrease of the water level at particular intervals. The main aim of this project is to conserve the water which are taken care by the sensors trying to optimize the use of water for irrigation using sensors. The output of these sensors are fed to the micro-controller. The presence of audrunio and other sensors has made the task easy to rely upon and also devices to measure moisture sensors. The proposed system makes use of audrunio and also the message app to indicate the different levels of the soil.

## II. EASE OF USE

### A. Need for Automated Irrigation

helps in the distribution of the water at regular intervals and in an equal manner so that every part of the plant receives in equal amount. It would also develop the performance of the crop and avoid unnecessary over watering of the crop. The use of atomized irrigation system would also reduce manual labor by just sticking to the message indication and also using of motors.

### B. Proposed System

The major problem that is being faced is the availability of the resources and many places fall into watercrises. In order to solve these problems soil moisture sensor, temperature sensor and the wifi module is being used. They are being connected to the audrino uno based on the different pin configuration. The centrifugal pump also has the function of letting out water to the soil at the mentioned time intervals. .

## III. HARDWARE AND SOFTWARE REQUIREMENT

### 1. Arduino:

Arduino is open source hardware and software development platform. It is used in IoT applications It gives theuser the flexibility to adapt and develop the design for their own project.



### 2. ESP8266 (Wi-Fi Module)

ESP8266 is a low-cost Wi-Fi module. It helps in the keeping track of the data and the uploading of the data to the internet.



### 3. Temperature Sensor

Audruino IDE is the software that is the connecting medium to the board and uploads the code to the board and ensures the functioning.

#### A. Motor Driver Board

A motor converts electrical energy to mechanical energy which requires high power. The microcontroller cannot provide such high power, so the motor driver board acts as an interface between microcontroller and motor.

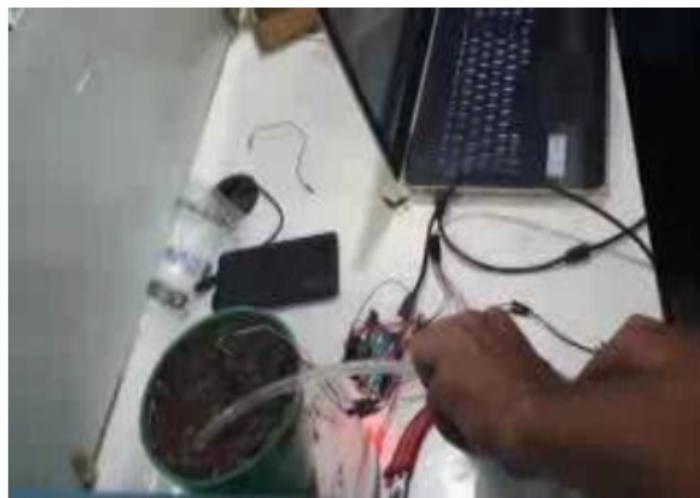


## B. Results and Observation

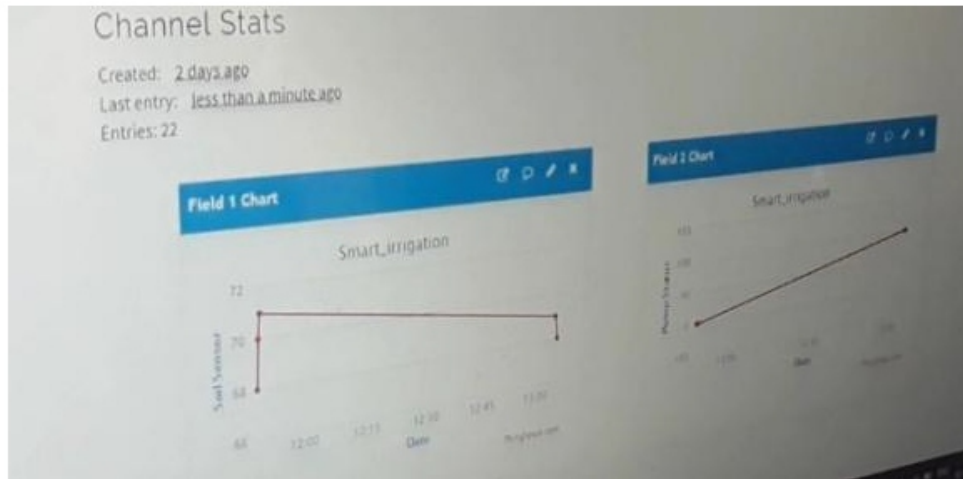
The proposed system is tested in various cases and at most of the cases. The soil moisture test the soil for all climatic conditions. The data in this case is demonstrated using ThingsSpeak. It is by Using this tool we get a graphical view in order to satisfy the different test cases. The data that is observed and then uploaded and then stored.



At first there is the basic functioning of the sensors and then the process begins with first the reading of the value if it is within the threshold value, then the centrifugal pump takes in the water and then passes it to the soil till the required level.



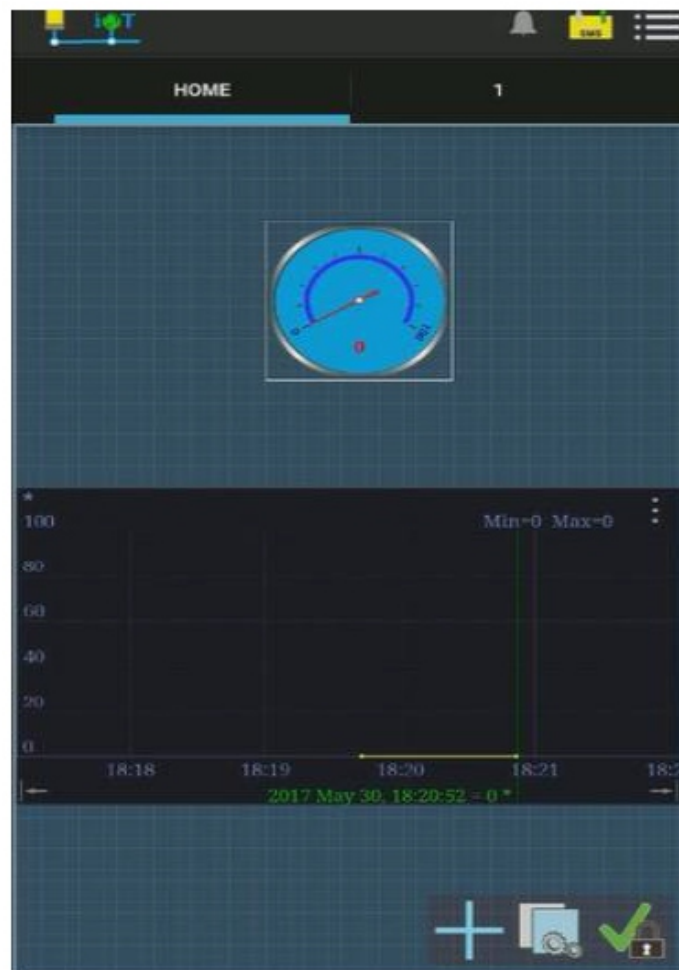
The soil moisture sensor value depends on the value of the resistance of the soil. It has a maximum threshold value of which it can accommodate and after which there would be a few changes that may take place.



This contains the sensor data and the status of the pump at the different intervals as well and this.

#### IV. VIRTUINO APPLICATION FOR INTIMATION

This application helps us to create an interface that helps us to give an intimation to the farmer regarding the level of the soil. It also gives one of the best interfaces and one of the most easy to use. The graphical representation in this application would be more precise and accurate and easy to understand as well.





In this application we follow a sequence of procedures that include a graph that monitors soil moisture data, indicator that triggers alarm, SMS based on moisture level.

## V. VIRTUINO APPLICATION FOR INTIMATION

The implemented smart irrigation system is cost effective and reliable. The automatic watering system reduces the wastage of water. Real-time update to the cloud helps to view the current water condition in the plant. This could be implemented and would also be an effective solution as well.

## REFERENCE

- [1] Hansen, V. E.; Israelson, O. W. ; Stringham, G. E., Book : *Irrigation principles and practices*. 4th edition. 1980 pp.417 pp.Descriptor(s): drainage, irrigation, soil water,Identifier(s) : soil moisture, watering
- [2] Walker, W. R. ; Skogerboe, G. V., Author Affiliation : Utah State Univ., Logan, UT 84322, USA. Book : *Surface irrigation. Theory and practice*. 1987pp.386pp, Descriptor(s): Irrigation, Irrigation systems, models, surface irrigation Identifier(s): watering
- [3] N Gondchawar, RS Kawitkar - *International Journal of International Journal of Advanced Research in Computer and Communication Engineering-2016*
- [4] Bright Keswani Ambarish G. Mohapatra Amarjeet Mohanty Ashish Khanna Joel J. P. C. Rodrigues Deepak Gupta Victor Hugo C. de Albuquerque
- [5] N Suma, SR Samson, S Saranya... *Journal on International Journal on Recent and Innovation Trends in Computing and Communication*
- [6] Ameglio T, Archer P, Cohen M, Valancogne C, Daudet FA, Dayau S, Cruiziat P.1998. Significance and limits in the use of predawn leaf water potential for tree irrigation. *Plant and Soil* 205,155–167.
- [7] Colaizzi PD, Barnes EM, Clarke TR, Choi CY, Waller PM, Haberland J, Kostrzewski M.2003. Water stress detection under high frequency sprinkler irrigation with water deficit index. *Journal of Irrigation and Drainage Engineering* 129,36–42.

# Data Breaches in IOT A Study and Solution by PUF Approach

<sup>1</sup> Prakhar Sahu, <sup>2</sup> S. K. Singh, <sup>3</sup> Pankaj Kumar

<sup>1</sup>Research Scholar, Amity Institute of Information Technology, Amity University Uttar Pradesh, Lucknow, India

<sup>2</sup>Professor, Amity Institute of Information Technology, Amity University Uttar Pradesh, Lucknow, India

<sup>3</sup>Shri Ram Swaroop Memorial College of Engineering and Management, Lucknow, Uttar Pradesh, India

E-mail: <sup>1</sup>prakhar.sahu@student.amity.edu, <sup>2</sup>sksingh1@amity.edu, <sup>3</sup>pk79jan@gmail.com

## ABSTRACT

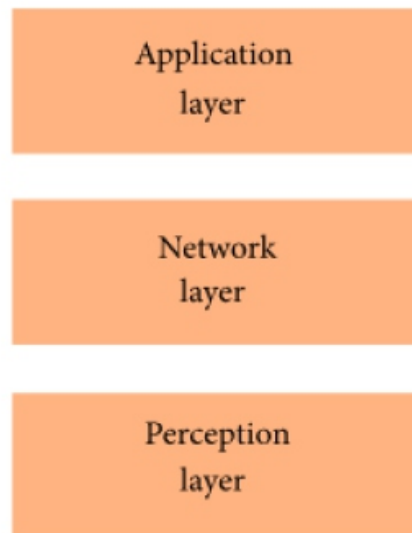
*Internet of things is heralded as one of trending technology of 21st century aims to revolutionize the way we live in many fields like Transportation, Health, and Entertainment etc. There is a quote "With great power comes great responsibility" fits here. Since, IoT devices are equipped with a sophisticated computational power responsible to capture user related data of everyday life. It creates a timeline for a user the way he lives and uses complex algorithms to extract information from the data it captured. This data contains many information like Location, Heart Rate, Sleep Patterns, Power Consumption, Device status and much more; which is very useful for a hacker or a company who is craving for betterment of their services. Since this data is private and bounded by digital laws no one has rights to use it for commercial purposes. A single user data shows information about a user but data of multiple users represents the trends of a particular geological region. This is why it becomes necessary to protect the user data from hackers who are inventing multiple ways to steal it from us, the users. This paper presents a brief overview of security and privacy of Internet of Things as well introduces PUF (Physically Unclonable Function) as a solution to protect user data from cybercriminals and further a suggestive model for IoT authentication using PUF is also demonstrated.*

**Keywords - IoT Security, IoT Privacy Issues, Industrial IoT, PUF, Chip based Authentication.**

## I. INTRODUCTION

The term "Internet of things" was first introduced by Kevin Ashton during the presentation for P&G on RFID based inventory management system in the Year 1999. Internet of Things also refers as IoT are the devices that are now connected to the internet continually collecting and sharing data. Pretty much any physical object can be transformed into an IoT device if it has capability to connect via internet. For example, a lightbulb that can be switched on via smartphone is an IoT device. An IoT device may be cute as a toy or intelligent as an autonomous car as well as complicated as a jet engine equipped with so many sensors responsibly sharing data and collecting information to ensure that it is functioning properly The structure of IoT is divided into three layers namely Perception Layer (Sensing), Network Layer (Data transfer) and the Application layer (Data storage and manipulation). In the perception layer, we meet technologies like Wireless Sensor Node, RFIDs and now near field communication or NFC. In general, a

WSN consist of a low power Microcontroller unit attached to a single or multiple sensor modules supporting low power wireless communication technology. In the second layer (Network layer), the sensor collected data/information is forwarded towards a remote infrastructure for storage, analysis, processing to extract valuable information. In the third layer (Application layer), has an important role to play. Applications layer is responsible for identifying and controlling IoT devices present in the network. It is also responsible for delivering applications specific services to the users and defines various applications in which Internet of Things may be deployed[1, 2, 3].



