A Survey of Cloud Computing and Mobile Cloud Computing: Architecture, Applications

Mrs. Anuradha Diwan
(anuradha.shinestar@gmail.com)

ABSTRACT
Together with an explosive growth of the mobile applications and emerging of cloud computing concept, mobile cloud computing (MCC) has been introduced to be a potential technology for mobile services. In this paper we are discussing about cloud computing, their types, we also explain why we use mobile cloud computing? What is mobile cloud computing? We also study the architecture of cloud computing and mobile cloud computing. This paper gives a survey of Cloud Computing and mobile cloud computing (MCC), which helps general readers have an overview of the MCC including the definition, architecture, and applications. This paper also contains Security issues in Mobile Cloud. Mobile Clouds can be categorized into two types: security issues for mobile users such as privacy concerns and data storage and access in Cloud such as authentication and access control.

Keywords- cloud computing, Mobile cloud computing, Security.

1. INTRODUCTION

These days cloud commuting is very famous and it is used to run various types of business application. We also explain the architecture of cloud commuting, mobile cloud commuting, To create cloud commuting environment internet, server hardware, cloud OS is required. The cloud OS is installed on the hardware. It provides interfaced between user and server and manages and distributes all resources of cloud systems. Some of the Cloud OS are eye OS, VMware Cloud Operating System, icloud, Cloud, Cornelios etc...

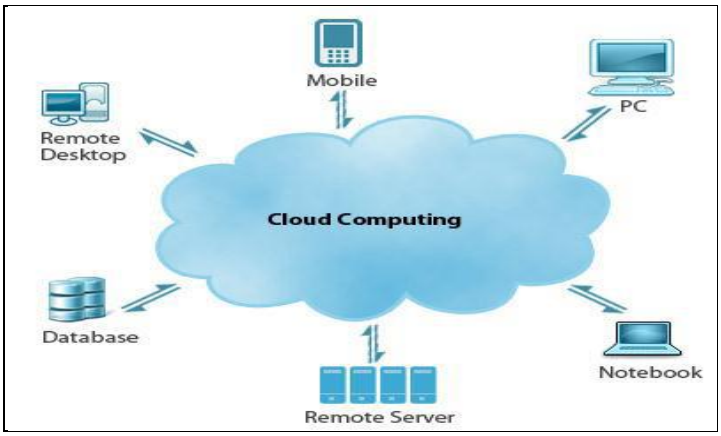


Fig no: 1

With the help of utilities provided by Cloud OS we can create virtual servers for our clients. Clients can login and use the allocated virtual server/resources. We can increase and decrease the resource usage by client. Cloud Computing is growing fast and companies are taking full advantage of the services provided by Cloud computing.

Mobile devices (e.g., smart phone, tablet pcs, etc) are increasingly becoming an essential part of human life as the most effective and convenient communication tools not bounded by time and place. Mobile users accumulate rich experience of various services from mobile applications (e.g., iPhone apps, Google apps, etc), which run on the devices and/or on remote servers via wireless networks. The rapid progress of mobile computing (MC) [1] becomes a powerful trend in the development of IT technology as well as commerce and industry fields. However, the mobile devices

are facing many challenges in their resources (e.g., battery life, storage, and bandwidth) and communications (e.g., mobility and security) [2]. The limited resources significantly impede the improvement of service qualities.

2. OVERVIEW OF CLOUD COMPUTING

Cloud Computing can be defined as a model which delivers applications as services (known as Software as a service or SaaS) over internet and providing hardware and system software for users to implement, deploy and maintain their custom-made applications and/or services [3]. The increased popularity and usage of mobile devices such as smart phones, laptops and tablets, mobile Clouds (MCC) are increasingly popular.

2.1 What is Cloud Computing

Cloud computing is a computing term or metaphor that evolved in the late 2000s, based on utility and consumption of computing resources. Cloud computing involves deploying groups of remote servers and software networks that allow centralized data storage and online access to computer services or resources. Clouds can be classified as public, private or hybrid.

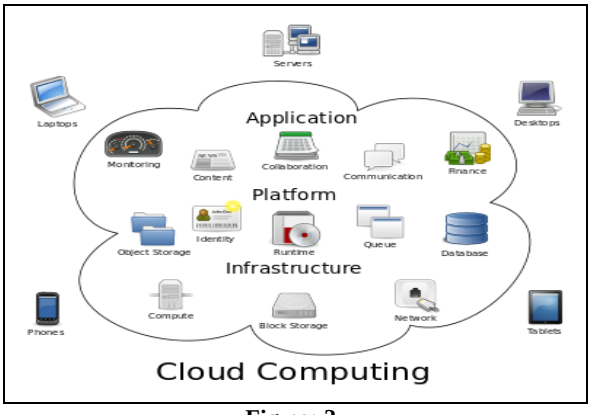


Fig no: 2

2.2 Benefits of cloud computing

In this section we will see the benefits of Cloud computing for businesses. The most important benefits of Cloud Computing this save money and time. If we required we demand for the required resource from the Cloud resource and pay to the service provider based on our usage.

Reduced Cost: Cloud Computing reduces our overall physical hardware and maintenance of these hardware. We want to just pay for what we use from the Cloud Computing resource pool. It does all work for us.

Scalability of System: we can easily request for more processing power from the resource pool at very minimum cost according to our requirement.

Automatic Updates of software: Cloud Computing company will automatically update the software if a new version is released.

Remote Access of the System: our employees and customers can access the data from anywhere around the world.

Disaster Relief: The Cloud Computing Company keeps the backup of data and ensures the proper functioning of the system.

Quick Customer Support: the Cloud Computing vendor provides quick customer support, which is essential for the functioning of your business.

Sufficient Storage: more space is available for storage of our data.

3. OVERVIEW OF MOBILE CLOUD COMPUTING

The term "mobile cloud computing" was introduced not long after the concept of "cloud computing" launched in mid-2007. It has been attracting the attentions of entrepreneurs as a profitable business option that reduces the development and running cost of mobile applications, of mobile users as a new technology to achieve rich

experience of a variety of mobile services at low cost, and of researchers as a promising solution for green IT [4]. This section provides an overview of MCC including definition, architecture, and advantages of MCC.

3.1 What is MCC

The Mobile Cloud Computing Forum defines MCC as follows [5]:

"Mobile Cloud Computing at its simplest, refers to an infrastructure where both the data storage and the data processing happen outside of the mobile device. Mobile cloud applications move the computing power and data storage away from mobile phones and into the cloud, bringing applications and mobile computing to not just smartphone users but a much broader range of mobile subscribers".

Mobile Cloud Computing (MCC) is the combination of cloud computing, mobile computing and wireless networks to bring rich computational resources to mobile users, network operators, as well as cloud computing providers. The ultimate goal of MCC is to enable execution of rich mobile applications on a plethora of mobile devices, with a rich user experience. MCC provides business opportunities for mobile network operators as well as cloud providers. More comprehensively, MCC can be defined as "a rich mobile computing technology that leverages unified elastic resources of varied clouds and network technologies toward unrestricted functionality, storage, and mobility to serve a multitude of mobile devices anywhere, anytime through the channel of Ethernet or Internet regardless of heterogeneous environments and platforms based on the pay-as-you-use principle."

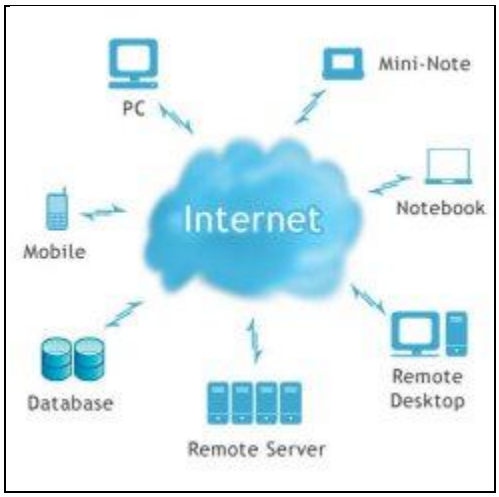


Fig. No. 3

3.2 Benefit of MCC

Cloud computing is known to be a promising solution for mobile computing due to many reasons (e.g., mobility, communication, and portability [6]). In the following, we describe how the cloud can be used to overcome obstacles in mobile computing, thereby pointing out advantages of MCC.

1) Extending battery lifetime: Battery is one of the main concerns for mobile devices. Several solutions have been proposed to enhance the CPU performance [7], [8] and to manage the disk and screen in an intelligent manner [9], [10] to reduce power consumption. However, these solutions require changes in the structure of mobile devices, or they require a new hardware that results in an increase of cost and may not be feasible for all mobile devices. Computation offloading technique is proposed with the objective to migrate the large computations and complex processing from resource-limited devices (i.e., mobile devices) to resourceful machines (i.e., servers in clouds). This avoids taking a long application execution time on mobile devices which results in large amount of power consumption

2) Improving data storage capacity and processing power: Storage capacity is also a constraint for mobile devices. MCC is developed to enable mobile users to store/access the large data on the cloud through wireless networks. First example is the Amazon Simple Storage Service (Amazon S3) [11] which supports file storage service.

3)Improving reliability: Storing data or running applications on clouds is an effective way to improve the reliability since the data and application are stored and backed up on a number of computers. This reduces the chance of data and application lost on the mobile devices.

4. ARCHITECTURE OF CLOUD COMPUTING

Cloud architecture, the systems architecture of the software systems involved in the delivery of cloud computing, typically involves multiple cloud components communicating with each other over a loose coupling mechanism such as a messaging queue. Elastic provision implies intelligence in the use of tight or loose coupling as applied to mechanisms such as these and others.

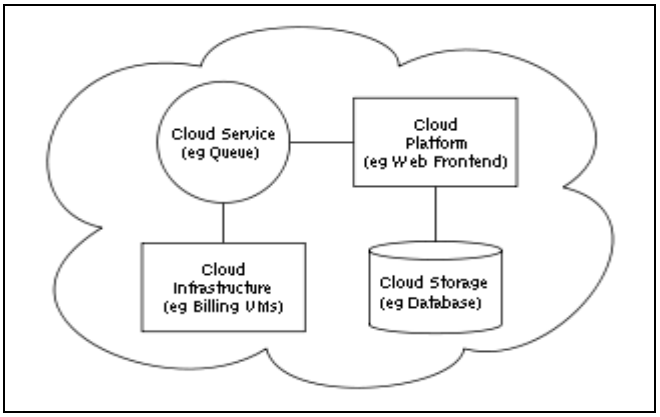


Fig no. 4 Cloud computing sample architecture

4. ARCHITECTURE OF MOBLE CLOUD COMPUTING

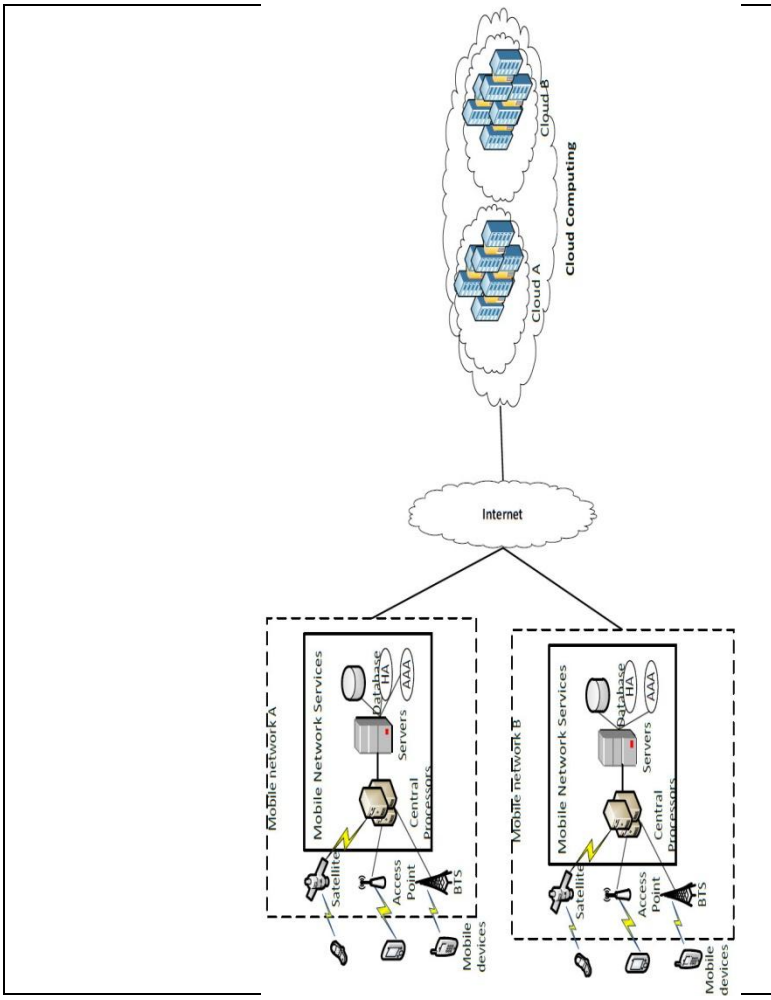


Fig no. 5 Mobile Cloud computing sample architecture

In Fig. 5, mobile devices are connected to the mobile networks via base stations (e.g., base transceiver station (BTS), access point, or satellite) that establish and control the connections (air links) and functional interfaces between the networks and mobile devices. Mobile users' requests and information (e.g., ID and location) are transmitted to the central processors that are connected to servers providing mobile network services. Here, mobile network operators can provide services to mobile

users as AAA (for authentication, authorization, and accounting) based on the home agent (HA) and subscribers' data stored in databases. After that, the subscribers' requests are delivered to a cloud through the Internet. In the cloud, cloud controllers process the requests to provide mobile users with the corresponding cloud services. These services are developed with the concepts of utility computing, virtualization, and service-oriented architecture (e.g.web, application, and database servers).

5. KEY SECURITY ISSUE IN MCC

MCC into two categories. The security for mobile users and the security for the data.

A. Security for mobile users

We now list some of the security issues from the user's perspective:

1) Security for mobile applications: Installing and running security applications such as McAfee and Norton are the easiest and simplest way to detect issues such as virus and worms in mobile applications. But, mobile devices have limited processing and battery power which makes it difficult for the mobile users to install and use such heavy anti-virus applications.

2) Privacy: Mobile users increasingly use location-based services due to the advantages of GPS positioning devices. Mobile users need to provide their private information such as location information to such services and it becomes potential threat to their privacy.

B. Security data on clouds

Users can store and access their data and applications in Cloud. When storing data in Cloud, several issues need to be addressed such as data authentication and integrity. Here, we list some of such issues [12]:

1) Integrity: Data integrity is an important concern by the mobile users. A typical solution to this should consider mobile specific issues such as energy consumption.

2) Authentication: Authentication is always an important issue when data resides in Cloud. Mobile clients need to be authenticated in an appropriate manner to access their data residing in the Cloud.

3) Digital Rights Management: The unstructured digital information such as videos have often being distributed illegally and are pirated. These contents need to be protected from illegal access in MCC context.

4) Confidentiality, preventing unauthorized users from gaining access to critical information of any particular user.

5) Availability, ensuring authorized users getting the access they require.

6) Legitimate, ensuring that only authorized users have access to services.

7) Accountability, ensuring that the users are held responsible for their security related activities by arranging the user and his/her activities are linked if and when necessary.

6) APPLICATION OF CC

In this section, we present some application of cloud computing

A) Cloud Computing For E-Learning

E-learning is a new trend in education that tries to make the best use of information technology (IT). Cloud computing is an attractive environment for students, faculty members and researchers. As an emerging IT, cloud computing can provide universities and research centers with powerful and cost-effective computational infrastructure. Students can connect to campus educational services through their personal mobile devices from anywhere. Faculty members can have efficient and flexible access to their course material in their class rooms. Researchers can find articles, models and run their experiments on the cloud faster than ever [13].

B) **Cloud Computing for ERP**
Traditional Enterprise Resource Planning (ERP)systems have some limitations. As the business grows inside an organization, different software applications may be needed to manage information in many areas such as human resources, payroll, finance and administration. Obviously, purchasing, installing and maintaining such multiple types of software applications represent a challenge for business growing.

C) **Cloud Computing For E- Government**
Traditional E-governance faces different challenges such as [14]:

- Resources cannot scale up and down with the demands that change over time. This may result in insufficient or redundant resources
- SW and HW have to be frequently upgraded and maintained which costs time and money
- New SW licences have to be purchased
- System should be available 24x7
- Limited data storage and recovery
- Need to provide secure environment with authentication and access control
- Lack of accountability

7) **APPLICATION OF MCC**
In this section, we present some application of mobile cloud computing

A) Service of Mobile Clouds - A number of researchers have introduced service clouds for mobile cloud computing and named Mobile service clouds. A lot of their model enables dynamic embodiment, installation, arrangement and rearrangement of services to be used by the mobile users.

B) Flexible application weblets'- A numbers of researchers created flexible applications that increase and enhance powerful smart phones, utilizing flexible computing resources from the cloud. A flexible application can have one or more weblets in it, while wallets have the most important feature of portability. Any given wallet can contribute in switched between both mobile and stationary devices.

C) Convenient Web Services- Meanwhile, other researchers suggest a suitable method for creating and developing mobile applications using cloud computing and Restful web services. Convenient web services which called Restful web services are so much simpler and flexible to use. The main aim is to offload computational capacity, storage and security for different kinds of mobile device to cloud by utilizing the restful web services [15].

8) Conclusion

Cloud computing is a new emerging technology that is expected to significantly change the field of IT in the next few years and lead it for the coming decades. This article has highlighted a comprehensive overview of mobile cloud computing. The suitable solutions for mobile cloud computing have also been discussed so that the readers can have a better understanding of the mobile cloud computing and its applications. On the other hand Security is an important issue for mobile users to use various features of Cloud Computing. these goals are achieved depends on the security policy adopted by the service providers.In this paper, we have discussed various security mechanisms existing in literature for mobile Clouds pertaining to user applications as well as for data storage on Cloud.

References:

[1] M. Satyanarayanan, "Mobile computing: the next decade," in Proceedings of the 1st ACM Workshop on Mobile Cloud Computing & Services: Social Networks and Beyond (MCS), June 2010.

[2] M. Satyanarayanan, "Fundamental challenges in mobile computing," in Proceedings of the 5th annual ACM symposium on Principles of distributed computing, pp. 1-7, May 1996.

[3] Armbrust Michael, Fox Armando, Griffith Rean, Joseph Anthony D., Katz Randy H., Konwinski Andrew, Lee Gunho, Patterson David A.,Rabkin Ariel, Stoica Ion, and Zaharia Matei. Above the clouds: A berkeley view of cloud computing. Technical Report UCB/EECS-2009-28, EECS Department, University of California, Berkeley, Feb 2009.

[4] M. Ali, "Green Cloud on the Horizon," in Proceedings of the 1st International Conference on Cloud Computing (CloudCom), pp. 451 - 459, December 2009.

[5] http://www.mobilecloudcomputingforum.com/

[6] G. H. Forman and J. Zahorjan, "The Challenges of Mobile Computing," IEEE Computer Society Magazine, April 1994.

[7] R. Kakerow, "Low power design methodologies for mobile communication," in Proceedings of IEEE International Conference on Computer Design: VLSI in Computers and Processors, pp. 8, January 2003.

[8] L. D. Paulson, "Low-Power Chips for High-Powered Handhelds," IEEE Computer Society Magazine, vol. 36, no. 1, pp. 21, January 2003.

[9] J. W. Davis, "Power benchmark strategy for systems employing power management," in Proceedings of the IEEE International Symposium on Electronics and the Environment, pp. 117, August 2002.

[10] R. N. Mayo and P. Ranganathan, "Energy Consumption in Mobile Devices: Why Future Systems Need RequirementsAware Energy Scale-Down," in Proceedings of the Workshop on Power-Aware Computing Systems, October 2003.

[11] http://aws.amazon.com/s3/

[12] Hoang T. Dinh, Chonho Lee, Dusit Niyato, and Ping Wang. A survey of mobile cloud computing: architecture, applications, and approaches. Wireless Communications and Mobile Computing, pages n/a–n/a, 2011.

[13] Abdulaah Alshwaier, Ahmed Youssef and Ahmed Emam "A New Trend for E-Learning in KSA Using Educational Cloud", Advanced Computing: An International Journal (ACIJ), Academy & Industry Research Collaboration Center (AIRCC), 2012.

[14] http://search.iiit.ac.in/uploads/CloudComputingForEGovernance.pdf

[15] P. Gupta and S. Gupta, "Mobile cloud computing: The future of cloud," International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering, vol. 1, issue 3, September, 2012.

A Comparative Study of Various Stream Cipher and Block Cipher Encryption Algorithm for Images

Dhatri Verma[1], Yogesh Rathore[2]
Dhatri.verma@gmail.com[1], yogeshrathore23@gmail.com [2]
M. Tech. Scholar Raipur Institute of Technology, Raipur[1], Asst. Prof. Raipur Institute of Technology, Raipur [2]

ABSTRACT
Security is an important issue in internet and network application. Nowadays various information is sent in the form of images such as online personal photograph album, medical imaging system, military image communication etc. Encryption is the technique that converts the original image into another image that is difficult to understand. In this paper we compared the various stream cipher and block cipher encryption algorithm according to their encryption quality and speed. We made comparison among most popular encryption algorithm namely DES, RC4, RC5, CAST, BLOWFISH, AES. To analyze the encryption quality and speed, correlation coefficient and throughput is used.

Keywords— encryption, correlation coefficient, stream cipher, Image security

1. INTRODUCTION
Secure image transmission over the internet is great demand by recent developments in digital image processing. So secure and reliable image encryption technique selection is an important factor for protection of data from unauthorized access, counterfeiting and tempering. In this regard various encryption algorithms have been proposed to secure the digital image content. In this paper we are analyzing the performance of various symmetric key algorithm applied on the images. In symmetric key encryption same key is used for encryption and decryption. There are two types of symmetric key encryption-1)stream cipher 2)block cipher. Stream cipher provides bit by bit or byte by byte encryption. There are various symmetric key encryption-1)DES 2)RC4 3)RC5 4)CAST128 5)BLOWFISH 6)AES. In this paper we are analyzing performance of the encryption algorithm according to encryption quality and speed.

In section 2 methodology we are describing the various symmetric key encryption algorithms. In section 3 we are showing the encryption quality and speed. Section 4 is the discussion about the result.

2. METHODOLOGY

2.1 DES (Data Encryption Standard)
Simplified DES, developed by Professor Edward chaefer of Santa Clara University. The algorithm is designed to encipher and decipher blocks of data consisting for 64 bits under control of a 64-bit key of which 56 bits are randomly generated and used directly by the algorithm[8][1]. The other 8 bits, which are not used by the algorithm, may be used for error detection. Its output is 64-bit block of ciphertext. Decryption takes 64-bit input of ciphertext analog with a 56-bit key and produces a 64-bit output of plaintext. The encryption process takes 16 rounds in which a round function, defined in terms the S-boxes, is applied over various subkeys of 56-bit input key, which are generated according to a well defined scheme. The diagram in Fig. 1 shows the

flowchart of DES. First we introduce the following notations: Let *L(x)* denote the left half of a 64-bit string x, let R(x) denote the right half of x, and let C(x) be given by

$C(x)=R(x)||L(x)$ ……..(1)

In other words C(x) changes the right and left halves of x. We explain this algorithm in the following steps:

1. An initial permutation, designated as IP, this applied to 64 bits of plaintext.

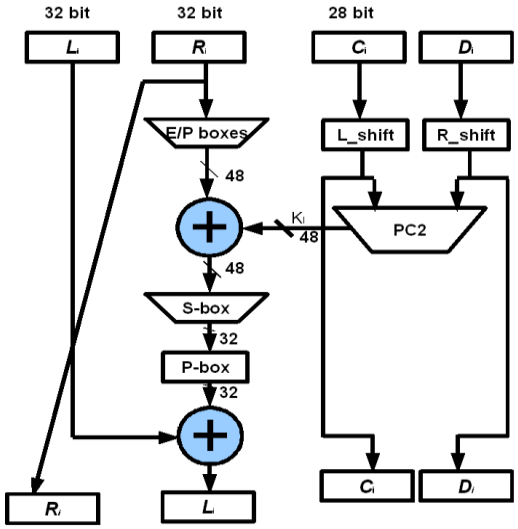


Figure 1: Data Encryption Standard Flowchart

2. This bits is split into two 32-bit halves designated L(left) and R(right).
3. At the same time, the first subkeys K1, a 48-bit string is generated.
4. The subkey K1 analog with the right halve R are used as inputs to the round function F(K;R(x)) to produce a 32-bit output, blow we explain briefly the steps of the round function F:

- Expand x from 32 bits to 48-bit, by using the expansion box E
- Apply the modulo 2 addition of E(x) and K, the output is also 48-bit.
- Where the later is concatenation of eight bit string Bi of length six, say
- $E(R(x)) \oplus K = B_1\ B_2\ B_3\ B_4\ B_5\ B_6\ B_7\ B_8$.

Enter each Bi into S-box where S-box is generated from a linear function, which takes six bits as an inputs and get four outputs.

- The output of the pervious step has a 32-bit length is entered into the permutation function P, which is defined as P box.

5. The output from the round function F is XOR-ed with the left half of the plaintext.
6. Finally, the left old half of the plaintext is replaced by the old right half, and the output of the XOR
replaces the old value of R. The function f represents this step.

$f_k(x)=(L(x)\oplus F(k,R(x))||R(x)$ …….(2)

where

$F_k(x)=P(S(E(R()x))\oplus L(x))$ ……..(3)

7. This completes one round of the DES. The same procedure is applied 15 more times, the only difference being the subkeys K2, K3,…,K15 generated by the subkey schedule are used as inputs to the round function f. Notice that when
FK16 is applied the right and left halves of the preoutput are not switched.
8. The last step of encryption is to reassemble the L and R output by the last round of fk16 of 64-bit string and apply the inverse of initial permutation IP-1

2.2 RC4

RC4 is a stream cipher, symmetric key algorithm[8][2]. The same algorithm is used for both encryption and decryption as the data stream is simply XORed with the generated key sequence. The keystream is completely independent of the plaintext used. It uses a variable length key from 1 to 256 bit to initialize a 256-bit state table. The state table is used for subsequent generation

of pseudo-random bits and then to generate a pseudo-random stream which is XORed with the plaintext to give the ciphertext. The algorithm can be broken into two stages: initialization, and operation. In the initialization stage the 256-bit state table, **S** is populated, using the key, K as a seed. Once the state table is setup, it continues to be modified in a regular pattern as data is encrypted. The initialization process can be summarized by the pseudo-code:

```
j = 0;
for i = 0 to 255:
S[i] = i;
for i = 0 to 255:
j = (j + S[i] + K[i]) mod 256;
swap S[i] and S[j];
```

It is important to notice here the swapping of the locations of the numbers 0 to 255 (each of which occurs only once) in the state table. The values of the state table are provided. Once the initialization process is completed, the operation process may be summarized as shown by the pseudo code below;

```
i = j = 0;
for (k = 0 to N-1) {
i = (i + 1) mod 256;
j = (j + S[i]) mod 256;
swap S[i] and S[j];
pr = S[ (S[i] + S[j]) mod 256]
output M[k] XOR pr
}
```

where, M[0..N−1] is the input message consisting of N bits.

