International Journal of Innovative Technology and Exploring Engineering (IJITEE)
ISSN: 2278-3075 (Online), Volume-8 Issue-10, August 2019

Enhanced Cyber Security for Big Data Challenges

S.Padmapriya,N.Partheeban,N.Kamal,A.Suresh,S.Arun

Abstract: ***In recent years mining of data from social media is attracting more attention due to the explosion in the growth of Big Data. In security, Big Data deals with collection of huge digital information for analyzing, visualizing and to draw the insights for the prediction & prevention of cyber attacks. The Big Data mined about an enterprise from the data cloud, if properly analyzed reveals the private information which is highly risky. Maintaining the privacy of users of social media is the major challenge with respect to the security issues. As the data is generally stored in a data cloud, a boundary of trust must be established between the social media users and the data bank owners. Hence there is requirement of developing an efficient protocol for sharing of data. To secure the sensitive information of the user, data mining can be used along with an effective algorithm. This paper proposes the technique of code inline parsing to make the data more secure from the attacks & cyber hacks along with the SQL injections such that the data on the social media is secured. The proposed method secures the platform of Big Data which protects the user's sensitive information.***

Keywords: Big Data , Privacy, Information Security, Social Media

I. INTRODUCTION

In the current era, information technology has attained rapid progress in industries and enterprises which made the term Big Data very popular. The expansion in data growth is very rapid as the data is generated from a variety of sources such as social media, pictures in digital format, digital videos, business record, etc. Management of this large amount of data known as Big Data is a challenging task. This data can gain revenue to the enterprises as proper analysis of this Big Data leads to proper understanding of the customer requirements to take decision on the strategic basis. On the other hand, hacking of the big data leads to serious threat as there is possibility of insertion of malicious software in the operating systems and the apps. Hence to secure the big data from the cyber threats enhanced method is proposed and implemented in this paper.

II. BIG DATA'S GROWTH

Day by day there is rapid increase in the amount of data generation as the number of users using social media such as whatsapp, facebook, twitter etc. are increasing rapidly.

Manuscript published on 30 August 2019.
***Correspondence Author(s)
S.Padmapriya, Professor, Prathyusha Engineering College chennai
N.Partheeban, Associate Professor, Vel Tech Rangarajan Dr.Sagunthala R&D Institute of Science and Technology Avadi. Chennai
N.Kamal, Professor, GRT Institute of Engineering and Technology Tiruttani
A.Suresh, Assistant Professor, Siddharth Institute of Engineering & Technology, Puttur, A.P.
S.Arun, Professor, Prathyusha Engineering College chennai

According to the analysis of IBM, 95% of the data of the world id generated in the last few years and still generation of data is continued at a rate of 2.5 quintillions of bytes data every data [2]. The great problem of big data is with its storage. Database of large sizes are required for storing big data and literally not possible to manage these huge database with the commonly used database management system.As the datasets ranges between terabytes & exabytes, securing its privacy and protection is its greatest challenge. Storage and analysis of big data provides a sense of reliability for the enterprises. This paper proposes enhanced method for securing big data.

III. SECURITY THREATS MEASURES

Securing the privacy of user's information is the major challenge in the point of view of security for the big data, such as attackers & hackers trying to access the information stored in the database about the user. There is breaching in the security of information due to the various security issues. A set of codes called as SQL injection are passed by the attackers & hackers to break the access of database. Default codes are used by the attackers through which the security of the database is broken. Around the globe, there is requirement for data as most of the companies face shortage of such skills. Hence to fill the gap in the skills of the workers companies avail online training for all the workers in order to meet the requirement. Therefore the security is breached by the various ways. The major aspects for maintaining the confidentiality are the user identification & authentication which indicates the right of accessing the information. The threats commonly encountered by the confidentiality of information are virus attacks, unauthenticated user activity, hackers, downloading of infected files.

IV. SECURITY OF BIG DATA

Many businesses use big data as it has broad prospect across the globe, in the field of marketing and technical research without looking for the prospect of security as generally there will not be any major concern with the new techniques for security. The big data breaches is huge as shown in Fig.1 similar to that of the technology with serious damage of reputation compared to present situation. Business can gain improved insight of capacity of customer by properly storing and analyzing the big data as these peta bytes of data includes content from the social media, weblogs and stream data. Reasonably classification is facilitated by the ownership of the big data.