**Figure 1: IoT Layers**

The "Internet of Things" brings up a next step in the progress of smart objects-interconnected things in which the line between the physical object and digital information about that object is blurred. A technology that started with the development of high-tech inventory management tags is now evolved and extended to include the collection and monitoring of an object's environment. In the current IoT vision, sensors give more and more things in our daily lives a "voice"; by capturing data, sensors enable things to become context-aware, providing more experiential information to help people and machines make relevant and valuable decisions[4]. This paper outlines the privacy challenges that arises during authentication and a suggestive model to overcome this challenge is proposed.

## **II. LITERATURE REVIEW**

In the era of the internet, Data is a key ingredient that plays a decisive role in the IT industry to get insights and to know recent trends for a user. For data security, PUF creates a barricade for cybercriminals. In[5], the author demonstrated multiple offline and online PUF for application in various fields with brief overview. The author of[6]proposed 5 lightweight low-cost authentication protocols for RFID based data exchanges. The author of[7]elaborated the working of PUF with IoT



environments for authentication further explained the possible attacks on PUFs. The authors of [8] used APUF for their work and explained other related technologies. The authors of [9] used a combination of PUF and MAC and proposed a lightweight protocol for authentication.

**1)IoT Applications:** Internet of Things (IoT) also called as Internet of Everything or Industrial Internet is a new technology term aimed to be a global network devices and machines capable to interact with each other without any human assistance or interaction. IoT is one of the most important area of future technology seeking attention from various industries, its value will be recognized once all the devices will be capable of interacting each other using sensors and analytics [10]. Few of the most common areas where the implementation of IoT will be fruitful is as follows:

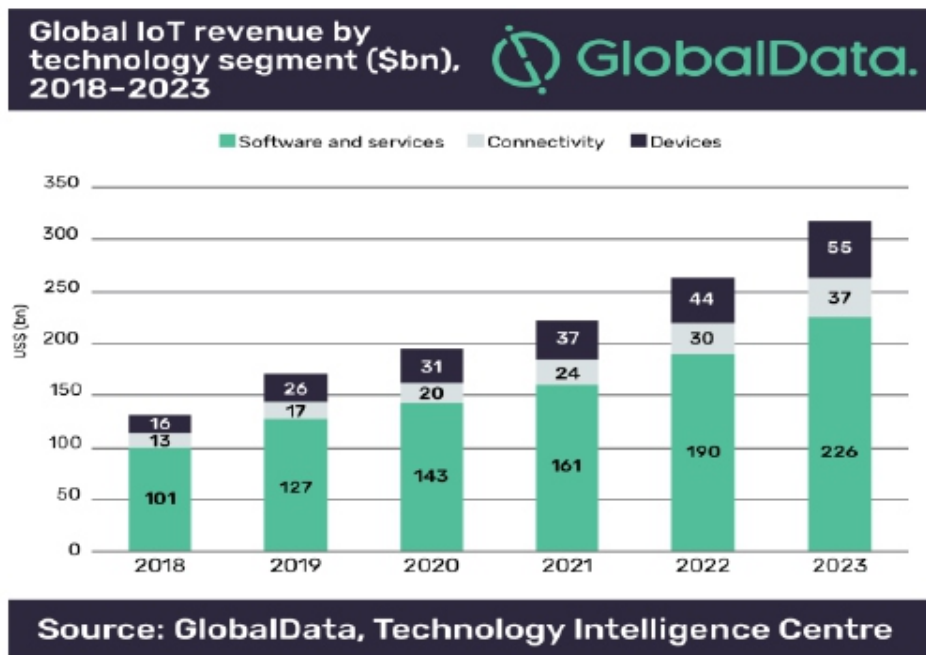
**2)Health Industries:** IoT technologies are very helpful in health industries. Implementation of IoT in health care industry will revolutionize the entire process of patient diagnosis. IoT will open more possibilities using body sensor nodes where the entire body will be diagnosed and help taking proper measures to ensure patient's safety. Recently Phillips Healthcare introduced a connected "Phillips Patient Monitoring" solution for the patients where a virtual interface is established to monitor patient's activities using smartphone and other wearable gadgets. It is a customizable IT solution which automatically connects with the supervisor according to patient requirements[11]. Scientists are also working on remote based surgery using 5G technologies and BSN where the patient will be attended and treated via robotic arms aims to provide intensive care to the patients in need in the absence of a qualified doctor at a particular remote area which has its own geographic limitations[12, 13].

**3)Inventory Management:** In Goods Supply Chain Management and Inventory management, Industrial internet of things plays an influential role which helps reducing cost at an initial level and it also improves the service quality of an organization. It's very hard to operate manually as it is low efficient and time consuming and it will cause huge losses if an error occurs. In general, Inventory Management usually refers by keep a record of every movement of Goods, Cargo and other type of supply related data to streamline the management in a warehouse or library[14].

**4)Agriculture:** Agriculture plays a vital role in the growth of any country's economy. It is the main source of Income in many developing countries and major part of the population relies on that. However not every country is full of natural resources to grow crops and other raw material due suitable environment available for them. In precision farming experts make sure that a perfect farming situation is met; however a proper infrastructure is required to meet that condition so that a particular crop can grow. Using IoT, a precision farming can do a lot more than providing a suitable environment to grow a

certain crops. Once implemented, IoT can maintain a record of many things like temperature, quality control, power management, resource management, inventory management, analytics and much more[15].

**5)Other Applications:** According to GlobalData, It is projected that the IoT industry will touch \$318 Billion mark by 2023 with annual growth rate (CAGR) of 20%[15]. Apart from above major application, IoT can also be implemented in Industrial Automation, Education Transformation, Home automation, Power management, Autonomous driving, Smart cities, Smart grids and many more. Its application is seamless and once implemented, it will be very useful to get insights and feedback for precise decision making.



**Figure 2: Global IoT Revenue Survey**

### III. PRIVACY CONCERNS

Since the amount of connected devices are growing rapidly, it is quite obvious to be aware and worried about the privacy concerns. IoT devices are continuously collecting data from us every day, creating logs and other entries in the cloud servers. The security of IoT is at risk as they use low power architecture and protocols due to limited computational power and other hardware related capabilities. This is why IoT devices are still a soft target for a hacker or other cyber criminals to sneak around. More information about privacy concerns is given below.

**1)Data capture and Preservation:** Once the data is captured, it goes to the cloud server or a local storage for processing to extract useful insights. When a large amount of data is available, it required to

be processed through data processing methods. These streaming data are about location, movement, vibration, temperature, humidity and even chemical changes in the air. Data mining tools can immediately invite corrective processes to address operational issues or inform competitors about strategic moves and searches about customer preference changes which can contribute to their short-term and long-term business activities[10].

**2)Authentication:** Authentication is the process of identifying users and devices in the network and providing access to authorized individuals and non-manned devices. Authentication is a way to reduce attacks on IoT systems such as Reply Attack, Man-in-the-Middle Attack, Impersonation Attack and Sybil Attack. Since users and devices create two-way communication in an IoT environment, there is a mutual communication between the device and the server. The device will send data to the server and receive control data transmitted by the server. That's why a mutual authentication is important in the IoT system to check the legitimacy of both the device and the server[16].

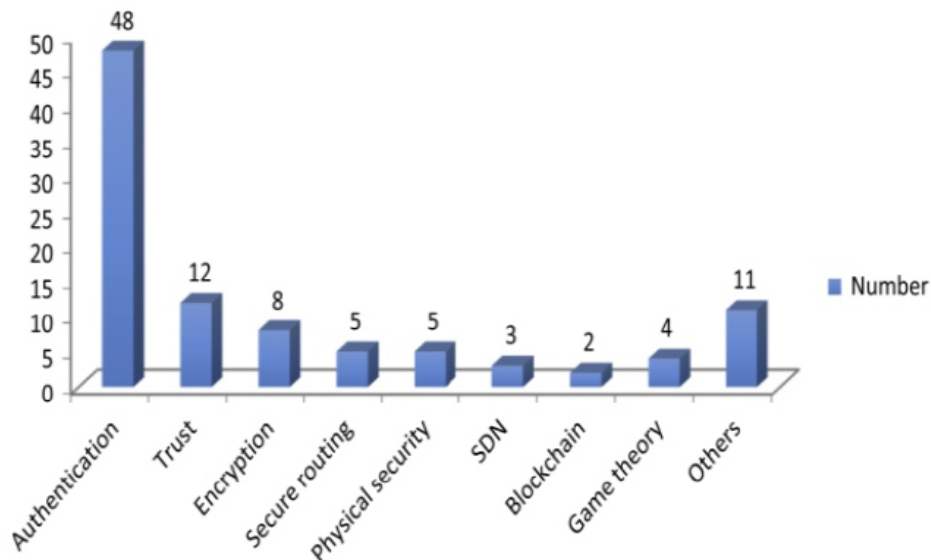
**3)Targeted Ads:** There are millions of smart vehicles roaming around us carrying multiple passengers at a time. Today, automobiles are not just a way to transfer people from Place A to Place B; they are connected and equipped with a complex computational system opening seamless possibilities to provide an environment required to invent, develop and replace existing services for a better future. Particularly, a passenger on board becomes a POI (Person of Interest) for ads agencies to gear up their businesses[17].

**4)Other Loopholes:** Other possible loopholes with IoT are 1. Lack of strict privacy laws, 2. Hardware and software upgrade challenges, 3. Location tracking, Profiling, 4. Unauthorized or Unwanted access, 5. Break in attacks, 6. Absence of strong encryptions etc.

#### **IV. POSSIBLE ATTACK POINTS OR DATA BREACH LAYERS OF IOT**

Internet of things is changing the way we are living, starting the trend of creating a connected atmosphere where every device is talking to each other. We are entering into an environment where everything or almost everything is connected to the internet. This is an interesting area to explore for Industries. Factories are using smart technologies that relies on sensor nodes to keep a record of machine performance. Cities are transforming into smart cities where everything from computing energy consumption of a street light to monitoring traffic is automated. Now, hospitals can monitor the health of high-risk patients remotely. This is a perfect environment for a hacker to invade our privacy and take control of connected devices. In December, 2015 a Ukrainian power grid was targeted by the hackers leaving people of a city in dark for 1 to 6 hours by just spamming power grid employees with infected attachments. It shows that Internet of Things are not fully secured and have some weak points that need to identify[18].

**1) Authentication and validation:** In IOT environment, there is a two way connection established between device and the servers. A device sends data to the server and receives control commands. Authentication is a stage where validity of the server and the device is checked. There is a huge demand of a light weight authentication and encryption technology in IoT aiming to provide a secure environment for communication and access control.



**Figure 3: Publication in IoT security from 2016 to 2018 in Elsevier, IEEE, Hindawi, Springer until June 2018.**

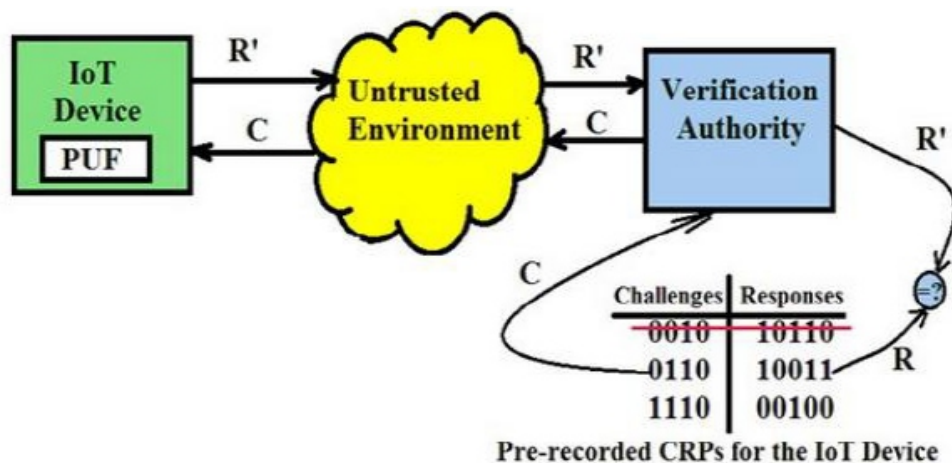
**2) Encryption:** To transmit data between two or more connected device, an encryption is needed. Due to limited development, mobility and low complexity, the cloud of things is hard to control. To ensure confidentiality, there are numerous encryption technologies are present. Although, the challenge is to make fast and low power consuming algorithms. Apart from that, an efficient key distribution system should be used for encryption scheme [18].