This algorithm produces a stream of pseudo-random values. The input stream is XORed with these values, bit by bit. The encryption and decryption process is the same as the data stream is simply XORed with the generated key sequence.

Some of the RC4 algorithm features can be summarized as:

1. Symmetric stream cipher
2. Variable key length
3. Very quick in software
4. Used for secured communications as in the encryption of traffic to and from secure web sites using the SSL protocol.

2.3 RC5

RC5 parameters are; a variable block size (w) a variable number of rounds (r) and a variable key size (k). Allowable choices for the block size (w) are 32,64 and128 bits[8][3]. The number of rounds can range from 0 to 255, while the key size can range from 0 bits to 2040 bits in size. RC5 has three modules:

1-key-expansion,
2-Encryption
3- Decryption units

The choice of r affects both encryption speed and security. The more number of rounds will increase the security but somehow slower down the encryption speed. The RC5 algorithm uses three primitive operations and their inverses.

(1) Addition/subtraction of words modulow2, where w is the word size.
(2) Bit-wise exclusive-or denoted by XOR.
(3) Rotation: the rotation of word m left by n bits is denoted by m<<<n.

The inverse operation is the rotation of word m right by n bits, denoted by m>>>n

In the key expansion module, the password key K is expanded to a much larger size using an expansion table (T).The size of table T is 2(r+1), where r is the number of rounds. The key-expansion process must be performed before encryption or decryption processes. The encryption process takes a plain text input and produces a cipher text as the output. The decryption process takes a cipher text as the input and produces a plain text as the output. In general, the same plaintext block will always encrypt to the same cipher text when using the same key in a block cipher whereas the same plaintext will encrypt to different cipher text in a

stream cipher. Both processes use the expanded key along with segments of the input message to produce their outputs. Figure2 shows the RC5 algorithm.

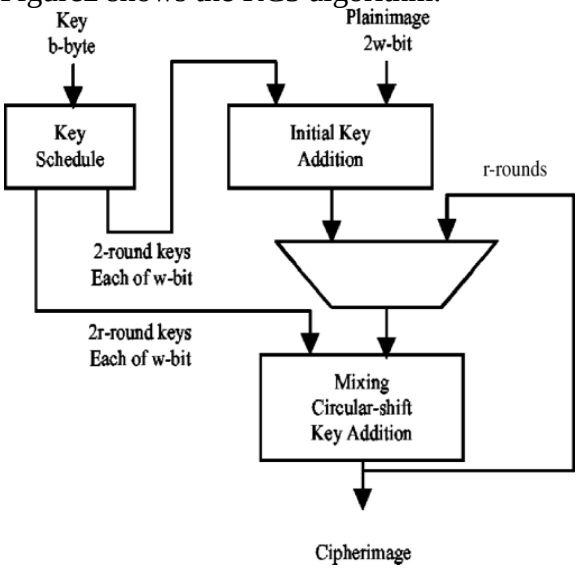


Figure 2:RC5 encryption algorithm

encryption algorithm

The two w-bit words inputs are denoted as A and B .

A = A + S[0];
B = B + S[1];
for i = 1 to r do
A = ((A ⊕B) <<< B) + S[2 * i];
B = ((B ⊕A) <<< A) + S[2 * i + 1];

decryption algorithm

The decryption algorithm can be easily derived from the encryption algorithm. The two w-bit word inputs are denoted as A and B.

for i = r down to 1 do
B = ((B –S[2 * i + 1]) >>> A) ⊕A;
A = ((A –S[2 * i] >>> B) ⊕B;
B = B -S[1];
A = A -S[0];

2.4 BLOWFISH

Blowfish was designed in 1994 by Bruce Schneier, it works on 64-bit units with key lengths from 32-bits up to 448-bits[4][8]. Each 64-bit block is divided into two 32-bit words, it encrypts every block by performing 16 rounds of encryption. Basically the algorithm consists of two parts: a key-expansion part and a data-encryption part. Key expansion converts a key of at most 448 bits into several subkey arrays totaling 4168 bytes.

data encryption:

It is having a function to iterate 16 times of network. Each round consists of key-dependent permutation and a key and data-dependent substitution. All operations are XORs and additions on 32-bit words. The only additional operations are four indexed array data lookup tables for each round.

Algorithm: Blowfish Encryption
Divide x into two 32-bit halves: xL, xR
For i = 1to 16:
xL = XL XOR Pi
xR = F(XL) XOR xR
Swap XL and xR
Swap XL and xR
(Undo the last swap.)
xR = xR XOR P17
xL = xL XOR P18
Recombine xL and xR

2.5 CAST

CAST is the first round finalist of AES competition. It is developed by Carlisle Adams and Stafford Taveres in canada, it uses 64-bit block for 64-bit and 128-bit key size variants and 128-bit block sizes for the 256-bit key version. The complete specification of CAST algorithm is given in [4][5][6]. It uses an f-function that produces a 32-bit output from a 32-bit input, and each round consists of modifying one 32-bit quarter of the block by XORing it with the f-function of another 32-bit quarter of the

block. There are 48 rounds in total, which are organized in groups of four, called quadrounds. Encryption begins with six forwards quadrounds, and then continues with six reversed quadrounds, which are reversed exactly as would be necessary for decryption. Means, for decrypting data, it is only necessary to change the order in which the subkeys are used. Figure 3 shows the CAST encryption algorithm.

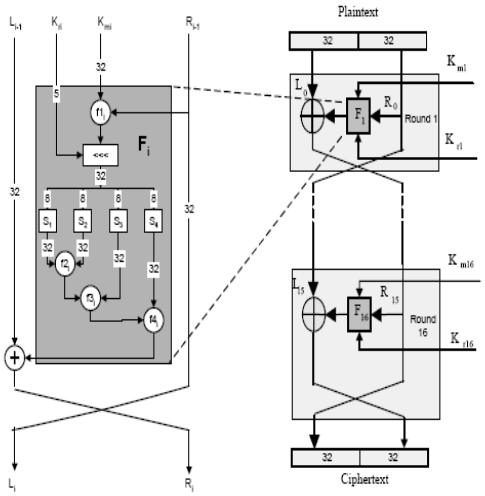


Figure 3: CAST encryption algorithm

2.6 AES(Advanced Encryption Standard)

The AES algorithm is a block-cipher operating on 128-bit data blocks supporting three different cipherkey lengths of 128, 192 and 256 bits. These three flavors of the AES algorithm are also referred to as AES-128, AES-192 and AES-256, for 128, 192, and 256-bit cipherkeys, respectively[7][8]. An AES encryption process consists of a number of encryption rounds (Nr) that depends on the length of the cipherkey. The standard calls for 10 rounds for AES-128, 12 rounds for a AES-192, and 14 rounds for a AES-256. During encryption, each round is composed of a set of four basic operations. The decryption process applies the inverse of these operations in reverse order. Figure 4 shows the basic structure of the AES encryption and decryption.

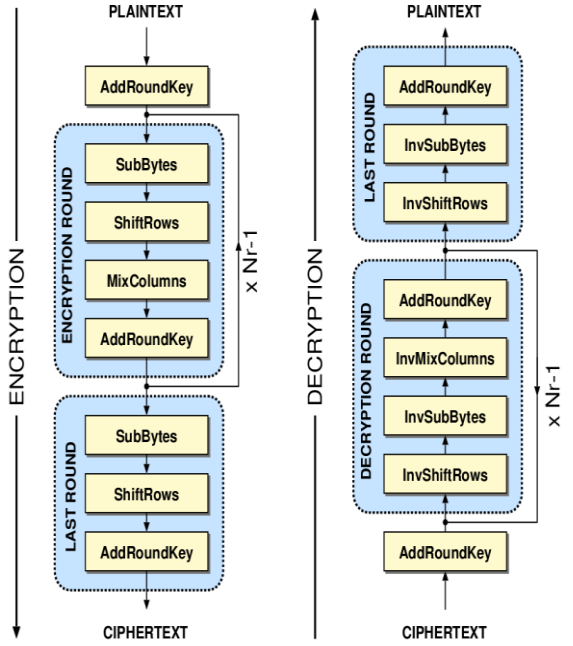


Figure 4: AES encryption algorithm

3 RESULTS

3.1 Encryption Quality Analysis

Correlation coefficient is one of the parameter for measuring encryption quality. If correlation coefficient is small then encryption quality is good.

correlation coefficient

The Correlation is a measure of the relationship between two variables. If the two variables are the plain image and cipher image, then they are in perfect correlation if they are highly dependent. In this case, the cipher image is the same as the plain image and the encryption process failed in hiding the details of the original image. If the correlation coefficient equals zero, then the

plain image and its cipher image are totally different. If the correlation coefficient equals -1, this means the cipher image is the negative of the plain image, so that the latter can be easily produced from the cipher image. Therefore, success of the encryption process corresponds to smaller absolute values of the correlation coefficient. The correlation coefficient is measured by the following equation

$$\text{The Correlation Coefficient} = \frac{\text{cov(x,y)}}{\sigma_x \sigma_y}$$

$$= \frac{\sum_{i=1}^{N}\big(x_i - E(x)\big)(y_i - E(y))}{\sqrt{(\sum_{i=1}^{N}(x_i - E(x))^2)}\sqrt{(\sum_{i=1}^{N}(y_i - E(y))^2)}}$$

Where $(E(x)=\frac{1}{N}\sum_{i=1}^{N} x_i)$ and x and y are gray-scale pixel values of the plain image and cipherimage.

Table 1 illustrate the measure of correlation coefficient [1][2][3][4][5][7]. Figure 5 illustrate the graph between encryption algorithm and their correlation coefficient.

Encryption algorithms	Correlation coefficient
DES	0.0419
RC4	0.0028
RC5	0.0049
CAST 128	0.0284
BLOWFISH	0.0348
AES	0.0170

Table 1: Encryption algorithms and their Correlation coefficient

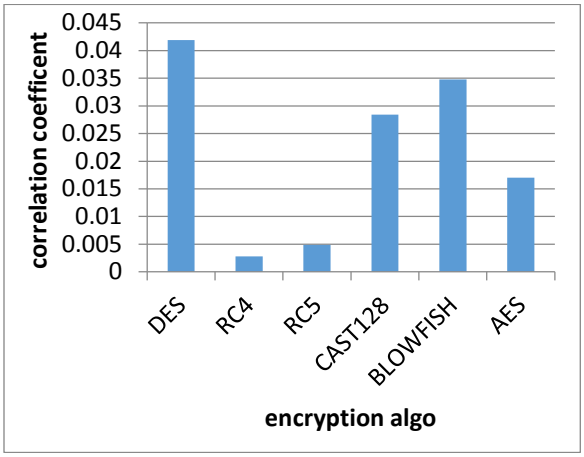


Figure 5: **Encryption quality analysis**

3.2 Speed Analysis

Here, our goal is to measure the Encryption speed of each algorithm for different packet sizes. Encryption time is used to calculate the throughput of an encryption scheme. It indicates the speed of encryption.

throughput

The throughput of the encryption scheme is calculated by dividing the total plaintext in Megabytes encrypted on the total encryption time for each algorithm.

Table 2 illustrate the throughput of various algorithm [9]. figure 6 illustrate the graph between encryption algorithm and their throughput.

Encryption algorithms	Throughput (MB/sec)
DES	32
RC4	126
RC5	75
CAST 128	55

BLOWFISH	58
AES	61

Table 2: Encryption algorithms and their Throughput

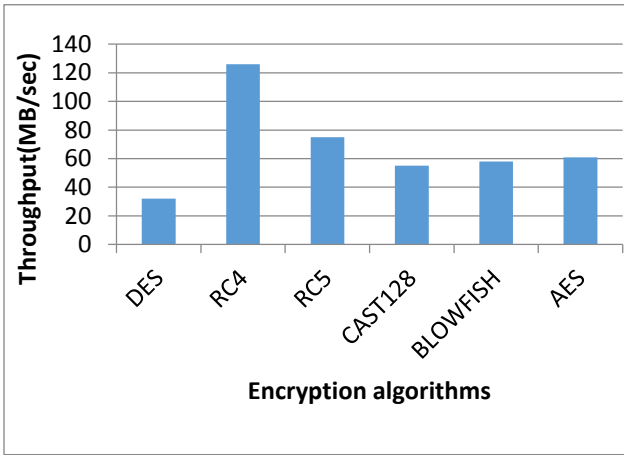


Figure 6: **Speed analysis of encryption algorithm**

4. DISCUSSION

As shown in graph we analyze that RC4 and RC5 gives better result as compared to other algorithm. RC4 and RC5 algorithm has minimum correlation coefficient and maximum throughput. If we compare between RC4 and RC5 then RC4 has minimum correlation coefficient and maximum throughput. That means RC4 is better that RC5 in encryption quality and speed.But according to [10] [11] [12] [13] [14] [15] various type of weakness has been found in RC4 like Bit flipping attack ,Key reconstruction from permutation, Biased output of RC4 etc.RC5 is more secure than RC4. So RC5 can be considered as near to a real time ,fast and secure symmetric encryption for digital imaging.

5. ACKNOWLEDGEMENT

Authors would like to thank Raipur Institute of Technology, Raipur for accomplishing the related work in this survey.

REFERENCES

[1]. Said F. El-Zoghdy, Yasser A. Nada, A. A. Abdo: " How Good Is The DES Algorithm In Image Ciphering?" . Int. J. Advanced Networking and Applications 796 Volume: 02, Issue: 05, Pages: 796- 803 (2011).

[2]. F *Shamsulddin Abdulsattar"* On the security of Bitmap Images using Scrambling based Encryption Method" *Journal of Engineering and Development, Vol. 13, No. 3,September (2009)*

[3]. Hossam El-din H. Ahmed, Hamdy M. Kalash, and Osama S. Farag Allah, "Implementation of RC5 Block Cipher Algorithm for Image Cryptosystems," International Journal of Information and Communication Engineering, 3:8, pp. 537-542, 2007

[4]. Shailaja S, Dr Krishnamurthy G N" Comparison of Blowfish and Cast-128 Algorithms Using Encryption Quality, Key Sensitivity and Correlation Coefficient Analysis" American Journal of Engineering Research (AJER) e-ISSN : 2320-0847 p-ISSN : 2320-0936 Volume-3, Issue-7, pp-161-166 www.ajer.org,2014

[5]. Krishnamurthy G N,Dr. V Ramaswamy" Encryption Quality Analysis and Security Evaluation of CAST-128 Algorithm and its Modified Version using Digital Images" International Journal of Network Security & Its Applications (IJNSA), Vol.1, No 1, April 2009

[6]. C.M.Adams, "The CAST-128 Encryption Algorithm," Request for Comments (RFC) 2144, NetworkWorking Group, Internet Engineering Task Force, May, 1997.

[7]. R.TAMILSELVI, DR.G.RAVINDRAN" Encryption Analysis and Security Using Modified Advanced Encryption Standard Based Algorithm in DICOM Images Using Entropy and Correlation" *3rd International Conference on Machine Learning and Computing,2011*

[8]. W. Stallings, Cryptography and Network Security Principles and Practices Fourth Edition, Pearson Education, Prentice Hall, 2009.

[9]. http://www.cryptopp.com/benchmarks.html (downloaded 01/04/2015)

[10]. G. Paul, S. Rathi and S.Maitra, "Non-negligible Bias of the First Output Byte of RC4 towards the First Three Bytes of the Secret Key", Proceedings of the International Workshop on Coding and Cryptography (WCC) 2007, pp. 285-294 and Designs, Codes and Cryptography Journal, pp. 123-134, vol. 49, no. 1-3, December 2008.

[11]. S. R. Fluhrer, I. Mantin, and A. Shamir, "Weaknesses in the key scheduling algorithm of RC4", In Serge Vaudenay and Amr M. Youssef,

editors, Selected Areas in Cryptography 2001, volume 2259 of Lecture Notes in Computer Science, pp. 1-24. Springer, 2001.

[12]. A. Klein, "Attacks on the RC4 stream cipher", volume 48 Issue 3, pp. 269 – 286, Designs, Codes and Cryptography, September 2008.

[13]. Souradyuti Paul and Bart Preneel "A New Weakness in the RC4 Keystream Generator and an Approach to Improve the Security of the Cipher" Appeared in Fast Software Encryption, FSE 2004, Lecture Notes in Computer Science, Springer-Verlag, pp. 245-259, 2004.

IMPLEMENTATION OF RFID IN SENSITIVE AREA OF CHHATTISGARH

Dr. Ashim Ranjan Sarkar
HOD, Dept. of CS & IT, Christ College, Jagdalpur, Bastar (C.G.)

Abstract

Radio Frequency Identification (RFID) is the radio-wave based technology to identify items automatically. One of the most popular signals is UHF (Ultra High Frequency). In UHF RFID systems Receiver is usually isolated from Transmitter by a circulator or a directional coupler. Since tags become more sensitive in order to improve reading distance, the relatively poor isolation of 20dB – 30dB limits the tag to reader link. A way to improve a directional coupler's isolation is to mismatch the unused port to generate a carrier cancelling signal. In our work we use an impedance network using digitally tunable capacitors. The advantage of this solution is that changes in the system – like antenna characteristics or cable length - can be recalibrated at any time by a controller. The objective of this paper is to present RFID technology & its applications, study various potential threats to security and privacy, and give an introduction to some suggested protocols for efficient security mechanisms in sensitive area of Chhattisgarh.

Introduction

Regarding "IMPLEMENTATION OF RFID IN SENSITIVE AREA OF CHHATTISGARH", some essential features are known: For instance that there must be a paradigm change from the – relatively – simple "identification of objects at a distance" which may suffice in the current supply chain projects, to the much more challenging "communication between objects" and the even more challenging "distributed intelligence (or Internet) of things", which implies that there must be a scalable, efficient, reliable, secure and trustworthy infrastructure, in order to link all involved objects.

Healthcare, independent living, mobility are known as "Internet of things for the citizen", putting the citizen at the centre of the new scenarios. In other words, acceptance- and "political acceptability" as well – will be of paramount importance for realising these future visions.

The full potential of intelligent environment will be fruitful if the user/consumer/citizen gives his proactive support: Transferring the "collision avoidance system" of civil aviation into millions of cars by the way of an Internet of cars will only lead to similar very positive results, if the public can be persuaded that no Big Brother shall be watching them, while their too frequent behavioural excesses are being brought under control.

The RFID research needs according to expertise within the following RFID application fields :

- Logistical tracking & tracing
- Production, monitoring & maintenance
- Product safety, quality & information
- Payment

For each RFID application field the state-of-the-art, a vision, and as result a gap analysis and the according research targets are depicted.

1 Logistical Tracking & Tracing

1.1 State of the art

Information is of primary importance in logistic processes. Improved transparency makes information on the location and availability of materials and equipment accessible to all authorized stakeholders regardless of their location. This development is promoted by the implementation of automatic identification and tracking systems. Classical inventory systems keep track of lots or batches, whereas RFID systems allow the unique identification of items and a shift to a 'per unit' or serial-level inventory system, which allows the tracking of individual items throughout the supply chain. This poses new requirements for the existing IT systems. Integration of the RFID system in existing IT systems, such as ERP systems, is critical to the success of the implementation. Different commercial applications are available.

In this context, RFID can be seen as an enabling technology for automatic tracking and tracing systems. To date, RFID has slowly started to replace and complement manual labels, trading stamps, barcodes as well as methods based on optical character recognition. However, barcode technology is by far the most widely adopted technology for identification in logistics applications.

RFID applications have mostly concerned separate, closed and in-house systems. The use of passive UHF RFID technology is increasing the most rapidly, due to the low price, good standardisation situation and sufficient performance (3-4 m reading range). Due to their increased use, UHF RFID tags which are designed for specific applications (e.g. mounting on metal) have also become commercially available. A growth in active RFID technology is foreseen in RTLS (Real Time Location Systems), due to improved standardisation, lower prices and the possibility for integration with existing IEEE 802.11 (WLAN) infrastructure.

More pressure and demand for new identification solutions are generated by the new modes of operation such as pull-based flow control, reducing inventories, legislation (tracking of food, pharmaceuticals etc.), cost reduction, automation, communication between partners and visibility of the supply chain, decreasing waste and improving security. The main factors that have slowed down RFID implementation are the investment costs which are regarded as high and the uncertain expectations for the return on investment (ROI). There is not enough or not specific enough information about the benefits, especially for manufacturers and suppliers of parts, who bear the cost of tagging. On the hardware side especially the UHF readers are rather expensive. Unsuccessful pilots with low readability rates have created an image of immature technology for passive UHF RFID technology and delayed implementation, although there are also reports of continuous improvements in the technology, especially when the latest firmware modifications are applied to readers.

1.2 Vision

Automatic identification develops towards systems, which can read the IDs of all products automatically without the need to stop the process or manual intervention. The technology used has to be able to read each of the individual tags. The long term vision in logistics is a system providing the necessary real-time information on the supply chain, this information being extensively utilised. RFID is the key technology to make this happen. Real-time information on the location, contents and conditions of individually identified shipments, products, transport units, and

transport vehicles can be gathered in a controlled manner. The collected data can be combined with the planning information and processed into appropriate information to be used at different stages of the process. Product history will become available via B2B networks. The information can be distributed effectively and in real time to the stakeholders.

RFID systems provide item and product visibility within the supply chain. This visibility can further be translated into actionable data and predictive changes with additional information attained through sensors.

1.3 Gap Analysis

Doubts about the maturity of technology and standards were still the biggest concerns of potential users. Also costs and benefits, which cannot be verified, slow down the implementation in logistics applications. According to the RFID industry, improvements in standardisation and performance in the recent years make RFID ready enough for implementations, but this message has not yet been assimilated by potential users.

In the logistics sector, the introduction of RFID technology is expected to start from the side of retail but a better target could be e.g. technical wholesale. Retail trade has also bad experiences from earlier pilots. Consumer Packaged Goods, consumer electronics, healthcare, pharmaceutical, aerospace and defence and high-tech industries will be the first adopters.

The organizational management has the key role in investment decisions and implementation. The benefits and ROI should be justified to them. In any case RFID will be implemented first in closed systems. Organisations should put effort into wider and more open logistics applications and multi actor supply chains.

2 Production, Monitoring & Maintenance

2.1 State-of-the-Art

Most of the RFID-based solutions used today in the field of manufacturing & maintenance are limited to tracking and tracing of parts or tools. Compared with the previous identification means, RFID provides a fast and automatic capability to identify parts without the need to see or contact the ID. However, the reading performance of RFID tags (reliability, range…) for several products/processes is not yet totally sufficient to guarantee the economic benefit of adopting RFID compared with the well established systems using barcode, 2D matrix or even name plates with written part or serial number.

Most of the RFID-based solutions are either "closed loop" (RFID is not used outside a company) or "semi closed-loop" (RFID is not used outside a supply chain), although "open loop" solutions (RFID is used anywhere during the product lifetime) would be of great use, especially for maintenance purposes. Monitoring of product usage and status during their service life is important to allow the development of condition-based or predictive maintenance services. RFID tags are a key element to this since they can wirelessly communicate the identity of a unique product to an information system where this identity can be merged with other data and processed. However, the available RFID technology has not yet the performances to allow industrial use: in terms of robustness, temperature capability, sensing and communication functions, miniaturisation.

2.2 Vision

Manufacturing and maintenance will benefit from RFID-based solutions developed for track and trace of components allowing better automation of supply chain and resource management. Beyond this, further

benefits from RFID-based solutions will be obtained by using the unique capability given by RFID tags to be read wirelessly and automatically without human triggering and processing in order to inform the surrounding information system in real time of its identity, through which a connection can be made to possibly significant data content directly "attached" to it. Note that in some scenarios, additional data will be recorded to an RFID tag, while in other scenarios, a simpler tag is used, carrying a unique ID or 'licence-plate', which is then used for identifying sources of information on the network, as well as retrieving relevant records, since the unique ID read from the tag can also be used as a database key.

The networking of information systems will then make it possible to know in real time, from anywhere, the configuration of an individual product throughout its lifetime, to access necessary information on this product and provide tailor-made added value services.

2.3 Gap Analysis

Before this vision can be realized, four main types of gaps need to be filled:

System integration

- To take full benefit of the RFID tags, key problems will be integration: to shift from isolated- RFID-solutions to global networked-solutions integrating the RFID paradigm into the diverse systems, while maintaining the quality of services of the ICT systems and the continuous workflow in the shop floor or in the field.
- This integration of RFID-enabled systems (e.g. tags for equipment, machines, operators, products, material, and consumables) in a global service network introduces new challenges regarding the support of diverse and heterogeneous data models, which are imposing diverse demands on data exchange and (intelligent) transformation, due to the decentralised and partly new requirements for data interpretation.
- Currently, RFID technology is handled as an add-on and not fully integrated in the overall product life-cycle, still imposing additional efforts to integrate RFID with the product itself (e.g. attaching RFID readers in mobile phones, mapping RFID specific data to manufacturing/ maintenance data) within the manufacturing or maintenance process.

Standards and regulation

- These RFID data models and ontologies can only be adopted in an efficient way if proper standards and standardisation roadmaps exist to give directions and confidence to all technology providers and end users.
- As mentioned above, RFID will enable remote monitoring and autonomous behaviour of products and machines. Generic solutions will be needed to assure the integrity of services, especially when considering security of users and workers, and the cost of downtime of products and machines. These solutions will need to be reinforced by standards and regulations.
- There are generally no models for data-sharing among multiple partners, taking into account intellectual property, data ownership and lifetime management, company confidentiality or user privacy requirements, and multi level access rights issues. Connecting in real time the identity and status of an individual component to an information network will certainly worsen this gap.
- In the same way worldwide standards need to converge and stabilize for the RFID technology itself (frequencies, power, read protocols etc…).

