International Journal of Advanced Research in Engineering Technology & Sciences ISSN: 2394-2819
June-2016 Volume 3, Issue-6
Email:editor@ijarets.org www.ijarets.org

A REVIEW ON BIG DATA ANALYSISON SECURITY USING RSA AND MODIFIED AES

Vandana Sehgal
M.Tech
Computer Science Department
Ganga Institute of Technology and Management
Kablana , Jhajjar,Haryana
Maharashi Dayanand Universit
Rohtak, Haryana

Dr. Yashpal Singh
Associate Professor
Computer Science Department
Ganga Institute of Technology and Management
Kablana , Jhajjar,Haryana
Maharashi Dayanand Universit
Rohtak, Haryana

ABSTRACT— The IT industry is always developing new technologies, and big data is the one of them. With the attracted more and more attention By 2012, the volume of data has increased from Terabyte level to Petabyte level. These larger volumes and new assets are known as Big Data. Big Data generate three issues like volume, variety, & velocity of data. Many challenges are faced to handle big data source. One of them issues is security of data. Security of necessary data does by using method of encryption and decryption. Encryption is a process which convert simple file in to encoded file that is difficult to access by unauthorized user. Researchers provide some algorithm for encryption like RSA, Hill cipher, AES, one bit rotation, etc. we implement RSA algorithm and get total time consume for encryption/decryption. Now some modification performs on AES and implement on same input. The time consumption by modify AES is compared with RSA time consumption. Now the size of input file increases and implements both algorithms. The time consumption will vary in both cases as input file is small and large in size. So there are two comparisons one is between RSA and Modify AES and other between both sizes of file.

KEYWORDS: Big Data, RSA Algorithm, AES, Volume, Cryptography etc

I. INTRODUCTION

A. Big Data

Big Data is the large and complex data that is difficult to use the traditional tools to store, manage, and analyze in an acceptable duration. Therefore, the Big Data needs a new processing model which has the better storage, decision-making, and analyzing abilities. This is the reason why the Big Data technology was born. The Big Data technology provides a new way to extract, interact, integrate, and analyze of Big Data. The Big Data strategy is aiming at mining the significant valuable data information behind the Big Data by specialized processing.

B. The Characteristics of Big Data

According to a research report from Gartner the growth of the data is three-dimensional, which is volume, velocity and variety. So far, there are many industries still use the 3Vs model to describe the Big Data. However, the Big Data is not only 3Vs but also has some other characteristics. The first one is the volume. As mentioned, the volume of Big Data has moved from Terabyte level to Petabyte level. The second one is the variety. Compared with the traditional easy to storage structured text data, there is now an increasing amount unstructured data that contains web logs, audio, video, pictures, and locations. Data no longer needs to be stored as traditional tables in databases or data warehouses but also stored as variable data types at the same time. To meet this requirement, it is urgent to improve the data storage abilities. Next is velocity. Velocity is the most significant feature to distinguish the Big Data and the traditional data mining. In the Age of Big Data, the volume of high concurrency access of users and submission data are huge.

C. Big Data Security

Cryptography is probably the most important aspect of communications security and is becoming increasingly important as a basic building block for computer security. Cryptography is the study of secret writing that deals with the all aspects of data security, authentication, digital signature, electronic money and other application. It concerned with the development of algorithms which can be used to provide security to the data in multimedia transmission, verify the correctness of the message at the recipient and form the basis of many technological solution to computer and communication security problem. In the other we can say cryptography is science of mathematical concept that is used to encrypt and decrypt the information, so no one can read it except the intended recipient.