**3) Physical Security:** To understand the physical security of an IoT device, we need to think about a situation where connected street lights uses AES encryption to deliver firmware updates. The process is to ensure that only authorized sources can deliver the updates in a secure manner. Now, a hacker can steal these credentials using side channel analysis (SCA) and later on use it to hijack the streetlights network. This attack was originally demonstrated. To attack and get credentials, a hacker need to be inside the coverage of the device. There will be no traces of SCA attack afterwards. These types of attacks may be more dangerous when SCA is used to compromise the bank cards, mobile devices or even medical equipment. Physical security is not a new concept, we have seen it in our daily lives from and the threat exists from many years. However, the risks are increasing with the use of IoT devices [19].

**4) Routing:** Secure routing is an essential part of a sensor network. Every sensor node depends upon a routing protocol to deliver and execute data between server and a device. If the routing protocol is weak, it may cause serious damage to the privacy of a user and leave traces for a hacker to get inside the network. If a hacker can capture packets from the network, he can see what commands are being sent and may release his own commands to manipulate a sensor or servers[20].

## V. PHYSICALLY UNCLONABLE FUNCTIONS AS A SOLUTION

There are millions of devices connected to the Internet communicating through networks. In IoT, security and reliability of data has a significant importance; to ensure that, backend infrastructure should be safe and verified. Since IoT is mobile, its hardware is highly vulnerable to many security threats like spams, device thefts, piracy, cloning, Intellectual Property thefts etc. that are increasing with rapid growth of IoT technologies becoming a most challenging issue to be resolved. It further affects the trustworthiness of IoT technologies and unwillingly causing monetary losses to the makers and the users both. To overcome this situation, PUF is presented as a reliable privacy tool for authentication and key storage. PUF is a chip level verification technology that doesn't provide a dedicated identification and authentication environment. The concept of PUF was introduced in the year 2001 as a one-way hardware-based cryptographic tool. PUF is usually embedded into Integrated Circuits to meet high security requirements at cryptographic level which is almost impossible to clone, replicate or predict even the exact manufacturing process is used. PUF works on a challenge-response pair instead of key-based authentication that is linked to the device's identity.

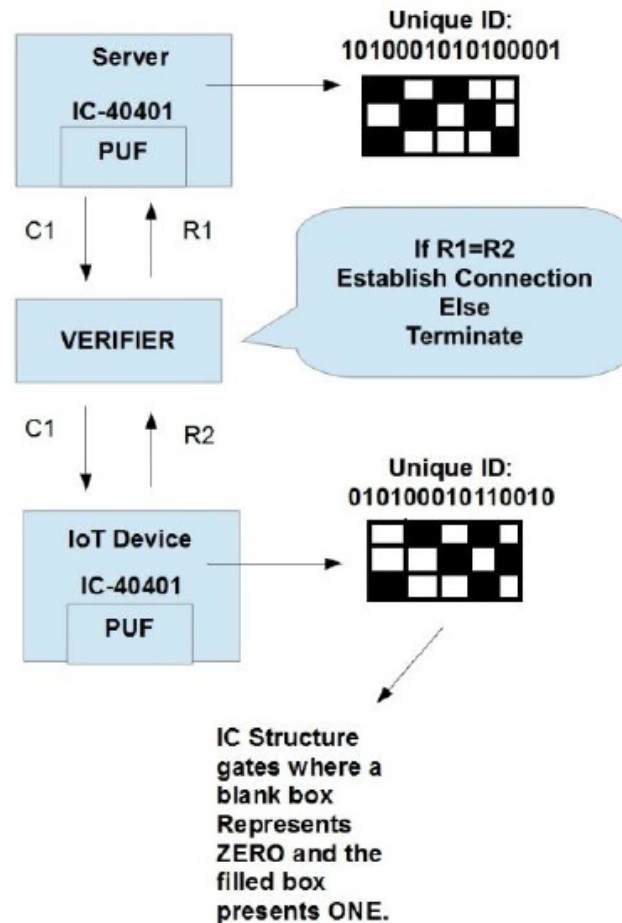


**Figure 4: Physical Unclonable Functions**

Weak PUF is limited to produce a few; in most cases, the electrical stimulus, called challenge, is applied on the physical structure to react as a response. This is a capricious process due to the complex interaction of stimulus with the physical microstructure of a device. PUFs usually have very low hardware overhead, so they are a key authentication ingredient for IoT's. PUF is already applied on devices for identification and

authentication, hardware-software binding, key storage, keyless secure communication etc. PUFs have two types, either it is strong or weak. While a strong PUF is capable of producing vast amount of challenge and response pairs but cases, only one[21, 22, 23].

### Suggestive Model:



**Figure 5: An initial overview of PUF implementation in IoT environment.**

The above model exhibits the initial overview of PUF implementation on IoT infrastructure for device authentication where PUF creates a challenge and response pair for a device to authenticate a new device. The pair of Challenge and Response is shared with the verifier and Challenge is further forwarded to the IoT node, the node replies back. If the response is correct, IoT device will be authenticated otherwise the connection will be terminated by the verifier. This model also tried to show the IC gates setup to bifurcate an IC structure which is physically uncloneable and unique due to architecture based identification which helps to create random PUF challenges are response pairs for authentication.



| PUF Protocols                          | Strength   | Weakness                               | Compatible to which PUF?                    | FPGA Compatibility         | Further Notice   |
|--|--|--|---|----------------------------|--|
| Early protocols                        | (Partly) Resilience against machine learning attacks | Cryptographic primitives utilization   | Integrated into different PUF architectures | No                         | Occupy considerable amount of computational resources                        |
| Mutual authentication protocol         | Mutual authentication feature                        | Assumption of perfect PUF architecture | Not specified                               | Need further investigation | Implementation overhead has not been reported                                |
| Obfuscated challenge response protocol | Resiliency against machine learning attacks          | Dependency on random number generation | Arbiter PUF                                 | Need further investigation | Computational resource utilization & FPGA compatibility need to be clarified |
| Lockdown protocol                      | Resiliency against machine learning attacks          | Need further investigation             | XOR Arbiter PUF                             | Need further investigation | Computational resource utilization needs to be clarified                     |

**Figure 6: Physical Unclonable Functions**

## VI. FUTURE WORK & CONCLUSION

As the use of IoT in our life is increasing, the security and privacy loopholes of IoT need to be reviewed to overcome the privacy and data theft related challenges. In 21st century, user data is just like a fuel, a single data represents an individual user but a large amount of data can show you a bigger picture. It can be used for you are against you to manipulate your thoughts, lifestyle and also to gain profits from you by targeting a user or a group of users. In this paper, the suggestive model demonstrates the PUF implementation on an IoT architecture for authentication. PUF can be a sophisticated solution for authentication where a device is only authenticated and allowed in the network if the challenge given by PUFs are satisfied. PUF ensures that only a genuine IoT node is allowed for access and the rest is ignored, however the protocols are required to be updated time by time as the attack methods are evolving drastically. If PUF methods are not updated, it will become vulnerable for a hacker to get access. We also believe that some work should be done to improve the PUF optimization as the limited computational power may cause latency. Other researchers also emphasize improving the efficiency of PUFs by Creating encryption keys, better random numbers or electronic signatures.

---

**REFERENCE**

- [1] Tzounis, N. Katsoulas, T. Bartzanas and C. Kitta, "Internet of Things in agriculture, recent advances and future challenges," *Biosystems Engineering*, vol. 164, pp. 31-48, 2017.
- [2] S. Ranger, "What is the IoT? Everything you need to know about the Internet of Things right now," *ZDNet*, February 2019. [Online]. Available: <https://www.zdnet.com/article/what-is-the-internet-of-things-everything-you-need-to-know-about-the-iot-right-now/>. [Accessed 17 August 2019].
- [3] J. Gubbi, R. Buyya, S. Marusic and M. Palaniswami, "Internet of Things (IoT): A vision, architectural elements, and future directions," *Future generation computer systems*, vol. 29, no. 7, pp. 1645-1660, 2013.
- [4] S. Burt, "IoT in Education: Landscape Review," 23 March 2018. [Online]. Available: <https://www.clarity-innovations.com/publications/iot-education-landscape-review>. [Accessed 17 August 2019].
- [5] R. Maes and I. Verbauwhede, *Physically Unclonable Functions: A Study on the State of the Art and Future Research Directions*, Berlin, Heidelberg: Springer, 2010, pp. 3-33.
- [6] Vajda and L. Buttyan, "Lightweight authentication protocols for low-cost RFID tags," *Security in Ubiquitous Computing—Ubicomp*, pp. 1-10, 2003.
- [7] D. Mukhopadhyay, "PUFs as promising tools for security in Internet of things," *IEEE Design & Test*, vol. 33, no. 3, pp. 103-115, 2016.
- [8] Y. Gao, L. Gefei, M. Hua, S. F. Al-Sarawi, O. Kavehei, D. Abbott and D. C. Ranasinghe, "Obfuscated challenge-response: A secure lightweight authentication mechanism for PUF-based pervasive devices," in *IEEE*, Sydney, NSW, 2016.
- [9] M. N. Aman, K. C. Chua and B. Sikdar, *Position paper: Physical unclonable functions for iot security*, ACM New York, NY, USA: *Proceedings of the 2nd ACM international workshop on IoT privacy, trust, and security*, 2016.
- [10] Lee and K. Lee, "The Internet of Things (IoT): Applications, investments, and challenges for enterprises," *Business Horizons*, vol. 58, no. 4, pp. 431-440, 2015.
- [11] Philips Healthcare, "Patient Monitoring Solutions: Philips Healthcare," *Philips Healthcare*, [Online]. Available: <https://www.philips.co.in/healthcare/solutions/patient-monitoring>. [Accessed 17 August 2019].
- [12] H. Laaki, Y. Miche and K. Tammi, "Prototyping a Digital Twin for Real Time Remote Control Over Mobile Networks: Application of Remote Surgery," *IEEE Access*, vol. 7, pp. 20325-20336, 2019.
- [13] P. Gope and T. Hwang, "BSN-Care: A Secure IoT-Based Modern Healthcare System Using Body Sensor Network," *IEEE Sensors Journal*, vol. 16, no. 5, p. 1368-1376, 2016.
- [14] W. Ding, *Study of smart warehouse management system based on the IOT*. In *Intelligence computation and evolutionary computation*, Berlin, Heidelberg: Springer, 2013, pp. 203-207.
- [15] M. Froese, "Global IoT market to reach \$318 billion by 2023, says GlobalData," *Windpower Engineering and Development*, 19 November 2018. [Online]. Available: <https://www.windpowerengineering.com/business-news-projects/global-iot-market-to-reach-318-billion-by-2023-says-globaldata/>. [Accessed 19 August 2019].
- [16] M. Binti, M. Noor and W. H. Hassan, "Current research on Internet of Things (IoT) security: A survey," *Computer Networks*, vol. 148, pp. 283-294, 2019.
- [17] H. Aksu, L. Babun, M. Conti, G. Tolomei and A. S. Uluagac, "Advertising in the iot era: Vision and challenges," *IEEE Communications Magazine*, vol. 56, no. 11, pp. 138-144, 2018.
- [18] S. Price, "These Are The Weakest Points in Your IoT Security," *IOT Central*, 11 May 2016. [Online]. Available: <https://www.iotcentral.io/blog/these-are-the-weakest-points-in-your-iot-security>. [Accessed 18 August 2019].
- [19] K. Marneweck, "The role of physical security in IoT," 14 March 2019. [Online]. Available: <https://community.arm.com/iot/b/blog/posts/the-role-of-physical-security-in-iot>. [Accessed 18 August 2019].
- [20] R. Ahmed, A. Matrawy and C. H. Lung, "Secure Routing in IoT: Evaluation of RPL Secure Mode under Attacks," 2019.
- [21] H. Wang, Y. Wang, T. Taleb and X. Jiang, "Special issue on security and privacy in network computing," *World Wide Web*, pp. 1-7, 2019.
- [22] Braeken, "(2018). PUF based authentication protocol for IoT," *Symmetry*, vol. 10, no. 8, p. 352, 2018.
- [23] Babaei and G. Schiele, "Physical Unclonable Functions in the Internet of Things: State of the Art and Open Challenges," *sensors*, vol. 19, no. 14, pp. 3208-3226, 2019.



## Recommendation Model for Cloud Service Selection using Fuzzy Logic

<sup>1</sup> Parthasarathy. M, <sup>2</sup> Aniruth. R, <sup>3</sup> Gokulakrishnan. R, <sup>4</sup> Krishnan. V. C, <sup>5</sup> Sundaraganapathy. L, <sup>6</sup> Subha. T.

<sup>1,2,3,4,5</sup>UG Student, Department of Information Technology, Sri Sai Ram Engineering College, Chennai.

<sup>6</sup>-Associate Professor, Department of Information Technology, Sri Sai Ram Engineering College, Chennai.