Confidence of stakeholders

- Before decision makers launch long term investments in a RFID-enabled solution, they need to get confidence in the

viability and durability of the solution. This is not the case today: RFID related technologies (hardware, middleware and system integration…) are evolving, standards are in progress and many legal questions remain unanswered.

- Several aspects of the business model of RFID-based systems are difficult to quantify – for example: evaluation and sharing of the added value (especially information sharing), actual practical performance of RFID tags, cost of maintaining a RFID-based system operating reliably in harsh industrial environments. This evaluation is made even more difficult by the lack of education, the difficult access to reliable information and to experts who are independent from technology solution providers.
- Difficulty to define a tag and attachment solution for the complete lifetime of a component.
- Difficulty to read tags in an environment strongly constrained by the process (metallic surfaces, complex shapes, limited space available for readers, etc.). The use of active tags is limited by the difficulty to install and maintain batteries during the required service life of the tag or part to which the tag is attached.
- Lack of sensing and communication capability of tags
- Lack of multi purpose readers (hardware and middleware) with the capability to interface to several kinds of tags with several applications. At the moment multi-purpose readers are only available as stationary gates; still missing are e.g. multi-purpose readers in mobile phones or planes.

3 Product Safety, Quality and Information

3.1 State of the art

At the current point in time, RFID technology is beginning to pervade the domain of trade with tags being introduced mostly to logistical units such as pallets, but not so much on the item level, yet. There are already several examples of pilot installations also targeted at consumers, most notably the Future Store initiative of the Metro group, which is pioneering the trials of RFID related applications for the benefits of the end consumers.

Manufacturing and production industries are in general still quite hesitant regarding the adoption of RFID technologies for reasons of missing interoperability standards and also because of lack of knowledge about best practices.

In general, most producers lack both the knowledge of the respective processes (esp. with regard to safety regulations) and the potential connections to related technologies that make sense for integration in the field of product safety and quality such as temperature or humidity sensors.

3.2 Vision

When we regard the issues of product safety and especially product quality, we have the clear objective to provide detailed information about the history of a product to the end consumer. This will help to create trust and transparency in sensitive product areas such as perishable or sensitive goods, but also pharmaceutical and luxury goods and high-value goods, some of which are highly composite products. Depending on the consumer's context and demands, correct and complete information has to be provided.

The implementation of RFID enabled data exchange infrastructures at a broad level will significantly and sustainably influence the importance and the future development of all related IT systems.

3.3 Gap Analysis

While the integration of RFID technologies with sensing infrastructures for product quality information has already been

implemented at several pilot installations in the trade domain, these are mostly prototypical proof-of-concept studies, and in the future, we can even expect to see substantial development of hardware-related aspects such as production of small but reliable sensors with sufficient noise rejection, as well as lightweight, efficient and robust power supplies.

Commercially available, standardized sensing infrastructures that scale to the requirements of large supply chains and can be used throughout the complete product lifecycle are still to emerge. Likewise, cross-industry standards are not yet fully available.

While many prerequisites (data models, transaction protocols etc.) are already present or could be developed with today's premises, their combination, standardization and establishment of best practices in communication across company borders still requires much effort and field experience.

3.4 Resulting Research Targets

In order to close the identified gaps and realize the RFID solutions vision, we need to advance the state of the art in several technological areas as well as solve standardization issues.

Regarding several of the security aspects that are a precondition to establishing cross-organizational RFID enabled businesses (e.g. who is allowed to access which event data, how is counterfeiting being addressed etc.), there is probably a need for respective legislations and also for appropriate security technologies to be at the heart of the IT systems for inter-organizational information exchange, to support authentication and enforce access control policies, as well as assuring the integrity and non-repudiation of the information that is exchanged.

As with many other application domains, the costs for the IT infrastructures, but mostly for the tags themselves, need to be reduced in order to move from pallet level tagging to smaller quantities of packaging and down to the item level in the long term. Related to this is research on novel tag technologies that work well with different materials such as metal or textiles and in general more work on the next generation of RFID tags that include or integrate sensing technologies and embedded computer platforms to implement the "smartness" necessary for certain advanced business processes.

4 Payment

Payment applications have many interesting new technologies and development directions. Closed-end payment schemes, interoperability of ticketing and payment instruments, mobile payment applications, more stable security solutions, electronic purse and e-cash are just a few ones to be mentioned. From this list we selected electronic cash as the research subject in focus, because it has the highest relevance due to its economic impact and its strong reliance on contactless technology.

5 Conclusion

In this paper, we have considered the current state of RFID application in some important application areas. We described potential benefits of an intensified use of RFID tags in these areas, and are coming up with an analysis of the causes that prevent this intensification. This analysis should help researchers and industry to direct their efforts to advance RFID technology itself and its application. In order to identify the most pressing problems that promise the most advancement, if overcome, at a reasonable investment, we tried to judge the issues encountered in the application fields according to additional criteria. First, it must be noted that a rating as it is done is necessarily subjective and highly influenced by the priorities and working areas of the authors. Therefore, our conclusions drawn

here should be taken by the reader as a supplement to provide a stronger basis of her or his own assessment. The second criterion is the immaturity level, thus prioritizing those issues that need a stronger investment to yield results. If it is easy to read the table the other way around and consider those issues first that require less investment for good results.

OBJECT ORIENTED DESIGN PATTERN DETECTION USING STATIC ANALYSIS IN SCALA SOFTWARE

Ajay Singh Thakur
Asst. Prof. Christ College Jagdalpur, Bastar(C.G.)

1. Abstract: Many fields use patterns in various ways: In music and literature, a pattern is the coherent structure or design of a song or book. In art, a pattern is the composition or plan of a work of graphic or plastic art. In architecture, a pattern is an architectural design or style. In psychology, a pattern is a thinking mechanism that is basic to the brain's operation, helping oneto perceive things quickly.
With each pattern, small piecework is standardized into a larger chunk or unit. Patterns become the building blocks for design and construction. Finding and applying patterns indicates progress in a field of human endeavor. Design patterns represent the best practices used by experienced object-oriented software developers. Design patterns are solutions to general problems that software developers faced during software development. These solutions were obtained by trial and error by numerous software developers over quite a substantial period of time.
Software design patterns abstract reusable object-oriented software design. Each pattern solves design problems that occur in every day software development. The detection of design patterns during the process of software reverse engineering can provide a better and faster understanding of the software system.

2. Introduction: In 1994, four authors Erich Gamma, Richard Helm, Ralph Johnson and John Vlissides published a book titled **Design Patterns - Elements of Reusable Object-Oriented Software** which initiated the concept of Design Pattern in Software development. These authors are collectively known as **Gang of Four (GOF)**. According to these authors design patterns are primarily based on the following principles of object orientated design.

- Program to an interface not an implementation
- Favor object composition over inheritance

Design Patterns have two main usages in software development.

Common platform for developers :-
Design patterns provide a standard terminology and are specific to particular
scenario. For example, a singleton design pattern signifies the use of single object
so all developers familiar with single design pattern will make use of single object
and they can tell each other that program is following a singleton pattern.

Best Practices :-
Design patterns have been evolved over a long period of time and they provide best solutions to certain problems faced during software development. Learning these patterns help unexperienced developers to learn software design in an easy and fast way.
As per the design pattern reference book **Design Patterns - Elements of**
Reusable Object-Oriented Software, there are 23 design patterns which can be
classified in three categories: Creational, Structural and Behavioral patterns.
1. **Creational Patterns** These design patterns provide a way to create objects
while hiding the creation logic, rather than instantiating objects directly
using new operator. This gives more flexibility to the program in deciding

which objects need to be created for a given use case.

2. **Structural Patterns** These design patterns concern class and object
composition. Concept of inheritance is used to compose interfaces and
define ways to compose objects to obtain new functionalities.

3. **Behavioral Patterns** These design patterns are specifically concerned
with communication between objects.[1]

2.1 Static Analysis

Static analysis, also called static code analysis, is a method of computer program debugging that is done by examining the code without executing the program. The process provides an understanding of the code structure, and can help to ensure that the code adheres to industry standards. Automated tools can assist programmers and developers in carrying out static analysis. The process of scrutinizing code by visual inspection alone (by looking at a printout, for example), without the assistance of automated tools, is sometimes called program understanding or program comprehension. The principal advantage of static analysis is the fact that it can reveal errors that do not manifest themselves until a disaster occurs weeks, months or years after release. Nevertheless, static analysis is only a first step in a comprehensive software quality control regime. After static analysis has been done, dynamic analysis is often performed in an effort to uncover subtle defects or vulnerabilities. In computer terminology, static means fixed, while dynamic means capable of action and/or change. Dynamic analysis involves the testing and evaluation of a program based on execution. Static and dynamic analysis, considered together, are sometimes referred to as glass-box testing.

2.2 Scala

Scala is a modern multi-paradigm programming language designed to express common programming patterns in a concise, elegant, and type-safe way. Scala has been created by Martin Odersky and he released the first version in 2003. Scala is an acronym for "Scalable Language". This means that Scala grows with you. You can play with it by typing one-line expressions and observing the results. But you can also rely on it for large mission critical systems, as many companies, including Twitter, LinkedIn, or Intel do. To some, Scala feels like a scripting language. Its syntax is concise and low ceremony; its types get out of the way because the compiler can infer them. There's a REPL and IDE worksheets for quick feedback. Developers like it so much that Scala won the ScriptBowl contest at the 2012 JavaOne conference. At the same time, Scala is the preferred workhorse language for many mission critical server systems. The generated code is on a par with Java's and its precise typing means that many problems are caught at compile-time rather than after deployment. At the root, the language's scalability is the result of a careful integration of object-oriented and functional language concepts.

3. Problem Definition

Software design patterns abstract reusable object-oriented software design. Each pattern solves design problems that occur in every day software development. The detection of design patterns during the process of software reverse engineering can provide a better and faster understanding of the software system. The first time the term design pattern was related to software development was in the book Design Patterns - Elements of Reusable Object-Oriented Software written by Erich Gamma, Richard Helm, Ralph Johnson and John Vlissides(1). The group of authors is also known as the Gang-of-Four or GoF. They discuss object-oriented design techniques that are based on their experience as software developers. They introduce 23 design patterns using a consistent format of information

for each pattern to provide a uniform structure. Using the intent, trade-offs and graphical notations for design patterns, software engineers can decide which design pattern solves their design problems. It also makes it easier to discuss design problems and solutions with colleagues by using a common vocabulary. It is also useful documenting which design patterns have been used in a software so that other developers will get a better overview of the software without having to read the source code in detail. This thesis is concerned with the problem of detecting software design patterns. We present an approach that will detect software design patterns using their static structure - as described in class diagrams - as well as their dynamic behaviour. The results can be used to verify the implementation of design patterns that were specified before the implementation phase. Having this additional information can be crucial during software maintenance. If well-designed software is poorly documented then this good design might be broken by a different developer that needs to add more functionality. These changes might introduce new bugs and problems and lead to degradation of the software. Therefore, it is important to document these design choices to improve the understanding of the software.

4. Literature Review

The literature review is organized by the following categories: System Dynamics, ObjectOriented Design, Design Patterns. The first three categories are intended to provide a basis of comparison using foundation statements about each of the disciplines

I. System Dynamics
II Object-Oriented Design
III Design Patterns

4.1 System Dynamics Background

Forrester set the cornerstone for the structure of System Dynamics and it has stood the test of time. In Principles of Systems, Forrester (1990) states that structure is essential if we are to effectively interrelate and interpret our observations in any field of knowledge: "Without an organizing structure, knowledge is a mere collection of observations, practices, and conflicting incidents". It is the structure of a subject that guides us in organizing information. "If one knows a structure or pattern on which he can depend, it helps him to interpret his observations. An observation may at first seem meaningless, but knowing that it must fit into one of a limited number of categories helps in the identification. Structure exists in many layers or hierarchies. Within any structure there can be substructures".

Likewise, Bruner (1960) tells us that it is the understanding of the structure of a subject that allows many other things to be related meaningfully. Bruner (1960) tells us that learning through the transfer of principles is dependent upon mastery of the structure of a subject. Understanding the fundamentals makes a subject comprehensible. Human memory is dependent upon structured patterns for recall. Understanding the specific case of a structure is a model for understanding other things like it that one may encounter. The constant re-examination of material's structure results in a narrowing of the gap between advanced and elementary knowledge.

4.2 Object-Oriented Background

The objected-oriented paradigm captures system and software engineering work product in frameworks of packages, classes, objects, and methods. The language of the customer is captured by Use Cases as a statement of requirement and concept of operation. Leveraging the objectoriented paradigm to System Dynamics models may lead to benefits such as better understood: models, software design, and reusable software model libraries

4.3 Design Patterns

Design patterns describe the key ideas in the system, Fowler and Scott (1997). Patterns help explain why a design is the way it is. The design pattern represents the fundamental algorithm being implemented by the software, an algorithm that is repeated in many other designs. Vlissides et al., (1995) characterize design patterns as a description of communicating objects and classes that are customized to solve a general design problem in a particular context. The design pattern names, abstracts and identifies the key aspects of a common design structure that makes it useful for creating a reusable object-oriented design. Design patterns describe simple and elegant solutions to specific problems, Vlissides et al., (1995). Design patterns capture designs that have developed and evolved over time; they reflect extensive redesign and recoding as developers have striven for greater reuse and flexibility in their software, Vlissides et al., (1995).Vlissides et al., (1995) say that a pattern has four essential elements:

I. The Pattern Name describes a design problem, its solutions, and consequences in a word or two. The name allows one to design at a higher level of abstraction. The vocabulary of pattern names facilitates dialog. The name enables thinking about good designs and communicating them and their trade-offs to others.

II. The Problem describes the criteria for when to apply the pattern. Occasionally, the problem will include a list of conditions that must be met before it makes sense to apply the pattern.

III. The Solution contains the elements that make up the design, their relationships, responsibilities, and collaborations. The solution is not a particular concrete design or implementation but a template that can be applied to many different situations.

IV. The Consequences are the results and trade-offs of applying the pattern. These are important for evaluating design alternatives and understanding the costs and benefits of applying the pattern.[1]

5. Approach

Design patterns have their own unique intent and are described with roles and responsibilities.In the source code, each role is commonly represented by a class andthe responsibilities are coded in the classes with attributes and methods. The patternsalso describe the collaboration between objects at runtime. In this thesis, weuse the roles, responsibilities and collaboration information to analyze applications,detect design patterns and rank the results by their classification. Understandingthe structure and intent of the software system will give the developer a fasteroverview of the whole system without going into details of the source code.

In order to detect design patterns in Scala software, we used several existing toolsand technologies.

5.1 Eclipse

We developed our software using Eclipse. Eclipse is an open-source software frameworkwritten in Java. In its default form it is a Java Integrated Development Environment(IDE), comprised of the Java Development Toolkit (JDT) and compiler(ECJ). Users can easily extend its capabilities by installing plug-ins written for theEclipse software framework, such as development toolkits for other programminglanguages, and can write and contribute their own plug-in modules. Over the yearsit has become the most popular IDE for developing Scala programs. The portabilityand large number of plug-ins make it the standard development platform for mostdevelopers.In our case we are using one additional plug-in for Eclipse. We use the plug in fromthe Test and Performance Tools Platform Project (TPTP) which allows us to buildtest and performance tools, such as debuggers, profilers and benchmarking applications.

6. Bibliography

1. Erich Gamma, Richard Helm, Ralph Johnson and John Vlissides “Design

Patterns - Elements of Reusable Object-Oriented Software".

2. Booch, G. *Object-Oriented Design with Applications.* Benjamin/Cummings, Redwood City, Ca., 1991.

3. Coad, P. and Yourdon, E. *Object-Oriented Analysis.* Second ed. Prentice Hall, Englewood Cliffs, N.J., 1991.

4. Coad, P. and Yourdon, E. *Object-Oriented Design.* Prentice Hall, Englewood Cliffs, N.J., 1991.

5. Goldberg, A. Information models, views, and controllers. *Dr. Dobb's J.* (July 1990).

6. Johnson, R. and Wirfs-Brock, R. Object-oriented frameworks. Tutorial notes. In *Proceedings of ACM OOPSLA* (1991).

7. Leibs, D. and Rubin, K. Reimplementing model-view-controller. *The Smalltalk Report* (Mar./Apr. 1992).

8. Rumbaugh, J., Blaha, M., Premerlani, W., Eddy, F. and Lorensen, W. *Object- Oriented Modeling and Design.* Prentice Hall, Englewood Cliffs, N.J., 1991.

9. *Webster's Third New International Dictionary.* Merriam Webster, Inc., 1986.

10. Wirfs-Brock, R., Wilkerson, B. and Wiener, L. *Designing Object-Oriented Software*. Prentice Hall, Englewood Cliffs, N.J., 1990.

11. Ahmed, U. (1997). A process for designing and modeling with components. Proceedings of the 15th International System Dynamics Society. Turkey: System Dynamics Society.

12. Alexander, C., Ishikawa, S., Silverstein, M., Jacobson, M., Fiksdahl-King, I., and Angel, S., (1977). A pattern language. New York: Oxford University Press.

13. Basnet, C., Farrington, P., Pratt, D., Kamath, M., Karacal, S., Beaumariage, T. (1990). Experiences in developing an object-oriented modeling environment for manufacturing systems. Proceedings of the 1990 Winter Simulation Conference. IEEE.

14. Bishak, D. & Roberts, S. (1991). Object-oriented simulation. Proceedings of the 1991 Winter Simulation Conference. IEEE.

15. Braude, E. (1998). Towards a standard class framework for discrete event simulation. Proceedings of 31st Annual Simulation Symposium.

16. Bruner, J. (1960). The process of education. Boston: Harvard University Press.

17. Corbin, D. (1994). Integrating archetypes and generic models into a framework for model conceptualism. Proceedings of the International Systems Dynamics Society. Sterling: System Dynamics Society.

18. Rudolf K. Keller, ReinhardSchauer, SebastienRobitaille, and Patrick. Pattern-based reverse-engineering of design components.

19. Kyle Brown. Automated design pattern detection in smalltalk.

International Seminar
" Emerging Trends in IT and Applied Science" on 28th to 30th March, 2015"
MATS School of Information Technology, MATS University, Raipur

A Review on Performance of different Mathematical Algorithms in Parallel Environment

Dilip Kumar Nayak,Mr. Avinash Dhole
Raipur Institute of technology, Chhatouna, Chhattisgarh 492101 India
1dilipnayak989@gmail.com
2Avi_dhole33@rediffmail.com

Abstract -In Present day, multicore systems have become popular as it provides parallelism and hence less delay, but multicore systems provide only hardware parallelism. In order to achieve best result we should use software parallelism also. To achieve software parallelism there are many programming model like OMP, MPI etc. Now a day's high performance computing mainly center's around parallel computing. Parallel computing is the ability to carry out multiple operations or tasks simultaneously. Ideally, parallel processing makes programs run faster because there are more engines (CPUs or cores) to run the program. Sequence and series like sin series ,cos series ,arithmetic progression, geometric progression ,harmonic progression are most frequently used in mathematical tools. Therefore the parallel computation is an efficient way to improve the performance. By putting some constraints on the data and taking the advantage of the hardware. The performance of the different sequential and serial algorithms can be significantly improved. This paper provides the review of various mathematical algorithms which has developed parallely.

***Keywords*—sequence and series, Geometric Progression(g.p.),Harmonic progression(h.p.) Arithmetic Progression (a.p.), Parallelism.**

I. INTRODUCTION

The field of numerical analysis predates the invention of modern computers by many centuries .Linear interpolation was already in use more than 2000 years ago. Invention of the computer also influenced the field of numerical analysis ,since now longer and more complicated calculation could be done [1].

Geometric Progression is a mathematical tool which is designed for solving many problems like repeating decimals , Archimedes' quadrature of the parabola, Fractal geometry, Zeno's paradoxes, Euclid, Economics ,power series in taylor theorem etc[10]

A geometric progression is a sequence of numbers where each term after the first is found by multiplying the previous term by a fixed number called the common ratio. The sequence

1, 3, 9, 27 . . .

is a geometric progression with first term 1 and common ratio 3. The common ratio could be a Fraction and it might be negative.

In general we can write a geometric progression(g.p.) as follows : a, ar, ar^2, ar^3 .the nth term of a g.p. is given by : $\mathbf{ar^{(n-1)}}$ The sum of the first n terms of a g.p. is $Sn=a(1-r^n)/(1-r)$ valid only if $r \neq 1$.The sum of the terms of a geometric progression is known as geometric series.[2]

Today's the parallel algorithms are focusing on multi-core systems. The design of parallel algorithm and performance measurement is the major issue on multicore environment. If one wishes to execute a single application faster, then the application must be divided into subtask or threads to deliver desired result.

II. VARIOUS MATHEMATICAL TECHNIQUES.

A. Mr D.S. Ruhela and Mr R.N.Jat, "Complexity & Performance Analysis of Parallel Algorithms of Numerical Quadrature Formulas on Multi Core system Using Open MP)".(2014)

The authors studies two version of numerical quadrature algorithms: sequential and parallel.In the experiments the execution times of both the sequential and parallel algorithms have been recorded to measure the performance (speedup) of parallel algorithm against sequential. The result obtained shows a vast difference in time required to execute the parallel algorithm and time taken by sequential algorithm. Based on their study They concluded that parallelizing serial algorithm using Open MP has increased the performance. For multi-core system Open MP provides a lot of performance increase and parallelization can be done with careful small changes. The parallel algorithm is approximately twice faster than the sequential and the speedup is Linear [1].

B. M. F Mridha, Mohammad Manzurul Islam, Syed Mohammad Oliur Rahman. "A New Approach of Performance Analysis of Certain Graph Algorithms "(2013).

In this paper authors presented a new parallel Prim algorithm that grows multiple trees in parallel. They made simple observations based on the cut property of the graph to grow MSTs in parallel. Their algorithm achieves reasonable speedup when it is compared with Serial Prim algorithm for dense graphs and sparse graphs. Breadth First Search and Prim's Algorithm's parallel implementations using CUDA/C was successfully done. The Speedup computed helped realize performance improvements by the use of parallel algorithms. In case of breadth first search algorithm in parallel when graph is sparse speed up is 2.0 while that of when graph is dense 1.9. As for as prim's is concerned speedup is at the minimum of 1.96 i.e 2.0 more when at least 2 threads are used.. [4].

C. *Mr. Nagraj and Mr. Kumarasvamy,"Serial and Parallel Implementation of Shortest Path Algorithm in Optimization Of Public Transport Travel",(2011)*

This paper suggests execution of 3 shortest path algorithms (Dijkstras's algorithm, Bellman Ford algorithm and Ant-Colony algorithm) serially and parallely (Using OMP).And shows the result that time cost of multithreaded parallel algorithm on dual core system are much faster than the serial algorithm. The parallel running speed can be improved with the increase of number of cores[5].

D. *Manwade K. B "Analysis of Parallel Merge Sort Algorithm "(2010)*

The algorithm has been tested on loosely coupled parallel machines and the performance of the algorithm has been observed. It has been found that the computational time of the algorithm varies logarithmically for varying number of processors scenario. Also it is found that for varying number of elements the computational time varies linearly. It is also found that the practical analysis closely matches with theoretical analysis [6].

E. *Mr. Sanjay Kumar Sharma and Dr. Kusum Gupta "Performance Analysis of Parallel Algorithms on Multi-core System using OpenMP",(2012).*

Authors have studied the typical behavior of sequential algorithms and identified the section of operation that can be executed in parallel. and presented the execution time of both serial and parallel algorithm for computation of Pi value. They concluded that the parallelizing serial algorithm using OpenMP has increased the performance and for multi-core system OpenMP provides a lot of performance increase and parallelization can be done with careful small changes. And at last the parallel algorithm is approximately twice faster than the sequential and the speedup is linear [7].

F. *Pranav Kulkarni and Sumit Pathare "Performance Analysis of Parallel Algorithm over Sequential using OpenMP"(2014)*

The author studied some algorithms like matrix multiplication and Floyd-Warshell Algorithm and found that the algorithms with small data set gives good performance when executed by a sequentially programming. But as data set increases performance of sequential execution falls down where parallel execution is used for large data set then it gives best results than sequential execution. [8].

G. *Kil Jae Kim, Seong Jin Cho and Jae-Wook Jeon "Parallel Quick Sort Algorithms Analysis using OpenMP 3.0 in Embedded System",(2011)*

Authors have studied the parallel quick sort algorithms and also analyze the parallel quick sort algorithms using OpenMP and also analyze the effect of parallelization considering memory size and number of cores. Then shows the result that parallel quick sort algorithm performance drops roughly 10~20% even if it has exactly same algorithm [9].

III. CONCLUSIONS

Many researchers have analyzed the serial and parallel effect of various mathematical algorithm. In this paper, we have summarized the effect of parallelism in existing algorithms. and Mismatch problem between software and hardware will be tackled because we will make use of multicore, and by using multicore we will try to get higher speed up, throughput and utilization etc.

REFERENCES

[1] D.S. Ruhela and R.N.Jat ." Complexity & Performance Analysis of Parallel Algorithms of Numerical Quadrature Formulas on Multi Core system Using Open MP" Volume 3 Issue 7 July, 2014 International Journal Of Engineering And Computer Science

[2] Sequences And Series Http://Ltcconline.Net/Greenl/Courses/103b/Seqseries/Seqser.Htmanization" Tmh.As Retrieve On Date 10/01/15.

[3] Http://Www.Wyzant.Com/Help/Math/Precalculus/Series And Sequences. As Retive On Date 10/01/15

[4] M. F Mridha, Mohammad Manzurul Islam, Syed Mohammad Oliur Rahman. " A New Approach of Performance Analysis of Certain Graph Algorithms "International Journal of Advanced Research in Computer and Communication Engineering Vol. 2, Issue 9, September 2013

[5] Mr. Nagraj and Mr. Kumarasvamy,"Serial and Parallel Implementation of Shortest Path Algorithm in Optimization Of Public Transport Travel",international jounal of computer science engineering and information technoo

[6] Manwade K. B "Analysis of Parallel Merge Sort Algorithm " *International Journal of Computer Applications (0975 - 8887) Volume 1 – No. 19 –(2010)*

[7] Mr. Sanjay Kumar Sharma and Dr. Kusum Gupta "Performance Analysis of Parallel Algorithms on Multi-core System using OpenMP", International Journal of Computer Science, Engineering and Information Technology (IJCSEIT), Vol.2, No.5, October 2012.

