International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056
Volume: 07 Issue: 09 | Sep 2020 www.irjet.net p-ISSN: 2395-0072

Expostulations and Scope of Big Data with Evolving Cybersecurity

Rahul Haripriya[1]

[1]*Department of Computer Science and Engineering, University Institute of Technology, RGPV, Bhopal, Madhya Pradesh, India*

Abstract - *This research paper brings out well-collected facts and possibilities about the threats and opportunities of Big Data technology when used with always evolving Cyber Security. As everyone is nowadays all time available online and are connected to the internet 24/7 the amount of data each person generates nowadays is really enormous and that is where the usage of Big Data comes as millions of people are generating data at a single point of time and storage and management of that much of data is not an easy task. At the same time the security of all the data is also important and so in terms of security of that enormous amount of data the security should also be considered with the help of Big Data hence this paper brings out the connection between the cybersecurity Big data and also brings out the possibilities and challenges of them working together. So this paper describes the major challenges of using Big Data in cybersecurity and opportunities of using Big Data with cybersecurity.*

Key Words: **Cybersecurity, Big Data, Data Privacy, Secure Database, Database Security.**

1. INTRODUCTION

Big Data mainly deals with very large and very complex data sets when there is a requirement of storage and management of very large-sized data. It is mostly utilized in keeping very large and complex records of organizations and individuals. The major purpose of Big Data is to provide meaningful information from a huge amount of complex stored data.

According to one of the most renowned firms in data analytics Cambridge Analytica a normal person has at least 5000 data points near him from where information about the person can be easily gained and mobile phones and computer systems are just 2 of those data points which we are aware of and when data analytics companies like them collect data from any normal person it is obvious they are processing the data with the help of Big Data and no doubt they are interfering our privacy and we are going to discover all those things involved in this process such as threats and possibilities of Big Data in cybersecurity and privacy.

All of our collected data could be used for the purpose of manipulating our search results on the web and even affect the marketing strategies of online ads as all the advertisement is based on what we search and store and not just that as we allow access of camera, microphone and location services the data that we are unaware of is also getting extracted by those firms using our calls and places we visit.

1.1 Cybersecurity

Whenever we are connected to the internet we have potential threats regarding our private data that it can be either accessed or misused by malicious users and the security of our personal data from such malicious users are termed as cybersecurity. Before you begin to format your paper, first write and save the content as a separate text file. Keep your text and graphic files separate until after the text has been formatted and styled. Do not use hard tabs, and limit use of hard returns to only one return at the end of a paragraph. Do not add any kind of pagination anywhere in the paper. Do not number text heads-the template will do that for you.

1.2 Big Data

Big Data refers to data which is enormous in amount and have a very complex size and structure which are meant to be mined in order to bring out any pattern of information from that data so bringing out useful information from complex and large-sized data is considered as big data analytics.

2. PROBLEM DEFINITION:

Millions of people are generating more and more data every second and each of those data elements is being observed by data specialists in the name of search optimization for users which is rising as a threat for every user connected with the

internet. So with the rise in the amount of data being generated, there is a need for raising the bar in terms of security as well.

Today's lifestyle has made every person from kids to adults 24/7 available online and that leads to major data sharing and also we need such significant security for our private data so that it could be secured 24/7 from the malicious attackers and cybercriminals. Our all-time availability has given a platform to cyber attackers to become a threat to our privacy.

3. OBJECTIVES:

As it is really common nowadays to share data among any network and also we have generated a habit of being with our smartphones and computers daily then to observe the effects of our data sharing and all-time mobile phone usage we have to achieve the following objectives:

- To study and compare various findings of challenges faced by Big Data in cybersecurity and also the opportunities for Big Data in cybersecurity.
- To determine various data points from which personal information of the user can be obtained by data analytics and marketing companies by performing basic tests.
- To determine potential sources of threats to our private data from malicious users as well as companies which are based on data analytics.
- To figure out various ways of data management using Big Data which could help in the enhancement of our data security.

4. LITERATURE SURVEY:

(a) There are big possibilities of utilization of Big Data in the field of cybersecurity and also the possible there are threats associated with our privacy which are of so many different types.[6] The main points of concern are the basic cybersecurity loopholes which are exploited by the malicious users and the further used for unauthorized access of our personal data. All the pros and cons of involving Big Data in cybersecurity.[1]

(b) There exists various possible technical and non-technical attacks on Big Data by many means such as attacking malicious codes and other methods of manipulating the security architecture of the Big Data and gaining access to the stored consumer data. As there various areas of applications of Big Data in different streams and that is why the security threats are not limited to a few number of sources instead there are many newly found possible sources can be there for storing information in Big Data. There are multiple challenges related to Big Data and countably a few possible solutions for security threats of Big Data.[2]

(c) There are various newly found innovative trends in Big Data which are found by various technology companies and also there are many latest softwares which can perform Big data operations. [8] The life cycle of the Big Data and how each step involved in it has its own significance. Various ideas of how the whole Big Data structure works and also how Big Data can be helpful in terms of cybersecurity.[3]

(d) There are so many security threats related to every kind of database that we use either in personal or organisational level and new databases such as MongoDB and Voldemort also have serious security concerns. New databases like HBase and Voldemort are utilized in maintaining the social media databases and they are most often under a massive risk because data from social media is considered as mostly valuable and contains personal information of millions of individuals.[4]

(e) Different databases have different pros and cons over each other. When compared in detail the advanced databases like HBase and classic databases like Hive are found similar in many aspects but at the same time very much different in capacities than each other in many other aspects. Each of them has a certain type of security threat based on different types of users like databases Hive is mostly involved in organizational usage and has corporate and finance related data meanwhile HBase is an opensource database which has various types of users from personal to organizational.[5]

5. PROPOSED METHODOLOGY

Performing various tests on smartphones and computers to figure out various data points from where our personal information is gained by any third party such as data analytics companies who later on modify our web browsing results as per our patterns and information and also enhance their advertisement in the name of search engine optimization.