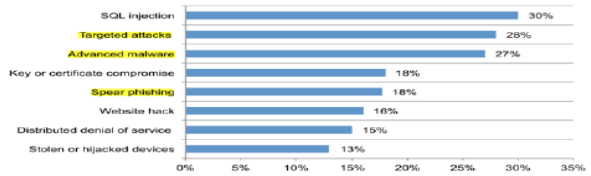

Fig 1.Breaching of Criminal or Malicious Software

Implementation of this concept has become the significant challenge as most of the organizations struggle for it. Identification of the outputs and the processes of the big data in its raw form is required for its implementation. Thus the ownership of the data will be different with that of ownership of the information possibly the IT may own the raw data and the outputs are taken responsible by the business units. Building of in-house environment of big data is likely to be done by few organizations hence big data and cloud will be linked inextricably. Huge amount of information is collected and processed by the organizations nowadays. More the storage of data it becomes more important to ensure the security of the data. Lack of security of data leads to huge losses to the organization financially and also damages the reputation of the company. Lacking of IT security in big data can even lead to worst than expected. Most of the organizations are aware that it's their responsibility to protect the data stored in cloud both in the perspective of commercial and regulatory basis. As the risk in cybercrime is increasing and the Internet's malicious activity prompts the organization to implement possible control measures for providing security which collects huge data compared to earlier cases. This situation provokes the application of big data for monitoring the security broader and deeper analysis is performed for protecting the important data of the organization.

V. PROPOSED SOLUTION

A. Protection of Privacy by Technical Solution

The security for the user data is enhanced by adding prefix & suffix to the information from the user so that the user data is encrypted making difficult for the hackers to understand. The encrypted data is saved in the database table and the original user data is saved in some other table. Security is enhanced as the data is stored in two tables with the encrypted database placed instead of original database. As the hackers access the data, they will find only the encrypted database which is of no use instead of the original database enhancing the security of the user data. The database in the present scenario is connected directly to the frontend through the database connectivity of single layer as shown in Fig. 2. Hackers or malicious software can pass through the database easily creating threat to the user data.

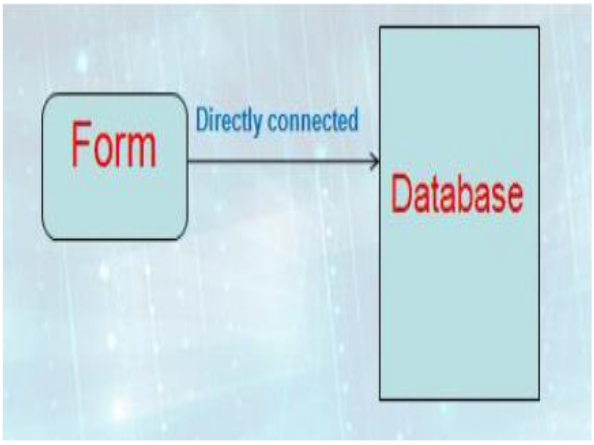

Fig.2 Connection Technique of Simple Database

The query in the proposed system has to pass through two functions as shown in Fig.3 known as the code parsing technique where the functions are related with each other and the hackers will be able to reach only the encrypted database which is stored in the place of the original database. Decryption of the database is not possible as only authenticated users know the key for the decryption process. Hence security of the user data is enhanced than the conventional techniques where the hackers access the database providing threat for the user data.

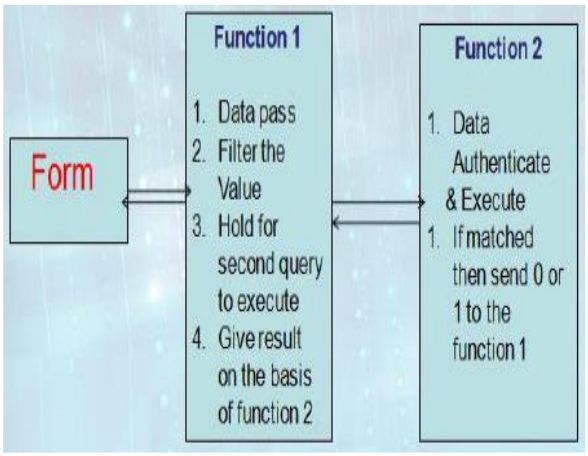

Fig 3.Proposed Connection Technique of the Enhanced Database

B. Protection of Privacy by Non Technical Solution:

Privacy protection by technical method provides security only for the privacy issues that occur technically but there is possibility for some privacy issues related to the non technical ways such as regulations & deals of the industries, laws which are also cause for the privacy threat of the user data [3]. Protecting the privacy of the user information is the vital responsibility of the legislation. Most of the countries have framed laws for protecting the privacy of user data such that personal data of the user can be preserved and also data related to the government for the security of the nation. Awareness is also created among the public about how to secure their data from hackers & malicious software.

Retrieval Number: J97290881019/19©BEIESP
DOI: 10.35940/ijitee.J9729.0881019
Journal Website: www.ijitee.org

Published By:
Blue Eyes Intelligence Engineering and Sciences Publication (BEIESP)

VI. CHALLENGES FOR THE SECURITY OF BIG DATA

In this section discussion will be done on the benefits of the enterprises and about the major challenges for the security of Big Data as shown in Fig 4.along with its privacy. For providing security and privacy for the user data some of the techniques with their solutions are discussed.