[8] Pranav Kulkarni and Sumit Pathare "Performance Analysis of Parallel Algorithm over Sequential using OpenMP" IOSR Journal of Computer Engineering (IOSR-JCE) Volume 16, Issue 2, Ver. X (Mar-Apr. 2014)

[9] Kil Jae Kim, Seong Jin Cho and Jae-Wook Jeon "Parallel Quick Sort Algorithms Analysis using OpenMP 3.0 in Embedded System", International Conference on Control, Automation and Systems Oct. 26-29, 2011 in KINTEX, Gyeonggi-do, Korea

[10] http://en.wikipedia.org/wiki/Geometric_series#Applications .As retrieve On Date 20-3-2015

\

International Seminar
" Emerging Trends in IT and Applied Science" on 28th to 30th March, 2015"
MATS School of Information Technology, MATS University, Raipur

DNA Computing And Its Application In The Field Of Information Security

Sangita Vishwakarma
Assistant Professor, Department of Computer Science
Maharaja Agrasen International College,
Raipur, (C.G.)
E-mail: sangscorpion@gmail.com

Abstract— **Silicon microprocessors have been the heart of computing for more than four decades. The manufacturers of computer chips are continuously making betterment in the speed and performance of microprocessor chips by integrating more and more devices onto the microprocessor and thus miniaturizing the chip. But there is a limit on this miniaturization and it has been predicted that Moore's law will cease to be obeyed and in near future they would need a new material that could complement the current computing speed and performance of the silicon chips along with equal or fewer complexities. Scientists have found the material that might become the foundation of next era of computing. And the material is DNA, the basic element of our genes. There are numerous benefits that DNA computing offers over conventional silicon based computing. They have enormous storage capacity which is much larger than that of the conventional computers and the enzymes and biological catalysts act as software for executing the required tasks. They exhibit massive parallelism that makes the computation of complex problems much faster than that can be done on conventional silicon based computers. DNA computing has achieved great success in almost every field it has been applied like biomedical, pharmaceutical, information security, cracking secret codes, etc. One such field is Information Security because security of data has always been the matter of great concern. The data being transmitted is under great threat of being attacked and the network carrying data does not have inherent security. Most of the modern cryptographic algorithms are broken, so the DNA computing has brought a new hope in the direction of development of unbreakable algorithms. In this paper the principles of DNA computing and use of DNA computing in the areas of secure data transmission has been outlined. Though DNA computing has almost been successful, the constraints on its implementation are very much demanding like high tech laboratories, labor intensive extrapolation, computational limitations, etc., that moves it far away from being efficiently implemented in today's security world. The aim of this paper is to give a detailed view of DNA computing that could provide a better environment for secure data transmission across networks and the challenges that this technology is facing will be discussed.**

Keywords—DNA, DNA Computing, Cyptography, Algorithms, Molecules.

I. INTRODUCTION

DNA computing is a novel interdisciplinary research area that simulates bio-molecular structure of DNA & computes by means of molecular biological technology. It combines the techniques of biology, chemistry, and mathematics and computer science. One of the main goals of this research area is to develop computers which will be biologically inspired & based on DNA molecules which might replace silicon based computers or at least complement them. In DNA computing, strands of DNA are used to represent data. When we compare the execution time of a DNA reaction to the speed of silicon based processor, we'll find that it is much slower but the massive parallelism feature present in it can be used to solve NP-complete and NP-hard problems. DNA computing was first demonstrated by Adleman in his study as a proof of concept that solved Hamiltonian Path problem. Since then many advancements have been made in this field. Molecules of DNA try different possibilities of a problem at once which is the main reason behind its parallelism. Computation with the assistance of DNA introduces a completely new paradigm in the field of computing. In the recent years it has become an exciting area of research but still there is a long way to implement DNA computing in real life. The scientists and researchers are continuously devoting their efforts in developing models and algorithms for DNA computers.

II. DNA & ITS MOLECULAR COMPONENTS

Before understanding the application of DNA in data security, its basic structure should be understood. Each organism on this planet is made up of same type of blueprint. The way in which this blueprint is coded differentiates one organism from the other. DNA

(Deoxyribonucleic Acid) is a nucleic acid found in the cell of every living organism that contains all the information and instructions for the growth of any organism and is passed from generation to generation. Its main role is to provide the storage medium for all genetic information that acts as the building block and major source of information for growth and development of any living organism.

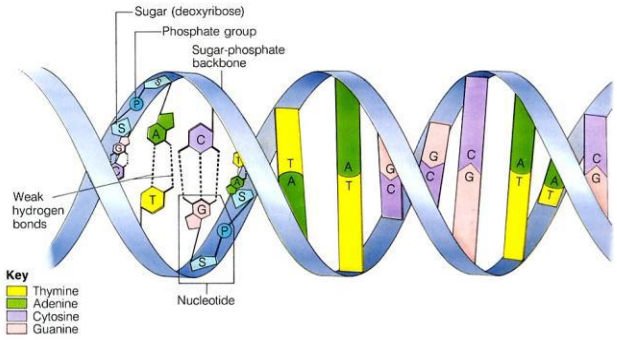


Fig.2. Basic DNA Structure

DNA is basically a polymer which is a collection of various monomers, each monomer is called nucleotide and each nucleotide contains a base. It is double stranded helix of these nucleotides. Each strand of DNA is a long polymer linking millions of nucleotides. A nucleotide consists of one of four nitrogen bases, a five carbon sugar and a phosphate group. There are four different nitrogen bases: Adenine, Guanine, Cytosine and Thymine abbreviated A, G, C and T, respectively. While modeling DNA mathematically, it is represented as X = {A, G, C, T}. All nucleotides differ from each other in terms of their bases. These nucleotides combine in such a way that Adenine is paired with Guanine resulting in Purines and Cytosine is paired with Thymine resulting in Pyrimidines. These combination of nucleotides in the extensively long polymer results in billions of combinations in DNA structure because of which there exists an extensively large variety of living things on this planet ranging from small (mammals as well as plants) to large.

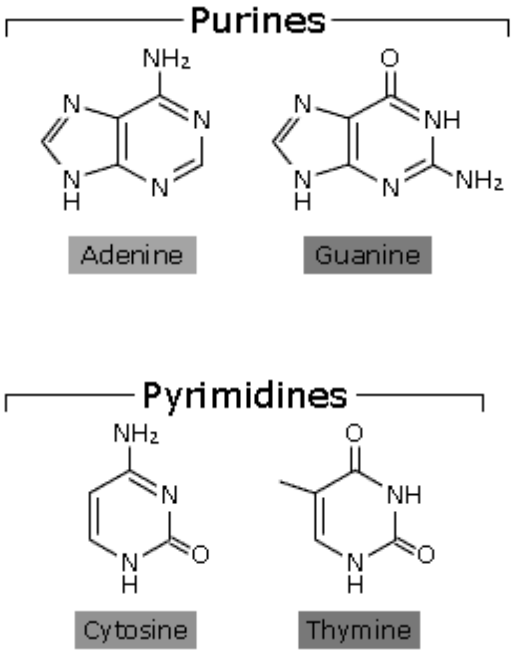


Fig.2. Combinations of Bases Forming Purines & Pyrimidines

The two strands of DNA run anti-parallel to each other. This ability of DNA to bind its pair of strands together forms the basis of its exploitation in various application and is known as Watson-Crick complementarily.

III. DNA COMPUTING

The field of DNA Computing has risen in the past decade. The double helix structure of DNA molecule and Watson-Crick rule form the main principle of DNA Computing. DNA computing or molecular computing are terms used to describe utilizing the inherent combinational properties of DNA for massively parallel computation. The idea is that with an appropriate setup and enough DNA, one can potentially solve huge

mathematical problems by parallel search. Basically this means that you can attempt every solution to a given problem until you came across the right one through random calculation. Utilizing DNA for this type of computation can be much faster than utilizing a conventional computer, for which massive parallelism would require large amounts of hardware, not simply more DNA. Leonard Adleman, a computer scientist at the University of Southern California was the first to pose the theory that the makeup of DNA and it's multitude of possible combining nucleotides could have application in brute force computational search techniques. In early 1994, Adleman put his theory of DNA computing to the test on a problem called the Hamiltonian Path problem or sometimes referred to as the Traveling Salesman Problem. The 'salesman' in this problem has a map of several cities that he must visit to sell his wares where these cities have only one-way streets between some but not all of them. The crux of the problem is that the salesman must find a route to travel that passes through each city (A through G) exactly once, with a designated beginning and end.

The salesman wants to make efficient use of his time and does not want to backtrack or double back on a path he has already taken previously.

IV. Challenges Posed by DNA Computing to Traditional Cryptography

The cryptographic algorithm is usually based on complex mathematical problems such as RSA algorithm. Once these mathematical formulae are broken, it gets easier to break the algorithms. But DNA computing provides a parallel processing at molecular level by introducing new data structures. It poses new challenges to the traditional cryptographic field. A number of algorithms have been proposed to attack a number of problems.

A. Challenges to DES

DES is a cipher which based on a Symmetric-key algorithm that uses a 56-bit key. DES is now considered to be insecure for many applications. This is chiefly due to the 56-bit key size being too small. Dan Boneh constructed DES liquid that can break DES within a day. It has been claimed that any symmetric system under 64 bits can be broken with this method. The process to solve this kind of problem is listed as follows: Firstly, encode appropriate binary codes, create initial DNA liquid which contains all possible keys; Secondly, carry out 16 wheels of encryption after pasted known plaintext strands respectively. Lastly, find the solution by searching. Though this idea was simple in theory, the practical operation and execution is not that easy because binary system is completely abstract.

B. Challenges to RSA

RSA is a public-key cryptographic algorithm. The security of the RSA cryptosystem is based on two mathematical problems: the problem of factoring large numbers and the RSA problem Weng-Long Chang have designed integer factorization way for utilizing DNA computing, which can break RSA. Beaver analysed 1000 bits RSA and concluded that to solve HPP problem required the acme number to be 106 at least, namely 10200000L liquid to be needed on the grounds of conservative estimation but it is infeasible. For this, Winfree came up with the idea of computation by self-assembled tiles since DNA tiles can be more easily programmed to incorporate the constraints of a given problem.

V. DNA Cryptography

In this paper, the research conducted by a number of authors related to the discipline of DNA Cryptography has been studied and has tried to find out the basics of DNA Cryptography that how DNA cryptography field emerged and how DNA computation can be used in cryptography for encrypting, storing and transmitting the information. It has been shown that how DNA cryptography uses DNA as the computational tool with molecular techniques to manipulate it with various algorithms for encryption.

A. Advantages of DNA Cryptography

The biggest advantage of cryptography is its secure nature although; it never needs to be transmitted to anyone.

1. Moreover, encrypting along with the DNA sequence makes data more secure. One gram of DNA contains

10^21 DNA bases = 10^8 tera-bytes of data. A few grams of DNA can hold all the data stored in world.

2. Since DNA is used for encryption, Signature authorization is not needed. DNA replaces the cause of Digital signatures.

3. Works in a massively parallel fashion: DNA is modified biochemically by a variety of enzymes, which are minute protein machines that read and process DNA according to nature's design. There is a wide variety and number of these "operational" proteins, which manipulate DNA on the molecular level. Just like a CPU has a basic set of operations like addition, bit-shifting, logical operators (AND, OR, NOT NOR), etc. that allow it to execute even the most complex calculations, DNA has cutting, copying, pasting, repairing, and many other capabilities.

4. Large storage: A gram of DNA contains about 10^21 DNA bases, or about 10^8 tera-bytes of data. Hence, a few grams of DNA have the capability of storing all the data stored in the world.

5. Input and output of the DNA data can be moved to conventional binary storage media by DNA chip arrays.

6. The main goal of the research of DNA cryptography is exploring characteristics of DNA molecule and reaction, establishing corresponding theories, discovering possible development directions, searching for simple methods of realizing DNA cryptography, and laying the basis for future development.

B. Limitations of DNA Cryptography

Apart from advantages, DNA cryptography has few disadvantages. They are:

1. Lack of the related theoretical basis.

2. Difficult to realize and very expensive to apply.

VI. CONCLUSION

DNA cryptography is basically hiding of data in terms of DNA sequences. This is done by using various DNA technologies with the biological tools. In this paper we summarized basics of DNA and basics of where DNA is found are discussed. We have also discussed the factors on which DNA cryptograph differs from traditional cryptography. Few of the advantages and disadvantages have been summarized. Later on the techniques used by the DNA computing to break the existing cryptographic algorithms have been studied. We want to conclude that although DNA computing has broadened the view of people towards natural phenomena of computing but it is still in its theoretical stage and moving it towards the practical stage will require huge time, enormous computation and large number of expertise. The future work in this field will consist of analyzing deeply the performance of all the DNA cryptographic techniques based on secure data transmission processes.

References

[11] Donald Nixon, "DNA and DNA Computing in Practices – Is the Future in our Genes?", Global Information Assurance Certification Paper

[12] J. Clerk Maxwell, "Integrating DNA Computing in International Data Encyption Algorithm (IDEA)", International Journal of Computer Applications (0975 – 8887), Volume 26– No.3, July 2011

[13] Sanjeev Dhawan, Alisha Saini, "Secure Data Transmission Techniques Based on DNA Cryptography", International Journal of Emerging Technologies in Computational and Applied Sciences (IJETCAS)

[14] Harneet Singh, Karan Chugh, Harsh Dhaka, A. K. Verma,,"DNA based Cryptography: An Approach to Secure Mobile Networks", International Journal of Computer Applications (0975 - 8887), Volume 1 – No.19.

[15] Guangzhao Cui, Cuiling Li, Haobin Li, Xiaoguang Li, "DNA Computing and Its Application to Information Security Field", 2009 Fifth International Conference on Natural Computation.

[16] Junzo Watada, "DNA Computing and its Application"

[17] Rohani binti abu Bakar, Junzo Watada, "DNA COMPUTING AND ITS APPLICATIONS: SURVEY", ICIC Express Letters, Volume 2, Number 1, March 2008

"Aeromycoflora of Rice Mill Tilda In Rainy Season"

Dr. (Mrs.) Sandhya S. Lanjewar, Govt. College, Kohka- Neora , Tilda, Chhattisgarh,493114,India
Tel.:+919425572802, Email: sandhya.lanjewar@gmail.com

Abstract:
The present study of aeromycoflora was done during rainy season (from Month: Jul'2011-Oct'2011). The frequently isolated fungi were Aspergillus ,Penicillium, Alternaria ,Fusarium, Cladosporium ,Mucor hemalis . The most dominant species of Aspergillus niger showed (5.98%) and Aspergillus terreus showed lowest percentage contribution of (0.30%). Fungi can cause allergies and respiratory disease, and the toxins it produces can wear down the immune system of leaving people, especially children, vulnerable to many illnesses. Spores monitoring also enables us to predict many fungal diseases, and provides the means to study the disease onset by phyto-pathogenic spores.
Key Words : Aeromycoflora, Rice Mill , Aspergillus ,PDA (Potato Dextrose Agar) , Rainfall

Introduction:
The spores are often liberated in the air in massive concentration and can remain airborne for a long time. The study of aerobiology has its bearing on various aspects of human health and welfare, chief of which are allergic and plant pathogenic. Some spores of the fungi are responsible for allergy, since the spores are inhaled and deposited on sensitive mucosa. Environmental aeromycology constitutes one of the major aspects mainly because of the dominance of fungal spores in the aerospora (Tilak.1991). Since fungal species constitutes the major component of airborne flora, the study of aeromycology is highly significant in this context.

Material & Methods :
For study of aeromycoflora over the plants, 5 petriplates containing PDA (potato ,dextrose, agar) media were used. The petriplates were exposed over Rice Mill area (Photo plate-1) for 5-10 minutes, then this petriplates were brought in to the laboratory and incubated at 28±10C for 5 to 7 days. After incubation period, number of colonies was counted, and identification done with the help of available literature and finally identified from authentic authority like : National centre of fungal taxonomy Delhi. For ecological studies, at the end of the incubation period percentage frequency and percentage contribution of isolated fungal flora was calculated (Jadhav *et al*., 1994, Sharma 2001,Saluja 2005,Lall 2008).

For ecological studies, at the end of the incubation period percentage frequency and percentage contribution of isolated fungal flora was calculated (Sharma 2001) with the help of the following formula:

$$\text{Percentage frequency} = \frac{\text{Number of observation in which a species appeared}}{\text{Total number of observations}} \times 100$$

$$\text{Percentage contribution} = \frac{\text{Total No. of colonies of a species in all observations taken together}}{\text{Total number of colonies of all species}} \times 100$$

Result & Discussion :

Aerobiological studies can be helpful for forecasting harmful fungal spores and its dissemination, which are responsible for many diseases on plants, animals and human beings. This work was carried to study the analysis of Aeromycoflora of rice mill area of Tilda, Chhattisgarh (India) during rainy season . Variation in the composition and concentration of fungal spores was observed in rice mill environment from season to season and month to month. The study of airborne fungal aerospora is of great importance in order to understand the deposition and dissemination of Rice mill area. The seasonal percentage contribution of Aeromycoflora observed during the investigation period. Temperature , rainfall and relative humidity affect sporulation and subsequent dispersal of fungal spores in the air however no statistical correlation of temperature precipitation and relative humidity level with airborne fungi was observed during the sampling duration.

Monthly Variation :

A seasonal variation in the diversity of fungal species was easily seen in the study area. During rainy season 48 fungal species (227 fungal colonies) belonging to 26 fungal genera was observed. Out of 48 fungal species, 4 fungal species (14 fungal colonies) belongs to 4 fungal genera of Zygomycotina, 4 fungal species (5 fungal colonies) belongs to 3 fungal genera of Ascomycotina, 39 fungal species (203 fungal colonies) belongs to 18 fungal genera of Anamorphic fungi, 1 fungal species (5 fungal colonies) belongs to 1 fungal genera of Mycelia Sterilia and Unknown fungi was absent during this rainy season.

Table

SEASONAL VARIATION OF AEROMYCOFLORA OF SELECTED SITES OF RICE MILL OF TILDA, DURING RAINY SEASON CHHATTISGARH (INDIA)

S. NO.	FUNGAL GROUPS	NO. OF FUNGAL SPECIES	NO. OF FUNGAL GENERA	TOTAL NO. OF FUNGAL COLONIES
1	Zygomycotina	04	04	14
2	Ascomycotina	04	03	05
3	Anamorphic Fungi	**39**	**18**	**203**
4	Mycelia sterilia	01	01	05
5	Unknown Fungi	-	-	-
Total		**48**	**26**	**227**

Graph : SEASONAL VARIATION OF AEROMYCOFLORA OF SITES DURING RAINY SEASON

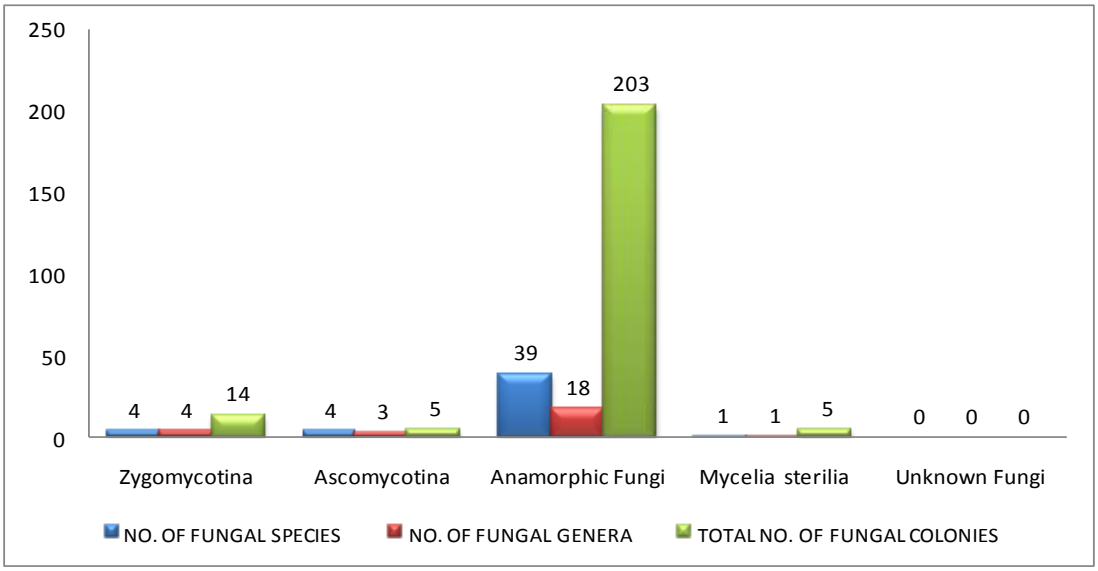
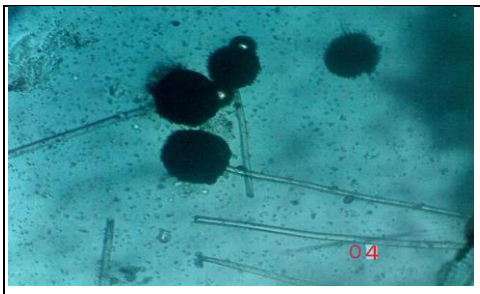
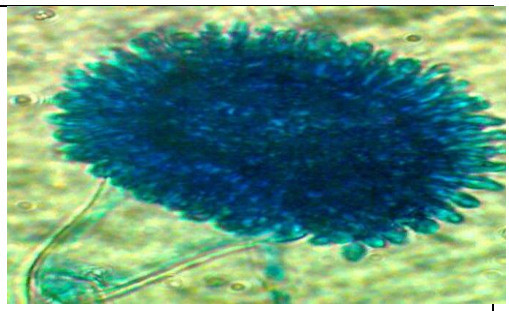


Photo plate 1: Study site:

	
Sampling Site	**Petriplates in Rice Mill**

Photo plate -1: Isolated fungi :

	
Aspergillus niger	**Aspergillus oryzae**

Conclusion:

Study of this kind is interdisciplinary in nature and has tremendous scope to find the significant effect in human health. The spores in air are the representatives of the members of microorganisms growing in that area. In this context, this study shall certainly enlighten the researchers and planners to make a better environment. Some pathogenic and harmful microorganisms detected through this study and methods can be developed to eradicate the same. Among them fungal forms were taken into consideration to find out the status of various types of allergic and pathogenic spores at various places and their role in causing health hazards to plants and human beings.

References :

1. **Lall, B. M. (2008) :** Studies of Indoor and outdoor Aeromycoflora of Dr. Bhimrao Ambedkar Hospital Raipur (C.G.). Ph.D. Thesis, Pt. R. S. U. Raipur (C. G.)
2. **Sharma K., 2001**, Studies of aeromycoflora in relation to leaf surface mycoflora of Ocimum sanctum L., Ph.D. Thesis, Pt. R.S. University, Raipur (C.G.)
3. **Sharma, K. (2010) :** Isolation of soil mycoflora of Katao near Gangtok, India Journal of Phytology 2(5)30-32
4. **Tilak, S.T. and Kulkarni, R.L. (1972)** Microbial content of air inside and outside the caves at Aurangabad. Current Science **23**, pp:850-851
5. **Tiwari K.L. and Jadhav S.K. and Kunjam, S. (2005)** Studies of aeromycoflora of slum area at Raipur (C.G.). 13th Nat. Conf. on Aerobiology., Nagpur.
6. **Tiwari K.L. and Jadhav S.K., 2004,** Studies of aeromycoflora over rice, field-1 at Balodabazar, Raipur (C.G.), India, Ecol. Env. And Cons. 10 (3), pp:387-390
7. **Tiwari K.L., Jadhav S.K., and Kunjam S., 2005,** Studies of aeromycoflora of Dairy Area, Flora and Fauna, 11 (2), pp:191-196
8. **Tiwari P., 1999.** Aerobiological studies of Raipur with special reference to fungal spores Ph.D Thesis. Pt. R.S. University, Raipur (M.P.)
9. **ST Tilak, 1982**,Aerobiology, Vaijayanti Prakashan, Aurangabad, 1-211.

International Seminar
" Emerging Trends in IT and Applied Science" on 28th to 30th March, 2015"
MATS School of Information Technology, MATS University, Raipur

Genetic Algorithm Applied in Network Intrusion Detection System

Rekha Singh, A.K.Tiwari, Seema Pathak

Rekha Singh, Asstt. Prof., Department of Computer Science, Disha College, Raipur, C.G.
Contact No: 9827103180, E-mail : rekhac.mca@gmail.com
A.K.Tiwari, Principal, Disha College, Raipur, C.G., Contact No: 9300491034, E-mail : ani1969_rpr@yahoo.com
Seema Pathak, Asstt. Prof., Department of Computer Science, Disha College, Raipur, C.G.
Contact No: 9009449718, E-mail : seemapathak3010@gmail.com

***Abstract*—The existing NIDSs involve various mechanisms in order to identify the patterns related to the network problems. In this paper we described the implementation of genetic algorithm using steady-state selection mechanism. We have performed various test cases in order to analyze the effect of varying iterations and varying initial chrome length along with different fitness functions.**
The analyzed result would be helpful to have quicker generation of reliable rule sets for Network Intrusion Detection System using less/more number of iterations, depending on the varying initial chrome length.

***Keywords*—Data Mining, Genetic Algorithm, Network Intrusion, Network Intrusion Detection System..**

INTRODUCTION

Intrusion Detection System is the process of monitoring the events occurring in a computer system or network and analyzing them for signs of possible incidents, which are violations or imminent threats of violation of computer security policies, acceptable use policies, or standard security practices. Incidents have many causes, such as malware (e.g., worms, spyware), attackers gaining unauthorized access to systems from the Internet, and authorized users of systems who misuse their privileges or attempt to gain additional privileges for which they are not authorized. Although many incidents are malicious in nature, many others are not; for example, a person might mistype the address of a computer and accidentally attempt to connect to a different system without authorization.

There exist a number of mechanisms in order to identify the patterns related to the network problems. One of them is Genetic Algorithm, in which some basic operations are performed in order to generate the patterns (also termed as Chromosomes or RULE SETS). These generated patterns can be used along with the audit dataset as input into the Network Intrusion Detection System, in order to achieve the kind of intruder in the audit dataset.