E-mail: <sup>1</sup>mvgopesh@gmail.com, <sup>2</sup>anikrishna3@gmail.com, <sup>3</sup>gokulshyam567@gmail.com, <sup>4</sup>krishnan.vcg@gmail.com, <sup>5</sup>techiinet@gmail.com, <sup>6</sup>subha.it@gmail.com

### ABSTRACT

*The Cloud computing is an enormously growing technology which offers storage, computation and various network resources as a service across the internet by the cloud service providers (CSPs). The users need to pay for what they use. The users need to be careful while finding and selecting the optimal service that suits their needs. The earlier retrieval systems and the recommendation approaches are dealt with fundamental service selections. Fuzzy logic has been widely adopted in many applications like information retrieval, control systems and in decision making. Web ontology language plays an important role in specifying user profiles and services for finding accurate recommendation of information or services. In this paper we propose a recommendation system for selecting the better service using fuzziness concept integrated with web ontology to reduce the user's burden.*

**Keywords - Cloud Computing, Fuzzy Logic, Web Ontology, Recommendation, Service, Storage, User**

### I. INTRODUCTION

By the definition of the National Institute of Standards and Technology [8], a simple way to get a common resource pool is through the service model called cloud computing. On-request access is the key behind this innovation and is utilized to register resource including systems, servers, stockpiling, applications, and so forth when contrasted with the other conventional models, distributed computing model can give boundless figuring assets on interest, in very compelling processing and capacity limit, brisk reaction to the flood of requests, simple extension of resource on interest, and charge dependent on the use. The net outcome is the expense of service arrangement can without much of a stretch be decreased, and the adaptability and versatility of framework can likewise be expanded. Real IT organizations have made their own cloud computing arrangements and cloud applications have grown up quickly in numerous regions.

In a cloud computing model, cloud based clients can acquire their needs as indicated by prerequisites without knowing the host and their provision. The expanding number of IT merchants is sent to the grid and is racked in space. The application offers stockpiling and calculation of assets as cloud host. As a final product, countless services are made available and accessible for the clients [1] for browsing.

Naturally it's a challenge for the users to select the apt provisions that meet their Quality of Service (QoS) needs in the service cycle. It starts from selection along with deployment to orchestration e.g. determining an optimal web service when making service selection identifying suitable virtual machine VM servers for deploying web service instances etc. [2] effective service recommendation techniques play an important role in helping users including developers in their decision-making processes for critical application developments and deployments [3], those applications include interactive games, real-time, social networks, data analytics, scientific computing, business, the internet of things (IoT) and other mobile apps as follows. The basic needs and requirement of apps are different cloud service consumers (CSC) along these lines need to profoundly locate the ideal provisions by finding enormous number of details about the provisions qualities and their properties Service Level Agreements (SLA) and so on. Besides in regards to the characteristics functionalities adaptability and interoperability of the provisions existing cloud service providers (CSP) offer a wide assortment of interfaces principles approaches and Service Level Agreement (SLA) parameters which prompts a great deal of challenges in retrieval of information about service understanding and investigation.

Subsequently, they are forced to prompt the needs and the extraordinary difficulties in the determination and recovery of cloud services through a perfect cloud service recommendation framework that is sought after for an assortment of CSC.

Online supportive networks like recommendation techniques, recommendation frameworks [3], [4], effectively propose a set of things that are ranked and constrained from every single accessible thing without the contribution of clients. They deal with the issues made by the alleged "information explosion" in an range of online applications like web based business [5], e-learning [7], and e-tourism, and in regions like the suggestion of news, motion pictures, books, recordings, assets, and land [9]. Before making a suggestion, background information is being utilized by recommendation frameworks, for example, historic information comprising of evaluations from clients, and input information, for instance characteristics of an item or rating given by clients, to start a suggestion; models and algorithm contribute in consolidating the two and creates a suggestion[9], [10].

In real situations, item features and user behavior are often impressionistic, vague, and not accurate [8] and item priorities of the users are frequently subjective and uncertain. It is been very difficult for a user to express his/her interest in an item with exact numbers. Fuzzy set theory effectively handles the fuzziness and uncertain issues in recommendation problems. Priorities of the users and item features have been illustrated as fuzzy sets in previous research [11]–[15], and recommendations to customers for the selection of suitable items are made with incomplete and uncertain information [15], [16].

Present research and recommendation system applications mainly focus on making recommendations to personal users. Clients are furnished with frameworks that find suggestions and along these lines customize the internet experiences of the client hence giving the capacity to find their activities. The fundamental goal of any recommendation framework is to direct the clients to discover needed data in straightforward possible path through helping them to find pages which they may not discover during the site visit. Extra related data, which clients may discover significant yet not contained on that specific site, can likewise be given. For such frameworks to be deployed, usage information must be gathered and analyzed. Subsequently, adjustment of sites strategically could be executed for upgraded and improved customer experience.

Generally, suggestion module depend on the general user details in the webpage and in this way it ends up straightforward for the end-clients or being customized with the intention to alter the web cognizance for a specific client. Suggestion module are the projects that endeavor to prescribe things to clients by foreseeing the interests of the clients in such a way that the given item is dependent on various kinds of details, including explicit insights concerning things, clients, and the communications among clients and items. The fundamental thought behind the suggestion module is that like-minded clients like same items.

## II. RELATED WORK

Ontology is a formal description of concepts that are often conceived as a set of, instances, functions, entities, properties and axioms. The Web Ontology Language (OWL) in this sense defines and instantiates ontologies in a manner that allows web agents to interpret and exchange information based on a common sense vocabulary [17]. A two layer hierarchical ontology model is expressed by them:

- A general upper ontology representing general concepts and ontological classes in smart spaces. It represents context as ontology instances with associated properties. We refer to this combination as context markups.
- A domain-specific ontology that represents contextual details about people, objects and items existing in smart homes.

Cognitive reasoning refers to combining the objective and subjective aspects of context to produce a recommendation list fitted to the situation.

Konstantinou et al. proposed an approach to plan, model, and deploy Cloud service compositions. In their approach, the deployment plan and the solution model for the composition in Cloud platform are developed by skilled users and executed by unskilled users. Likewise, in our system, a set of

compatibility constraints from experts were captured which would be utilized to simplify the process of deployment for end users by eliminating invalid compositions solutions. But it lacks an approach for appliance selection and their placement on the Cloud which is offered by our work.

Similarly Chieu et al. proposed the use of composite appliances to automate the deployment of integrated solutions. However in their work, QoS objectives are not considered. The other work involves utilization of Intuitionist Intuitionist Fuzzy Set (IFS) is used in other works. They are used for ranking service compositions in the context of Grid and SOA . Users' constraints such as compatibility is not dealt and the execution time is not acceptable when the problem is NP-hard (like our service composition problem. Furthermore,our approach improves the composition solution diversity and convergence and decreases the execution time in comparison with the work which considered evolutionary approach such as NSGA-II for service composition .

Unified Cloud Interface (UCI) gives ontology<sup>4</sup> model to construct Amazon EC2 services. Mosaic undertaking is proposed to create multi-Cloud oriented applications. In it, Cloud ontology is essential, and indicates the necessities of the applications for resources of the regarding SLAs and QoS prerequisites. It is being used to give a typical access to the Cloud benefits in Cloud federations. However, these ontologies do not keep focus on modeling of the compatibility of Cloud services. There are many existing methodologies that are fit for taking care of the incompatibilities. Be that as it may, a large portion of them just determined the similarity of Input and Output (I/O) of administrations and did not consider the contradictions that are brought about by strategies and different variables that are not identified with service functionalities.

While there are other studies that keeps focus on selection of appliance and deployment problem, we are not aware of any work that provides a mechanism for deploying and composing multiple virtual appliances with focus on QoS aware ranking and automatic compatibility checking.

CoCoOn is an infrastructure service model which comprises both non-functional and functional specifications of cloud virtual machine (VM) and storage resource aspects; it still doesn't include service information across wider Resource abstraction levels. Although Cloud Ontology is able to express the provisions of a variety of cloud services, it only discloses some basic elements regarding the diverse provision levels and functions. In fact, for models that already exists, the cloud computing and services concept specifications are maintained evenly across service function categories and multiple abstraction levels.

Indeed, except mosaic, very less ontology touch the explicit details of both non-functional and functional properties of diverse cloud provision types. Besides, none of the other ontology methods attempt to express the aspects of service agility or the appropriate specifications through fuzzy extensions; none of the current practices supports collaborative model editing for the field.

Existing discovery systems, tools and service recommendations are seen limited in terms of their overall, flexibility, applicability and comprehensiveness. Some [11] are found to be focusing on IaaS centric service recommendation. Specifically, CSDS presents a system of discovery for VM services according to search parameters such as frequency, memory and storage size, virtual CPU architecture, network parameter, operating system(OS), etc. Cloud Recommender [11] offers improved functions which accept both non-functional and functional service properties as recommendation requirements.

### III. PROPOSED RECOMMENDATION MODEL

#### 3.1 Ontology design

In this framework the primary cloud administrations are represented as a class esteem relationship. In this model the various provisions of cloud are declared less than one of the specific classes. It keeps up relations among the sub children of its primary class, for example, characteristic, service and features. The ontology structures of these qualities are appeared in Fig.1. The qualities, for example, unwavering quality, adaptability, security, accessibility (sub substance of administration attributes) observing, relocation, backing of programming language (sub element of service features) are given as information properties. The choice of provision is affirmed as a service work and indicates the distinctive cloud service models and the sending models as article properties. It very well may be deciphered as service is conveyed as IaaS; administration is sent as private cloud; and so forth.

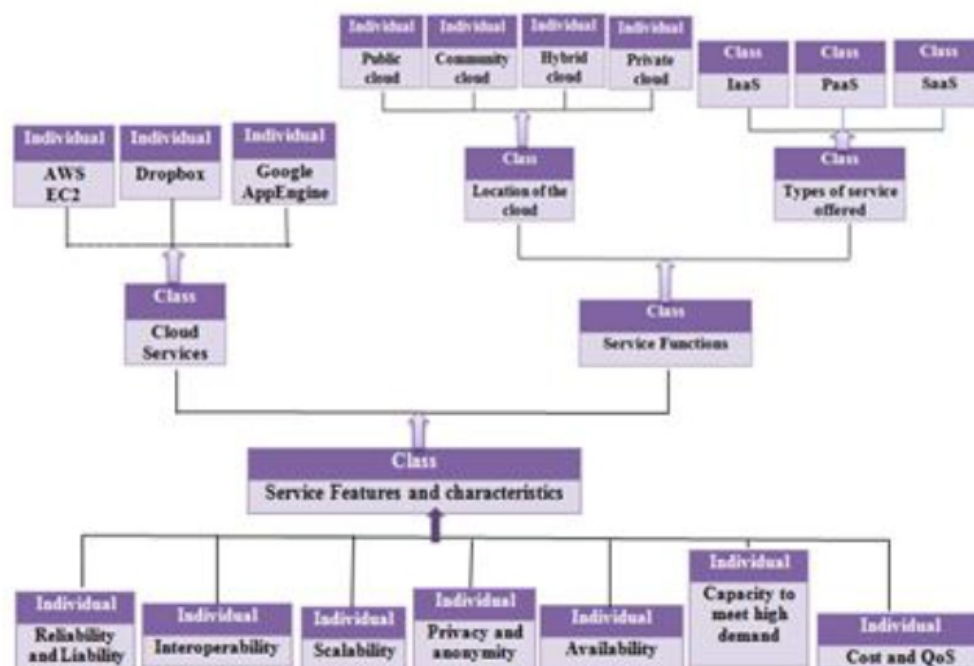


Fig.1. Ontology Design of services

According to the cloud administrations are coordinated in nature which isn't considered in the past work. Models are adjustment of administration, provisioning of assets, interoperability, observing and arrangement support. The diverse cloud administrations, for example, IaaS, SaaS and PaaS are connected with one another here and there to give an administration. In this way, readiness attempts to fuse administration qualities, highlights and capacities in a practical and non-practical way.

### 3.2 Integration of Fuzzy Logic

Difficulties in management of cloud service from client's perspective are talked about and an answer for moderating this issue has been inferred utilizing fuzzy logic coordinated with ontology design. First case manages clients who are clueless to cloud provision and in a spot to choose a provision. This calculation encourages them to choose a cloud specialist organization which gives high caliber of service, dependent on the trust level appraised by clients. Second case manages intermediate clients who have officially experienced cloud however not with various cloud specialist organizations. These middle level clients are clients who think well about service features and attributes. They are in a spot to relocate the features by starting with one specialistco-op to the next dependent one that is on their provision checking (External – input from specialists, Internal – their very own understanding). Third case manages clients who have encountered provisions from various specialist organizations and have recommendations to choose great quality specialist co-op. In view of the provision monitored, the choice is taken whether to proceed with a similar provision or to relocate among different specialist organizations. Fig 2 Speak to ontology development dependent on class object properties.

