ISSN (e): 2250 – 3005 || Volume, 08 || Issue, 2|| February – 2018 ||
International Journal of Computational Engineering Research (IJCER)

Security Challenges in Big Data

Tolla L N .Varaprasad[1], A. KiranKumar[2]
[1] Assistant Professor, Dept of CSE , Medak College of Engineering and Technology, Siddipet, Telangana, India.
[2] Assistant Professor, Dept of CSE, CMR Technical Campus, Hyderabad, Telangana, India.
Corresponding author: Tolla L N .Varaprasad1

ABSTRACT
Big Data is a huge volume of data from variety of sources the data can be fetched like face book, internet websites, Google search engines etc. The term Big Data is a large-scale data management and analysis technologies that cross the boundaries of formal data processing technologies. Big Data is expressed from traditional issues in four ways: the quantity of data, the quality of data generation and transmission, and the types of structured and unstructured data and the rate of data production. This paper focused in how the incorporation of Big Data is changing security analytics by providing new technologies and challenges for large amount of structured and unstructured data. Big Data is to take a holistic vision at security. Big Data consideration is the identification of the different data sources, the origin and creators of data, as well as who is authorized to access the data. It is important to conduct a correct partition to identify complex data, and align with the origin information security policy in terms of enforcing access control and data handling policies.
Keywords: Big Data, Security, Privacy, Velocity, Volume, Variety, Social Applications.

I. INTRODUCTION

he Big Data is a present trend applied to manage data records whose data size is beyond the ability of commonly used software tools to capture and manage that amount of data. The amount of data to be analyzed is expected to double everyone years. All these data are very often unstructured and from various origins such as social media, sensors, scientific applications, video and image archives, search engines, indexing, health care data records, business policies and system backups. Big data is more and more attention since the number of devices connected to "Internet of Things". In this way, security and privacy issues can be potentially strengthen by the volume, variety, and wide area allotment of the system infrastructure to support Big Data applications. Data privacy is important the concept about which normal people are most relevant, but it should also be one of the highest associations for the companies that use Big Data Tools. A Big Data mechanism usually contains a huge amount of individual information that companies use in order to get a benefit from that data sets. However, we should question ourselves where the limit concerning the use of that information is. Companies should not have complete freedom to use that information without our knowing facts, although they also need to obtain some ease from the use of that data. There are various techniques and mechanisms with which to protect the privacy of the data, and also allow organizations to make a profit from it have therefore been developed, and attempt to solve this problem in possible different ways. The Big Data can be divided into two main classifications: Systems which provide some features and operational capabilities for real word in real time applications, transactional/interactive situations where data is gathered and stored. The other type is systems that facilitate analysis capabilities for complex analysis of the data that has been stored and maintained. Big data analytics address to the process of gathering, cumulating and analyzing large sets of data to find patterns and other useful information. With the help of Big Data analytics, organizations use the huge amounts of data made available to them to discover patterns and fetch useful information. Big Data analysis not only helps us to guide the information contained in the data but also discover the information that is most certainty to the organization and future business decisions. Today's data comes from various origins. And it is still an undertaking to process and transform data among systems. However, it is useful to connect and correlate relationships.

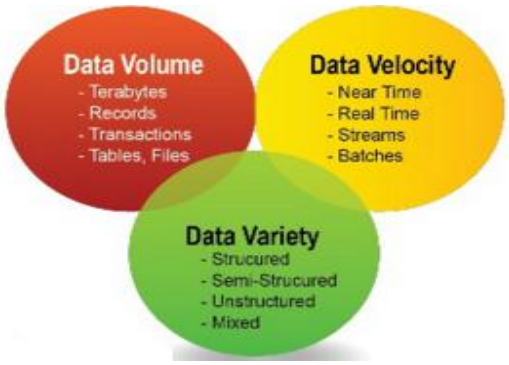
II. CHARACTERISTICS IN BIG DATA

Instead, most experienced professional define big data in terms of the three Vs. You have big data if your data stores have the following characteristics:

- **Volume:** Big data is any collection of data that is very large in quantity with different size that the companies that opts it taking challenges related to storing or processing it. In practical approach, trends like

online transactions, mobility, social media and the Internet of Things are producing so much information, that nearly every organization probably meets this rule.

- **Velocity:** If your company is generating new data sets at a high rate and needs to respond in reality, you have the velocity associated with big data. Most of the companies that are involved in online transactions, social media and Internet of things to meet this rule for big data.
- **Variety:** If your data exists in many different formats, it has the variety that leads different kind of data associated with big data. This data may be structured and unstructured.



The 3vs of Big Data

III. SECURING BIG DATA

Security and privacy issues are enhanced by the volume, variety, and velocity of Big Data. The range of data sources, formats, and data streams, integrated with the streaming nature of data gain and high volume create unique security risks. Big Data Security Challenges there is various challenges to protect big data that can compromise its security. One thing we mind that these challenges are by no means limited to certain platforms. They also pertain to the cloud, databases and warehousing. When you launch your big data platform in the cloud, take nothing for granted. Work closely with experts to overcome these sorts of challenges with strong security service level agreements. Advanced analytic tools for structured and unstructured big data and non relational databases are newer technologies in smart development. It can be complicated for privacy and security software and processes to preserve these new technologies. Sophisticated security tools technically protect data storage. However, they may not have the same behavior on data output from different analytics tools to different locations. Big data specialists may conclude to mine data without privileges or notification. Whether the conversation is curiosity or criminal profit, your security tools need to examine and manipulate on suspicious cases no matter where it comes from which location.