NETWORK INTRUSION DETECTION SYSTEM

H. Introduction

An intrusion detection system (IDS) is software that automates the intrusion detection process. An intrusion prevention system (IPS) is software that has all the capabilities of an intrusion detection system and can also attempt to stop possible incidents. This section provides an overview of IDS and IPS technologies as a foundation for the rest of the publication. It first explains how IDS and IPS technologies can be used. Next, it describes the key functions that IDS and IPS technologies perform and the detection methodologies that they use. Finally, it provides an overview of the major classes of IDS and IPS technologies.

I. Components of NIDS

An intrusion detection system normally consists of three functional components:

The first component of an intrusion detection system, also known as the event generator, is a data source. Data sources can be categorized into four categories namely Host-based monitors, Network-based monitors, Application-based monitors and Target-based monitors.

The second component of an intrusion detection system is known as the analysis engine. This component takes information from the data source and examines the data for symptoms of attacks or other policy violations.

The analysis engine can use one or both of the following analyzing approaches:

1) Misuse/Signature-Based Detection

This type of detection engine detects intrusions that follow well-known patterns of attacks (or signatures) that exploit known software vulnerabilities. The main limitation of this approach is that it only looks for the known weaknesses and may not care about detecting unknown future intrusions.

2) Anomaly/Statistical Detection

An anomaly based detection engine will search for something rare or unusual. They analyses system event streams, using statistical techniques to find patterns of activity that appears to be abnormal. The primary disadvantages of this system are that they are highly expensive and they can recognize an intrusive behavior as normal behavior because of insufficient data.

The third component of an intrusion detection system is the response manager. In basic term, the response manager will only act when inaccuracies (possible intrusion attacks) are found on the system, by informing someone or something in the form of a response.

J. Types of Network Attacks

Denial of Service (DoS)

A DoS attack is a type of attack in which the hacker makes a computing or memory resources too busy or too full to serve legitimate networking requests and hence denying users access to a machine e.g. apache, smurf, neptune, pingof death, back, mail bomb, UDP storm etc. are all DoS attacks.

Remote to User Attacks (R2L)

A remote to user attack is an attack in which a user sends packets to a machine over the internet, which s/he does not have access to in order to expose the machines vulnerabilities and exploit privileges which a local user would have on the computer.

User to Root Attacks (U2R)

These attacks are exploitations in which the hacker starts off on the system with a normal user account and attempts to abuse vulnerabilities in the system in order to gain super user privileges.

Probing

Probing is an attack in which the hacker scans a machine or a networking device in order to determine weaknesses or vulnerabilities that may later be exploited so as to compromise the system. This technique is commonly used in data mining.

K. Types of NIDS

1) Host Based Intrusion Detection (HIDS)

HIDSs evaluate information found on a single or multiple host systems, including contents of operating systems, system and application files.

2) Network Based Intrusion Detection (NIDS)

NIDSs evaluate information captured from network communications, analyzing the stream of packets which travel across the network..

PAGE GENETIC ALGORITHM (GA)

Genetic algorithm is a family of computational models based on principles of evolution and natural selection. This algorithm converts the problem specific domain into a model, using a chromosome-like data structure. Chromosome-like data structure is evolved by performing selection, recombination, and mutation operators over the various chromosomes.

Components of GA

There are mainly 6 components in Genetic Algorithm System.

Evaluation function (or fitness function)

A fitness function is a particular type of objective function that is used to summarize, as a single figure of merit, how close a given design solution is to achieving the set aims.

Population size, *Crossover Rate & Mutation Rate*

Population size: Good population size is about 20-30, however sometimes sizes 50-100 are reported as best. Some research also shows that best population size depends on encoding, on size of encoded string. It means, if you have chromosome with 32 bits, the population should be say 32, but surely two times more than the best population size for chromosome with 16 bits.

Crossover rate: Crossover rate generally should be high, about 80%-95%. (However some results show that for some problems crossover rate about 60% is the best.)

Mutation rate: On the other side, mutation rate should be very low. Best rates reported are about 0.5%-1%.

Encoding Mechanism

a) Binary Encoding: In binary encoding every chromosome is a string of bits, 0 or 1.

b) Permutation Encoding: In this encoding mechanism, every chromosome is a string of numbers, which represents number in a sequence

c) Value Encoding: In value encoding, every chromosome is a string of some values. Values can be anything connected to problem, form numbers, real numbers or chars to some complicated objects.

d) Tree Encoding: In tree encoding every chromosome is a tree of some objects, such as functions or commands in programming language.

Parent Selection Mechanism

a) Roulette Wheel Selection: Parents are selected according to their fitness. The better the chromosomes are, the more chances to be selected they have.

MATS Journal of Engineering and Applied Science, Volume 1, Issue 3, ISSN 2394 - 0549 (Disclaimer - The authors are solely responsible for the contents of the research paper compiled in this seminar proceeding. The publishers or editors do not take any responsibility for the same in any manner. Errors, If any, are purely unintentional.

b) Rank Selection: Rank selection first ranks the population and then every chromosome receives fitness from this ranking.

c) Steady-State Selection: For every generation a few (good - with high fitness) chromosomes are selected for creating a new offspring. Then some (bad - with low fitness) chromosomes are removed and the new offspring is placed in their place. The rest of population survives to new generation.

d) Elitism: Elitism is name of method, which first copies the best chromosome (or a few best chromosomes) to new population. The rest is done in classical way.

Variation Operators (crossover and mutation)

a) Crossover selects genes from parent chromosomes and creates a new offspring.

b) Mutation changes randomly the new offspring.

There are various ways to perform cross-over & mutation, as per the kind of encoding implementation listed using a table:

Table. 1. Variation Operators Table

Encoding Type	Cross-Over Type	Mutation Type
Binary	Single Point Cross-Over Two Point Cross-Over Uniform Cross-Over Arithmetic Cross-Over	Bit Inversion
Permutation	Single Point Cross-Over	Order Changing
Value	Same as for Binary	Add/Subtract or Replace by Random number
Tree	Tree Cross-Over	Changing Operator, number

Survivor Selection Mechanism (replacement)

The survivor selection mechanism is normally based on the type of "parent selection mechanism" been selected. This operation is basically used for deciding which of the existing chromosome should be replaced with the new generated (off-spring) chromosome.

PROPOSED SYSTEM

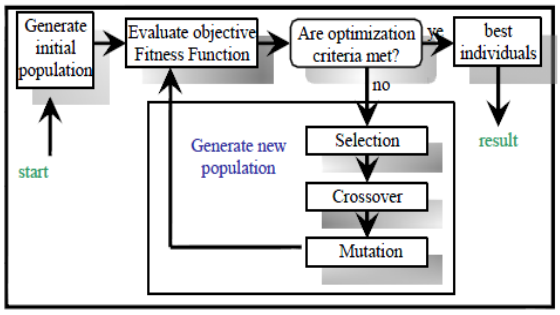
We have chosen Genetic Algorithm to make our own intrusion detection system. This section gives an overview of the algorithm and the system. These algorithms convert the problem in a specific domain into a model by using a chromosome-like data structure and evolve the chromosomes using selection, recombination, and mutation operators. The range of the applications that can make use of genetic algorithm is quite broad. In computer security applications, it is mainly used for finding optimal solutions to a specific problem.

The process of a genetic algorithm usually begins having a randomly selected set of chromosomes of a specific size, acting as an input.

These chromosomes are representations of the problem to be solved. According to the attributes of the problem, different positions of each chromosome are encoded as bits, characters, or numbers. These positions are referred to as genes and are changed randomly within a range during evolution. The set of chromosomes during a stage of evolution are called a population. An evaluation function is used to calculate the "goodness" of each chromosome. During evaluation, two basic operators, crossover and mutation, are used to simulate the natural reproduction and mutation of species. The selection of chromosomes for survival and combination is biased towards the fittest chromosomes.

Figure 3 shows the structure of a simple genetic algorithm. It starts with a randomly generated population, evolves through selection, recombination (crossover), and mutation. Finally, the best individual (chromosome) is picked out as the final result once the optimization criterion is met.

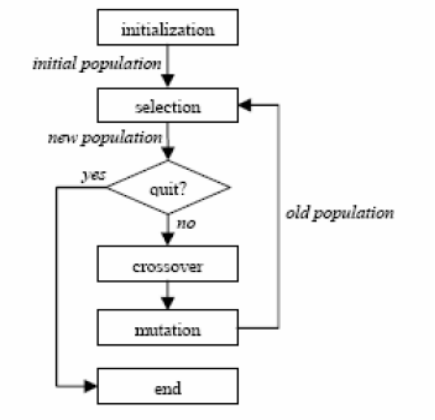
Figure. 3. Coding in GA



Flow Chart for the proposed Algorithm

MATS Journal of Engineering and Applied Science, Volume 1, Issue 3, ISSN 2394 - 0549 (Disclaimer - The authors are solely responsible for the contents of the research paper compiled in this seminar proceeding. The publishers or editors do not take any responsibility for the same in any manner. Errors, If any, are purely unintentional.

Figure. 4. Flow Chart for the proposed Algorithm



Proposed Algorithm

Algorithm:

1. Input the value for population size and Iteration Count
2. Initialize the new generated chromosome as per binary encoding
3. Set the genes (chg, gb) to the value "false".
4. Calculate fitness value for each generated initial chromosome
5. Now start on a loop for specified iteration.
6. For each iteration perform below operations:
   - **Sorting:** Sort the chromosome set as per fitness value
   - **Selection:** Select the 2 chromosomes such that they have the highest fitness value
   - **Crossover:** Generate new off-spring using the combination of the selected chromosomes (single point crossover is performed, selecting the crossover point randomly)
   - Assign the new generated chromosome, replacing the chromosome having least fitness value in the chromosome set
   - Set the gene(chg) value to "true", for this chromosome, indicating that the chromosome has been modified (not as initial)
   - **Mutation:** Change the value for any single gene for the generated off-spring (randomly)
   - Calculate fitness value for the new off-spring
   - Set value for gene(gb), if its fitness value is at least half the highest fitness value in the set chromosome
7. Display the chromosome set obtained after the specified iterations, once sorted again.

EXPERIMENT

For a sample test, we have taken a chromosome structure as below:

- 6 genes representing the network data
- 1 gene representing the change in chromosome (1=changed, 0=unchanged)
- 1 gene representing the chromosome satisfying the condition : fitness value > (highest fitness value)/2 (1=satisfied, 0=unsatisfied)

We have under gone two fitness functions (x^2-5x and x^2+12x-5), along with the population size of 4 & 15.

We have performed up to 14 iterations for each population size, each fitness functions.

Table. 2. Outcome of the Experiment

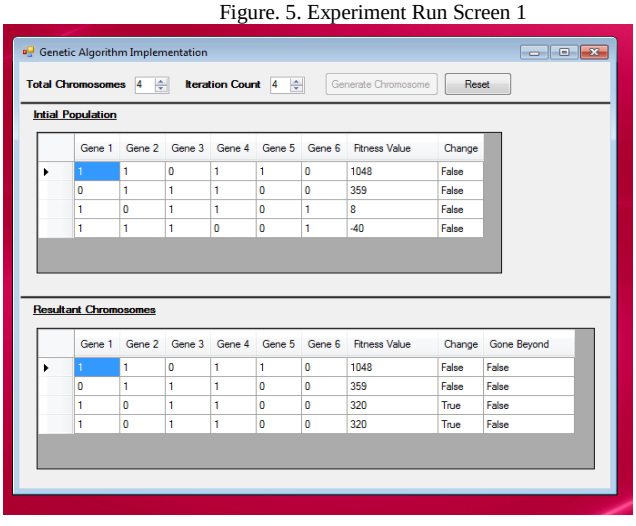
Population Size	Total Number of Iterations	
	Group-1	Group-2
4	31	42
15	57	145

The resultant output is as shown in classified into 2 groups depending on the iteration cycle: Group-1: 1-7 iteration, Group-2:8-14 iterations.

Above table, provides the total number of changed chromosomes for each chromosome size (4 & 15), depending on the each classified group.

Following figures shows the experiments.

Figure. 5. Experiment Run Screen 1



MATS Journal of Engineering and Applied Science, Volume 1, Issue 3, ISSN 2394 - 0549 (Disclaimer - The authors are solely responsible for the contents of the research paper compiled in this seminar proceeding. The publishers or editors do not take any responsibility for the same in any manner. Errors, If any, are purely unintentional.

Figure. 6. Experiment Run Screen 2

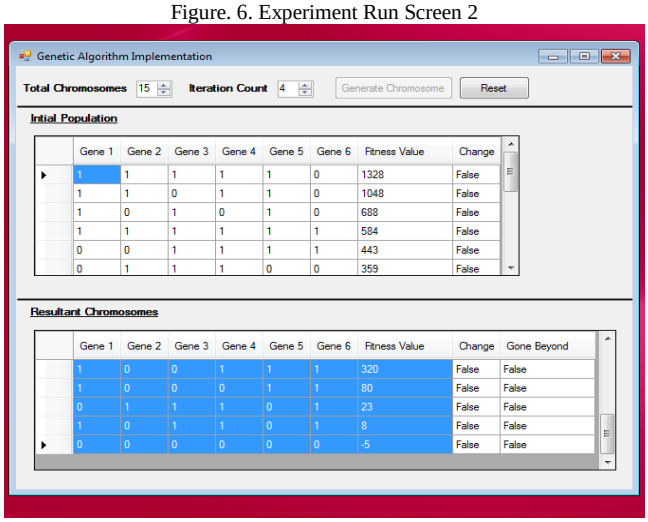
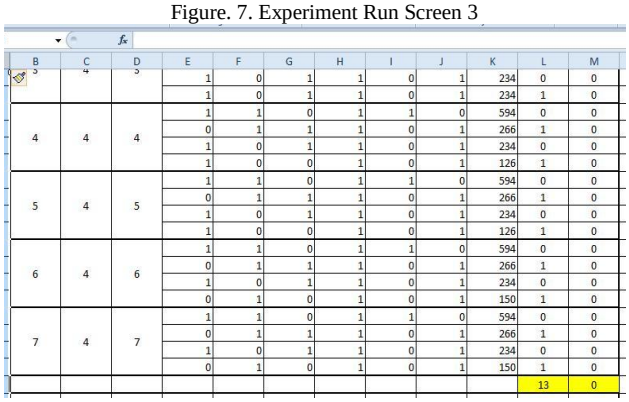
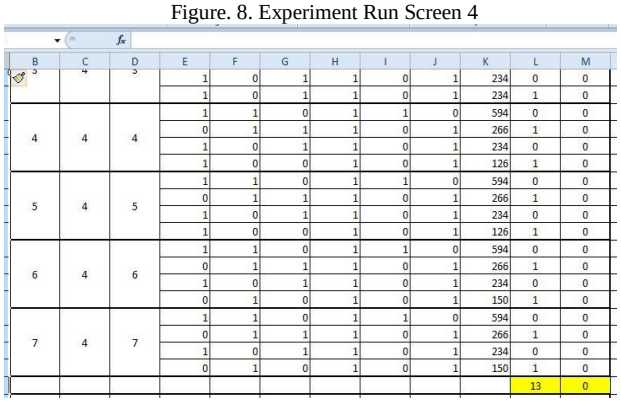


Figure. 7. Experiment Run Screen 3



			1	0	1	1	0	1	234	0	0
			1	0	1	1	0	1	234	1	0
			1	1	0	1	1	0	594	0	0
4	4	4	0	1	1	1	0	1	266	1	0
			1	0	1	1	0	1	234	0	0
			1	0	0	1	0	1	126	1	0
			1	1	0	1	1	0	594	0	0
5	4	5	0	1	1	1	0	1	266	1	0
			1	0	1	1	0	1	234	0	0
			1	0	0	1	0	1	126	1	0
			1	1	0	1	1	0	594	0	0
6	4	6	0	1	1	1	0	1	266	1	0
			1	0	1	1	0	1	234	0	0
			0	1	0	1	0	1	150	1	0
			1	1	0	1	1	0	594	0	0
7	4	7	0	1	1	1	0	1	266	1	0
			1	0	1	1	0	1	234	0	0
			0	1	0	1	0	1	150	1	0
										13	0

Figure. 8. Experiment Run Screen 4



			1	0	1	1	0	1	234	0	0
			1	0	1	1	0	1	234	1	0
			1	1	0	1	1	0	594	0	0
4	4	4	0	1	1	1	0	1	266	1	0
			1	0	1	1	0	1	234	0	0
			1	0	0	1	0	1	126	1	0
			1	1	0	1	1	0	594	0	0
5	4	5	0	1	1	1	0	1	266	1	0
			1	0	1	1	0	1	234	0	0
			1	0	0	1	0	1	126	1	0
			1	1	0	1	1	0	594	0	0
6	4	6	0	1	1	1	0	1	266	1	0
			1	0	1	1	0	1	234	0	0
			0	1	0	1	0	1	150	1	0
			1	1	0	1	1	0	594	0	0
7	4	7	0	1	1	1	0	1	266	1	0
			1	0	1	1	0	1	234	0	0
			0	1	0	1	0	1	150	1	0
										13	0

CONCLUSION

From the above experiment, we have finally reached to two different conclusions, for the case of genetic algorithm being implemented using steady–state selection mechanism.

Conclusion-1

If there are less number of initial chromosomes, the no. of iteration to be performed in order to have new reliable chromosomes, will be less.

Similarly, for higher number of initial chromosomes, it requires more number of iterations to be performed to have new reliable chromosomes.

Table. 3. Conclusion Table

Population Size	Total Number of Iterations (% change)			
	Group-1	Group-2		
4	55.36%	75%	Likely Similar	Conclusion 1
15	27.14%	69.05%	Much Differing	
	Nearly Half	**Slightly changed**		

Conclusion-2

If there is less number of iterations, then the no. of new reliable chromosomes being generated goes on decreasing with an increase in length of initial chromosomes being given.

Similarly, if we keep on increasing the iterations then the resultant new reliable chromosomes scope is not much affected due to varying chromosome length.

REFERENCES

The heading of the References section must not be numbered. All reference items must be in 8 pt font. Please use Regular and Italic styles to distinguish different fields as shown in the References section. Number the reference items consecutively in square brackets (e.g. [1]).

[1] Ren Hui Gong, Mohammad Zulkernine, Purang Abolmaesumi, "A Software Implementation of a Genetic Algorithm Based Approach to Network Intrusion Detection", School of Computing, Queen's University, Kingston, Ontario, Canada, 2005.

[2] Mohammad Sazzadul Hoque1, Md. Abdul Mukit2 and Md. Abu Naser Bikas, "AN IMPLEMENTATION OF INTRUSION DETECTION SYSTEM USING GENETIC ALGORITHM", International Journal of Network Security & Its Applications (IJNSA), Vol.4, No.2, March 2012

[3] Adhitya Chittur, "Model Generation for an Intrusion Detection System Using Genetic Algorithms", Ossining High School Ossining, NY, November 27, 2001

[4] Brian Lavender, "Implementation of Genetic Algorithms into SNORT, a Network Intrusion Detection System"

[5] Wei Li, "Using Genetic Algorithm for Network Intrusion Detection", Department of Computer Science and

Engineering Mississippi State University, Mississippi State, MS 39762

[6] Data Mining for Network Intrusion Detection: Eric Bloedorn, Alan D. Christiansen, William Hill, Clement Skorupka, Lisa M. Talbot, Jonathan Tivel The MITRE Corporation1820 Dolley Madison Blvd.McLean, VA 22102 (703) 983-5274

[7] Genetic Algorithm Tom V. Mathew Assistant Professor, Department of Civil Engineering, Indian Institute of Technology Bombay, Mumbai-400076.

[8] M. Revathi 1 1Department Of Information Technology, Bharathiar University, Coimbatore, Tamilnadu, India T.Ramesh2 Assistant Professor Department Of Information Technology, Bharathiar University, Coimbatore, Tamilnadu, India

[9] Modeling An Intrusion Detection System Using Data Mining And Genetic Algorithms Based On Fuzzy Logic by G.V.S.N.R.V. Prasad Y.Dhanalakshmi Dr.V.Vijaya Kumar Dr I.Ramesh Babu Professor Scholar Professor &Dean Professor Dept. of CSE Dept of CSE Dept. of CSE &IT Dept. of CSE Gudlavalleru Engg.College A.N.U G.I.E.T A.N.U Gudlavalleru Guntur Rajamandry Guntur

[10] Importance of Intrusion Detection System (IDS) by Asmaa Shaker Ashoor (Department computer science, Pune University) Prof. Sharad Gore (Head department statistic, Pune University)

[11] Guide to Intrusion Detection and Prevention Systems: http://csrc.nist.gov/publications/nistpubs/800-94/SP800-94.pdf

[12] Understanding Intrusion Detection Systems: http://www.sans.org/reading_room/whitepapers/detection/understanding-intrusion-detection-systems_337

[13] Using Genetic Algorithm for Network Intrusion Detection: http://www.security.cse.msstate.edu/docs/Publications/wli/DOECSG2004.pdf

[14] Understanding-Intrusion-Detection-Systems_337 Intelligent Network Intrusion Detection Using DT and BN Classification Techniques

[15] A genetic algorithms based approach for conflicts resolution in requirement Bowman, M., Debray, S. K., and Peterson, L. L. 1993. Reasoning about naming systems.

MATS Journal of Engineering and Applied Science, Volume 1, Issue 3, ISSN 2394 - 0549 (Disclaimer - The authors are solely responsible for the contents of the research paper compiled in this seminar proceeding. The publishers or editors do not take any responsibility for the same in any manner. Errors, If any, are purely unintentional.

Immobilization of fungal protease using calcium alginate beads

Namrata Sharma[1], Sapna Rai[2] and Sanchali Padhye[2]

[1]Dept. of Biochemistry, MGMM, Jabalpur (M.P)

[2]Dept. of Microbiology, MGMM, Jabalpur (M.P)

Abstract

Immobilized enzymes are widely used for variety of applications. The immobilized enzymes can be separated from the reaction mixture and reused and also immobilized in order to prevent the enzyme from being exposed to harsh conditions such as high temperature, surfactants and oxidizing agents etc. Proteases are the most important group of enzyme from an industrial point of view. They have a wide variety of applications in many field. In this study, protease was extracted from various fungal species. The extracted proteases were immobilized on calcium alginate beads and the specific activities of the immobilized enzymes were estimated and compared. It was found that immobilized protease activity was more in *Aspergillus* strain AS#7. The optimum pH was also determined and was found to be 10 for AS#7 strain. The protease activity with respect to incubation time were also analyzed. For *Aspergillus* strain AS#7 the activity was 10 min and in case of AS#6 and AS# 1 it was at 15 min . *Aspergillus* strain AS#2 had highest activity at the 20 min.

Key words – *Aspergillus sp, immobilization, protease, specific activity.*

Introduction

Enzymes are biocatalyst and catalyze most of the metabolic reactions in living organisms. Proteases constitute a very large group of enzymes that degrade protein into small peptides and amino acids. Protease have a wide range of applications such as leather processing , meat tenderization ,detergent formulation , baking, brewing, cheese manufacturing and soya sauce production, protein hydrolysate, waste treatment, silk industry, recovery of silver from waste photographic films, as well as analytical tool in basic research and have high commercial value(Godfrey and West,1996).Among the bulk of industrial enzyme, protease constitute around 60% of the total worldwide sales(Kunamneni et al., 2003). For commercial application of protease different methods have been used to reduce cost and increase the utilization of protease, one of which is immobilization. The property of immobilized enzymes which is of greatest industrial importance is that ease with which they can be separated from reaction mixtures. Hence, the reaction can be stopped by physical removal of immobilized enzyme, without requiring such procedure as heat inactivation. Furthermore the enzyme will still be active and largely uncontaminated, so can be used again (Palmer, 2004). The properties of immobilized enzyme preparations are governed by the properties of both enzyme and the carrier material. Calcium alginate is the most widely used matrix for entrapment of enzyme. Entrapment with insoluble calcium alginate is recognized as a

rapid, nontoxic, inexpensive and versatile method for immobilization of enzyme as well as cell(Fraser and Bickerstaff,1997).

In the present investigation protease from four different strains of *Aspergillus* species were partially purified and were immobilized in

MATS Journal of Engineering and Applied Science, Volume 1, Issue 3, ISSN 2394 - 0549

calcium alginate beads. Then their optimum pH and incubation time were studied.

MATERIALS AND METHODS

Micro organism and maintenance of culture -

The fungal strains used throughout this study were isolated from soil samples of Jabalpur area. These isolates were identified as of *Aspergillus sp,* on the basis of their morphological and microscopic identification. The *Aspergillus* strainsAS#1, AS#2, AS#6 and AS#7 were maintained on potato dextrose agar plates at 28 $\pm 2^0$C. Spores for inoculums were prepared from 5-7 days old cultures by sterile cork borer in Tween - 80 solution.

Growth Conditions -

The medium used for protease production by *Aspergillus sp,* was Yeast Extract Broth, composed of Yeast Extract, 0.5%, KCl 2%, Peptone, 2%, Sucrose 2%, Casein, 1.0%, pH 7.5. YE broth was autoclaved at 121^0C for 15 mins. Broth was inoculated with two *Aspergillus* strains AS#1,AS#2, AS#6 and AS#7 and incubated in a rotary shaker at 150 rpm for 96h at 28 $\pm 2^0$C, in separate 250 ml Erlenmeyer flasks with working volume of 100 ml. The cultures were centrifuged at 10,000 rpm for 10 min at 4^0C to remove fungal mycelia and supernatants were used as the crude enzyme solution.

Assay of protease activity -

Protease activity was measured by the method of Anson (1938), using casein as a substrate. A control lacking the enzyme was included in each assay.

One unit of protease hydrolyzed casein to produce color equivalent to 1.0µmol (181µg) of tyrosine per minute at pH 7.5 at 37^0C. The enzyme activity was expressed as U/ml.

Determination of Protein content -

The protein content of fraction, obtained after ammonium sulphate precipitation was determined by the method of Lowry (1951) using BSA as a substrate.

Ammonium Sulphate fractionation -

Solid ammonium sulphate was added to the crude extract to 0-50% saturation. The precipitate was collected by centrifugation, dissolved in minimal volume of Tris-HCl buffer (pH 7.8) and desalted by using prepacked desalting column.