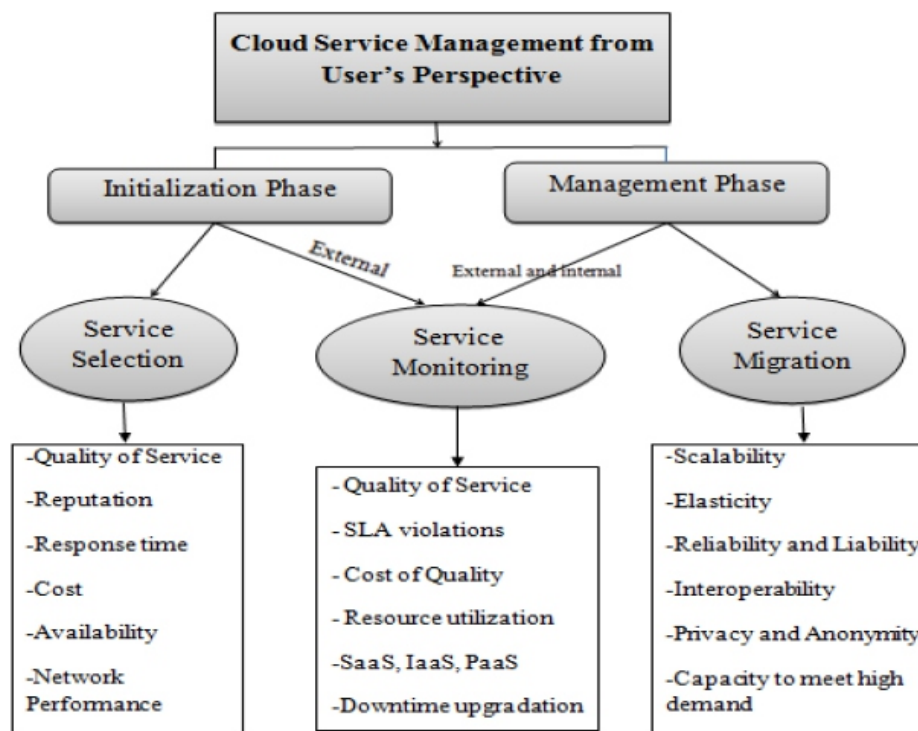


Fig.2. Class object property ontology construction



### 3.2.1 Fuzzy cloud service

This algorithm will assist the client with selecting the best selection of administrations dependent on the evaluations given by various degrees of client. We propose an answer dependent on fuzzy logic coordinated with ontology. Fuzzy logic allocates truth esteems for the variable which range somewhere in the range of 0 and 1. The numbers are genuine and are considered as fuzzy weight.

### 3.2.2 Fuzziness implementation

Assign fuzzy weights to all the parameters based on its priority. Let fuzzy weight of a parameter be „W“, trust rated by customers be „T“ and number of customers who rated the service be „N“

Set minimum value called trust to rate good Quality of Service provided by the company. Get trust values for the cloud service from the customers.

Let new-user be „U“, Intermediate (the one who using the cloud currently) be „I“ and the Expert (who used more than one service) be „E“. Now Calculate the Quality of Service using the formula given in equation 9.

$$\text{FUZZY}_{QoS} = \frac{(W_k * U_{i-n} * T_{i-n}) + (W_k * I_{i-n} * T_{i-n}) + (W_k * E_{i-n} * T_{i-n})}{(W_k * U_{i-n}) + (W_k * I_{i-n}) + (W_k * E_{i-n})} \quad (9)$$

Calculate trust level rated by new users, intermediate users and expert users are given in table 1.

| Trust level rating by different users |                                |                                |
|---------------------------------------|--------------------------------|--------------------------------|
| New User                              | Intermediate                   | Expert                         |
| $\sum_{i=1}^n W_k * U_i * T_i$        | $\sum_{i=0}^n W_k * I_i * T_i$ | $\sum_{i=0}^n W_k * E_i * T_i$ |

**Table 1. Trust Level Rating of different User**

Where  $I_{i-n}$  denotes the  $i^{\text{th}}$  number of Intermediate user who rated the service from the list of „N“ users, likewise  $E_i$  and  $U_i$  ,,

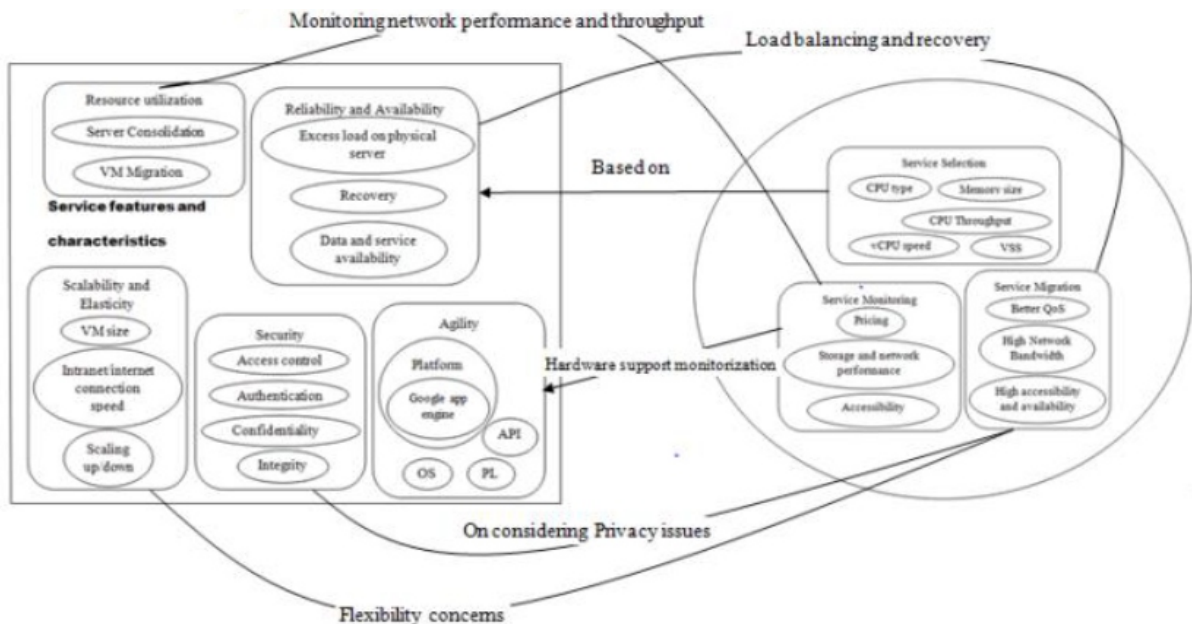
$W_k$ “ denotes the parameter weight (here we considered 7 parameters hence „k“ varies from 1 to 7),  $T^{i-n}$  denotes the trust level rated  $i_{\text{th}}$  customer. The result is taken into consideration for rating the Quality of Service (QoS) being provided by the Cloud Service Management.

**IV. RECOMMENDATION SYSTEM**

The proposed suggestion system is shown in figure 3. This system contains client element, where distinctive degree of client information is acknowledged for computing fuzzy weight. Provision rating enable the clients to rate the provision dependent on their insight, constant utilization of the provision, need and trust. A proposal motor examines every one of the qualities and creates results portraying the provisions, diverse specialist co-ops and their rating. At last, the consequence of the proposed procedure is sent back to the client.



**Fig. 3. Recommendation Model implementation**



**Fig.3. Data property ontology diagram**



## V. CONCLUSION

Thus in this work we have propounded an outcome to play out a suggestion examination to choose the ideal cloud provision. A suggestion framework is created to rate the provision dependent on clients intrigue and the use of provision. It is executed by doling out fuzzy loads, figuring trust esteems for various degrees of clients. Web ontology is incorporated in this work to accomplish an ideal arrangement. Our work results in finding the best and ideal provision.

## REFERENCE

- [1] Burstorm, "Think vertical Layered tech official blog." [Online]. Available:<http://www.layeredtech.com/blog/thing-vertical/>
- [2] M.Zhang, R. Ranjan, A. Haller, D. Georgakopoulos, and P. Strazdins, "Investigating decision support techniques for automating cloud service selection," in *Proc. IEEE 4th Int. Conf. CloudCom*, 2012, pp. 759–764.
- [3] M. Zhang et al., "Investigating techniques for automating the selection of cloud infrastructure services," *Int. J. Next-Gener. Comput.* vol. 4, no. 3, pp. 1–18, 2013.
- [4] G. Adomavicius and A. Tuzhilin, "Toward the next generation of recommender systems: A survey of the state-of-the-art and possible extensions," *IEEE Trans. Knowledge Data Eng.*, vol. 17, no. 6, pp. 734–749, Jun. 2005.
- [5] X.Guo and J. Lu, "Intelligent e-government services with personalized recommendation techniques," in *Proc. Int. J. Intell. Syst.*, 2007, vol. 22, pp. 401–417.
- [6] P. Markellou, I. Mousourouli, S. Sirmakessis, and A. Tsakalidis, "Personalized e-commerce recommendations," in *Proc. IEEE Int. Conf. e-Bus. Eng.*, Beijing, China, 2005, pp. 245–252.
- [7] J.Lu, "personalized e-learning material recommender system," in *proc. 2nd Int. Conf. Inf. Technol. Appl.*, 2004, pp. 374-379.
- [8] P.Mell and T.Grance, "The NIST Definition of Cloud Computing", *NIST Special Publication - 800145*, 2011; <http://csrc.nist.gov/publication/nistpubs/800-145>.
- [9] X. Yuan, J. H. Lee, S. J. Kim, and Y. H. Kim, "Toward a user-oriented recommendation system for real estate websites," *Inf. Syst.*, vol. 38, pp. 231–243, 2013.
- [10] R. Burke, "Hybrid recommender systems: Survey and experiments," *User Modeling User-Adapted Interact.*, vol. 12, pp. 331–370, 2002.
- [11] A. Zenebe and A. F. Norcio, "Representation, similarity measures and aggregation methods using fuzzy sets for content-based recommender systems," *Fuzzy Sets Syst.*, vol. 160, pp. 76–94, 2009.
- [12] A. Zenebe, L. Zhou, and A. F. Norcio, "User preferences discovery using fuzzy models," *Fuzzy Sets Syst.*, vol. 161, pp. 3044–3063, 2010.
- [13] C. M. Chen and L. J. Duh, "Personalized web-based tutoring system based on fuzzy item response theory," *Expert Syst. Appl.*, vol. 34, pp. 2298–2315, 2008.
- [14] C. Cornelis, J. Lu, X. Guo, and G. Zhang, "One-and-only item recommendation with fuzzy logic techniques," *Inf. Sci.*, vol. 177, pp. 4906–4921, 2007.
- [15] C. Porcel, A. G. Lopez-Herrera, and E. Herrera-Viedma, "A recommender system for research resources based on fuzzy linguistic modeling," *Expert Syst. Appl.*, vol. 36, pp. 5173–5183, 2009.
- [16] Dina Hussein and Son N.Han, Institut Mines-Telecom, Gyu Myoung Lee, Liverpool John Moores University, Noel Crespi, Institut Mines-Telecom, "Social Cloud-Based Cognitive Reasoning for Task-Oriented Recommendation," *IEEE Cloud Computing*, 2325-6095/15/\$31.00 @2015 IEEE, November/December 2015.
- [17] A.V. Konstantinou, T. Eilam, M. Kalantar, A.A. Totok, W. Arnold, and E. Snible, "An Architecture for Virtual Solution Composition and Deployment in Infrastructure Clouds," *Proc. Third Int'l Work-shop Virtualization Technologies in Distributed Computing*, pp. 9-18, 2009.



# Use of Big Data and Data Mining to Enhance Productivity in the Business

<sup>1</sup>M. G. P. M. Samarasinghe,<sup>2</sup> K. A. Dilini. T. Kulawansa

<sup>1</sup>Faculty of Information Technology, University of Moratuwa, Sri Lanka Department of Computational Mathematics, University of Moratuwa, Sri Lanka  
E-mail: <sup>1</sup>pmsamarasinghe95@gmail.com, <sup>2</sup>dilini\_kulawansa@hotmail.com

## **ABSTRACT**

*Business is the power of economics globally and human civilization. In the today's business has competitive market in worldwide and companies must collaborate and communicate with customers, suppliers and other parties to get important decisions about the business and improve the business. Big data and data mining techniques are also major topics under the development of business process. in a business profit calculation and profit function is very important as it helps to determine the financial benefits from a business by considering the amount of total revenue, expenses, cost of production and taxes. Also, productivity is a major parameter and indicator (determinant of cost efficiency) in here which shows then state of a business, basically it calculated using average output (per period) divide by total costs (incurred, resources) in that period. here resources refer to capital, energy, material, personal etc. So, the productivity can lead the business for several directions like greater profit or lower profit. According to that for increase the productivity in a business must need a proper guidance about the business environment. In that case the main purpose of this review paper is introducing sophisticated method by using big data and data mining techniques to enhance the business productivity in a business. With the help of big data analysis and the data mining methods can improve the performance of business and productivity.*

**Keywords – Business Productivity, Enhancement, Data Mining, Big Data, Profit, Tools and Techniques**

## **I. INTRODUCTION**

In today big data is high level concept deal with many fields, among them that is very essential for the business industry. In the big data concept refers to large number of data sets and that are used to handle large datasets. Actually, the word Big data means data sets which have a huge volume as a result of that traditional and existing database management and analyzing techniques can't be applied with this data sets. [22]. now a days there are some data chunks(sets) which are huge, so existing software tools, storage techniques can't manage process or store that much amount of data within the expecting time limit [6]. fundamental requirements for big data analysis is similar to traditional database analysis. By keeping that on mind still it is necessary to handle these data sets with huge volume, in correct and expecting speed of processing without any losses at each level, so it has emerged some new significant challenges when looking for answers. Main aim of large data set analysis is to extract patterns and

connections to explore the heavy volume efficiently but conventional methods will not be able to do this task. There are mainly three defining properties or dimensions of big data as volume, variety and velocity.

