IJCSE International Journal of Computer Sciences and Engineering Open Access
Research Paper Volume-6, Issue-1 E-ISSN: 2347-2693

An Proposal of Securing the Data at Application Level using Enhanced Schmidt Samoa in Big Data

Narayana Galla[1*], Padmavathamma Mokkala[2]

[1*] Dept. of Computer Science, Rayalaseema University, Kurnool, India
[2]Dept. of Computer Science, S.V.University, Tirupati, India

[*]*Corresponding Author: phd.narayana@gmail.com, Tel.: +91 9492410600*

Available online at: www.ijcseonline.org

Received: 15/Dec/2017, Revised: 26/Dec/2017, Accepted: 17/Jan/2018, Published: 31/Jan/2018

***Abstract*—** In the Big data world securing the sensitive data become more complex and time consuming process. In the big data sharing of sensitive, it exacerbates the threat of sensitive data falling into the un-authorized. To combat this sensitive data threat, enterprises turn to cryptosystem. In the cryptosystem encryption is the process of encoding sensitive data so that only authorized or privileged parties can decrypt and read the sensitive data applying this methodology in application level we provide complete security on the sensitive data.

***Keywords*—** Cryptography – Policy – Data Encryption - Privileged User – Enhanced Schmidt Samoa

I. INTRODUCTION

In the new modern distributed big data[1],[2],[3] environment the organizations and individuals are more connected to digitally than ever before. In the Digital world the government\Companies collecting the massive data of their resource. For the day to day active this big data will help a lot, but it may not have the fundamental assets of securing the sensitive data is missing. If a security breach occurs to big data, it would result in even more serious legal repercussions and reputational damage than at present.

In this new modern world many companies are using the technology to store the sensitive[4], [5] and non-sensitive data which may be petabytes. As a result, information classification becomes even more critical. In classification of sensitive data and encrypting the sensitive data is very essential. Not only security but also data privacy challenges existing industries and federal organizations. With the increase in the use of big data in business, many companies are wrestling with privacy issues on the sensitive data.

Data privacy [1][2] is a liability, this must be on privacy defensive on sensitive data. But unlike security, privacy on sensitive data should be considered as an asset. There should be a balance between data privacy and security on sensitive data.

II. RELATED WORK

Data sources for information fed into a Big Data implementation inevitably contain either sensitive, protected information or key intellectual property. This information is distributed throughout the Big Data implementation. That entire sensitive data should be protected. Today's big data environments often include both sensitive and no sensitive data (including anonymous data). Hackers can correlate de-anonymized[6] data sets to identify people and their preferences. Generally speaking, outsiders are prevented from accessing big data environments by traditional perimeter security at the boundaries of a private network. However, with today's sophisticated break-in strategies, perimeter security is no longer adequate. Criminals often try to lift health information, credit card numbers, and other vital information in order to sell it on the black market. No company wants its data to be compromised or its systems to be breached. However, most traditional IT security practices aren't strong enough to resist the new types of malware, phishing schemes, netbots, and SQL injection attacks unleashed by cybercriminal organizations for sensitive data.

Security Issues with Hadoop[7], [8], [9], [10] Many of today's big data projects incorporate Apache Hadoop, an open-source framework for storing and processing big data in a distributed fashion. Business analysts load data into Hadoop to detect patterns and extract insights from structured, semi-structured, and unstructured data. Unfortunately, not all organizations have strong data security in place for these activities. There may be personally identifiable information and intellectual property loaded into these data sets. Initially developed as a way to distribute big data processing jobs among many clustered servers, the Hadoop architecture wasn't built with security in mind. Namely, it lacks access controls on the data, including password controls, file and database authorization, and auditing. As such, it doesn't comply with important industry standards such as the

Insurance Portability and Accountability Act (IPAA) and the Payment Card Data Security Standard (PCDSS) [11] ,[12].

Sometimes supplementary group of users can access sensitive data. So we need to provide the privileges user can access sensitive data. Applying the Policy for classification of sensitive data after classification we use our proposed model to encrypt the sensitive data. Using this it will overcome the time and space complexity. Our proposed model ensure that authorized users can only access the sensitive data that they are entitled to access and also the protection of data in the rest and transit mode.