Entrapment -

The partially purified protease from AS#1,AS#2, AS#6 and AS#7 strains were immobilized in the calcium alginate beads through entrapment by the method of Banerjee et al.,(1984). Bovine serum albumin (5mg) was added to 5.0 ml of enzyme solution. To this 1.5% of sodium alginate was added and stirred gently. The entrapment was carried out by dropping the mixture through a glass pipette into 50 ml of 2.0% (w/v) $CaCl_2$ solution. The beads so formed were left for 1h in calcium chloride solution and then stored in 0.1 M Tris HCl buffer, pH 9.0 at 4^0C.

Effect of pH and incubation time on the immobilized protease activity.

Effect of pH and incubation time on the immobilized protease was determined under standard assay conditions using casein as a substrate. Protease activity was studied in the pH range from 5-11 for immobilized form protease and their activity was measured at various incubation time (5-25 min). Per assay tube 5 beads of immobilized enzyme were added.

Results and Discussion

Partial purification- Purification of enzyme is required for better understanding of the functions of the enzymes (Sandhya et al., 2005).Partial purification of the enzyme by ammonium sulphate precipitation followed by desalting through prepacked desalting column of dextrin, resulted in nearly

MATS Journal of Engineering and Applied Science, Volume 1, Issue 3, ISSN 2394 - 0549 (Disclaimer - The authors are solely responsible for the contents of the research paper compiled in this seminar proceeding. The publishers or editors do not take any responsibility for the same in any manner. Errors, If any, are purely unintentional.

19.2 fold increase in the specific activity of the protease from AS#7strain, 7.5 fold increase in the specific activity of the enzyme from AS#6 strain whereas specific activity of the protease from AS#1 strain showed fold increase. Specific activity of protease fromAS#2 strain was quite low.

Effect of pH and incubation time on immobilized protease activity-

Alginate entrapped enzyme was assayed at different incubation time and pH ranging from 5-25 min and pH 5-10, respectively.

The longer an enzyme is incubated with its substrate, the greater the amount of product that will be formed. However, the rate of formation of product is not a simple linear function of the time of incubation. All proteins suffer from denaturation and hence loss of catalytic activity, with time. Some enzymes, especially in partially purified preparations, may be noticeably unstable, losing a significant amount of activity over the period of incubation. If the activity of the enzyme is such that much of the substrate is used up during the incubation, then, even if the concentration of substrate added was great enough to ensure saturation of the enzyme at the beginning of the experiment, it will become inadequate as the incubation proceeds, and the formation of product will decrease.

Entrapped protease from AS#7 strain showed maximum activity at 10 min of incubation time and in case of AS#6 strain and AS#1 strain it was 15 min. AS#2 had highest activity at the 20 min of incubation (Table-1). Enzyme have properties that are quite pH sensitive because of their proteinaceous nature. pH can affect activity by changing the charge on an amino acid residue which is functional in substrate binding or catalysis. The optimum pH value of the immobilized protease for AS#7 strain is 10 and for AS#6 it is 9. Immobilized protease from AS#1 and AS # 2 strains showed optimum pH at 8(Table-2). A fall in the hydrolysis rate on either side of the optimum value is due to decrease in affinity of enzyme for the substrate (Godfrey and Reichelt,1983).

Conclusion

This study suggests that alginate beads of 1-2 mm diameter were a suitable support for protease immobilization. Impact of pH and incubation time on immobilized protease activity showed that protease from *Aspergillus* strain AS#7 is better. In future, other parameters of immobilized protease can also be investigated for maximum production at a cheaper rate.

Acknowledgement

Authors are very thankful to Management and Principal of Mata Gujri Mahia Mahavidyalaya, Jabalpur for Providing the Lab Facility.

References

1.Anson ML, The estimation of pepsin trypsin, papain and cathepsin with hemoglobin, *J. Gen. Physiol, 22,* 1938, 79-89.
2. Banerjee M, A. Chakravarty and S.K. Mujumdar, Characteristics of yeast α galactosidase immobilized on calcium alginate gels *App. Microbiol Biotechnol, 20,* 1984, 271-274.
3. Fraser J.E. and G.F. Bickerstaff, Entrapment of enzymes and cells in calcium alginate. In immobilization liasation of enzymes and cells *Humana press,* 1997, 61-66.
4. Godfrey T. and J. Reichelt, Industrial enzymology *(Nature Press, New York, 1983),* 1-7.
5. Kunamneni A., P. Ellaiah and DS Prasad,Purification and partial characterization of thermostable serine alkaline protease for a newly isolated *Bacillus subtilis* PE-11,AAPS PharSciTech,2003;4(4):440-448.
6. Lowry O.H., N.J. Rosebrough, A.L. Farr and R.J. Randall, Protein measurement with Folin Phenol reagent, *J.Biol Chem. 193,* 1951, 265-275.

7. Palmer T., Enzymes : Biochemistry, Biotechnolgy, Clincial Chemistry. (Affiliated *East-West Press Pvt. Ltd*., New Delhi, 2004, 363).

8. Sandhya C., A. Sumantha, G. Szakacs and A. Pandey, Comparative evaluation of neutral protease production by *Aspergillus Oxyzae* in submerged and solid state fermentation, *Process Biochem, 40,* 2005, 2689-2694.

Table-1:Effect of incubation time on immobilized protease activity (in U/ml)

	Aspergillus Sp.			
Incubation time(in min)	**AS#1**	**AS#2**	**AS#6**	**AS#7**
5	98	90	114	157
10	110	82	138	177
15	145	106	161	149
20	126	122	110	134
25	106	98	102	114

Table2:Effect of pH on immobilized protease activity (in U/ml)

pH	***Aspergillus Sp.***			
	AS#1	**AS#2**	**AS#6**	**AS#7**
5	106	47	90	110
6	110	67	118	114
7	118	110	140	134
8	141	134	160	153
9	134	106	152	161
10	114	82	131	173

MATS Journal of Engineering and Applied Science, Volume 1, Issue 3, ISSN 2394 - 0549 (Disclaimer - The authors are solely responsible for the contents of the research paper compiled in this seminar proceeding. The publishers or editors do not take any responsibility for the same in any manner. Errors, If any, are purely unintentional.

Role of Computers in Law Enforcement: A Review

Neeraj Prakash Rai
Dept. of P. G. Studies and Research in Law, R.D.V.V.Jabalpur

ABSTRACT

Computers play an important role in almost all aspects of life today since they are used in business, in medical-related fields, in education and even in law enforcement. The use of computers in law enforcement has changed and developed rapidly, especially in recent years. Computers are used to store information, analyze particular objects found at crime scenes, and help in collecting information about criminals as well as victims. Computers are used to hold databases of information. The use of computers in researching legal matters and new ways of dealing with crime has become a common activity among law enforcement personnel. The Internet is rich in helpful resources that help boost law enforcement. Exchange of vital information through various government agencies is also made easier through the use of computers. Technology is advancing in the areas of communication, computer systems, weapons, brain wave sensors, density scanners, amplified realism, biometrics, vision enhancers, and many more. Developments in technology will greatly improve the effectiveness and efficiency of law enforcement personnel.

INTRODUCTION AND REVIEW

The use of computers in law enforcement has changed and developed rapidly, especially in recent years. Computers are used to hold databases of information, to run sophisticated software that can recognize faces or identify fingerprints and to connect to the Web, an avenue for communication and a rich source of intelligence. As well as desktop computers, law enforcement personnel also use mobile devices, such as laptops and tablets, to do their job (1).Computer technology allows law enforcement services to store and retrieve vast amounts of data. This information can include details of incident reports, criminals' descriptions, fingerprints and other identifying marks. It can also include descriptions and registrations of vehicles involved in criminal activity. Another crucial pool of information is DNA data taken from suspects. DNA databases allow samples of DNA taken from suspects to be matched with samples taken from crime scenes.

Computers are an invaluable tool for communication between individuals, departments and law enforcement agencies. Documents, photographs and other material can be sent almost instantaneously from one location to another, saving valuable time. Email is a good example: Encrypted emails can be used to send important data securely while mitigating the risk that the information they contain will fall into the wrong hands.

Mobile computing devices -- laptops, notebook computers and tablet PCs -- are very useful to law enforcement. Armed with a laptop, a police officer can take notes, access records or contact colleagues in other

MATS Journal of Engineering and Applied Science, Volume 1, Issue 3, ISSN 2394 - 0549 (Disclaimer - The authors are solely responsible for the contents of the research paper compiled in this seminar proceeding. The publishers or editors do not take any responsibility for the same in any manner. Errors, If any, are purely unintentional.

districts, all without leaving a vehicle. Mobile devices can be used to check the identity or other credentials of individuals at the scene of a crime, as well as recording and tracking vital data such as vehicle license plates. Computers can also be used to track the position of GPS devices, helping law enforcement officers to find vehicles.

The Internet is used by law enforcement agencies in innumerable regards. Web sites can be used by law enforcement agencies to educate and inform the public, appeal for information or alert people to ongoing situations such as a missing child or a felon at large. Because criminals often use the Internet to share information, it can be very useful in crime prevention and detection. (1)

Law enforcement agencies must also use the Internet when tackling online crime. This can include the sharing of illegal material, such as pirated commercial movies or music. "Phishing" and other forms of identity theft that use email or the Internet must also be addressed using computer technology, as must attacks using viruses and hacking attacks. Law enforcements from different countries must often work together to tackle cyber crime (2).

Here, in the 21st century, technology is advancing in the areas of; communication, computer systems, weapons, brain wave sensors, density scanners, amplified realism, biometrics, vision enhancers, and many more. Developments in technology will supply police departments with viable equipment that will greatly improve the effectiveness and efficiency of law enforcement personnel. Scientists within the Counterdrug Technology Assessment Center (CTAC) are operating with government agencies in the development of new technological devices that are going to be used by law enforcement agencies .

Augmented reality (AR) is a powerful new technology that is being developed. AR will provide situational awareness by projecting images into a person's real world vision. This device could aide law enforcement officers in several ways: (3). Identification Friend or Foe technology, worn by every police officer to reduce or eliminate friendly fire casualties by visually, audibly highlighting fellow police officers both on and off duty. Display of officer location, activity and status information projected on a 3-dimensional map of the community.

Modern systems of law enforcement are aimed at providing order and security, while at the same time maintaining liberty in an open society. With the rapid expansion of the use of computers, the Internet, and other information technologies, these dual objectives of policing are even more precarious, for the development of computer technology has also brought about novel opportunities to engage in illegitimate conduct. Because of the spread of computer technologies across the globe, moreover, the threats to computer security transcend the boundaries of individual countries. Pertinent law enforcement efforts, consequently, also have important international dimensions. This chapter addresses recent developments in the policing of computer security threats, paying special attention to national and international legal systems and their respective methods of enforcement.

From a legal and law enforcement viewpoint, measures against computer security threats pose problems of jurisdictional authority. National legal systems and their enforcement agencies are formally bound to nationally defined borders, whereas even a single transmission

MATS Journal of Engineering and Applied Science, Volume 1, Issue 3, ISSN 2394 - 0549 (Disclaimer - The authors are solely responsible for the contents of the research paper compiled in this seminar proceeding. The publishers or editors do not take any responsibility for the same in any manner. Errors, If any, are purely unintentional.

of computerized information over a network may pass through a dozen or more types of carriers, such as telephone companies, satellite networks, and Internet service providers, thereby crossing numerous territorial borders and legal systems (4). The cross-border nature of threats to computer security justifies the need for international cooperation and the development of global frameworks of law and law enforcement. In this chapter, we review the most important law enforcement efforts that have been taken at selected national and international levels to respond to the challenges affecting computer security.

Law enforcement is an important and necessary component among the efforts to maintain computer security. Because of the rapid and widespread expansion of computerized technologies and because of the border-transcending nature of computers linked through networks, the policing of threats against computer security presents a challenge to traditional means of crime detection and investigation on an international scale. Existing notions of jurisdictional authority have to be redefined to meet the global needs of information security. Trying to avert cybercrimes and the economic and social harm they can cause, many nations across the world have developed new legislation. Extending these legislative efforts are international systems of law, such as the European Convention on Cyber-Crime, to respond to the need for international legal cooperation and more adequately address cybercrimes and related cross-borders threats against computer security.

Without adequate law enforcement, laws remain ineffective. In the case of computer security, law enforcement agencies have instituted specialized computer crime teams to focus on the ways in which crimes can be perpetrated against or with the aid of computers. As with their accompanying legal systems, pertinent law enforcement activities often extend beyond the reach of jurisdictional boundaries, whether via cooperation among the police forces of different nations or through unilaterally enacted police actions abroad. International police operations pose special problems of coordination among the law enforcement agencies of various countries and they also lead us to rethink the need for police to preserve liberty and legitimate computer transactions while seeking to police computer crimes effectively.

Law enforcement efforts against threats to computer security do not respond merely to technological developments, but also take shape in specific sociohistorical circumstances. Since the terrorist attacks of September 11, 2001, many dimensions of law enforcement have undergone considerable changes, not only in terms of counterterrorism strategies but also with respect to other aspects of crime and crime control (5). The policing of computer security issues has also been altered since 9/11 because scores of systems relating to security, means of transportation and communication, and other public facilities rely heavily on computerized systems (6,7,8,). Given contemporary society's heavy reliance on computers, it is possible, for instance, for a terrorist group or individual to hack into the computers that oversee the subway system of a city or the railway network of a country. Following the attacks of 9/11, interest in and concern for computer security has skyrocketed, especially in connection with cyberterrorism. To be sure, cyberterrorism does not fully equate with cybercrime, but there is some overlap. For example, the initial stages of the offenses

MATS Journal of Engineering and Applied Science, Volume 1, Issue 3, ISSN 2394 - 0549 (Disclaimer - The authors are solely responsible for the contents of the research paper compiled in this seminar proceeding. The publishers or editors do not take any responsibility for the same in any manner. Errors, If any, are purely unintentional.

may be similar (e.g., sending out a computer virus), so that the response from a law enforcement viewpoint can be similar as well. But cybercrime and cyberterrorism differ in the harm they may cause and the motivation that is involved. In practice, however, the legislative responses— on both the national level and the international level—often confuse between the two offenses and have thus sped up the development of new means to police cybercrimes.

The strongest indicator of the changes affecting cyber-related matters in the post-9/11 era is the Uniting and Strengthening America by Providing Appropriate Tools Required to Intercept and Obstruct Terrorism (USA PATRIOT) Act in the United States (9,10). Among other provisions, the act gives authorities new powers by means of expanded options for wiretaps and technological systems of cyber surveillance. Relatedly, also, the National Cyber Security Division was created in the Department of Homeland Security in June 2003 as part of the National Strategy to Secure Cyberspace. Similar such new laws and the means to enforce them are now being set up in many other countries. Cyberlaw and law enforcement are a rapidly expanding reality. Ongoing developments indicate that, after several years of slowly responding to the threat of cybercrime, the events of 9/11 have served as an important catalyst to step up efforts to provide computer security through law and law enforcement. Although most cybercrimes do not relate to terrorism, the terrorist events of 9/11 may have provided the strongest impulse to develop new coordinated means against all types of cybercrime.

Legal challenges deterring and punishing computer criminals requires a legal structure that will support detection and prosecution of offenders. Yet the laws defining computer offences and the legal tools needed to investigate criminals using the internet often lag behind social and technological changes, creating legal challenges to law enforcement agencies. In India, however the I.T. Act 2000 has been enacted in pursuance to the General Assembly of United Nations resolution A/RES/51/162 dated the 30th Jan 1997. The I.T. Act 2000 elaborates upon digital signature, electronic governance attribution, acknowledge and dispatch of electronic records, secure electronic records etc. It specifies that Electronic Signatures will be valid and legally enforceable only if the e-transactions are done through ―public key cryptography‖ the Act delineates two separate types of penal provisions: contraventions and information technology offences. While contravention results in monetary penalty, I.T. Offences result in the offender being imprisoned or fined or both. Amendments have also been made to the IPC, Indian Evidence Act the Banker‘s Book Evidence Act and R.B.I. Act to facilitate investigation and prosecution of cyber crime. 48 (c) O

Transnational drug trafficking is a global concern with drug traffickers using latest technologies to avoid detection at land borders, airports and seaports by Customs officers. Traffickers use innocuous couriers and adopt novel modus operandi to smuggle drugs - inter alia. The Indian Customs, in addition to its primary function of revenue collection also actively takes preventive action against smuggling of contraband and narcotic drugs using modern information technology. Customs officers collect intelligence information from various sources to interdict narcotic drugs under the Narcotic Drugs and Psychotropic Substances Act. The National Academy of Customs,

MATS Journal of Engineering and Applied Science, Volume 1, Issue 3, ISSN 2394 - 0549 (Disclaimer - The authors are solely responsible for the contents of the research paper compiled in this seminar proceeding. The publishers or editors do not take any responsibility for the same in any manner. Errors, If any, are purely unintentional.

Excise and Narcotics (NACEN), is the key Indian Government institution which imparts training to all officers of the Customs, Central Excise, Service tax and Narcotics departments (11).

CONCLUSION

The 21st century has shed a whole new light on policing. Many causes for crime and disorder are now being originated outside of our borders. This fact demands new and innovative technology to be available for use by our police officers. International crime fighting has been added to policing responsibilities. Our police officers are trying to cope with many more aspect to crime and criminal than ever before Law enforcement is an important and necessary component among the efforts to maintain computer security. Because of the rapid and widespread expansion of computerized technologies and because of the border-transcending nature of computers linked through networks, the policing of threats against computer security presents a challenge to traditional means of crime detection and investigation on an international scale. Existing notions of jurisdictional authority have to be redefined to meet the global needs of information security. Trying to avert cybercrimes and the economic and social harm they can cause, many nations across the world have developed new legislation. Extending these legislative efforts are international systems of law, such as the European Convention on Cyber-Crime, to respond to the need for international legal cooperation and more adequately address cybercrimes and related cross-borders threats against computer security.

REFERENCES

1. Clare Edwards 2012 How Is Computer Technology Used in Law Enforcement? Opposing views Science by Demand Media.

2. How Technology Has Changed Law Enforcement Criminology Essay www.ukessays.com

3. Cowper & Buerger, 2003 Improving Our View of the World: Police and Augmented Reality Technology.European Journal of Criminology

4. Aldesco, A. I. (2002). The demise of anonymity: A constitutional challenge to the convention on cyber crime. Loyola of Los Angeles Entertainment Law Review, 23, 81–123.

5. Deflem, M. (2002b). Technology and the internationalization of policing: A comparative-historical perspective. Justice Quarterly, 19, 453–475.

6. Birnhack, M. D., & Elkin Koren, N. (2003). The invisible handshake: The reemergence of the state in the digital environment. Virginia Journal of Law and Technology, 8.

7. Brenner & Goodman, 2002 Brenner, S. W., & Goodman, M. D. (2002). In defense of cyberterrorism: An argument for anticipating cyber attacks. University of Illinois Journal of Law, Technology & Policy, 2002, 1–57.89.

8. Raghavan, T. M. (2003). In fear of cyberterrorism: an analysis of the congressional response. University of Illinois Journal of Law, Technology & Policy, 2003, 297–312.

9. Technology Law Journal. (2004). Cyberlaw: additional developments. Berkeley Technology Law Journal, 19, 543–553.

10. Copeland, R. A. (2004). War on terrorism or war on constitutional rights? Blurring the lines of intelligence gathering in post-September 11 America. Texas Tech Law Review, 35, 1–31.

11. Law Enforcement, Drug and Precursor Control. Annual Report 2014 www.unodc.org/.../customs-academy-to-roll-cbt-in-drug-law-enforcement India: National Customs Academy rolling out UNODC developed Computer Based Training in drug law enforcement

MATS Journal of Engineering and Applied Science, Volume 1, Issue 3, ISSN 2394 - 0549 (Disclaimer - The authors are solely responsible for the contents of the research paper compiled in this seminar proceeding. The publishers or editors do not take any responsibility for the same in any manner. Errors, If any, are purely unintentional.

Screening of potential α-amylase producers from the soil

Sanchali Padhye[1], Sapna Rai [1]and Namrata Sharma[2]
[1]Department of Microbiology, M.G.M.M.Jabalpur, M.P.
[2]Department of Biochemistry, M.G.M.M.Jabalpur, M.P.
E mail: padhyesanchali @yahoo.com

ABSTRACT:
Alpha-amylases are now gaining importance in biopharmaceutical applications. However, their application in food and starch based industries is the major market and thus the demand of alpha-amylases would always be high in these sectors. Amylases from microbial sources, especially fungi, have gained much attention because of the availability and high productivity of fungi, which are also amenable to genetic manipulation. In the present study, fungi have been isolated from the soil samples taken from garden soil, play ground soil and the garbage soil and subsequently plated on the PDA medium. Their amylase production potential was tested by plating them on starch agar medium. From the results obtained it is concluded that amylase producing fungi could be grown on potato dextrose agar medium and could be successfully isolated using starch agar medium. Amongst the isolated fungi *Rhizopus arrhizus* and *Aspergillus flavus* were found to be the largest amylase producers. *A. flavus* is one of the most potential α-amylase producers; it was selected for observing the effect of various cultural parameters on the growth of amylolytic fungi. The results indicate that adding buffer to the fermentation broth certainly increased the yield of the test fungus.

Keywords: amylase, fungi

INTRODUCTION:

Many soil dwelling bacteria and fungi produce amylase. Bacteria like *Bacillus amyloliquefacicns , B. caldolyticus, B. coagulans, B. lichenformii, B. subtilis, Escherichia spp.., Lactobacillus spp.., Micrococcus spp., Pseudomonas spp., etc produce amylase.* While, fungi like Aspergillus, *Candida, Mucor, Neurospora, Rhizopus, Penicillium*, etc produce amylases. Fungi are eukaryotic organisms, which appear either in the form of thread-like mycelium or are cylindrical yeasts. Certain fungi produce amylase, which breaks down starch. Thus, the organism derives energy by this process on the one hand, and the waste starch is removed from the environment on the other. This enzyme has many industrial applications and is used for removing starch from the environment. Fungal amylase has a greater industrial application than bacterial amylase. The Amylase producing fungi are ubiquitous and could be easily isolated from the soil. They degrade starch and related polymers to yield products characteristic of individual amylolytic enzymes. The term amylase was used originally to designate enzymes capable of hydrolyzing alpha-1, 4-glucosidic bonds of amylose, amylopectin, glycogen and their degradation products **(Bernfeld, 1955**). They act by hydrolyzing bonds between adjacent glucose units, yielding products characteristic of the particular enzyme involved.

Amylase is present in human saliva, where it begins the chemical process of digestion. All amylases are glycoside hydrolases and act on α-1,4-glycosidic bonds. They break down large molecules (starch) into smaller ones, so the intestine could absorb them. Amylases can be divided into two categories: Endoamylases and Exoamylases. Endoamylases catalyze hydrolysis in a random manner in the interior of the starch molecule producing linear and branched oligosaccharides of various chain lengths. Exoamylases act from the non-reducing end successively resulting in short end products (Gupta et al., 2003).

Amylases can also be classified into α-Amylase, β-Amylase and γ-Amylase. The most important type of amylases is the alpha-amylase. It aids the breakdown of starch to maltose. Alpha-amylase hydrolyzes bonds between glucose repeats.

Starch is a heterogeneous polysaccharide composed of two high molecular weight entities called amylose and amylopectin. These two polymers have different structures and physical properties. Starch may be separated into its two components (amylose and amylopectin) by addition of a polar solvent, e.g. n-butanol, to a dispersion of starch. The insoluble amylose complex can then be separated from soluble amylopectin fraction. Amylose is composed of linear chains of α -1, 4 linked D-glucose

MATS Journal of Engineering and Applied Science, Volume 1, Issue 3, ISSN 2394 - 0549 (Disclaimer - The authors are solely responsible for the contents of the research paper compiled in this seminar proceeding. The publishers or editors do not take any responsibility for the same in any manner. Errors, If any, are purely unintentional.

residues. Hence it is extensively degraded by α ively degraded by 5555555555555555555555555-amylase. Amylopectin may account for 75 to 85% of most starches. The hydrolysis of starch may be carried out using either acid or enzyme as catalyst. The enzymatic hydrolysis of starch has been practiced on an industrial scale for many years and is gradually replacing the traditional acid hydrolysis process. Thus, amylases constitute a class of industrial enzymes having approximately 25% of the enzyme market (Rao et al., 1998). Demand for novel amylases worldwide is increasing day by day, as the application spectra of these enzymes are spreading in various industrial sectors. The classical approach is the isolation of microbial species, which produce novel enzyme from exotic environments and would offer a competitive advantage over the existing products.

REVIEW OF LITERATURE:

Amylases are produced by a variety of living organisms, ranging from bacteria to plants and humans. Bacteria and fungi secrete amylases to the outside of their cells to carry out extra-cellular digestion. Although, the use of amylases, alpha amylases in particular, in starch liquefaction and other starch based industries has been prevalent for many decades and a number of microbial sources exists for the efficient production of this enzyme, the commercial production is limited to only a few selected strains of fungi and bacteria. Moreover, the demand for these enzymes is further limited with specific applications as in the food industry, wherein fungal alpha-amylase are preferred over other microbial sources due to their more accepted GRAS status. Filamentous fungi are attractive organisms for production of useful proteins and biologically active secondary metabolites. These fungi produce high levels of polysaccharide-degrading enzymes and are frequently used for the production of industrial enzymes. Fungi have high secretion capacity and are effective hosts for the production of foreign proteins (Tsukagoshi et al., 2001).

Recent data show that fungal alpha-amylases from Aspergillus oryzae and Aspergillus niger (which hydrolyse internal alpha-l, 4 glycosidic bonds in starch, producing maltose and maltotriose), cyclodextrin glucosyltransferases from Bacillus circulans (which produce cyclodextrins from starch), and oligo-l, 6-alpha-glucosidase from Bacillus cereus (Bce) all possess the same basic structure.Amylase could be obtained from bacteria and fungi by different methods.Obtaining amylase from fungi is very simple.Due to their diversity, fungi have been recognized as a source of new enzymes with useful and/or novel characteristics (Bakri et al., 2009).

