ISSN:0975-9646

Challenges to Big Data Security and Privacy

Manju Kaushik[#1], Aditi Jain[#2]

[1]*Associate Professor, Computer Science and Engineering, JECRC University, Jaipur*
[2]*Research Scholar, Software engineering, JECRC University*

***Abstract*-- Now a days, it is really noteworthy that Government and Private organization are pursuing greater insights for potentials to enhance human behavior at its maximum frequency. Scientist in real world are now thinking that potential source of "Big Data" is transforming world and giving new horizons to the Private and Government sectors and even in new generations of 2020 we will approximately leading with 35 zeta bytes of data storing capacities in data warehouse. Organizations are now finding the best way to collect data from various sources and analyzing that data to take out maximum of it so to have deeper insight into business trends and patterns. Big data is too big to deal with and every association wants to use maximum of t his data. They must think about Big Data security issues and should implement security steps on right time, so that big data can come out with its big problems. This paper will focus on these security and privacy issues and major challenges to secure this big use of big data of this data.**

Keywords: **Big Data, Security, Privacy**

I. INTRODUCTION

In recent years, dramatic change has noticed in our ability to collect data from various sources, sensors, devices, in different format, form independent or connected applications. Scientist in real world are now thin king that potential source of "Big Data" is transforming world and giving new horizons to the Private and Government sectors and even in new generations of 2020, we will approximately leading with 35 zeta bytes of data storing capacities in data warehouse. Big data collector will collect data for m various domains such as health care domain, social media & networking, phone &TV, web data, financial data, genome data, retails, public service, environmental monitoring, scientific activity etc. Big Data is now used to convey large sorts of concept, as huge data quantity, real time data, social media data, and much more. In industries, it is required to learn mo re about how to handle such huge amount of data and extract values for m it. These industrial datasets mostly contains social media content which has customer's information at maximum. Every organization and customer wants their information to be hidden from outside world. Thus makes data security and privacy a big issue for big data. These factors are now becoming so important that it must be necessarily implemented at right place so to keep the organizations protecting this netting of data. Big data cannot be managed by our current methodologies or data mining tool.

Due to low bandwidth, big data doesn't fits in current architecture. Thus it is necessary to introduce new architecture and tools for big data security. This paper further describes big data security issues and privacy factors.

II. DEFINING BIG DATA

Big data refers to large data sets that are challenging to store, search, share, visualize and analyze. It is high volume, high velocity, high variety information assets that demand cost effective, innovative forms of information processing for enhance insight and decision making. It is a new generation of technology and architectures designed to extract value economically from very large volume of a wide variety of data by enabling high velocity capture, discovery and analysis. The data is too big, moves too fast and does not fit the structures of obtainable data base architecture. Big data moves across 5dimensions: volume, velocity, variety, value, and veracity. Key intent of Big Data is to find hidden value through intelligent filtering of low density and high volume of data. The use of Big Data - large pools of data that can be brought together and analyzed to discern patterns and make better decisions will become the basis of competition and growth for individual firms, enhancing productivity and creating significant value for the world economy by reducing waste and increasing the quality of products and services [1]. It is the need of every growing organization to manage continuously increasing data and retrieve useful information for m that huge database. Data analysis and finding out patterns from that data helps organizations in good decision making .

Table 1. Characteristics of big data V5

Characteristics	Description
Volume	Scale of data and processing need. It calls for scalable storage and distributed approach to querying
Velocity	How fast data is being produced and changed. The speed with which data is received, understood and processed
Veracity	The quality and provenance of the information in the face of data un-certainty from many places
Variety	Data in different format and from various sources. Difficult to integrate
Value	Identifying which data is valuable then transformed and analyzed.

III. SECURITY AND PRIVACY IN BIG DATA

Big Data is an extremely well-liked talking point, but here we are only discussing about security. Big data is only around 5v's discussed in section 2. Each of these is growing at an outstanding rate and has required a shift in how security vendors manage threats. From a security perspective, there are two distinct issues: securing the organization and its customers' information in a Big Data context and using Big Data techniques to analyze, and even predict, security incidents. [2]. Organizations must think about Big Data security issues and should implement security steps on right time, so that big data can come out with its big problems. To handle challenges in security, several aspects must be examined.

- Why big data architecture is different?
- Why it is different operationally
- Open issues

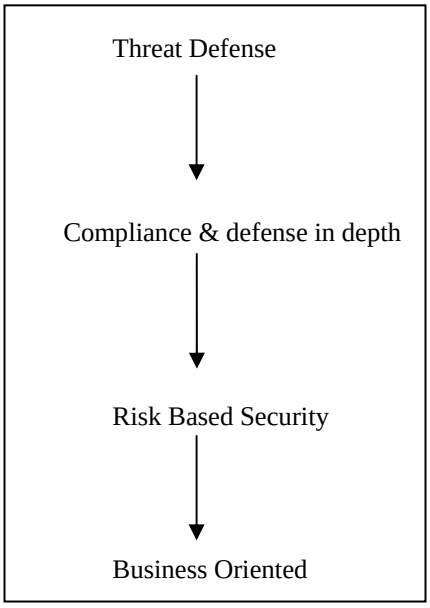

Fig.1 Stages of big data security awareness

IV. CHALLENGES TO BIG DATA SECURITY

With the increasing use of internet, we can see that the world's population is steadily coming online. This leads to increase in cybercrime also. Thus becomes essential for every organization, social media sites and other vendors to provide protection to users from these cybercrime. Security and privacy is itself a big data issue. Security at each level of architecture is required to preserve data from attacks at each level. Big data has a big issue of storage, computation, data mining, analysis, predictions, transactions and many more and at each point security must be implemented. Big data security challenges are classified into 4 categories [4]

1. Infrastructure security
   - Secure computations in distributed programming frameworks.
   - Security best practices for non-relational data stores.
2. Data Privacy
   - Scalable and compo sable privacy-preserving data mining and analytics.
   - Cryptographically enforced data centric security.
   - Granular access control.
3. Data Management
   - Secure data storage and transactions logs.
   - Granular audits.
   - Data Provenance
4. Integrity and reactive security
   - End-point input validation/filtering.
   - Real-Time security monitoring.

V. CONCLUSION

In our fast growing and connected world, big data is the topic among researchers. It is very critical to think about aspects of security and its importance along with processing of massive amounts of data. Security starts with the collection of data and there is no end of security levels. To apply security at each level, it is necessary to understand the data and security policies of the organization. This paper provided a brief idea of big data and its privacy and security issues.

REFERENCES

1. The New Competitive Advantage", IVEY Business Journal, 76(4), 1, (2014)
2. Computerweekely.com[online], Available:http://www.computerweekely.com/How-to-tackle-big-data-from-a-security-point-of-view.
3. "Securing Big Data: Security recommendations for Hadoop and NoSQL Environments" Vormetric,Inc.(2013)
4. "Top Ten Big Data Securiy and Privacy Challenges", Cloud Security alliance,(2012)