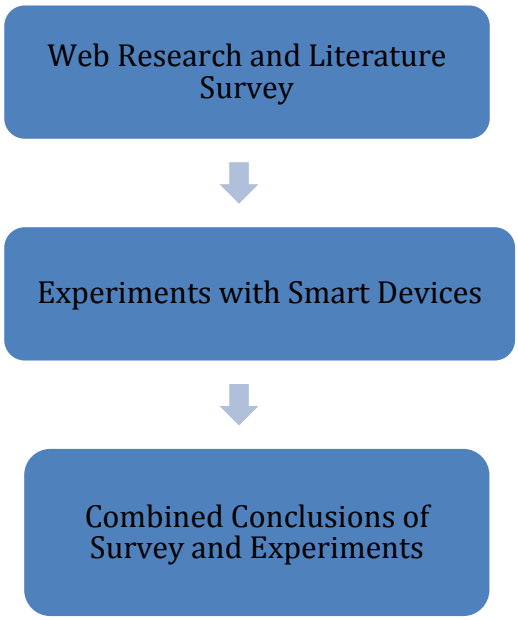
We are proposing the following tests:

(a) **Voice Test:** In this test, we will speak some things near the smart devices repeatedly for some time and will observe either the device listens to us or not even while we are not surfing on the browser by later on observing the advertisements and search results as voice recognition is a potential data point for any data collection firm.
(b) **Picture Test**: In this, we will try to capture some pictures using our smartphone camera and we will observe if we get any relevant advertisements to those pictures on the web as gallery access is considered as a potential data point of any user.
(c) **Location Test**: In this test, we will keep our location access service of our smartphone turned on and will frequently visit some particular locations for some days and will observe that either it affects our web search results or not.

Collection and analysis of data related to the performance and limitations of various tools of Big Data analysis:

(a) Apache Hadoop: It is a software framework which is used for the clustered file system and management of big data. It is basically an open-source framework which is written in Java and provides cross-platform support.
(b) IBM Security QRadar: It Prevents the stored data from potential threats proactively. It also helps in finding out hidden relationships between the data using analytics techniques to a controllable set of prioritized incidents.
(c) Apache Spark: It is a fast engine for data processing on a large scale and also it is an open-source cluster computing framework. It is used for analysis and answer finding of data.

Also collecting secondary data by performing web research and literature survey of various renowned journals, magazines, research papers, survey papers, books and other helpful websites based on Big Data and its's usage in cybersecurity along with challenges and issues of all possible attacks to our data in cybersecurity. And then after the collection of all the data, we will generate a comparative study of all the facts and information which will be gathered during the survey.



CONCLUSION

As per the proposed methodologies we can come to a conclusion that there are loads of opportunities for Big Data to be utilized in the field of cybersecurity also there are numerous attacks which can be considered as threats. After going through all the mentioned processes like the described experiments and also going deep with all the available research material for survey purpose we can consider that if all the things are done properly and ethically then major flaws in the cybersecurity system can be easily figured out and also some of the unknown secrets of the data analytics companies and social media websites such as Google and Facebook can also be discovered.

REFERENCES

[1] Challenges and Opportunities of Big Data with enhancing Cyber Security by Gajanan S. Kumbhar and Dr Ajit S. Ghodke [IJSCSEIT 2018 Vol.3]

[2] Big Data analytics in Cyber Security by Ashish Bajpai[JCSE 2018 Vol.6]

[3] Enhanced Cyber Security for Big Data Challenges by S.Padampriya, N.Partheeban, N.Kamal, A.Suresh, S.Arun [IJITEE 2019 Vol.8]

[4] A Survey on Security Issues in Big Data and NoSQL by Ebrahim Shahafizadeh and Mohammad Ali Nematbaksh [ASCIJ July 2015 Vol.4]

[5] Survey Paper on Big Data by C Lakshmi and V.V Nagendra Kumar [IJARCSSE 2016 Vol.6]

[6] Big Data in Computer Cybersecurity Systems by Amani Mobarak Al Madahkah [IJSCNS 2016 Vol.16]

[7] Enhanced Cybersecurity for Big Data Challenges S.Padampriya, N.Pratheebhan, N.Kamal, A.Suresh and S.Arun [IJITEE 2019 Vol.8]

[8] The Importance of Big Data Anakytics in the Field of Cyber Security by A.M Chandrashekhar, Hariprasad.M and Manjynatha.A [IJSRD 2016 Vol.3]

BIOGRAPHIES

Rahul Haripriya is a research scholar and a M Tech final year student in UIT, RGPV, Bhopal Madhya Pradesh, India