However, enzymes from fungal and bacterial sources have dominated applications in industrial sectors (Pandey et al., 2000). Several Bacillus sp. and thermostable Actinomycetes including Thermomonospora and Thermoactinomyces are versatile producers of the α-amylases. The genus Bacillus produces a large variety of extracellular enzymes, of which amylases is of significant industrial importance. An extremely thermostable α-amylase is available from the mesophile B. licheniformis. Fungi, due to their diversity, have been recognized as a source of new enzymes with useful and/or novel characteristics (Bakri et al., 2009). Amylases from microbial sources, especially fungi (Aspergillus spp.), have gained much attention because of the availability and high productivity of fungi, which are also amenable to genetic manipulation (Kathiresan and Manivannan, 2006,Sidkey et al., 2010), it is oftenly selected for the purpose of production, purification and investigating properties of the enzyme.

Among physical parameters, pH of the growth medium plays an important role by inducing morphological changes in microbes and in enzyme secretion. The pH change observed during the growth of microbes also affects product stability in the medium (Rani Gupta et al., 2003). The influence of temperature on amylase production is related to the growth of the organism. Hence, the optimum temperature depends on whether the culture is mesophilic or thermophilic. Among the fungi, most amylase production studies have been done with mesophilic fungi within the temperature range of 25–37 °C. Thermostable amylase is especially employed for industrial applications.

The new potential of using microorganisms as biotechnological sources of industrially relevant enzymes has stimulated renewed interest in the exploration of extracellular enzymatic activity in several microorganisms. (A Saleem et al., 2014) Fungal species have been studied a lot for the production of alpha amylase because of the low cost of substrates used for the production of alpha amylases(Jahir Alam Khan*and Sachin Kumar Yadav 2011).

The major markets for amylases are food industries for the preparation of sweeteners and syrups. The pigments produced by Penicillium (PP-V and PP-R) and Monascus (monascorubrine and monascuscorubramine) are structurally similar. Lovastatins or monacolins produced by Penicillium, Monascus, Aspergillus and Rihzopus inhibits

MATS Journal of Engineering and Applied Science, Volume 1, Issue 3, ISSN 2394 - 0549 (Disclaimer - The authors are solely responsible for the contents of the research paper compiled in this seminar proceeding. The publishers or editors do not take any responsibility for the same in any manner. Errors, If any, are purely unintentional.

cholesterol biosynthesis by binding to catalytic site of HMG-CoA reductase a key enzyme in cholesterol biosynthesis and scavenged DPPH radicals (Dhale et al., 2007). Fungal amylases are used for hydrolyzing carbohydrate, protein and other constitutes of soy beans and wheat into peptides, amino acids, sugars and other low molecular weight compounds in soy sauce production (Negi and Banerjee, 2009).

With the advent of new frontiers in biotechnology, the spectrum of amylase applications has expanded into many new fields such as clinical, medicinal and analytical chemistry (Pandey et al., 2000). The demand for amylase is increasing day by day because of its magnificent potentiality in the above mentioned industrial sectors taking into consideration that the properties of α-amylases such as thermostability and pH profile should match the application.

MATERIALS AND METHOD:

Sample collection:

Soil samples were collected for isolating amylolytic fungi from the following places:

A) **Garden soil**: from own garden, located at Hari Singh Colony, Marhatal, Jabalpur.
B) **Play ground soil**: from the play ground of Hari Singh Colony, Marhatal, Jabalpur.
C) **Garbage soil**: from the garbage dump of the vegetable market located at Niwarganj, Jabalpur.

The collected soil samples were used for preparing dilution series

Seeding of dilutions on media:

Potato dextrose agar medium was prepared and autoclaved at 121 lbs pressure for 15 minutes. The pH was adjusted to 5.6 and chloramphenicol was added to it, before pouring it into plates. 0.1 ml of 10^{-10} was spreaded on the culture The plates were then incubated at 30^{O}C for 72 hours.

Identification and preservation of isolated fungi:

The plates were observed for growth, after 72 hours of incubation. Both, macroscopic and microscopic observations of the isolated fungi were taken and recorded. The isolated fungi were identified by studying their characteristics. The identified fungi were preserved by transferring them in slants containing potato dextrose agar medium, incubated at 30 ± 2^{O}C and were then maintained at 4^{O}C.

Estimating the amount of amylase produced by different fungi identified and comparing them on this basis:

(A) Estimating the growth of amylase producing fungi on Starch Agar media:

Fungi isolated from different soil sources were grown on starch PDA. They were incubated at 30 ± 2 ^{o}C for 72 hours.

Percentage of different fungi was calculated as follows:

Growth percent (%) of a fungi = Number of colonies of that fungi

VII. TOTAL NUMBER OF COLONIES OBTAINED

(B) Study of amylase production by different isolated fungi (Amylase test):

The isolated fungi were subjected to amylase test. After sporulation, the fungi were tested to find the extent of amylase production. Iodine was added to fungi grown on starch agar medium. Iodine reacts with starch to form a blue color. If on adding iodine, a blue color appears around the colonies, it indicates that starch is present even in the medium surrounding the colonies. This implies that the fungi have not produced enough amylase to break down the starch. The impact of various cultural parameters (pH, temperature and addition of buffers) on its growth was also observed.

Determination of the effect of pH:

The test fungus, i.e., Aspergillus flavus was innoculated into fermentation broth and the pH was adjusted to 5.0, 6.5 and 9.0 respectively. It was then incubated at 30+2 oC for 72 hours to observe the effect of pH on fungal growth.

Thermotolerence Test:

Fermentation broth was inoculated with the test fungus. It was kept at 20oC, 30oC and 40oC for 72 hours. The fungal growth was observed and measured after 72 hours.

Determination of effect of buffer:

Buffers, i.e., Tris borate EDTA buffer and 3(N-morpholino) propane sulfonic acid buffer were respectively added to fermentation broth. To this, the test fungus was added and incubated for 72 hours at 30+2 oC. The growth was measured after 72 hours.

MATS Journal of Engineering and Applied Science, Volume 1, Issue 3, ISSN 2394 - 0549 (Disclaimer - The authors are solely responsible for the contents of the research paper compiled in this seminar proceeding. The publishers or editors do not take any responsibility for the same in any manner. Errors, If any, are purely unintentional.

Table No.1: Number of different fungi isolated from various soil samples (garden, play ground & garbage soils):

Fungi	Fungi from garden soil	Fungi from play ground soil	Fungi from garbage soil	Total number of fungi
Absidia sp.	---	2	---	2
VIII. ASPERGILLUS FLAVUS	2	---	1	3
Aspergillus fumigatus	---	3	---	3
Aspergillus niger	2	1	2	5
Aspergillus sp.	1	1	---	2
Cladosporium sp.	---	---	1	1
Curvularia lunata	---	2	---	2
Fusarium sp.	---	2	---	2
Pythium sp.	---	---	3	3
Rhizopus arrhizus	3	---	---	3
A. ***TOTAL***	8	11	7	26

Table No.2: Frequency of fungi isolated from different soil samples:

Fungi	Frequency of fungi from garden soil (%)	Frequency of fungi from play ground soil(%)	Frequency of fungi from garbage soil(%)
Absidia sp.	---	18.18	---
IX. ASPERGILLUS FLAVUS	25	---	14.28
Aspergillus fumigatus	---	27.27	---
Aspergillus niger	25	9.09	28.57
Aspergillus sp.	12.5	9.09	---
Cladosporium sp.	---	---	14.28
Curvularia lunata	---	18.18	---
Fusarium sp.	---	18.18	---
Pythium sp.	---	---	42.85
Rhizopus arrhizus	37.5	---	---

MATS Journal of Engineering and Applied Science, Volume 1, Issue 3, ISSN 2394 - 0549 (Disclaimer - The authors are solely responsible for the contents of the research paper compiled in this seminar proceeding. The publishers or editors do not take any responsibility for the same in any manner. Errors, If any, are purely unintentional.

Table No.3: Production of amylase (Amylase Test):

A. ***Fungi***	**Color after adding iodine**	*1)* ***Inference***
X. ASPERGILLUS FLAVUS	XI. MEDIA SURROUNDING COLONIES=> NOT BLUE. XII. REMAINING MEDIA => VERY LIGHT BLUE.	Amylase production: **Abundant**
Aspergillus fumigatus	Media surrounding colonies=> not blue. Remaining media => blue.	Amylase production: **Average**
Aspergillus niger	Media surrounding colonies=> light blue Remaining media => dark blue	Amylase production: **Below average**
Pythium sp.	Whole media turned blue	Amylase production: **Negligible**
Rhizopus arrhizus	XIII. NO BLUE COLOR AT ALL	Amylase production: **Maximum**

Based on the results, the amylase producing ability of the tested fungi is in the following order:

R. arrhizus > A. flavus > A. fumigatus > A. niger > Pythium sp.

Table No.4: Effect of various cultural parameters on growth of *Aspergillus flavus*

Parameters		**XIV. DAYS OF INCUBATION**	
		2nd day of incubation	**7th day incubation (wet mycelial weight in gm.)**
Temperature	20 °C	+	3.802
	30 °C	+++	5.225
	$40^{O}C$	++	5.822
pH	5.0	+	1.756
	6.5	+++	3.917
	9.0	++	2.370
Buffers	Fermentation broth	+	3.518
	Tris borate EDTA buffer	+++	5.206
	3(N-morpholino)propane sulfonic acid buffer	++	5.964

RESULTS AND DISCUSSIONS:

The fungi isolated were identified, by observing their macroscopic and microscopic observations. The fungi isolated from various soil samples and their frequencies are expressed in table no.1 and 2. The common fungus isolated from garden soil was *Rhizopus arrhizus* followed by A*spergillus flavus* and *Aspergillus niger.* From play ground soil the common fungus isolated were A*spergillus flavus*, A*spergillus fumigatus. Fusarium sp., Absidia sp., Chlamydosporum sp.* and *Curvularia lunata.* In garbage soil Dominant species were Aspergillus *flavus and* A*spergillus niger* . The other species isolated include: *Cladosporum sp. and Pythium sp*. the results of their amylase production ability are summarised in Table No.3.

MATS Journal of Engineering and Applied Science, Volume 1, Issue 3, ISSN 2394 - 0549 (Disclaimer - The authors are solely responsible for the contents of the research paper compiled in this seminar proceeding. The publishers or editors do not take any responsibility for the same in any manner. Errors, If any, are purely unintentional.

These dominant species were: A*spergillus flavus*, A*spergillus fumigates*, A*spergillus niger* and *Rhizopus arrhizus*

Table No.3 shows the results of Iodine test. According to it, *Rhizopus arrhizus* showed no blue color, *Aspergillus flavus* showed a very light blue color. While, *Aspergillus fumigatus* and *Aspergillus niger* showed a dark blue color. The amount of amylase production by the test fungi (in decreasing order) is:

Rhizopus arrhizus* > *Aspergillus flavus* > *Aspergillus fumigatus* > *Aspergillus niger

The top two amylase producers i.e., *Rhizopus arrhizus* and *Aspergillus flavus* were grown in fermentation media at 30^{O}C for seven days. Their yield were:

***Rhizopus arrhizus* = 9.18 grams (wet weight)**
***Aspergillus flavus* = 8.82 grams (wet weight)**

pH of the growth medium plays an important role by inducing morphological changes in microbes and in enzyme secretion. Table No. 4 shows that maximum growth of *Aspergillus flavus* occured at pH 6.5. Lesser growth was recorded at the pH of 9.0 While, least growth was found at the pH of 5.0.

Temperature plays a very important role in the growth of microorganisms. Temperature optimum for amylase was found to be in the range of 25-37°C for the mesophilic fungi **as reported to Rani Gupta et al., 2003. Kathiresan and Manivannan, 2005** found the optimum temperature being 30^{O}C in case of *Penicillium fellutanum*. The influence of temperature on amylase production is related to the growth of the organism. Among the fungi, most amylase production studies have been done with mesophilic fungi within the temperature range of 25–37 °C.

In the present study, the test organism, *Aspergillus flavus* showed adequate growth at 30^{O}C. While, maximum growth was achieved at 40^{O}C. Addition of buffers to the fermentation media was found to have augmented the growth of the test organism, *Aspergillus flavus*. The growth yield in TEB buffer i.e (Tris borate EDTA buffer), buffer was higher as compared to the MOPS i.e. 3(N-morpholino) propane sulfonic acid buffer.

CONCLUSION:

Fungi isolated isolated from the soil samples Rhizopus arrhizus and Aspergillus flavus were the largest amylase producers. As reported by Sidkey et al., 2010, A. flavus is one of the most potential α-amylase producer it was selected for observing the effect of various cultural parameters (temperature, pH and addition of buffer) on the growth of amylolytic fungi.

Temperature plays a significant role in the activity of the produced α-amylase. In the present study, it was found that after the second day of incubation, maximum growth occured at 30oC, while after the seventh day of incubation, maximum growth was recorded at 40oC. Many investigators have studied this parameter and our results are in complete accordance to some of them. Significant growth of A flavus occurs at 30oC and reached maximum level at 45oC similar to that reported by Z.A. Omala and S.A. Sabry, 1989. The optimum temperature of the purified α-amylase was 30oC from each of A. flavus (Abou-Zeid, 1997), A. niger JGI 24 (Varalakshmi et al., 2009), Penicillium camemberti PL21 (Nouadri et al., 2010); 40 oC from A. niger NRRL-337 (Mahmoud et al., 1978) and 45oC from strain marine Streptomyces sp. D1 (Chakraborty et al.,2009).

The present study revealed that maximum growth of A. flavus occured at the pH of 6.5. In comparison with our result, A. flavus recorded maximum growth at initial pH of 7 (Z.A. Omala and S.A. Sabry,1989). Maximum α-amylases activity was obtained at pH 7.0 for A. flavus (Abou-Zeid, 1997), pH 6.0 for A. flavus (Khoo et al., 1994; Sidkey et al., 1997), pH 6.2 for A. flavus var columinaris (El-Safey and Ammar, 2004), pH 6 from A. flavus (Sidkey et al., 1997), pH 4.3 from A. niger NRRL-337 (Mahmoud et al., 1978).

The results indicate that adding buffer to the fermentation broth certainly increased the yield of the test fungus. Thus, from the present study, it is concluded that amylase producing fungi could be successfully isolated from soil. Further, it is proved that the cultural parameters like temperature, pH and addition of buffers to the growth medium do remarkably impact the growth of amylase producing fungi.

ACKNOWLEDGEMENT

We are thankful to Mata Gujri Mahila Mahavidyalaya for providing the lab facilities to accomplish present work.

References:

1. A. Saleem , Mohsen K.H. Ebrahim(2014) Production of amylase by fungi isolated from legume seeds collected in Almadinah Almunawwarah, Saudi Arabia. *Journal of Taibah University for Science* 8(2):90-97

2. Bairoch A. (2000). "The enzyme database in 2000". *Nucleic Acids Research* 28 (1): 304–305

3. Bernfeld P(1986). "Amylases, alpha and beta*". Met/i. Enzymology* 1:149-58, 1955.

4. Birgit Schwermann, Karsten Pfau, Birgit Liliensiek, Manfred Schleyer, Fischer Thomas and Evert P. Baker (1994) "Purification, properties and structural aspects of a thermoacidophilic alpha-amylase from *Alicyclobacillus acidocaldarius"* - *European Journal of Biochemistry* Vol 226 Issue 3 94 0915/3.

5. Cech T (2000). "Structural biology. The ribosome is a ribozyme". *Science* Vol 289 (5481): page 878–9.

6. Crueger, W. and Crueger, A. 1989. "A Textbook of Industrial Microbiology".2nd ed, *Sinauer Association.,Inc. Sunderland* pp 59-63

7. Dhale. A. Mohan and Vijay-Raj A. S. (2007) "Pigment and amylase production in *Penicillium* sp NIOM-02 and its radical scavenging activity" *International Journal of Food Science and Technology, vol.44(12); 2424-2430.*

8. Ghalanbor, Z; et al. (2008). "Binding of Tris to Bacillus licheniformis alpha-amylase can affect its starch hydrolysis activity.". *Protein Peptide Lett.* 15 (2): 212–214.

9. Gupta, A., Gupta V.K.,Thomas., Modi D.R. and Yadava, L.P. (2008). "Production and characterization of α-amylase from *Aspergillus niger*." *Biotechnology*, Vol7: 551-556.

10. Gupta, Rani and Mohapatra, Harapriya (2003). "Microbial biomass: An economical alternative for removal of heavy metals from waste water*." Indian Journal of Experimental Biology, 41 (9):* 945-966.

11. Hill, Robert and Needham, Joseph (1970), "The Chemistry of Life: Eight Lectures on the History of Biochemistry" (*London, England: Cambridge University Press,, page 17).*

12. Jahir Alam Khan*and Sachin Kumar Yadav (2011). Amylases by *Aspergillus niger* using cheaper substrates employing solid state fermentation. *International Journal of Plant, Animal and Environmental Science.*1(3): 100-108.

13. Jenshinn Lin[1], Yeong-Shenn Lin[2], Sho-Tin Kuo[1], Chii-Ming Jiang[3], Ming-Chang Wu[1] (2009)"Purification of -amylase from human saliva by super paramagnetic particles";' *Journal of the science of food and agriculture*. Vol 89, 4-574-578.

14. Kathiresan K. and Manivannan S. (2005)"Amylase production by *Penicillium fellutanum* isolated from mangrove rhizosphere soil*" African Journal of Biotechnology (10: 829-832, 2006).*

15. Lilley D (2005). "Structure, folding and mechanisms of ribozymes*". Curr Opin Struct Biol* 15 (3): 313–23.

16. Mapp CE (May 2001). "Agents, old and new, causing occupational asthma". *Occup Environ Med* 58 (5): 354–60, 290.

17. Noto, Yuka; Tetsumi Sato, Mihoko Kudo, Kiyoshi Kurata, and Kazuyoshi Hirota (2005). "The Relationship Between Salivary Biomarkers and State-Trait Anxiety Inventory Score Under Mental Arithmetic Stress: A Pilot Study." *Anesthesia & Analgesia* Vol: 101, Issue: 6, Pages: 1873-1876.

18. Pandey, Ashok;Nigam, Poonam; Soccol, Carlos R.; Soccol, Vanete T; Singh, Dalel and Radjiskumar ,Mohan (2000) "Advances in microbial amylases" * *Biotechnology and Applied Biochemistry* Volume: 31 (Pt 2), Issue: 2, Pages: 135-152.

19. Perry, GH, et al. (2007). "Diet and evolution of human amylase gene copy number variation," *Nature Genetics* 39:1256-1260.

20. Prakasham R.S., Subba Rao Ch., Sreenivas Rao R. and Sarma P.N.(2007) "Enhancement of acid amylase production by an isolated Aspergillus wamori",

MATS Journal of Engineering and Applied Science, Volume 1, Issue 3, ISSN 2394 - 0549 (Disclaimer - The authors are solely responsible for the contents of the research paper compiled in this seminar proceeding. The publishers or editors do not take any responsibility for the same in any manner. Errors, If any, are purely unintentional.

21. *Journal of Applied Microbiology* 102 204–211.
22. Prasanna V. Aiyer (2005) "Amylase and their applications*." African Journal of Biotechnology* Vol. 4 (13), pp. 1525-1529.
23. Quillaguamán J, Hashim S, Bento F, Mattiasson B, Hatti-Kaul R(2005) "Poly(beta-hydroxybutyrate) production by a moderate halophile, Halomonas boliviensis LC1 using starch hydrolysate as substrate*." Journal of applied microbiology* ; 99(1):151-7.
24. Rao, Uma; Marui,Junichro; Kato, Masashi; Tsukagoshi; Norihiro Tetsuo (2002) "Regulation of the xylanase gene, *cgxA*, from *Chaetomium gracile* by transcriptional factors, XlnR and AnRP" *Biotechnology Letters* Vol 24 ,no.13, 1089-1096.
25. Reddy N.S., Nimmagadda, Annapoorna and Sambasiva Rao K.R.S. (2003) "An overview of the microbial α-amylase family" *African Journal of Biotechnology* Vol. 2 (12), pp. 645-648, ISSN 1684-5315.
26. Safarikova, M; Royb, I., Gupta M.N; Safarik I. (2003) "Magnetic alginate microparticles for purification of alpha-amylases" *Journal of Biotechnology* 105: 255-260.
27. Sidkey, Nagwa M.; Abo-Shadi, Maha A.;Balahmar, Reham; Sabry, Reham ; Badrany Ghadeer (2011) "Purification and characterization of α-amylase from a newly isolated *Aspergillus flavus* F2Mbb" *International Research Journal of Microbiology Vol. 2(3) pp. 096-103.*
28. Silverman, Richard B. (2002) "The Organic Chemistry of Enzyme-catalyzed Reactions*",. Academic Press, London, England*, 2nd ed, page 1.
29. Udani J, Hardy M, Madsen DC (2004). "Blocking carbohydrate absorption and weight loss: a clinical trial using Phase 2 brand proprietary fractioned white bean extract". *Altern Med Rev* 4 (1): 63–9.
30. Voet, D., & Voet, J. G. (2005). Biochemistry. (2e éd.). *Bruxelles: De Boeck.* 1583 p.
31. Yasser Bakri, Masson Magali and Philippe Thonart "Isolation and identification of a new fungal strain for amylase biosynthesis." *Polish journal of microbiology* Vol: 58, Issue: 3, Pages: 269-273.

MATS Journal of Engineering and Applied Science, Volume 1, Issue 3, ISSN 2394 - 0549 (Disclaimer - The authors are solely responsible for the contents of the research paper compiled in this seminar proceeding. The publishers or editors do not take any responsibility for the same in any manner. Errors, If any, are purely unintentional.

Outlier Analysis on Financial Markets: A Survey

Gitesh Kumar Markandey[1]
M. Tech Scholar, Department of Computer Science & Engineering,
Rungta College of Engineering & Technology,
Bhilai CG. India.
Email : gitkumar@gmail.com

Devesh Narayan[2]
Asst. Professor, Department of Computer Science & Engineering
Rungta College of Engineering & Technology,
Bhilai CG.India
Email : devesh.narayan@rungta.ac.in

Abstract Data mining is the extraction of hidden predictive information from big databases. It is a new technology which gives great potentiality to help companies focus on the most important and useful information in their data warehouses. Data mining uses statistical, machine learning and visualization techniques to discovery and present knowledge in a form which is easily comprehensible to humans. Various popular data mining tools are available today. Data mining tools predict trends for the future and behaviors this allows businesses to make proactive and lossless decisions. Data mining tools can solve business problems that traditionally were too time consuming to resolve. They extract databases for hidden patterns and finding predictive information that experts may miss because it lies outside their expectations. In this paper we aim to provide a structural approach in outliers of stock markets data. This paper presents various attributes related to the outlier detection methodologies and its specific approach of analysis.

Index Terms—Data mining, Outlier detection, Temporal data

1.INTRODUCTION

Data Mining is a process of extraction of Knowledge base from the databases & the nontrivial extraction of implicit, potentially useful and information from data in databases. data mining is a part of the finding the useful knowledge in databases.

Data mining finds the relevant patterns in a database by using defined approaches, algorithms and methods to look into current and historical data that can then be analyzed to predict the future trends. Because Data mining allows us to see the future trends and behaviors by reading through databases for hidden patterns, it allows organizations to make proactive, knowledge based decisions and answer questions that were previously too time-consuming to resolve. Business organizations that wish to use data mining tools can purchase mining programs designed for existing software and hardware platforms. These platforms can be integrated into new products and systems in online context. They can build their own custom mining solution.

The output of a Data Mining exercise into another computer system like neural network. It is a quite common and can give the mined data more value because the data mining tool gathers the database, while the second program (e.g., the neural network) makes decisions based on the data collected for any purpose. Different types of data mining tools are available in the market, each with their own strengths and weaknesses. The auditors need to be aware of the different kinds of data mining tools available and recommend the purchase of a

MATS Journal of Engineering and Applied Science, Volume 1, Issue 3, ISSN 2394 - 0549 (Disclaimer - The authors are solely responsible for the contents of the research paper compiled in this seminar proceeding. The publishers or editors do not take any responsibility for the same in any manner. Errors, If any, are purely unintentional.

tool that matches the business organization's current detective needs.

M.J. Zhou & J.C. Tao [9] has proposed the outlier data mining and commonly used outlier mining methods, on this basis of outlier mining algorithm based on attribute entropy (OMABAE). Firstly, the concept of attribute entropy is introduced to calculate attribute entropy of each attribute, and constructs the attribute matrix of entropy . Secondly, the object entropy of each object is computed according to the attribute matrix of entropy and finally outlier will be detected by comparing the object deviation degree with entropy threshold. The result shows that this algorithm can detect outlier efficiently.

2.Stock Market Analysis

2.1Stock Fraud & The Manipulators

Fraud in Stock usually takes place when brokers try to manipulate their customers into trading stocks without regard for the customers' own real interests. Stock fraud can be at a company level or it can be committed by a single stockbroker. Corporate insiders, underwriters, brokers, large shareholders and market makers are likely to be manipulators.

2.2 Why Stock Fraud Detection is Necessary

Several fraud detection methods are available for the fields like credit card, telecommunications, network intrusion detections etc. But the area of stock market fraud detection is still behind. Since stock market enhances the economic development of a country greatly, this field has a vital need for efficient security system. Also the amount of money involved in stock market is very big and huge. So, appropriate fraud detection system is essential. For example, in Australia, 64 per cent of people's superannuation, like their retirement savings, is invested in security deposits. Investment in stock market is high in almost all the countries. If we don't protect against the ability of people to manipulate those securities, then implicitly, we're open to attack, or we're allowing open to attack a country's very wealth Stock fraud may not be very frequent but when it occurs the amount of loss is abundant.

A. Rahmani & S. Afra [7] has proposed the concepts on Outlier detection. It has a large variety of applications ranging from detecting intrusion in a computer network, to forecasting tornados and hurricanes in weather data and to identifying indicators of potential crisis in stock market etc. The problem of finding outliers in sequential data has been widely studied in the data mining literature and many techniques have been developed to tackle the problem in various application domains. Many of these techniques rely on the peculiar characteristics of a specific type of data to detect the outliers. As a result, It cannot be easily applied to different types of data in other application domains; they should at least be tuned and customized to adapt to the new domain. They also may need certain amount of training data to build their models. The work described by A. Rahmani & S. Afra [7] tackle the problem by proposing a graph-based approach for the discovery of contextual outliers in sequential data.