C. Code Explanation:

The proposed method is implemented by the following which is explained as follows. A single case is encoded by the transition function that is infected. The state variable 'infected' takes a new value when an event that is infected occurs. The value can be either TRUE or FALSE. This transition is an exogenous as the variable is altered by the event that occurs externally. The exogenous change is associated with the variable exempt whereas the variable policy is termed as endogenous.

Code:

```
## SET UP TRANSITION FUNCTIONS

@transition

def exempt(self):

    self.case(occurred(self.event),self.event)

@transition

def infected(self):

    self.case(occurred(self.event),self.event)

@transition

def policy(self):

    # If exempt, redirect pkt to gardenwall;rewrite dstip to 10.0.0.3

    self.case(test_and_true(V('exempt'),V('infected')),
ectToGardenWall()))

    # If infected, drop pkt

    self.case(is_true(V('infected')),C(drop))

    # Else, identity -> forward pkt

    self.default(C(identity))
```

Generally the hackers who make use of the techniques to hack user data employ low visible, slow pace techniques [17] in order to avoid the user to detect them.

In the platform of , for the security of big data, this approach is utilized [13].

- Mining of information by the IT specialist unethically gather the user's personal data without the permission of the user or without any notification.
- Solutions for security of the data is becoming difficult as the databases that are non-relational are evolved actively.

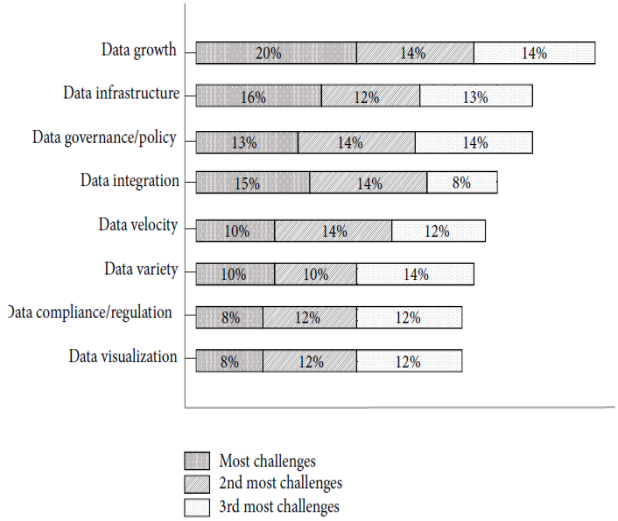

Fig 4.Big Data Challenges

- The protection is of single level for the computation of the most of the distributed systems which is generally not recommended.
- Detailed audits are recommended for the big data as the amount of information is huge but not performed routinely.
- Monitoring and tracking of the origin of the big data is not done consistently as the amount of data is huge.
- To maintain the confidentiality level most of the organizations do not divide the access control.
- The system receiving the huge amount of data must be reliable in maintaining the data but generally it does not occur.
- Security of the connections and encryption of the access control has become inaccessible & dates for the relying IT specialist.
- Security measures are required additionally for the data transfer automatically which are not available generally.

VII. BIG DATA APPLICATION AREAS

A. Big Data in Healthcare:

The pharmaceutical companies and the companies that produce medical devices make use of big data for practicing research and development. Elimination of challenges in the sector of health care is possible with the big data even before creating problem [12]. The doctors are able to analyze the history of the patient which helps them to provide proper service based on the condition of the patient. The health centers traditionally maintain the data in a structured manner such the results of laboratory test, clinical data, and diagnostic results. Hence great impact will be there in the field of clinical research due to the advancements of big data in the future.

B. Governance:

An important role is to be played by the big data in all the process and undertakings of the government. Challenges such as security, privacy, ownership of data, stewardship of data arises in the governance field.

- The loopholes and flaws can be identified by the analysis of the data

- Decisions about various issues can be taken based on the big data.
- In most of the sectors, prevention of illegal practices is possible.
- Sectors such as healthcare, research, defense, education can be improved.
- Investment of budget in areas of benefits.

C. Industries of Consumer Goods: Different sources generate data of huge volume in the companies of consumer goods such as bill details, transaction details, feedback forms from the customer etc [14]. Organization and analysis of this data is to be done for the profit of the company in a systematic way, hence big data makes the company to get better insight for taking decisions timely.

VIII. ENHANCED SECURITY SOLUTION FOR BIG DATA

New opportunities can deliver innovations & optimizations that can breakthrough and create solutions for the data variety, executives, manager, marketing companies, scientist who can plan and take decisions.