1.1 Security trends

A security attack can be viewed as access to a confidential data. An attack can be occurred with the intention of reading the data or modifying the data. This report made a general consensus that the Internet needs more and better security, and it identified key areas for security mechanisms. Among these were the need to control of network traffic, the need to secure end to end user traffic using authentication, secure the network infrastructure from unauthorized monitoring and encryption mechanisms [1]. With the emergence of information technology, the attacks on the Internet and Internet-attached systems have grown more sophisticated and drastically while the amount of expertise and awareness needed to undertake an attack has declined. Critical infrastructures gradually more rely on the Internet for operations. Individual users rely on the security of the Internet, email, the Web, and Web-based applications to a greater extent than ever. Thus, a wide range of technologies and tools are needed to counter the growing threat. The OSI security architecture focuses on security attacks, mechanisms, and services. These can be defined briefly as follows:

1.1.1 Passive Attack

A passive attack on a cryptosystem is one in which the cryptanalyst cannot interact with any of the parties involved, attempting to break the system solely based upon observed data (i.e. the cipher text). In this type of attack, the main aim of the opponents is to obtain the information. It means they never harm the recourses or modify the information, they just read the information. It is very hard to detect this type of attack because passive attacks cannot be sensed by the decrypted message. There are two types of passive attacks, release of message content and traffic analysis.

1.1.2 Active Attack

Active attacks are used to modify the encrypted information or the creation of false stream which can change the meaning of the decrypted message [2]. Four types of active attack are these: Masquerade, Replay, Modification of the messages, and Denial of service. It uses two keys; public key and private key, sender encrypts the message by using the public key of receiver and receiver decrypt the message by using his private key. So there is no need of key transfer, which increases its security. In this technique both parties have their set of key, public key and private key. Public key is public to all in the network can be access by everyone while private key is private to its owner.

1.2 AES

The Advanced Encryption Standard (AES) is formal encryption method adopted by the National Institute of Standards and Technology of the US Government, and is accepted worldwide. This paper introduces AES and key management, and discusses some important topics related to a good data security strategy. In 1997 the National Institute of Standards and Technology (NIST), a branch of the US government, started a process to identify a replacement for the Data Encryption Standard (DES). It was generally recognized that DES was not secure because of advances in computer processing power. The goal of NIST was to define a replacement for DES that could be used for non-military information security applications by US government agencies. Of course, it was recognized that commercial and other non-government users would benefit from the work of NIST and that the work would be generally adopted as a commercial standard.

The NIST invited cryptography and data security specialists from around the world to participate in the discussion and selection process. Five encryption algorithms were adopted for study. Through a process of consensus the encryption algorithm proposed by the Belgium cryptographers Joan Daeman and Vincent Rijmen was selected. Prior to selection Daeman and Rijmen used the name Rijndael (derived from their names) for the algorithm. After adoption the encryption algorithm was given the name Advanced Encryption Standard (AES) which is in common use today. In 2000 the NIST formally adopted the AES encryption algorithm and published it as a federal standard under the designation FIPS-197. The full FIPS-197 standard is available on the NIST web site (see the Resources section below). As expected, many providers of encryption software and hardware have incorporated AES encryption into their products.

The AES encryption algorithm is a block cipher that uses an encryption key and a several rounds of encryption. A block cipher is an encryption algorithm that works on a single block of data at a time. In the case of standard AES encryption the block is 128 bits, or 16 bytes, in length. The term “rounds” refers to the way in which the encryption algorithm mixes the data re-encrypting it ten to fourteen times depending on the length of the key. This is described in the Wikipedia article on AES encryption. The AES algorithm itself is not a computer program or computer source code. It is a mathematical description of a process of obscuring data. A number of people have created source code implementations of AES encryption, including the original authors.

II. RELATED WORK

When applied for wireless sensor networks. While in Asymmetric Encryption, two keys are used. The SCADA communication takes place over radio, modem, or devoted serial lines. The internet SCADA facility has brought several advantages in terms of control, data generation and presentation. With these advantages, come the security issues about web SCADA. Masadeh, Turab (2010) this paper encryption algorithm are compared on the basis of wireless network. Encryption techniques play a main role in wireless network security systems. However, these schemes consume a significant amount of computing resources such as CPU time, and packet size [3]. This can be extended to many rounds. we generalize RSA encryption method in order to be implemented in the general linear group on the ring of integer mod n. The encryption method has no restriction in encryption and decryption order and is claimed to be efficient, scalable and dynamic.