III. METHODOLOGY

In our proposed approach secure model will provide company can restrict the sensitive data access and data theft which leads potential threat of the company. To overcome this issue we are proposing the privilege user access control on sensitive data at application level.

RISK Level	Time Complexity		Security Level
	Data Reading	Data Writing	
Full Disk Encryption	Time Intense	Time Intense	Semi-Moderate
File Level Encryption	Time Intense	Time Intense	Semi-Moderate
Application Encryption-Privileged Users	Moderate	Moderate	Moderate

Encrypting the sensitive in application will give more secure at transit phase. Which is better approach than disk and file encryption [13]. Below table will shows the advantages of application level encryption

RISK	Full Disk Encryption	File Level Encryption	Application Encryption-Privileged Users
Data unrecoverable when drive stolen or lost from data center	Yes	Yes	Yes
Data made inaccessible to root and system admins	No	Yes	Yes
Data made inaccessible to admins	No	Yes	Yes
Create access logs for threat analytics	No	No	Yes
Unstructured data , config files, logs protected from theft	Yes	Yes	Yes

In application level encryption we are purposing Key Generation & Policy Management, Encryptioning the Sensitive Data, Decrypting the Sensitive Data for authorized users, privileged user access control management

Policy Management

In this policy management approach will apply the standard policy such as Insurance Portability and Accountability Act (IPAA) and the Payment Card Data Security Standard (PCDSS) etc., using this policy user can classify the sensitive and non-sensitive data after classification of the sensitive data. Our proposed encryption process will encrypted those data and stored into the big data environment.

Key Generation Phase

In this phase our proposed system will generate the key privileged user's will get the users key, using this key user can encrypt and decrypt the sensitive data. To generating the Key Generation we can use the public key cryptosystem like RSA,Enhanced Schmidt Samoa etc., Policy management will classify the sensitive data from the file so sensitive data can't be tampered or hacked from other users such as Admin, Cloud Provider & Outsource Administrators of Cloud.

Encryption Phase

In this phase after classified sensitive data will encrypted and stored in the Big Data so that non-privileged users cannot be read or altered the secure data, secure data can't be tampered or hacked from other users such as Admin, Cloud Provider & Outsource Administrators of Cloud.

Decryption Phase

In this phase only privileged users can decrypted the sensitive data which is encrypted earlier phase. So security will provided in the application level which will more at transit level.

*Enhanced Schmidt Samoa algorithm:*We use the Enhanced Schmidt Samoa algorithm as a basis to provide data-centric security for Sensitive shared data:

Enhanced New variant Schmidt Samoa (i,msg)

Comment: Generating Radom prime number , generating public key & private key generation , using these public and private key cipher and decipher the message m

Begin

p ← radom prime number
q ← random prime number which is not distinct of p
i ← input power of p value which is grater than 2 (i>2)

msg ← message

$$pq = p * q$$

$$public\ key\ N = p^i * q$$
$$private\ key\ d = N^{-1}\ mod\ lcm(p-1, q-1)$$

Return $public\ key\ , private\ key$

End

Encryption Phase

In this phase after classified sensitive data will encrypted and stored in the Big Data so that non-privileged users cannot be read or altered the secure data, secure data can't be tampered or hacked from other users such as Admin, Cloud Provider & Outsource Administrators of Cloud.

$$\text{Cipher message } c = (msg^N) \bmod N$$

Decryption Phase

In this phase only privileged users can decrypted the sensitive data which is encrypted earlier phase. So security will provided in the application level which will more at transit level.

- Decrypt = $c^d \bmod pq \equiv$ Plain text msg

In the work flow we are elaborating the process step by step

Work Flow

Step 1:-User's data having sensitive and non-sensitive data transferring to the App Server's using the Standard policy

Step 2:-Data is moving\transferring to the Big Data[5][6] cluster's through App Server, while transferring the data through App Server we need encrypt the sensitive data using the Key & Policy Management. Key Management will generate keys and distributing to the group or user's using the private using Enhanced Schmidt Samoa algorithm as shown above.