When considering a business, big data analysis refers to checking and exploring huge chunks of data sets to extract and tap the patterns, connections and trends, and also to get a better idea about the supply chain. Mainly that helps to keep the manufacturing and productivity in right track and to also to find out the trends of employees and customer relationships. In modern era there are data related to businesses which haven't considered thoroughly. Recently as there are many advanced tools with reasonable costs organizations and businesses tend to release their complex databases to recognize useful connections and patterns which help them to make optimum decisions. Analysis of big data not only limited to structured information: customer relationship management data(CRM), purchase history data etc. in fact unstructured data like data from social media platforms, data from blogsites, videos, twitter feeds and other sources also can be considered. Likewise, both structured and unstructured data can be sent through the big data analysis to extract better solutions and decisions. Nowadays most of the large companies use big data analysis to secure their place in the market.

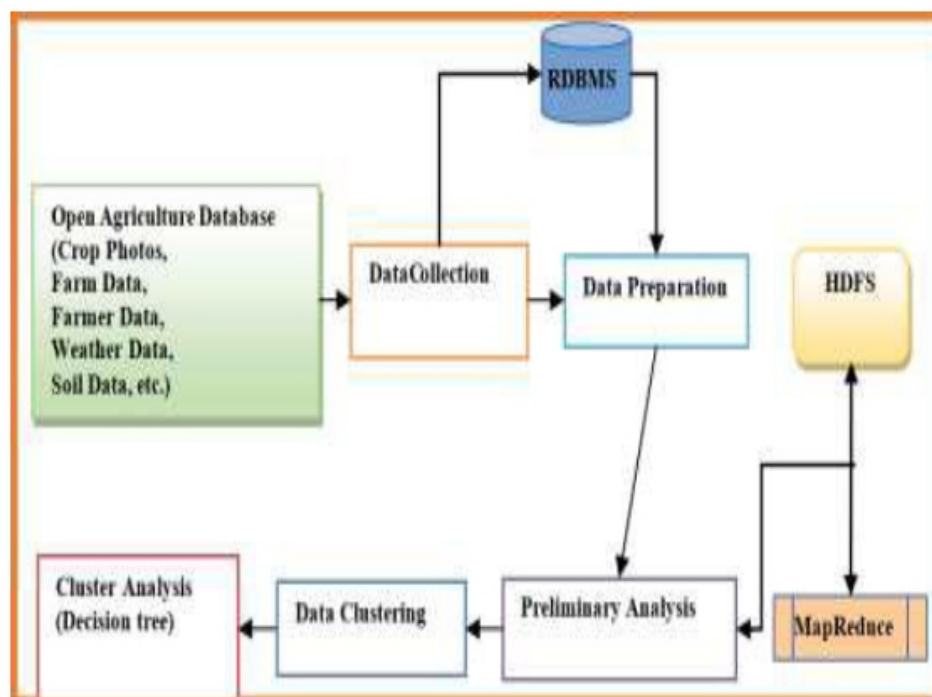


**Figure 1**

Using data analysis methods to increase productivity has become a very important for companies as it drive them to innovate and keep the competition with good economic value. Basically, professionals in database systems use big data analysis to figure patterns from unstructured data, demand connections, data chunk patterns etc. In order to boost the productivity, the big data analysis has become a major

companion. Nowadays fresh starting companies also uses data analysis strategies to compete with the well-organized & experienced organizations. By analyzing the unstructured data extract from blogs, videos and social media platforms using big data analysis techniques helps to improves the production quality. mixing big data analytics with advanced data extraction methods tends businesses to implement models for prediction. Productivity and satisfaction of a business can be drastically boost up by speed up them by using the suitable highly accurate big data analytic techniques. database marketing can be enhanced for a business by applying data analysis techniques in production. literally analysis of data shows new ways for develop a better and efficient campaign in marketing. Then the ultimate result will lead to develop a better productive set by enhancing the ROI of the business or company. increase in productivity will be massively help to enhance the profit and cash flow into the organization.

Using the big data analytical framework can get proper information about the business and that help to identify the large data sets and get proper decisions. As the example following shows the crop sales data in big data analytical framework [20]



**Figure 2**

And also the big data support for businesses to make accurate and precise decisions about the demand in the market with the help of predictive analytics platforms. So in that analysts can apply special algorithms in company's data and can have an ability to make real time predictions. So that making decisions and the predictions can transfer the future development of the business. To increase the productivity nowadays, many different business industry firms use HR analysis methods. but for HR

managers it is a huge challenging task to extract necessary amount of data to improve the performance. Also, to day data is the key for many successful organizations to implement strategies as well as long-term planning. So use of the big data analytics that is very efficient to do. And also big data for office productivity support in analyzing employee's behaviors and give result for estimate the supply and demand of the business and that directly support to enhance productivity in business. So due to this reason nowadays most of the organizations switch to big data to improve the productivity in their businesses. In order to estimate the parameters such as confidence of customer, choices real time big data analysis can be very much helpful. As this helps to add more weight or power to predictive analysis. Beside from that for users, data with high frequency enables the chance to test the theories in real time.

Data mining is also the main topic under the business improvements. With development of the today world data mining is most important and advanced topic in developing fast. Simply data mining can be identified as analyzing through large data sets to seek out patterns in data to develop connections and use those connections to solve problems. Recently data mining tools are vastly use for figure out trends in business sector.

As described, what is data mining above it is a very important tool nowadays. In recent times necessity and the demand of data in industry is growing in rapid rates so along with it demand for data scientists and analysts have been also growing. by using these techniques to analyzing data can be accurately converted into more meaningful information. Actually, this process is very much important and helpful for an organization. Reason is that from this mined information smart and precise decisions can be made, market can be predicted carefully and smartly and also it helps to understand the behavior of the consumers. So, data mining can help business to success if it ready to use these data driven strategies. [13]



**Figure 3**



It is recommended organizations to use data mining techniques today though it doesn't give any impact at start. Reason is that data mining can be used to identify the market sectors when production is happening. In other words, it helps to identify which product is need to make for which age group in which area of the society likewise. So it can use to minimize the unnecessary costs in production, marketing and helps to run a better advertising propaganda to Have a better and optimum communication with the customers. Data mining is very good asset and analyst can accurately use those tools to achieve above mentioned communication channel. So according to that support to enhance the business productivity with direct to the greater profits. So that data mining methods are used by many organizations for their business to process and convert raw data into useful information. Here many advanced software platforms use for figure out patterns in large data sets so that businesses can extract various useful and touching information about consumers in order to develop more accurate and efficient strategies in marketing and also to increase the productivity while keeping the costs at low level.

## **II. OVERVIEW OF USING BIG DATA AND DATA MINING ENHANCE THE PRODDUCTIVITY IN THE BUSINESS**

In modern world big data and data mining are two most rapidly developing technologies and that is cause for development and the enhancement the productivity in business processes [25]. So big data identify as large number of data sets and those data cannot be handle by using the traditional computing techniques. It is made out of Structured data, semi structured data and un structured data. [3]

Big data should be identified by 3vs.[21] [27] [23]

**Volume** - absolute size or scale of information which processed help to define big data systems. And the difference is big data sets have massive volume than traditional data sets. So additional processing and storing is needed at stage of analysis. As usually basic single computers can't handle this much of load it is very challenging when it comes to pooling, allocating and coordinating the available resources from computers(groups).so here tasks can be divided into pieces by using algorithms and cluster management techniques.

**Velocity** - another fact that big data changes from traditional data sets is the speed which that information travels inside a system. Inside a system, usually data is frequently flow into it by various sources and that data is processed in real time to produce outputs in order to update the system. so by velocity businesses can predict the growth of big data sets and the time which takes information to meet with sources and destinations. so data is frequently added analyzed processed to keep the order and data flow of the



system at relevant status. in order to protect from failures in the data pipeline robust systems with all available components which uses above mentioned ideas are needed [2]. Variety - Usually big data problems are uniquely due to the processing of wide range of both sources. usually data can be taken from the systems like social media platforms, applications, external APIs, sensors etc. by using big data analysis vast amount of data set can be handled and focus them into a single consolidated system regardless of format and the types of the media. That means media like videos, images can be ingested with text files log files etc. so in here big data systems stores the data in similar to their raw state. so developing will be only done at the stage of processing. Other than above discussed properties, there are some common additions are veracity, variability and value. So those characteristics are supposed to business areas for get decisions according to that. Usually big data is the application which uses advanced techniques to handle and processes massive sets of data. As these data sets are very large in scale and complex traditional and on-hand database management tools can't handle and process them. In today famous open source big data tools for use the for handle the large data sets as follows.

- Apache Hadoop –has the ability to handle and process mass scale data sets.
- Apache Spark –capable of very fast in-memory data processing.
- apache Storm –support any programming language capable of data stream processing with real time framework.
- Apache Cassandra – very efficient tool for handle structured data sets.
- MongoDB - support many programming languages this is open source NoSQL database based. Provide many tools with cross platform capabilities.
- R Programming Environment –provide wide range of processing tools and space for data visualization and statistical analysis.
- Neo4j –open source platform with rich tools for handle data and find patterns.
- Apache SAMOA –one of another tools from Apache to deal with big data. Samoa mainly use for develop streaming algorithms for efficient and accurate Big Data mining. Including Classification, Normalization, Regression, Clustering and Programming primitives in order to building custom algorithms.

So using big data technologies, can gain more benefits from it.

1. To enhance the performance
2. To boost the precise decision making
3. To achieve competitive advantage and make smart hiring choices
4. To achieve revenue gains
5. Improved Supplier Management
6. Better Order Fulfillment
7. Maximize Customer Value
8. Cost reduction
9. Enhancing the Advertising
10. Effective product management
11. Improved Service Level Performance

Recently data mining has become an interesting area in information industry as it can use for extract hidden knowledge from huge data sets data mining can be used to explore and visualize the data in a way that anyone can understand, which extracted from the areas such as statistics, machine learning, database management, science. The objective of data mining methods is to extract the hidden deep information and knowledge from the massive volumes of data sets. And the explored information can help business to make very good, accurate and precise decisions and it cause for enhance the productivity. simply advanced data mining systems does have the capability to generate unthinkable number of rules and patterns. More importantly these patterns represent useful knowledge. Exploration can be guided using the nature of the patterns (whether subjective or objective). The below things are support to enhance productivity in business. [5].