RELATED WORK TABLE

1. The SCADA communication takes place over radio, modem, or devoted serial lines. The internet SCADA facility has brought several advantages in terms of control, data generation and presentation. With these advantages, come the security issues about web SCADA. Masadeh, Turab (2010) this paper encryption algorithm are compared on the basis of wireless network.	2010

2. To remedy the wireless network security issue, a novel work has been deployed to secure the transmitted data over wireless network and examine a method for analyzing trade-off between efficiency and security.	2011
3. A comparison has been conducted for those encryption techniques at different settings for each method such as different sizes of data blocks, different platforms and different encryption/decryption speed. They offered proof of concept by applying a definite privacy homomorphism for sensor network.	2013
4. In this paper we show that a particular order preserving encryption technique achieve the above mentioned energy benefits and give when used to support comparison operations over encrypted texts for wireless sensor networks. The technique is shown to have reasonable memory and computation. In this paper, comparison between Encryption techniques as used in Communication between SCADA Components is discussed.	2014
5. To secure data or information by a modified RSA cryptosystem based on 'n' prime. This is a new technique to provide maximum security for data over the network. It is involved encryption, decryption, and key generation. Prime number used in a modified RSA cryptosystem to provide security over the networks.	2015

To secure data or information by a modified RSA cryptosystem based on 'n' prime. This is a new technique to provide maximum security for data over the network. It is involved encryption, decryption,

and key generation. Prime number used in a modified RSA cryptosystem to provide security over the networks. In this technique we used 'n' prime number which is not easily breakable. 'n' prime numbers are not easily decompose. This technique provides more efficiency and reliability over the networks. In this paper we are
Used a modified RSA cryptosystem algorithm to handle 'n' prime numbers and provides security [6].

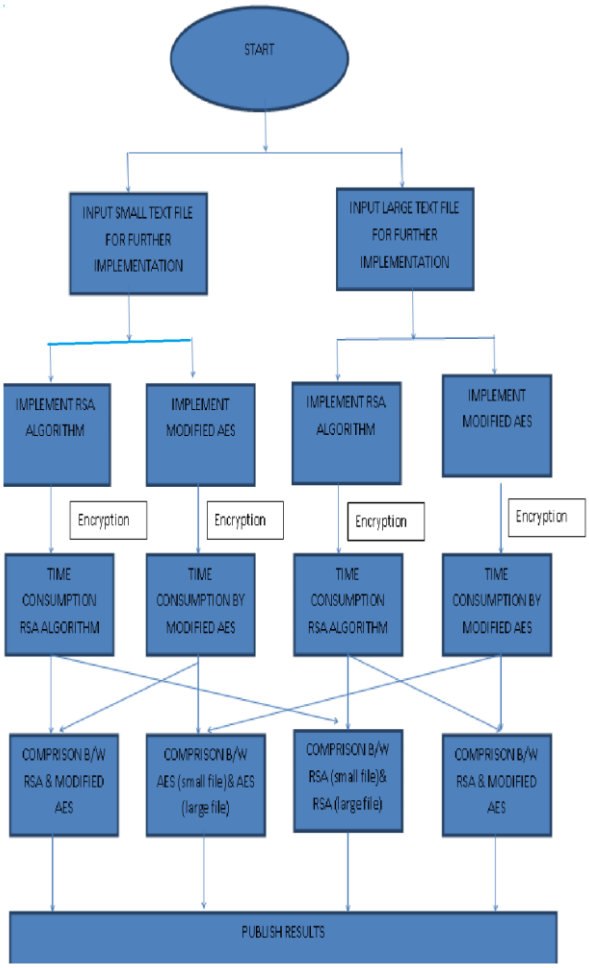
III. CURRENT RESEARCH TRANDS

Main motive of our work is providing secure communication between sender and receiver. Attack by unauthorized person is main problem in communication through networking. So this type attack is avoided by using some encryption algorithm. RSA is most popular technique used by sender for encryption of data. But as time passes new algorithms replace old algorithms due to some best features. So modification in AES will replace RSA in case of time consumption in encryption/decryption.
Decode that file. So this type of transaction motivates us to work on it.