IV. PRIVACY AND SECURITY WITH BIG DATA

- **Data security in repository**

Data and data record logs used to be kept in architectural storage media. As data size increased measures and accessibility became a major point hence automated progress for big data storage came to the fore. It doesn't keep track of where the data are stored unlike in previous architectural storage media where Information Technology managers aware of it which data consisted where and when. This gave thought to many upcoming challenges for data security storage. Storage service providers frequently explore for clues that help them in association of user activities and data records and get to know certain characteristics. As the data holder stores the cipher text in an automated storage system and shares the private key to each user, he gives the exact privileges to access data of certain level to certain users, he being unauthorized to take the data. The service provider can bring about roll back attack on users in case of a different user environment. He may pass on outdated versions of data while the updated ones are already uploaded in the repository. Data tampering and data loss resulted by dangerous users often results in controversy between the data storage provider or amongst users

- **Cryptographically imposed data**

We have two fundamental aspects of handling visibility of data to personal, organizations and systems. The system-based approach provides a larger attacking scope. There are many attacks like morphing data and access attack that diverts access control implementations and access the data. Securing data end-to-end by encryption

leads a much smaller well-organized attacking scope. For a cryptographic protocol for searching and filtering encrypted data the adversary should not be able to learn anything about the encrypted data beyond the corresponding predicate, whether satisfied or not. The cryptographic procedure must also ensure that helping must not be able to manipulating data that came from the source for this may well be false hence affecting combining of data.

V. CONCLUSION

This paper concludes with an explanation of the research carried out in order to problem finding, challenges and issues related to security in Big Data, and how fact finders are dealing with these set of problems. This objective was conceived by following the systematic study of methodology and methods, which allowed us to find the papers related to our main goal. Many authors, therefore, focus their research on protecting data, especially with respect to privacy, but privacy it is not the only security problem that can be found in a Big Data system, the traditional architecture itself and how to protect a Big Data tools also a huge involvement researchers, identified a lack of investigations in the field of data management, and data storage in the data repository with respect to organizations. Finally the Big Data technology seems to be reaching a sophisticated stage, and that is the cause why there have been a number of studies progressed in the ago.

ACKNOWLEDGEMENT

The authors express their sincere gratitude to all those participants who participated in the analysis of their body composition, otherwise; without their cooperation this study would not have been possible. In addition to this; authors would like to mention the helpful guidance of M. Chalapathi Rao R & D Co-Ordinator. CMRTC, Hyderabad.

REFERENCES

[1]. [1] Nada Elgendy, Ahmed Elragal ,"Big Data Analytics: A Literature Review Paper", ICDM, LNAI 8557, pp. 214–227, 2014
[2]. [2] Bhawna Gupta , Dr. KiranJyoti ,"Big Data Analytics with Hadoop to analyze Targeted Attacks on Enterprise Data", (IJCSIT) International Journal of Computer Science and Information Technologies, Vol. 5 (3) , 2014, 3867-38702014
[3]. [3] Weiyi Shang , Zhen Ming Jiang , Hadi Hemmati , Bram Adams , Ahmed E. Hassan , Patrick Martin "Assisting Developers of Big Data Analytics Applications When Deploying on Hadoop Clouds",IEEE 978-1- 4673-3076-3/13 IEEE,2013
[4]. [4] FENG Deng-Guo, ZHANG Min, LI Hao. Big Data Security and Privacy Protection. Chinese Journal of Computers, 2014,37(1):246-258.
[5]. [5] MA Li-chuan, PEI Qing-qi, LENG Hao, LI Hong-ning. Survey of Security Issues in Big Data. Radio Communications Technology.
[6]. [6] Hu Kun, Liu Di, Liu Minghui. Research on Security Connotation and Response Strategies for Big Data. Telecommunications Science, 2014(2):112-117,122.
[7]. [7] WANG Yu-long, ZENG Meng-qi. Big Data Security based on Hadoop Architecture. Information Security and Communications Privacy, 2014(7):83-86.

BIOGRAPHIES

Tolla L N.Varaprasad presently working as a Assistant Professor in Medak College of Engineering and Technology, Siddipet Dist,Telangana,India.He received his M.Tech(CSE) degree Sree Dattha Institute Of Engineering & Science,Hyd in the year 2009. He received B.Tech. (CSE) degree from Anurag Engineering College in the year 2006. His fields of interests are BigData, DataMining,Cloud Computing, and Computer Networks etc.

Adepu Kiran Kumar presently working as a Assistant Professor in CMR Technical campus, Hyderabad, Telangana, INDIA. He received his M.Tech (CSE) degree from Medak Engineering College in the year 2014. He received B.Tech. (CSE) degree from Jayamukhi Institute of Technological Sciences, in the year 2005. His fields of interests are Cloud Computing, Web Technologies, Big Data, etc.

Tolla L N .Varaprasad1 "Security Challenges in Big Data" International Journal of Computational Engineering Research (IJCER), vol. 08, no. 02, 2018, pp. 23–25.