International Journal of Scientific Research in Computer Science, Engineering and Information Technology
© 2018 IJSRCSEIT | Volume 3 | Issue 1 | ISSN : 2456-3307

Challenges and Opportunities of Bigdata with Enhancing Cyber Security

Gajanan S. Kumbhar*[1], Dr. Ajit S. Ghodke[2]

*[1]PhD. Research Scholar, Department of Computer Science, Tilak Maharashtra Vidyapeeth, Pune, Maharashtra, India

[2]Associate Professor, Department of Computer Science, Tilak Maharashtra Vidyapeeth,Pune,Maharashtra, India

ABSTRACT

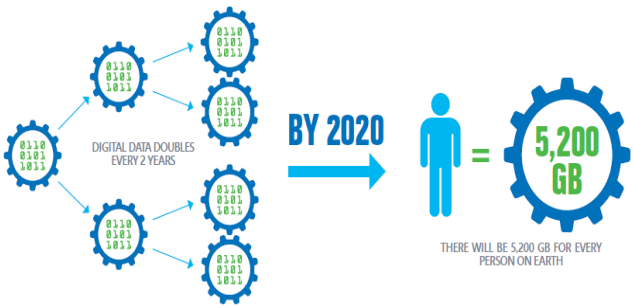
This Research paper offers review of challenges and opportunities of Big data with cyber security. All digital data created, replicated or consumed is growing by a factor of 30, doubling every two years. By 2020, there will be over 40 trillion gigabytes of digital data or 5200 gigabytes for every person on earth. The more data you produce and store, the more organized crime is ready to ingest. Now a day's 24/7 online lifestyle gives enormous opportunities to reach anyone from anywhere but this gives chances to cyber criminals. Cyber criminals use big data to learn more about infected machines, breached databases and compromised information systems. They use it to spot trends, failures and successes and to make their next attack more effective. The three V's of big data can be used to build a framework, which will integrate security issues, by design, using big data architecture.

Keywords: Big data, Cyber Security, Cyber Crimes, Challenges, Opportunities.

I. INTRODUCTION

Big data most commonly refers to more and more large and complex data sets. Big data not only refers to the volume of data but also data's variety and the velocity at which it is created, linked & altered. This is because of dramatically expanding universe of sensors, information technology services and connected devices, which all produces more and more data.

All digital data created, replicated or consumed is growing by a factor of 30, doubling every two years. By 2020, there will be over 40 trillion gigabytes of digital data or 5200 gigabytes for every person on earth. [1]



Source: Big Data: Seizing Opportunities, Preserving Value," White House Big Data Report, May 2014.,Aka.ms/WhiteHouse-BigData.

The more data you produce and store, the more organized crime is ready to ingest. The best example of this is Attack in India in Mumbai city on 2008, the attackers utilized cyber space for targeting venues and they simply processed the big data for their plan.

II. STATEMENT OF PROBLEM

Today across the world data is being created with large volume and variety of data, which is reached to extraordinary levels, and it will continue to accelerate. Protecting the information of individuals and organizations from online threats has to be a high priority is needed.

Now a days 24/7 online lifestyle gives enormous opportunities to reach anyone from anywhere but this gives chances to cyber criminals.

III. HYPOTHESIS OF THE STUDY

A. To study the risk perspectives associated with Bigdata and Cyber Security, which will give better results in data protection for organizations.
B. To study how big data will helps to fight against cyber-crimes towards predicting and in preventing them.

IV. REVIEW OF THE RELEVANT LITERATURE

A. Organizations are being encouraged to transition to intelligence-driven security for a broader view of risk and vulnerabilities. This requires analyzing external threat intelligence feeds, cloud-based calendars and documents, social network activity logs, website-generated information feeds and other non-traditional sources of security information.[4]

B. Cyber criminals have developed plug-in to query databases to transfer certain information, like credit card numbers, bank URLs or social security numbers into separate databases that they have full access. In addition to creating ways to mine big data for illicit gain, cyber criminals are also using it to monitor their processes and improve their own efficiency. They use big data to learn more about infected machines, breached databases and compromised information systems. They use it to spot trends, failures and successes and to make their next attack more effective.[4]

C. Rules-based IDS/IPS, SIEMS that analyze log data and network capture tools all have a role to play. Big data can be used to augment these defense strategies by providing fast and actionable information to the network defenders. Integrating network operations data and security product data from a layered defense strategy will provide an integrated data source that can capture event correlations or relationships where individually the risk appears low, but when analyzed in aggregate paint a clearer picture of cyber risk.[5]

V. THE METHODOLOGY COMPRISING

This study is based on secondary data collected from well-known articles of journals, books, outstanding websites.

