

---

**A REPORT  
ON  
BIG DATA SECURITY**

---

**MISS. ROKEYA AKTER**  
Student ID : 160210

**FERDOUSI HAQUE**  
Student ID : 160228

16 April, 2018

## Abstract

Without physically connection, nowadays we get connected with each other through 'cyber socity'.So, data is generated by many kind of devices and humans simultaneously. Now the coming problem is the security of this data. Big data is too big ,too fast and too diverse that does not compile with traditional database system.

Big data is defined as massive data sets having large, more varied and complex structure with the difficulties of storing ,analyzing and visualizing for further processes.hence big data security means the security of big data.

In this field, lots of work have been done.About this topic we try to read few papers.

In this paper,we firstly discuss about the huge benefits and challenges of security in big data.Then we present some possible methods and techniques to ensure big data security.

KEYWORDS : big data , security.

## I.Introduction

A widely recognized definition belongs to IDC: big data technologies describe a new generation of technologies and architectures, designed to economically extract value from very large volumes of a wide variety of data, by enabling the high-velocity capture, discovery or analysis [1] We live in the Age of Big Data[2]. For example , From 2005 to 2020, the amount of data is predicted to increase 300 times, from 130 exabytes to 40,000 exabytes[3].

In a 2001 research report, challenges and opportunities brought by increased data are defined as a 3Vs model, i.e., Variety, Velocity and Volume. According to ISO 2015, Big data can be defined as the standard 5Vs such as volume, velocity, variety, veracity, and value. IoT devices are directly connected to the internet and continuously generating data. It is not necessary to produce the same formatted data it can have various format as well as speed. So, collecting and storing large amount of data is one side of coin but the other side of coin is to protect massive amount of data from unauthorized access which is more challenging.

Big Data also created six security issues such as

- 1.confidentiality,
- 2.integrity,
- 3.availability,
- 4.monitoring and auditing,
- 5.key management,
- and 6.data privacy.

Fig1.Big Data Technology

Here, Section II gives characteristics of Big Data. Overviews of Big Data challenges are discussed in section III. Section IV reviews security issues of Big Data that need to solve or readdress and big data security techniques in section V.

## II. CHARACTERISTICS OF BIG DATA

We can categorize characteristics of big data in two types such as

A. HACE Theorem : HACE means Heterogeneous, Autonomous, Complex and Evolving

Fig2.HACE Theorem

Exploring Big Data is identical to combining information from various heterogeneous sources that indicate variety of data.

B. The 5Vs of Big Data : Big Data can also be characterized by 5Vs - volume, velocity, variety, veracity, and value.

Fig3.5V's Big Data

Volume relates with enormous growing amount of data from different sources. Speed of data indicates velocity of Big Data that defines types of processing needed for Big Data such as batch processing, real-time processing and interactive processing. Heterogeneity of data sources avail variety of data in different formats such as structured data, semi-structured data and unstructured data

### III. CHALLENGES OF BIG DATA

From all the papers, we find the challenges of big data, are given in the below

- A. Heterogeneity
- B. Data Life Cycle Management
- C. Data Processing
- D. Scalability
- E. Security and Privacy
- F. Data preparation
- G. Efficient distributed storage and search
- H. Effective online data analysis
- I. Effective machine learning techniques for Big Data mining
- J. Semantic lifting techniques
- K. Programming models
- L. Data Visualization

### IV. SECURITY OF BIG DATA

1. IN literature : Generally, data is favored as secure, when CIA (Confidentiality, Integrity, Availability) is satisfied [4]. PAIN (Privacy, Authentication, Integrity, Non-repudiation) is another measure to ensure data security and privacy benchmark [5] Hence, favoring the literature study, main security and privacy issues of Big Data are confidentiality, integrity, availability, monitoring and auditing, key management and data privacy that are expatiated in consequence of literature.

- A. Confidentiality
- B. Integrity
- C. Availability
- D. Monitoring and auditing
- E. Key management
- F. Data privacy

2. IN analysis for security : Analyzing log files, network route and the system events for forensics and encroachment detection have traditionally been a remarkable problem. Advanced Big Data technology such as database related to Hadoop system and stream processing are validating the storage and analysis of very large mixed data at an unprecedented scale and speed. These types of technologies will transform by the security analytics:

- 1. Gathering data at very large scale from many external and internal enterprise sources such as susceptibility databases.
- 2. Performing broad analytics on the data.
- 3. Giving a build up view of security related information and accomplishing real time analysis of streaming data.

## V.BIG DATA SECURITY TECHNIQUES

To ensure security there are some techniques, We can categorise the techniques of big data security according to [6] :

1. oral and written pledges - The most common solution to ensure security and privacy may be oral and written pledges. However, history has shown that this method is flawed..

2.cryptography - More advanced technological solution is cryptography. The famous encryption schemes have AES and RSA.Recent revelations show that the National Security Administration (NSA) may have already found ways to break or circumvent existing Internet encryption schemes .

3. Tracking, monitoring or auditing software - Tracking, monitoring or auditing software is developed to provide a history of data flow and network access by an individual user in order to ensure compliance with securityrelated. The limitation of this technology is that it is difficult and costly to implement on a large scale or with distributed data systems and users because it requires dedicated staff to read and interpret the findings, and the software can be exploited to monitor individual behavior rather than protecting data.

4. integrated Rule-Oriented Data (iRODS) - A novel technological named the integrated Rule-Oriented Data (iRODS) is proposed to be the solution to ensure security and privacy in big data.

## Conclusion

In this paper, an overview of big data's content, characteristics, methods, advantages and challenges and discusses security concern have been reviewed. Processing and computing infrastructure extra secure. Security is now a big data problem because the data that has security context is huge. Some possible methods and techniques to ensure security and privacy in Big Data have been discussed above. Some problems are already solved and some are about to be solved and some solved issues wants further attention to modify the algorithms to maximize the accuracy and speed. Hence, further research is required to develop a system. we hope that our exhaustive survey will help to develop better security and solutions.

## References

1. Gantz J, Reinsel D. Extracting value from chaos [J]. IDC iView, 2011: 1-12
2. Lohr S. The age of big data [J]. New York Times, 2012, 11
3. Gantz J, Reinsel D. The digital universe in 2020: Big data, bigger digital shadows, and biggest growth in the Far East [J]. IDC iView: IDC Analyze the Future, 2012.
4. L. Xu and W. Shi, "Security Theories and Practices for Big Data", Big Data Concepts, Theories, and Applications, 2016, pp. 157-192.
5. S. Sudarsan, R. Jetley and S. Ramaswamy, "Security and Privacy of Big Data", Studies in Big Data, 2015, pp. 121-136.
6. Matturdi Bardi, ZHOU Xianwei<sup>2</sup>, LI Shuai<sup>2</sup>, LIN Fuhong, Big Data Security and Privacy: A Review, pp.1-7
7. Neetu Chandhari, Dr. Satyajee Srivastava, Big Data Security Issues and challenges, 2016, pp.1-5
8. Seref SAGIROGLU and Duygu SINANC, Big Data: A Review, pp.1-4
9. Nikunj Joshi, Bintu Kadhiwala, Big Data Security and Privacy Issues A Survey, 2017, pp.1-5