Step 3:-Privileged user's Key and Policy classification (IPAA\ PCDSS) sensitive data is encrypting and storing in the Big Data clusters

Step 4:-while accessing sensitive data, primarily the system will check user's Key and their policy in Key Management and Policy Management after successful authentication privileged users can decrypt the sensitive data. If non-privileged user's (Admin's, Root user's, Cloud Provider / Outsource Administrators) trying to access the sensitive data they will receive the encrypted data

IV. PERFORMANCE COMPARISON OF ENHANCED SCHMIDT-SAMOA WITH RSA

In RSA , Elliptic Curve and Enhanced Schmidt Samoa public and private keys can be chosen of approximately equal lengths bit size. Table 1. Provides corresponding RSA, Elliptic curve and Enhanced Schmidt Samoa for security levels (k) of 128 bits , 256 bits , 512, 1024 and 2048 etc.

Security Level bits	RSA	elliptic curve	ESS
128	106	108	46
256	157	163	104
512	717	730	161
1024	7496	6991	636
2048	9900	10900	3285

TABLE : COMPARISON OF DIFFERENT SECURITY LEVELS TOTAL TIME TAKING PROCESS OF CRYPTOSYSTEM

V. CONCLUSION AND FUTURE WORK

In this paper we have implemented Enhanced Schmidt Samoa algorithm for encrypt the sensitive data to the file for privileged user's after applying the policy classification. Using the above model it's hard to hack or tamper the sensitive data for non-privileged user's such user's (Admin's, Root users, Cloud Provider / Outsource Administrators) [15]. From the results we obtained it is proved that Enhanced Schmidt Samoa gives more protection only authorized user can retrieve the encrypted data and decrypt it.

REFERENCES

[1] http://www.sas.com/en_us/insights/big-data/what-is-big-data.html

[2] https://globalecco.org/big-data-insider-threats-and-international-intelligence-sharing

[3] "Sensitive Information" (definition) Aug. 23, 1996. Retrieved Feb. 9 2013.

[4] "DEPARTMENT OF INDUSTRY: PERSONAL INFORMATION PROTECTION AND ELECTRONIC DOCUMENTS ACT" Canada Gazette, Apr. 03 2002. Retrieved Feb. 9 2013.

[5] http://motherboard.vice.com/read/even-tor-cant-save-small-time-hackers

[6] https://www.qubole.com/blog/big-data/hadoop-security-issues/

[7] https://securosis.com/assets/library/reports/Securing_Hadoop_Final_V2.pdf

[8] https://securosis.com/blog/securing-hadoop-architectural-security-issues

[9] http://www.bmc.com/blogs/big-data-security-issues-challenges-for-2016/

[10] https://en.wikipedia.org/wiki/Health_Insurance_Portability_and_Accountability_Act

[11] http://searchdatamanagement.techtarget.com/definition/HIPAA

[12] http://blog.vormetric.com/2015/06/23/locking-down-data-full-disk-encryption-vs-file-level-encryption/

[13] Performance analysis of Jordan Totient RSA (JkRSA) and NTRU, International Journal of Scientific & Engineering Research, Volume 5, Issue 3, March-2014 1099 ISSN 2229-5518

[14] https://www.vormetric.com/data-security-solutions/use-cases/privileged-user

Authors Profile

Narayana Galla he is a Research Scholar in the Computer Science Department, Rayalaseema University, Kurnool, He received Master of Computer Science (M.Sc) degree in 2009 from SV Universiry, Tirupati, AP, India and also received M.Tech degree in 2014 from JNTU Kakinada. His research interests are Big Data, Cryptosystems and Computer Networks etc.

*Prof. M .Padmavathamma***,** Head of the Department, Dept.of Computer Science, S.V.University, tirupati. She is an eminent professor and awarded nearly 18 Ph.D and 11 M.Phil., under her supervision. She got Best Teachers Award in 2014 from Govt of Andhra Pradesh, India. She delivered various keynote addresses on Network Security & Cryptography at various locations like Kuwait, China, Mauritius , Malaysia, and Singapore .She is an editorial member and reviewer for various well-known journals. Her area of research includes Network Security, Privacy Preserving Data mining, Cloud Computing, Artificial Intelligence and Image Processing.