#### A. Business intelligence and sales prediction by data mining algorithms.

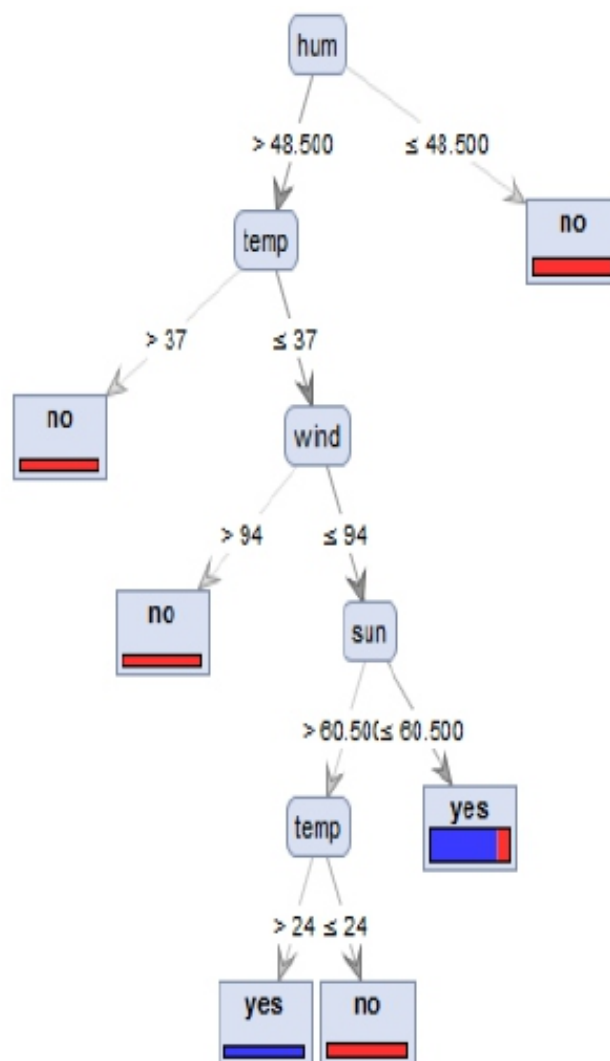


Figure 4

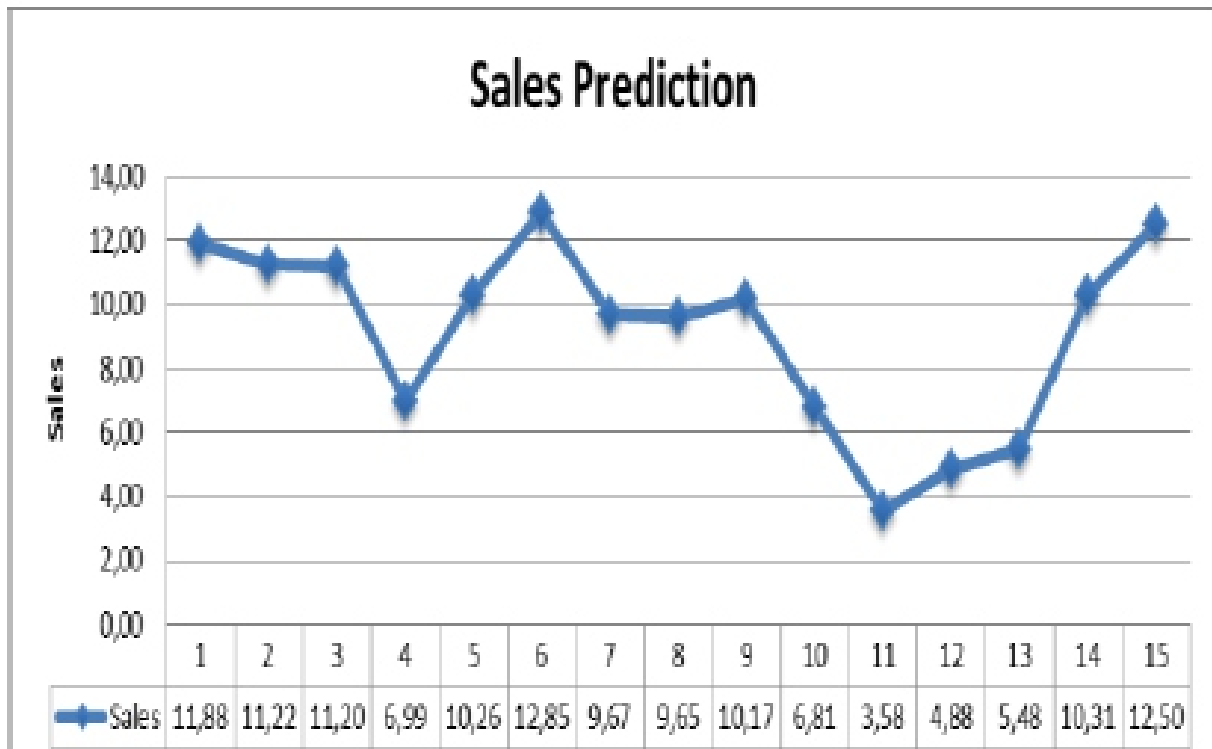


Figure 5

**B. Web mining analysis of social trends.****C. Business productivity can improve using the data mining algorithms [5].****D. Logistic algorithms embedded in the prototype platform.**

There are many types data mining methods and processes developed and under developing which uses the fundamental concepts.

the subject of data mining is frequently evolving and there are many advanced and comprehensive techniques which have been developed by using data mining concepts. Literally they provide businesses and companies a huge benefit by giving chance to analyze and explore data to produce important and applicable future trends. Listed below are the few from the methods developed by the professional in the field [4] [13] [15] [17] [18] [26].

1. Classification: apply to fetch the knowledge in data and metadata
2. Clustering: mining method to figure out data that similar to other. So that this help to figure the similarities and differences among the raw data.
3. Regression: method for develop and search for the connections among the variables/data.
4. Association Rules: use to fetch the hidden patterns in a data set.
5. Outer detection: use to identify the data which differs from the expected patterns.
6. Sequential Patterns: use to figure out the same or similar patterns in a given data set.
7. Prediction: use and analysis of past events in order to generate and predict the future.

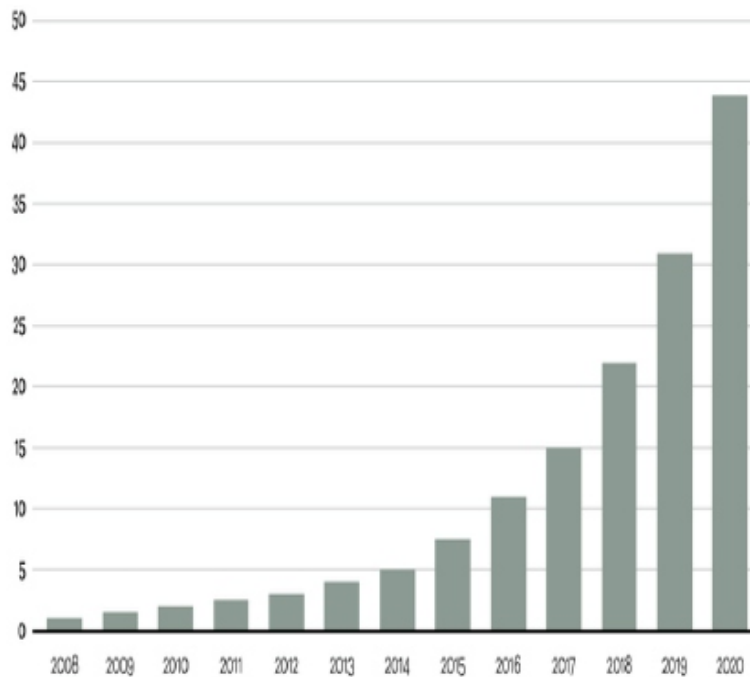
As today, massive amount of data is generated in a glimpse it is needed to apply most convenient and efficient tools for handle and process the amount of data and apply mining algorithms in within given time frame without any losses or errors [5][13][16][19].

Tools like - Rapid Miner, Orange, Weka, KNIME, Sisense, SSDT (SQL Server Data Tools), Apache Mahout, Oracle Data Mining, Rattle, Data Melt, IBM Cognos, IBM SPSS Modeler, SAS Data Mining, Teradata, Board, Dundas BI, Spark It necessary to improve the productivity of a business to achieve better profits. Productivity can be improved while lowering the costs if correct manufacturing strategies, market plans and propaganda is using by the business this also improve the competitiveness. Basically, efficiency of production can be determine using the productivity. that means by achieving high productivity greater profits can be achieved by the businesses or individual.As by providing more services and goods to customers helps to achieve higher profits productivity growth is very important to a business. Using the traditional computer systems that is very complex task but sing that above big data and data mining tools it make more efficiency and effectively also that lead to enhance the business productivity in proper way that the competitors who are in the industry. When using the that big data and data mining, can get a better understanding about the business. After get good knowledge and then identify the core competencies of business, according to that can improve the business productivity and increase profit obtain from the business.

### **III. MAJOR RESEACHERS IN USING BIG DATA AND DATA MINING ENHANCE THE PRODUCTIVITY**

There are various research's conducted which related with big data and data mining in the field of business for purpose of enhance the productivity and to obtain higher profits.

According to the Anupama Jha, Dr.Meenu Dave and Dr.Supriya Madan in india the research along with the big data and data mining techniques. So that case in the develop of the world day by day the data sets of all the fields are growth very fast. That is common for many areas and among them business industry in one thing. Compare with past or the traditional data, today that have been changed [3].



**Figure 6: Data growth by year**

And also the research was done by Dr.R. Shankar<sup>1</sup> and Dr.S. Duraisamy in india along with the data mining. That included the tools and techniques used in the data mining concept. So that is include techniques of data mining and tools as well as the application of data mining and the trends of it [13].

According to the research Nada Elgendy and Ahmed Elragal in Egypt, that is concluding with the basically big data analytics and the characteristics of it, tools and the methods also the storage of big data and the better management, analytical process, big data and decision making, supply chain performance, quality management and performance combing with big data. So those are support for the field of business for enhancing the productivity in business [6].

The research about the business informatics done by the P. K. Paul<sup>1</sup>, P. S. Aithal<sup>2</sup>, A. Bhuimali in India, is basically concluding with the big data and business management and the improvement of the business in proper way and enhance the productivity if the business and reach for the profit margins [8]. According to that big data is one of the major concept for business process development. In many ways big data analytical techniques are help to achieve that business goals. In the that concept is more powerful for business field.

Dursun Delena and Sudha Ramb in UK done a research about the challenges and opportunities in business analytics, that include the main challenges in business analytics as well as taxonomy for business analytics and historical view of analytical terminology [24].

According to the research done by S.D.Gheware<sup>1</sup>, A.S.Kejkar<sup>2</sup>, S.M.Tondare<sup>3</sup> in India related with the data mining. That including data mining tasks, technologies, tools, challenges and the application related to the data mining [18].

#### **IV. IDENTIFY THE CURRENT ISSUES IN USE OF BIG DATA AND DATA MINING TO ENHANCE THE BUSINESS PRODUCTIVITY**

In the field of business, when using the big data and data mining techniques for handle the massive volumes of data and get proper predictions from it, include some limitation and the difficulties also.

**A. Limited storage** - So the storage is most important thing under that big data and the data mining concepts. So large volume of data need high capacity of the storage [1] [10] [11] [14] [18]. So the Authors of the above mentioned research papers are discussed about the storage issue mainly.

**B. Incomplete data and the issue of having sufficient amount of quality data** - There is a big problem as most of the data fetched is not accurate and not collected in a proper way [18].

**C. High cost and volume-** That need high cost for handle the all the data in proper way. As well as management cost is needed for that. extracting the necessary and useful information from the big data managing the volume and cost [8] [14].

**D. The limited big data professionals** - The review paper about Big Data: Issues, Challenges and Techniques in Business Intelligence, discuss that limited human power.

#### **V. APPLICATION OF USING BIG DATA AND DATA MINING**

With development of the world big data and the data mining analysis are used in many fields. Education industry, health care, government sector, media and entertainment industry, weather patterns, transport, banking and finance, medicine, insurance, sales and marketing etc. among the business field is most important one using data mining and the big data for get better productivity through the help of those things [17].

#### **VI. DISCUSSION**

Business field is most important and the major fact in of economics all over the world. So the main objective of the business is maximizing profit. According to that businesses need to proper knowledge for manage it as well as effective ways to identify the surrounding business world in each and every business who are in the business world. On that case for the purpose of the productivity enhancement of

the business, the concepts of big data and the data mining given great support for enhance the business productivity and if cause to increase the profit rapidly.

## **VII. FUTURE DIRECTIONS OF USING BIG DATA AND DATA MINING ENHANCE THE PRODUCTIVITY**

At the present many researchers have done related concepts of the big data and the data mining with business enhancements. So other than those developments and the improvements, as the future directions hope to do algorithm to analyze online customers properly. So that be a good effect for the enhance the business productivity. And also explore various data access requirements of different applications and design and implementation of programming language abstractions to deal with parallelism [1].

Using the interesting machine learning concepts and tools in big data analysis to acquire optimum results. As many of the big data analysis deal with data processing, algorithm implementation, and optimization so machine learning can surely adopt with this field [1].

In future it is need to achieve includes fake identity detection using online social networks (this helps to push ranking the personalities) to gain advantage in competition. developing the business economics by enhancing the supply chains, and innovation capabilities associated with in agriculture learning the basics and the spread patterns of crop diseases from plant genomics data, by analyzing human genome and next generation sequencing data it can use to put an advance insight into the human diseases [7].

In future it is necessary to develop an application more accuracy and efficiency. Ex tool can be design to support many fields instead of one specific area. Development will be a complex procedure but indeed it will help to develop an efficient final product [12].

As in modern world many of the fields and social environments have digitalized as similar to business, speed of generation of big data volumes have increased drastically so efficient methods have to be implemented in order to get in pace with that mass volume of data sets and also, these approaches should also able applied to business, science, finance fields etc. without limitations [21].

## **ACKNOWLEDGEMENT**

My special thanks goes to my parents and also my brother and sister for their great support and motivation during this work to achieve my task. And also, I thank for everyone who helped me to success this review paper from starting to the end.