The developed algorithm offers a higher degree of flexibility and requires less amount of information about the nature of the analyzed data compared to the previous approaches described in the literature. In order to validate the approach, A. Rahmani & S. Afra [7] has proposed experiments on stock market and weather data; they compared the results with the results from

MATS Journal of Engineering and Applied Science, Volume 1, Issue 3, ISSN 2394 - 0549 (Disclaimer - The authors are solely responsible for the contents of the research paper compiled in this seminar proceeding. The publishers or editors do not take any responsibility for the same in any manner. Errors, If any, are purely unintentional.

our previous work. The analysis of the results demonstrate that the algorithm proposed by A. Rahmani & S. Afra [7] is successful and effective in detecting outliers in data from different domains, financial and the other meteorological.

3.Graph-based Outlier Detection

A.Rahmani & S. Afra [7] presented a graph-based outlier detection algorithm. Our algorithm uses the sliding window concept in combination with a MST-based clustering algorithm to cluster the instances of a given dataset. It then finds outliers by identifying the minority clusters in different regions of the data. Results from the conducted experiment show that our proposed algorithm is successful in detecting outliers in sequential data. However, the accuracy and efficiency of the proposed approach highly depends on the
appropriate configuration of the input parameters and thresholds.

The goal is to reduce the worst case run time of our algorithm to O (nlgn) or less, which has been reported in the literature as the run time of many similar outlier detection algorithms.

4.Stock Price Manipulations

D. Diaz & B. Theodoulidis [6] has proposed the concepts of challenges relating to applying data mining techniques to detect stock price manipulations and extends the previous results by incorporating the analysis of intraday trade prices in addition to closing prices for the investigation oftrade-based manipulations. This work extends previous results on the topic by analysing empirical evidence in normal and manipulated hourly data and the particular characteristics of intraday trades within suspicious hours. The analytical models described in the paper reinforce the results of previous market manipulation studies that are based on traditional statistical and econometrical methods providing an alternative portfolio of methods and techniques originating from the data mining and knowledge discovery areas. With the application of the analytical oriented approach described in this paper, it is possible to identify new fraud pattern characteristics encoded as decision trees which can be readily employed in fraud detection systems.

D. Diaz & B. Theodoulidis [6] also proposed a number of policy recommendations towards increasing the effectiveness of the operational processes executed by stock exchange fraud departments and regulatory authorities.

The Development of a data mining approach for the market manipulation problem that focuses explicitly on presenting structures in financial data in a simpler and comprehensible way In this sense, the analytical model is not only powerful in classifying trades, but also identifies new patterns associated with market manipulations in

both hourly and second by second levels.

The study confirms that quarter-ends and year ends, as well as closing hours, are common preconditions for the manipulations, confirming the existence of higher liquidity, returns and volatility associated with the manipulated sample In fact, a high presence of returns, volatility and volume outliers is directly related with the manipulation sample, nonetheless, not all those outliers are directly linked with manipulations due to the moderation effect of new information addition, when liquidity and volatility are within

MATS Journal of Engineering and Applied Science, Volume 1, Issue 3, ISSN 2394 - 0549 (Disclaimer - The authors are solely responsible for the contents of the research paper compiled in this seminar proceeding. The publishers or editors do not take any responsibility for the same in any manner. Errors, If any, are purely unintentional.

'normal' ranges, an important proportion of suspicious intraday trades presents abnormal returns. Moreover, when returns are within normal ranges, isolated jumps in liquidity are associated with suspicious trades in more than 20% of the cases. In the same line, isolated jumps in volatility are associated with suspicious blocks in a similar proportion of the cases. In terms of policy recommendations, these patterns confirm the importance of monitoring closely the volume of trading and the volatility of returns as sudden changes in these indicators are frequently associated with manipulations.

Despite the extensive work reported in this concept, there are still challenges that need to be addressed through further investigation. This work shares several of the caveats of previous work regarding a selection bias towards poor manipulations. This stems from the fact that only those cases in which at least a suspect of manipulation is present, either by a formal SEC or Press investigation, have been used. It is possible, therefore that cases may have been ignored in which manipulations occur but were not observed or reported, and cases in which the SEC did indeed carry out an investigation but this did not result in any action against the fraudsters due to lack of evidence or budget constraints. Regarding the coverage of the data sets, working only with one year of precredit-crunch crisis data can potentially skew the outcomes, considering there could have been structural changes in the way the financial markets and fraudsters work that are not taken into consideration. Furthermore manipulations involving the use of the internet, like pump-and-dump schemes using spam emails or other sophisticated strategies have not been considered. Further research will address insider trading as applied to both derivatives and the securities markets and also consider particular trade-based manipulations that do not require the transactions to be fulfilled. In this case, it is possible to alter prices and volume through ill-intentioned limit market orders placed near or at the closing of the daily transactions Finally, future work will also consider different ways of studying the patterns in quotes data and linking them with the analysis of social networks of fraudsters.

M. Gupta, J. Gao [11] presented the work on the statistics branch The outlier detection for time series data has been studied for decades. Recently, with advances in software technology as well as in hardware technology, there has been a large body of work on outlier detection from a computational perspective within the computer science field. The advances in hardware technology have enabled the availability of various forms of temporal data collection mechanisms, and advances in software technology field have enabled lots of data management mechanisms. This is the reason of growth of different kinds of data sets such as data streams,spatio-temporal data, time series data, distributed streams, temporal networks, generated by the applications. There is a need for an organized and detailed study of the work done in the area of outlier detection with respect to such temporal datasets.

M. Gupta, J. Gao [11] presented a comprehensive and structured overview of a large set of interesting outlier definitions for various forms of temporal data, application scenarios and novel techniques in which specific definitions and techniques have been widely used.

5.Financial Markets

A change in the stock market or a pattern within a specific window such as the

MATS Journal of Engineering and Applied Science, Volume 1, Issue 3, ISSN 2394 - 0549 (Disclaimer - The authors are solely responsible for the contents of the research paper compiled in this seminar proceeding. The publishers or editors do not take any responsibility for the same in any manner. Errors, If any, are purely unintentional.

flash crash of May 6, 2010 is an anomalous event which needs to be detected early in order to avoid and prevent extensive disruption of markets because of possible weaknesses in trading systems.

Various temporal economic datasets have been studied with respect to outlier detection. Gupta et al. [14] identify country outliers based on the unusual change in the constituents of the GDP (Consumption, Net Exports, Investment, Public Expenditure) across time, using temporal community outlier detection methods .They also find U.S. states as outliers from the Budget dataset with respect to anomalous change in the distribution of spending with respect to different components. Otey et al. [16] study the U.S.

Census Bureau's Income data set to detect outliers on the basis of unusual combinations of demographic attribute values. They also studied the U.S. Congressional Voting Data to find outliers. An example is a Republican congressman who voted significantly differently from his party on four bills. They use distance based outlier detection for distributed temporal data to obtain such outliers. Zhu et al. [15] perform outlier subsequence detection from the NYSE IBM stock time series data using Shifted Wavelet Trees.

In this work, M. Gupta, J. Gao [11] presented an organized overview of the various techniques proposed for outlier detection on temporal data. Modeling temporal data is a challenging task due to the dynamic nature and complex evolutionary patterns in the data. In the past, there are a wide variety of models developed to capture different facets in temporal data outlier detection. This survey organized the discussion along different data types, presented various outlier definitions, and briefly introduced the corresponding techniques.

Finally, M. Gupta, J. Gao [11] presented the work on various applications for which these techniques have been successfully used and discussed in the papers. This survey provides a number of insights and lessons as follows. The methods for different data types are not easy to generalize to one another, though some of them may have similarity in the framework at the broader level. For example, change detection in continuous time series and discrete time series both require forecasting methods. However, the specific kind of forecasting method is extremely different in the two scenarios (regression models in one vs Markov Models in the other). Most window based models are currently offline, whereas online methods do exist for point based models. Therefore, there is significant opportunity for research in the former. While the number of formulations of the temporal outlier detection problem are diverse, they are generally motivated by the most common applications which are encountered in the literature. Many recent applications, especially those corresponding to novel data types in the context of web-based applications, may result in combinations of facets, which have not been explored before. There are numerous formulations of temporal outlier detection, which have not been sufficiently explored. This is because of the many different combinations of facets, which can be used for defining temporal outlier detection problems. Complex data types such as social streams in which two different data types are present in combination (text and structure) have also not been studied.

MATS Journal of Engineering and Applied Science, Volume 1, Issue 3, ISSN 2394 - 0549 (Disclaimer - The authors are solely responsible for the contents of the research paper compiled in this seminar proceeding. The publishers or editors do not take any responsibility for the same in any manner. Errors, If any, are purely unintentional.

This comprehensive survey provides a good insight into various techniques that have been applied on multiple forms of temporal data and can be used to identify problem settings that can be worked on in the future. Finally, for further reading, we direct the reader to a recent book on outlier analysis [13] and [12] for a tutorial version of the survey.

The main approach is to find high-frequency trading from the perspective of computational science and presents a pattern recognition model for predicting the changes in price of assets of stock market . The technique is based on thefeature-weighted Euclidean distance to the centroid of a training cluster. A set of micro technical indicators is used in this setting, which is traditionally employed by professional scalpers. We also describe various procedures for normalization of feature points, computation of weights of features, classification of test points and removal of outliers. The complexity of computation received at each quote is proportional to the number of features. Also, processing of indicators is parallelizable and, thus, suitable in domains ofhigh-frequency . Experiments for different prediction time intervals and confidence thresholds are presented. Predictions which were made 10 to 2000 milliseconds before a price change gave an accuracy that ranged monotonically from 97% to 75%. Finally, it was observed an empirical relation between Euclidean distance in the feature space and prediction accuracy.

A method for stock market analysis has been developed which is based on computation of the distance to the centroid of a set of feature vectors. The centroid is used to represent an empirical combination of values of indicators that signal a price change. Even with naıve indicators,the model demonstrates good performance, . For future work, we can consider additional market indicators to be added to the model (including those computed from correlated products) to increase the interval of prediction .However , This would trigger a slight increase in the processing time. A careful analysis is required for finding the balance between the processing time, the prediction interval and the prediction accuracy . It is expected that the model presented can be used to achieve balance.

On more direction for development of the model is a study of the dynamics of distances. It was observed that price changes occur not only when the test point moves closer to the closer cluster, they sometimes occur only when the test point quickly moves farther from the farther cluster. To analyze these patterns, the sequence of distance values can be represented as a function. This function can then be approximated. The parameters of approximation are used to describe the relationship between the sequences and hence, to makewell-grounded decisions.

Another direction is investigation of grouping of points within a cluster and dividing them in sub-clusters. Then a prediction is based on proximity of a test point to one of thesub-clusters.

6.Prediction model

C. N. Babu & B. E.Reddy [8] has proposed the methodology. They focused on the concept that the accuratelong-term prediction of time series data is a very useful research challenge in diversified fields especially in the financial sector. As

MATS Journal of Engineering and Applied Science, Volume 1, Issue 3, ISSN 2394 - 0549 (Disclaimer - The authors are solely responsible for the contents of the research paper compiled in this seminar proceeding. The publishers or editors do not take any responsibility for the same in any manner. Errors, If any, are purely unintentional.

financial time series data (TSD) are volatile in nature, prediction of such TSD is a major research problem in time series data mining. The two challenges encountered are, preserving the data trend across the forecast horizon and maintaining high accuracy of prediction . Various linear traditional models like auto regressive integrated moving average (ARIMA) and generalized auto regressive conditional heteroscedastic (GARCH) preserve the data trend to some extent at the cost of prediction accuracy. Many non- linear models like ANN maintain accuracy of prediction but data trend is sacrificed. We devise a linear hybrid model which preserves both the data trend and also maintains prediction accuracy.

A pre-processing based on moving-average (MA) filter assists this hybrid ARIMA–GARCH model. This model uses a unique partitioning and interpolation (PI) technique for improving the accuracy of prediction. The proposed model is tested on selected data of NSE India stock market . The performance of this model is compared with some existing and traditional models, which shows that the proposed model outperforms the others in terms, for multi-step ahead prediction, the proposed model outperforms the others in terms.

Conclusion And Lessons Learned

The rapid development of data technology, as Internet growth, creates a large overload data for the business community. To find useful information hidden in them, data mining is came into being. Data mining is a process to extract potential information from a large amount of data. In general, it can be divided into four categories: related and dependent relationship discovery, type determination, category description, and outlier data mining.

Outliers are frequently treated as noise that needs to be removed from a data set in order for a specific model or algorithm to succeed. So, outlier is always cancelled or neglected simply. However, scholars gradually realize that certain outlier probably is the real reflection of normal data. So outlier mining becomes an important aspect of data mining. All these methodology presents an outlier data mining algorithm based on attributes. Experimental results show that this algorithm, compared with the traditional ones, has better recall ratio and precision ratio. So it is more suitable for massive data. However, this new algorithm is limited to numerical data sets. Further study shall be concerned about how to detect outliers innon-numerical data sets.

Financial time series is (at least at the tick level)non-periodic. The opening effect is a very big issue because you can't simply use the last day's data for initialization purpose even though you'd really like to (because otherwise you have nothing). The External events might cause the new day's opening to differ dramatically both in volatility from the previous day and also the absolute level on new day's data. The irregular frequency of incoming data Near open and close of the day the amount of datapoints/second can be 1-20 times higher than the average of the day and it also affects regularly sampled data.

The "outliers" in financial data exhibit some specific patterns that could be detected with specific techniques not applicable in other domains and I'm -in part-looking for those specific techniques. In more extreme cases (e.g. the flash crash) the

MATS Journal of Engineering and Applied Science, Volume 1, Issue 3, ISSN 2394 - 0549 (Disclaimer - The authors are solely responsible for the contents of the research paper compiled in this seminar proceeding. The publishers or editors do not take any responsibility for the same in any manner. Errors, If any, are purely unintentional.

outliers might amount to more than 75% of the data over longer intervals (> 10 minutes). In addition, the (high) frequency of incoming data contains some information about the outlier aspect of the situation

The future work would be to develop algorithms to automatically estimate the optimum parameters and thresholds for different datasets based on the statistical characteristics of each dataset. The run time of the algorithm should be improved.

References:

[1]V. Chandola, A. Banerjee, and V. Kumar, "Anomaly detection: A survey," *ACM Comput. Surv.*, vol. 41, no. 3, pp. 15:1–15:58, 2009.

[2]C. C. Aggarwal, *Outlier Analysis*. New York, NY, USA: Springer, 2013.

[3]V. J. Hodge and J. Austin, "A survey of outlier detection methodologies," *Artif. Intell. Rev.*, vol. 22, no. 2, pp. 85–126, 2004.

[4]C. C. Aggarwal and P. S. Yu, "Outlier detection for high dimensional data," *SIGMOD Rec.*, vol. 30, pp. 37–46, May 2001.

[5]Y. Zhang, N. Meratnia, and P. J. M. Havinga, "Outlier detection techniques for wireless sensor networks: A survey," Centre Telemat.
Inform. Technol. Univ. Twente, Enschede, The Netherlands, Tech. Rep. TR-CTIT-08-59, Oct. 2008.

[6]Analysis of stock market manipulations using knowledge discovery techniques applied to intraday trade prices D. Diaz, B. Theodoulidis, P. Sampaio, Manchester Business School, The University of Manchester,M13 9SS, United Kingdom Expert Systems with Applications 38 (2011) 12757–12771

[7]"Graph-based approach for outlier detection in sequential data and its application on stock market and weather data" by
A. Rahmani, S. Afra , O. Zarou 4 March 2014 Knowledge-Based Systems 61 (2014) 89–97

[8]Prediction of selected Indian stock using a partitioning–interpolation based ARIMA–GARCH model C. N. Babu, B. E. Reddy Dept. of Computer Science & Engg., JNT University College of Engineering, Anantapuramu, India 18 September 2014

[9] An Outlier Mining Algorithm Based on Attribute Entropy by M.J. Zhou, J. C. Tao, Procedia Environmental Sciences 11 (2011) 132 – 138

[10]Distance-Based High-Frequency Trading T. Felker1, V. Mazalov and S. M. Watt at ICCS 2014. 14th International Conference on Computational Science

[11]M. Gupta, J. Gao "Outlier Detection for Temporal Data: A Survey " IEEE TRANSACTIONS ON KNOWLEDGE AND DATA
ENGINEERING, VOL. 26, NO. 9, SEPTEMBER 2014

[12]M. Gupta, J. Gao, C. C. Aggarwal, and J. Han, "Outlier detection for temporal data," in *Proc. 13th SIAM Int. Conf. SDM*, 2013.

[13]C. C. Aggarwal, *Outlier Analysis*. New York, NY, USA: Springer, 2013.

[14]M. Gupta, J. Gao, Y. Sun, and J. Han, "Community trend outlier detection using soft temporal pattern mining," in *Proc. ECML PKDD*, Bristol, U.K., 2012, pp. 692–708.

[15]Y. Zhu and D. Shasha, "Efficient elastic burst detection in data streams," in *Proc. 9th ACM Int. Conf KDD*, New York, NY, USA, 2003, pp. 336–345.

[16]M. E. Otey, A. Ghoting, and S. Parthasarathy, "Fast distributed outlier detection in mixed-attribute data sets," *Data Mining Knowl. Discov.*, vol. 12, nos. 2–3, pp. 203–228, May 2006.

A Review on Parallel Implementation of Heuristic Algorithms

Jaynendra Patel,Mr. Avinash Dhole
Raipur Institute of technology, Chhatouna, Chhattisgarh 492101 India
[1]jaynendra.patel1989@gmail.com
[2]Avi_dhole33@rediffmail.com

Abstract—**Now a day's high performance computing mainly center's around parallel computing. Parallel computing is the ability to carry out multiple operations or tasks simultaneously. Ideally, parallel processing makes programs run faster because there are more engines (CPUs or cores) to run the program. Searching is one of the fundamental operations which is frequently required in theory and applications of computer Science. And in computer science, artificial intelligence has added a new dimension. AI uses frequently Heuristic searching currently the algorithms for the serial heuristic search has reached the time limitation. Therefore the parallel computation is an efficient way to improve the performance. By putting some constraints on the data and taking the advantage of the hardware, the performance of the heuristic search algorithms can be significantly improved. This paper provides the review of various heuristic search algorithms which has developed parallely.**

***Keywords*— heuristic search algorithms, N-puzzle problem, AI, Parallelism.**

IV. INTRODUCTION

In present days parallel and distributed computing has become a significant area in Computer Science. Parallel processing is the ability to carry out multiple operations or tasks simultaneously. Parallel processing is also called parallel computing and in computer science, artificial intelligence (AI) has added a new dimension. AI began in earnest with the emergence of modern computer during 1940s and 1950s [1]. Heuristic is a technique designed for solving a AI problem more quickly when

- Classic methods are too slow, or
- For finding an approximate solution when classic methods fail to find any exact solution [1].

There are various heuristic search algorithms like: Hill climbing, Constrains satisfaction, Best first search, Beam First Search , A* Method etc [1].

The N-Puzzle is a board game for a single player. we can use A* algorithm for searching in N-Puzzle problem.It consists of (N^2- 1) numbered squared tiles in random order, and one blank space. The object of the puzzle is to rearrange the tiles in order by making sliding moves that use the empty space, using the fewest moves. Moves of the puzzle are made by sliding an adjacent tile into the empty space. Only tiles that are horizontally or vertically adjacent to the blank space (not diagonally adjacent) may be moved.

With the invention of multi-core technology, the parallel algorithm can get the benefits to improve the performance of

the application. Multi-core technologies supports multithreading to executing multiple threads in parallel and hence the performance of the applications can be improved.[2]

V. VARIOUS SEARCHING TECHNIQUES.

L. Mr. Nagraj and Mr. Kumarasvamy,"Serial and Parallel Implementation of Shortest Path Algorithm inOptimization Of Public Transport Travel",(2011).

This paper suggests execution of 3 shortest path algorithms (Dijkstras's algorithm, Bellman Ford algorithm and Ant-Colony algorithm) serially and parallely (Using OMP). And shows the result that time cost of multithreaded parallel algorithm on dual core system are much faster than the serial algorithm. The parallel running speed can be improved with the increase of number of cores [3].

M. Mr. Panagiotis D. Michailidis and Mr. Konstantinos G. Margaritis "Implementing Parallel LU Factorization with Pipelining on a MultiCore using OpenMP",(2010).

In this paper they have presents reasonably analytical model to predict the performance of the Linear Algebra (LU) decomposition method with the pipelining technique. They have also developed parallel algorithms for LU decomposing. In this paper author analysis three OpenMP parallel algorithms for the LU factorization method. These parallel algorithms are based on the three different data distribution schemes among the available threads such as row block, row cyclic and pipeline. [4].

N. Nadira Jasika, Naida Alispahic, Arslanagic Elma, et-al. "Dijkstra's shortest path algorithm serial and parallel execution performance analysis"(2010)

Authors have studied Dijkstra's Shortest path algorithm serial environment and parallel environment. Dijkstra's algorithm can be applied to graphs with a various number of vertices and edges. Dijkstra's shortest path algorithm is

MATS Journal of Engineering and Applied Science, Volume 1, Issue 3, ISSN 2394 - 0549 (Disclaimer - The authors are solely responsible for the contents of the research paper compiled in this seminar proceeding. The publishers or editors do not take any responsibility for the same in any manner. Errors, If any, are purely unintentional.

implemented and presented, and the performances of it parallel and serial executions are compared. The algorithm implementation was parallelized using OpenMP (Open Multi-Processing). The results show that, because of Dijkstra's algorithm in itself is sequential, and difficult to parallelize, average speed-up ratio achieved by parallelization is only10% [5].

O. Kil Jae Kim, Seong Jin Cho and Jae-Wook Jeon "Parallel Quick Sort Algorithms Analysis using OpenMP 3.0 in Embedded System",(2011)

Authors have studied the parallel quick sort algorithms and also analyze the parallel quick sort algorithms using OpenMP and also analyze the effect of parallelization considering memory size and number of cores. Then shows the result that parallel quick sort algorithm performance drops roughly 10~20% even if it has exactly same algorithm [8].

P. Mr. Sanjay Kumar Sharma and Dr. Kusum Gupta "Performance Analysis of Parallel Algorithms on Multi-core System using OpenMP",(2012).

Authors have studied the typical behaviour of sequential algorithms and identified the section of operation that can be executed in parallel. and presented the execution time of both serial and parallel algorithm for computation of Pi value. They concluded that the parallelizing serial algorithm using OpenMP has increased the performance and for multi-core system OpenMP provides a lot of performance increase and parallelization can be done with careful small changes. And at last the parallel algorithm is approximately twice faster than the sequential and the speedup is linear [6].

Q. Maida Arnautovic, Maida Curić et-al, "Parallelization of the Ant Colony Optimization for the Shortest Path Problem using OpenMP and CUDA"(2013).

Author studied about Parallelization of the Ant Colony Optimization for the Shortest Path Problem using OpenMP and CUDA. In this paper two ways of parallelizing ant colony algorithm are presented. This paper aims to apply parallel implementation of ant colony optimization algorithm to the well known shortest path problem using CUDA on GPU and OpenMP on shared memory parallel computer architecture. Results show significant speedup compared to the sequential version of the ant colony algorithm. The first presented parallelization method was OpenMP parallelization technique which is a way of parallelizing sequential code using multiple threads [7].

R. Eid Albalawi, Parimala Thulasiraman and Ruppa Thulasiram."Task Level Parallelization of All Pair Shortest Path Algorithm in OpenMP 3.0".(2013)

Author studied about Task Level Parallelization of All Pair Shortest Path Algorithm in OpenMP 3.0. it has been shown that OpenMP produces poor performance for irregular applications. In 2008, the OpenMP 3.0 version introduced new features such as "tasks" to handle irregular computations .In this paper, consider one graph problem, the all pair shortest path problem and its implementation in OpenMP 3.0. Result show that for large number of vertices, They have taken linear equation problem the algorithm running on OpenMP 3.0 surpasses the one on OpenMP 2.5 by 1.6 times [9].

VI. CONCLUSIONS

Many researchers have analyzed the serial and parallel effect of various searching algorithm. In this paper, we have summarized the effect of parallelism in existing algorithms. and Mismatch problem between software and hardware will be tackled because we will make use of multicore, and by using multicore we will try to get higher speed up, throughput and utilization etc.

REFERENCES

[18] Patterson D. W., "Introduction to Artificial Intelligence And Expert Systems", pp 1, 178-184.

[19] Gallivan K. A., Plemmons R. J., and Sameh A. H.,"Parallel algorithms for dense linear algebra computations," SIAM Rev., vol. 32, pp. 54-135, March 1990.

[20] Nagaraj G. & Dr Y S Kumarswami, "Serial and parallel implememtation of shortest path algorithm in the optimization of Public transport travel" International Journal of Computer Science, Engineering and Information Technology, Vol.1,Issue 1(2011) 72-87.

[21] Mr. Panagiotis D. Michailidis and Mr. Konstantinos G. Margaritis" Implementing Parallel LU Factorization with Pipelining on a MultiCore using OpenMP" 2010 13th IEEE International Conference on Computational Science and Engineering.

[22] Jasika Nadira et-al. "Dijkstra's shortest path algorithm serial and parallel execution performance analysis" MIPRO 2012, May 21-25,2012, Opatija, Croatia..

[23] Kil Jae Kim1, Seong Jin Cho2 and Jae-Wook Jeon3" Parallel Quick Sort Algorithms Analysis using OpenMP 3.0 in Embedded System" 2011 11th International Conference on Control, Automation and Systems Oct. 26-29, 2011 in KINTEX, Gyeonggi-do,Korea.

[24] Sharma S. K. , Gupta K. "Performance Analysis of Parallel Algorithms on Multi-core System using OpenMP" International Journal of Computer Science, Engineering and Information Technology (IJCSEIT), Vol.2, No.5, October 2012

[25] Arnautovic Maida et-al." Parallelization of the Ant Colony Optimization for the Shortest Path Problem using OpenMP and CUDA" MIPRO 2013/SP.

[26] Albalawi Eid et-al." Task Level Parallelization of All Pair Shortest Path Algorithm in OpenMP 3.0" 2nd International Conference on Advances in Computer Science and Engineering (CSE 2013).

MATS Journal of Engineering and Applied Science, Volume 1, Issue 3, ISSN 2394 - 0549 (Disclaimer - The authors are solely responsible for the contents of the research paper compiled in this seminar proceeding. The publishers or editors do not take any responsibility for the same in any manner. Errors, If any, are purely unintentional.