IV. OBJECTIVE

1. It is necessary to study the concept of cryptography in Big Data. How a big data challenge security and how can we face this challenge?
2. Encryption process is applied on a file so we take two file whose volume varies but variety and velocity remain constant.
3. RSA algorithm is applied on those files one by one and time consumption in both cases is compared. This comparison finds out effect of volume variation on encryption process.
4. Now AES is modified and applied on those files one by one and time consumption in both cases is compared. This comparison finds out effect of volume variation on encryption process.
5. A table generated in which comparison for RSA and AES is shown for both files.

V. FLOWCHART OF RESEARCH METHODOLOGY


START
INPUT SMALL TEXT FILE FOR FURTHER IMPLEMENTATION
INPUT LARGE TEXT FILE FOR FURTHER IMPLEMENTATION
IMPLEMENT RSA ALGORITHM
IMPLEMENT MODIFIED AES
IMPLEMENT RSA ALGORITHM
IMPLEMENT MODIFIED AES
Encryption
Encryption
Encryption
Encryption
TIME CONSUMPTION RSA ALGORITHM
TIME CONSUMPTION BY MODIFIED AES
TIME CONSUMPTION RSA ALGORITHM
TIME CONSUMPTION BY MODIFIED AES
COMPRISON B/W RSA & MODIFIED AES
COMPRISON B/W AES (small file) & AES (large file)
COMPRISON B/W RSA (small file)& RSA (large file)
COMPRISON B/W RSA & MODIFIED AES
PUBLISH RESULTS

VI. CONCLUSION AND FUTURE SCOPE

To successful completion of the encryption and decryption of the given text files for all the algorithms. Along with this we have the completion of cryptanalysis for the two algorithms of modified AES& RSA encryption techniques. Comparison performed by these algorithms on text file. Size of text file varies to count effect of volume due to big data in comparative results. In future the proposed algorithm can be applied on audio and video files. The variety and velocity of big data is constant in this work so in future we can vary these issues also.

REFERANCES

1. Cryptography and Network Security Principles and Practices, Fourth Edition By William Stallings.
2. Masadeh, S.R. Aljawarneh,S.; Turab, N.; Abuerrub, A.M, "A comparison of data encryption algorithms with the proposed algorithm: Wireless security", Networked Computing and Advanced Information Management (NCM), 2010 Sixth International Conference.
3. M.; Caytiles, R.; Gelogo, Y.; Tai-hoon Kim,"Comparison of Encryption Schemes as Used in Communication between SCADA Components Ubiquitous Computing and Multimedia Applications (UCMA)", 2011 International Conference on Robles, R.-J. Dept. of Multimedia Eng., Hannam Univ., Daejeon, South Korea Balitanas,
4. Yu Li; Wenming Qiu; Awada, U. ; Keqiu Li,,(Dec 2012)," Big Data Processing in Cloud Computing Environments"
5. Garlasu, D.; Sandulescu, V; Halcu, I. ; Neculoiu, G. ;,(17-19 Jan. 2013),"A Big Data implementation based on Grid Computing", Grid Computing
6. Sagiroglu, S.; Sinanc, D. ,(20-24 May 2013),"Big Data: A Review"
7. Grosso, P. ; de Laat, C. ; Membrey, P.,(20-24 May 2013)," Addressing big data issues in Scientific Data Infrastructure"
8. Kogge, P.M.,(20-24 May,2013), "Big data, deep data, and the effect of system architectures on performance"
9. Szczuka, Marcin,(24-28 June,2013)," How deep data becomes big data"
10. Zhu, X.; Wu, G.; Ding, W.,(26 June,2013)," Data Mining with Big Data"
11. Zhang, Du,(16-18 July,2013)," Inconsistencies in big data"
12. Elisa Bertino. ―Big Data- Opportunities and Challenges‖.Vanya Diwan, Shubhra Malhotra,Rachna Jain. ―Cloud Security Solutions: Comparison Among Various Algorithms‖, volume 4, issue 4, April 2014.