- Source of Data: Exploitation of the advantages of the big data is possible with the leverage of different forms of data both unstructured and structured data that is in the form of different file types.
- : The final fruit of big data is the output, which helps the innovation and optimization of the business. The information in the database is presented in the form of reports or dashboards which can be accessed based on the demand of the queries. The asset that is most sensitive is represented by the big data in some of the businesses with intelligence providing the differentiation in the critical competition but creates heavy competition if struck into hand of the hackers.
- Frameworks of Big data: More amounts of data has to be managed by the environment of the big data whatever the system it is using such as mongo DB, Hadoop, Teradata, NoSQL. Assets that are sensitive does not rely on the nodes of the big data but also as system logs, file configurations, logs of error etc.

In the environment of big data, replication of data is done routinely and migration occurs among number of nodes, along with which information that is sensitive is stored in the system logs, disk caches, file configurations etc. Encryption that is transparent can protect data while providing encryption that is provided privilege for the control of user access provides security for the data of intelligence. The reasonable way for improving the big data believed by the experts of cloud computing is by the expanding continually the industry of antivirus [16]. Better security against the big data threat is provided by the vendors of antivirus which provides solutions. For the improvement of the security of big data following are recommended.

- Security of application must be focused than the security of the device
- Servers and devices which contains critical data has to be isolated
- Protection both proactive and reactive has to be provided
- Management of the event and real time information security has to be introduced.

IX. RESULTS & DISCUSSION

The proposed method is implemented to solve the security and privacy problems of big data on the user data posted on the social network. Code is designed such that infected threat is rooted out from the real data. The system proposed is interoperable as the rights were to the improvement of the user privacy. User can protect their privacy by themselves by proper use of the method. Users are able to control unwanted comment on their post either by other users or other social networks. IT industries are influenced big data to a greater extent in the today's scenario. Data generated by the machines enabled through sensors, cloud computing, satellite data, data from social media requires to take important decisions in order to take the business to higher levels. Organizations can be changed by the potential of big data. Big data plays an important role in the processing of data as the amount of data is huge [15]. The code parsing technique creates two databases separately for the original data and the encrypted data. This work can be extended by maintaining different databases for the recording of original data and encrypted data.

X. CONCLUSION

Different attacking techniques are used by the hackers for inserting the malicious software in the operating systems and the application software where a large amount of data is stored. This paper enhances the security of the big data by securing the user data using the code parsing technique. The proposed method finds way for the enterprises to gain profit by properly securing the company information. As the encrypted database replaces the original database hackers literally find it difficult to insert malicious software or hack the information.

REFERENCES

1. G Geethakumari Big Data Analysis for Implementation of Enterprise Data Security , IRACST - International Journal of Computer Science and Information Technology & Security (IJCSITS), Vol. 2, No.4, August 2012
2. Cloud Security Alliance Big Data for Security Intelligence September 2013
3. LEI XU, CHUNXIAO JIANG, (Member, IEEE), JIAN WANG, (Member, IEEE) Information Security in Big Data: Privacy and Data Mining, Received September 21, 2014, accepted October 4, 2014, date of publication October 9, 2014, date of current version October 20, 2014.
4. Mr. Mohammad Raziuddin & Prof. T.Venkata Ramana Literature Survey in Data Mining with Big Data International Journal of Advanced Engineering and Global Technology I Vol-03, Issue-04, April 2015.
5. Review Article Big Data: Survey, Technologies, Opportunities, and Challenges , Volume 2014 http://dx.doi.org/10.1155/2014/712826
6. Roger Schell Security A Big Question for Big Data 2013 IEEE International Conference on Big Data.
7. IBM Big Data at the speed of Business, http://www-01.ibm.com/software/data/bigdata/2012.
8. http://bigdataarchitecture.com/
9. http://en.wikipedia.org/wiki/Apache_ Hadoop

10. Big Data is the Future of Healthcare Bill Hamilton cognizant 20-20 insights 2012
11. Sam curry —big data fuels intelligence –driven security RSA Security brief 2013
12. Big Data for Health Javier Andreu-Perez, Carmen C. Y. Poon, Robert D. Merrifield, Stephen T. C. Wong, and Guang-Zhong Yang, Fellow, IEEE
13. http://www.datacenterknowledge.com/archives/2016/01/19/nine-main-challenges-big-data-security/
14. Big Data, Black Book: Covers Hadoop 2, MapReduce, Hive, YARN, Pig, R and Data Visualization
15. Big Data A New World of Opportunities, NESSI White Paper, December 2012.
16. Samiddha Mukherjee, Ravi Shaw Big Data Concepts, Applications, Challenges and Future Scope Vol. 5, Issue 2 , February 2016
17. CLOUD SECURITY ALLIANCE Expanded Top Ten Big Data Security and Privacy Challenges, April 2013