---

**REFERENCE**

- [1] D. P. Acharjya, Kauser Ahmed P. "A Survey on Big Data Analytics: Challenges, Open Research Issues and Tools". (IJACSA) *International Journal of Advanced Computer Science and Applications*, Vol. 7, No. 2, 2016.
- [2] Jens Baum 1, Christoph Laroque 1,\*, Benjamin Oeser 1, Anders Skoogh 2 and Mukund Subramaniyan 2. "Applications of Big Data analytics and Related Technologies in Maintenance". *Machines* 2018, 6, 54; doi:10.3390/machines6040054.
- [3] Anupama Jha, Dr. Meenu Dave, Dr. Supriya Madan. "A Review on the Study and Analysis of Big Data using Data Mining Techniques". *International journal of latest trends in engineering and technology*. August 2018.
- [4] Dina Fawzy1\*, Sherin Moussa and Nagwa Badr. "The Evolution of Data Mining Techniques to Big Data Analytics: An Extensive Study with Application to Renewable Energy Data Analytics". *Asian Journal of Applied Sciences (ISSN: 2321–089) Volume 04 – Issue 03, June 2016*.
- [5] Alessandro Massaro1,\*, Valeria Vitti1, Angelo Galiano1, Alessandro Morelli2." *Business Intelligence Improved by Data Mining Algorithms and Big Data Systems: An Overview of Different Tools Applied in Industrial Research*". *Computer Science and Information Technology* 7(1): 1-21, 2019 <http://www.hrpub.org> DOI: 10.13189/csit.2019.070101
- [6] Nada Elgendy and Ahmed Elragal. "Big Data Analytics: A Literature Review Paper". *Conference Paper in Lecture Notes in Computer Science · August 2014* DOI: 10.1007/978-3-319-08976-8\_16.
- [7] Mudasir Ahmad Wani1, Suraiya Jabin1. "Big Data: Issues, Challenges and Techniques in Business Intelligence". *Conference Paper · <https://www.researchgate.net/publication/309556100> December 2015*.
- [8] P. K. Paul1, P. S. Aithal2, A. Bhumali3. "Business Informatics: With Special Reference to Big Data as an emerging Area". <https://ssrn.com/abstract=3181045> April 2018.
- [9] Hiba Alsghaier, Mohammed Akour, Issa Shehabat, Samah Aldiabat. "THE IMPACT OF BIG DATA ANALYTICS ON BUSINESS COMPETITIVENESS". *Proceedings of the New Trends in Information Technology (NTIT-2017). The University of Jordan, Amman, Jordan. 25-27 April 2017*.
- [10] Jinchuan CHEN, Yueguo CHEN, Xiaoyong DU, Cuiping LI, Jiaheng LU, Suyun ZHAO, Xuan ZHOU. "Big data challenge: a data management perspective". *Front. Comput. Sci.*, 2013, 7(2): 157–164 DOI 10.1007/s11704-013-3903-7.
- [11] Pavithra B 1, Dr. Niranjnathurthy M2, Kamal Shaker J3, Martien Sylvester Mani F3. "The Study of Big Data Analytics in E-Commerce". *IJARCCCE International Journal of Advanced Research in Computer and Communication Engineering ICRITCSA M S Ramaiah Institute of Technology, Bangalore Vol. 5, Special Issue 2, October 2016*.
- [12] Kalpana Rangra, Dr. K. L. Bansal, "Comparative Study of Data Mining Tools" Volume 4, Issue 6, June 2014 ISSN: 2277 128X *International Journal of Advanced Research in Computer Science and Software Engineering Research Paper Available online at: [www.ijarcsse.com](http://www.ijarcsse.com)*.
- [13] Dr.R. Shankar1 and Dr.S. Duraisamy2. "Analysis of Data Mining Tasks, Techniques, Tools, Applications And Trends". *IOSR Journal of Computer Engineering (IOSR-JCE) e-ISSN: 2278-0661, p-ISSN: 2278-8727, Volume 20, Issue 5, Ver. II (Sep - Oct 2018), PP 12-19 [www.iosrjournals.org](http://www.iosrjournals.org)*.
- [14] Mashooque A. Memon1,\*, Safeullah Soomro2, Awais K. Jumani3 and Muneer A. Kartio3." *Big Data Analytics and Its Applications*". *Annals of Emerging Technologies in Computing (AETiC) Vol. 1, No. 1, 2017*.
- [15] J.Mary Dallfin Bruxella1, S.Sadhana2, S.Geetha3. "Categorization of Data Mining Tools Based on Their Types". *J.Mary Dallfin Bruxella et al, International Journal of Computer Science and Mobile Computing, Vol.3 Issue.3, March- 2014, pg. 445-452 © 2014, IJCSMC All Rights Reserved 445. [www.ijcsmc.com](http://www.ijcsmc.com)*.
- [16] Mrs. Parminder Kaur1., Dr. Qamar Parvez Rana2. " Comparison of Various Tools for Data Mining ". *International Journal of Engineering Research & Technology (IJERT) ISSN: 2278-0181 IJERTV3IS100246 [www.ijert.org](http://www.ijert.org) Vol. 3 Issue 10, October- 2014*.
- [17] Mr. S. P. Deshpande1 and Dr. V. M. Thakare2. "DATA MINING SYSTEM AND APPLICATIONS: A REVIEW" *International Journal of Distributed and Parallel systems (IJDPS) Vol.1, No.1, September 2010*.
- [18] S.D.Gheware1, A.S.Kejkar2, S.M.Tondare3. "Data Mining: Task, Tools, Techniques and Applications". *International Journal of Advanced Research in Computer and Communication Engineering Vol. 3, Issue 10, October 2014*,
- [19] Sadiq Hussain. "Survey on Current Trends and Techniques of Data Mining Research" . *London Journal of Research in Computer Science and Technology Volume 17 | Issue 1 | Compilation 1.0*.
- [20] Ch. Chandra Sekhar1, J. UdayKumar1, B. Kishor Kumar1 and Ch. Sekhar2. "Effective use of Big Data Analytics in Crop planning to increase Agriculture Production in India". *International Journal of Advanced Science and Technology Vol.113 (2018), pp.31-40. <http://dx.doi.org/10.14257/ijast.2018.113.04>*.
- [21] Shirish Jeble, Sneha Kumari, Yogesh Patil. " Role of Big Data in Decision Making". *OPERATIONS AND SUPPLY CHAIN MANAGEMENT Vol. 11, No. 1, 2018, pp. 36–44. ISSN 1979-3561 | EISSN 2759-9363*.

- [22] Dong-Hui Jin and Hyun-Jung Kim. "Integrated Understanding of Big Data, Big Data Analysis, and Business Intelligence: A Case Study of Logistics". [www.mdpi.com/journal/sustainability](http://www.mdpi.com/journal/sustainability).
- [23] Peter O'Donovan\*, Kevin Leahy, Ken Bruton and Dominic T. J. O'Sullivan. "Big data in manufacturing: a systematic mapping study". O'Donovan et al. *Journal of Big Data* (2015) 2:20 DOI 10.1186/s40537-015-0028-x.
- [24] Dursun Delena and Sudha Ramb. "Research challenges and opportunities in business analytics". *JOURNAL OF BUSINESS ANALYTICS* 2018, VOL. 1, NO. 1, 2–12, <https://doi.org/10.1080/2573234X.2018.1507324>.
- [25] Hsinchun Chen, Roger H. L. Chiang, Veda C. Storey, "BUSINESS INTELLIGENCE AND ANALYTICS: FROM BIG DATA TO BIG IMPACT". *MIS Quarterly* Vol. 36 No. 4/December 2012.
- [26] Mustafa Ergün1, "USING THE TECHNIQUES OF DATA MINING AND TEXT MINING IN EDUCATIONAL RESEARCH". *Yıl:2017,C:6,S:12(180-189)*.
- [27] Oliver Müller, Maria Fay, Jan vom Brocke. "The effect of big data and analytics on firm performance: An econometric analysis considering industry characteristics". Müller, O., Fay, M., vom Brocke, J. (2018) *The effect of big data and analytics on firm performance: An econometric analysis considering industry characteristics*..

# Instructions for Authors

## Essentials for Publishing in this Journal

- 1 Submitted articles should not have been previously published or be currently under consideration for publication elsewhere.
- 2 Conference papers may only be submitted if the paper has been completely re-written (taken to mean more than 50%) and the author has cleared any necessary permission with the copyright owner if it has been previously copyrighted.
- 3 All our articles are refereed through a double-blind process.
- 4 All authors must declare they have read and agreed to the content of the submitted article and must sign a declaration correspond to the originality of the article.

## Submission Process

All articles for this journal must be submitted using our online submissions system. <http://enrichedpub.com/> . Please use the Submit Your Article link in the Author Service area.

---

## Manuscript Guidelines

The instructions to authors about the article preparation for publication in the Manuscripts are submitted online, through the e-Ur (Electronic editing) system, developed by **Enriched Publications Pvt. Ltd.** The article should contain the abstract with keywords, introduction, body, conclusion, references and the summary in English language (without heading and subheading enumeration). The article length should not exceed 16 pages of A4 paper format.

### Title

The title should be informative. It is in both Journal's and author's best interest to use terms suitable. For indexing and word search. If there are no such terms in the title, the author is strongly advised to add a subtitle. The title should be given in English as well. The titles precede the abstract and the summary in an appropriate language.

### Letterhead Title

The letterhead title is given at a top of each page for easier identification of article copies in an Electronic form in particular. It contains the author's surname and first name initial .article title, journal title and collation (year, volume, and issue, first and last page). The journal and article titles can be given in a shortened form.

### Author's Name

Full name(s) of author(s) should be used. It is advisable to give the middle initial. Names are given in their original form.

### Contact Details

The postal address or the e-mail address of the author (usually of the first one if there are more Authors) is given in the footnote at the bottom of the first page.

### Type of Articles

Classification of articles is a duty of the editorial staff and is of special importance. Referees and the members of the editorial staff, or section editors, can propose a category, but the editor-in-chief has the sole responsibility for their classification. Journal articles are classified as follows:

#### Scientific articles:

1. Original scientific paper (giving the previously unpublished results of the author's own research based on management methods).
2. Survey paper (giving an original, detailed and critical view of a research problem or an area to which the author has made a contribution visible through his self-citation);
3. Short or preliminary communication (original management paper of full format but of a smaller extent or of a preliminary character);
4. Scientific critique or forum (discussion on a particular scientific topic, based exclusively on management argumentation) and commentaries. Exceptionally, in particular areas, a scientific paper in the Journal can be in a form of a monograph or a critical edition of scientific data (historical, archival, lexicographic, bibliographic, data survey, etc.) which were unknown or hardly accessible for scientific research.

### **Professional articles:**

1. Professional paper (contribution offering experience useful for improvement of professional practice but not necessarily based on scientific methods);
2. Informative contribution (editorial, commentary, etc.);
3. Review (of a book, software, case study, scientific event, etc.)

### **Language**

The article should be in English. The grammar and style of the article should be of good quality. The systematized text should be without abbreviations (except standard ones). All measurements must be in SI units. The sequence of formulae is denoted in Arabic numerals in parentheses on the right-hand side.

### **Abstract and Summary**

An abstract is a concise informative presentation of the article content for fast and accurate Evaluation of its relevance. It is both in the Editorial Office's and the author's best interest for an abstract to contain terms often used for indexing and article search. The abstract describes the purpose of the study and the methods, outlines the findings and state the conclusions. A 100- to 250-Word abstract should be placed between the title and the keywords with the body text to follow. Besides an abstract are advised to have a summary in English, at the end of the article, after the Reference list. The summary should be structured and long up to 1/10 of the article length (it is more extensive than the abstract).

### **Keywords**

Keywords are terms or phrases showing adequately the article content for indexing and search purposes. They should be allocated heaving in mind widely accepted international sources (index, dictionary or thesaurus), such as the Web of Science keyword list for science in general. The higher their usage frequency is the better. Up to 10 keywords immediately follow the abstract and the summary, in respective languages.

### **Acknowledgements**

The name and the number of the project or programmed within which the article was realized is given in a separate note at the bottom of the first page together with the name of the institution which financially supported the project or programmed.

### **Tables and Illustrations**

All the captions should be in the original language as well as in English, together with the texts in illustrations if possible. Tables are typed in the same style as the text and are denoted by numerals at the top. Photographs and drawings, placed appropriately in the text, should be clear, precise and suitable for reproduction. Drawings should be created in Word or Corel.

### **Citation in the Text**

Citation in the text must be uniform. When citing references in the text, use the reference number set in square brackets from the Reference list at the end of the article.

### **Footnotes**

Footnotes are given at the bottom of the page with the text they refer to. They can contain less relevant details, additional explanations or used sources (e.g. scientific material, manuals). They cannot replace the cited literature.

The article should be accompanied with a cover letter with the information about the author(s): surname, middle initial, first name, and citizen personal number, rank, title, e-mail address, and affiliation address, home address including municipality, phone number in the office and at home (or a mobile phone number). The cover letter should state the type of the article and tell which illustrations are original and which are not.

### **Address of the Editorial Office:**

**Enriched Publications Pvt. Ltd.**  
S-9, IInd FLOOR, MLU POCKET,  
MANISH ABHINAV PLAZA-II, ABOVE FEDERAL BANK,  
PLOT NO-5, SECTOR -5, DWARKA, NEW DELHI, INDIA-110075,  
PHONE: - + (91)-(11)-45525005