A. Challenges and Opportunities of Big Data :[4][14]

- In the era of big data, awareness is the first line of defense against cybercrime. As one recent survey revealed, most cyber security, professionals know that they need to worry about big data, but they do not always clearly understand what it means.
- Organizations should integrate customized processes and technical solutions geared to their specific risks and requirements to collect process, store, analyze and share data.
- Integrating big data analytics into a solid infrastructure to provide and develop security solutions is essential – as is employing an expert IT staff to deploy them.
- Strengthening cyber security teams with highly skilled data scientists and analytics experts may become increasingly essential.
- Future investments in technology should lean toward flexible, analytics-based solutions that can change as business requirements and security threats evolve.

B. Challenges and Opportunities of Cyber security:[4][14]

According to Robert Eastman [11], most current cyber security threats can be categorized into the following broad categories:

a) Advanced Persistent Threats (APT)

An APT is a set of quiet and continuous computer hacking processes, often coordinated by humans targeting a specific entity. An APT usually aims organizations or nations for business or political purposes.

b) Insider Data Theft

An insider threat is a malicious threat to an institute that comes from people within the institute, such as employees, contractors or business associates, who have inside information concerning the institute‘s security practices and data. Data theft is done with the intent to compromise privacy or gain confidential information.

c) Distributed Denial of Service (DDoS)

In computing, a denial-of-service (DoS) attack is an attempt to make a network resource unavailable to its intended users, by suspending services of a host connected to the Internet. In a DDoS, the attack source is more than one, often thous ands of unique IP addresses.

d) Trojan Attacks

A Trojan horse is any malicious computer program, which appears as useful, routine, or interesting in order to influence a victim to install it. Some form of social engineering commonly spreads Trojans.

e) Phishing

Phishing is an attempt to acquire sensitive information such as usernames, passwords, and credit card details, by impersonating as a trustworthy entity.

f) External Software Introduction including Malware

Malware is any software used to disturb computer operations, gather sensitive information, and gain access to private computer systems. Sometimes, it can be used to display unwanted advertising.

g) SQL Injection

SQL injection is a code injection technique. It is used to attack data-driven applications. Malicious SQL statements are inserted into an entry field for execution. An SQL Injection can destroy your database.

h) Zero day Attacks

Zero day susceptibility refers to a security hole in software that is not known to the vendor. This hole is then misused by hackers before the vendor becomes aware and tries to fix it this exploit is called a zero day attack.

i) URL Redirection or Parameter Tampering

The web parameter tampering is constructed on the Manipulation of parameters exchanged between client and server to modify application data, such as user credentials and permissions, price and quantity of items, etc. Generally, this information is stored in cookies, hidden form fields, or URL Query Strings. The threat actors for the above categories can be classified as insider, opportunist, and accidental user.

VI. METHODS OF DATA ANALYSIS

1. Apache Spark

Apache Spark is a fast engine for data processing on a large scale. It is an open source cluster-computing framework. Apache Spark can help cyber security officers analyze data and answer questions:

- ✓ Which internal servers of the company are trying to connect to internationally based servers?
- ✓ Has user‘s access pattern to internal resources changed over time?
- ✓ Which users exhibit irregular patterns of behavior such as connecting using non-standard ports?

Spark powered big data discovery solutions can be used to detect anomalies and outliers within large datasets. Visualization techniques help when petabytes of data is to be analyzed.

2. Fortscale Services

Fortscale is a big data solution against APT attacks. APT attacks can take place over a stretched period while the victim organization remains ignorant about the invasion. According to Fortscale, big data analysis

is a suitable approach for APT detection. Fortscale uses Cloudera Hadoop distribution to address big data challenges, and examine network traffic data to check for invasions if any.

3. IBM Security QRadar

This tool uses big data capabilities to help keep pace with advanced threats and prevent attacks proactively. It helps reveal hidden relationships within large amounts of security data, using analytics to reduce billions of security events to a controllable set of prioritized incidents. It uses the following features of Big Data solution:

- ✓ Real-time correlation and anomaly detection of security data, which is diverse in nature.
- ✓ High-speed querying of security intelligence data.
- ✓ Flexible big data analytics across structured as well as unstructured data
- ✓ Graphical front-end tool for visualizing as well as exploring big data.

VII. CONCLUSION

Instead of making use of old and traditional cyber security methods and techniques, big data with behavioral analytics offers the best opportunity to improve information security. The three V's of big data can be used to build a framework, which will integrate security issues, by design, using big data architecture.

VIII. REFERENCES

[1]. A. A. Cardenas, P. K. Manadhata, S. P. Rajan, Big Data Analytics for Security, IEEE Security & Privacy,11 (6), 2013, pp. 74 -76.

[2]. Enhancing Cyber security with Big Data: Challenges & OpportunitiesDecember 2, 2016 by Emmeline Short.

[3]. Elisa Bertino, E. (2014). Security with Privacy - Opportunities and Challenges.

[4]. http://www.villanovau.com/resources/bi/for-cyber-security-big-data-offers-advantages-challenges [Accessed on 20 February 2018]

[5]. http://www.siliconindia.com/magazine_articles/Cyber_Security_in_the_Era_of_Big_Data-UGUC712003788.html [Accessed on 20 February 2018]

[6]. ICTACT Journal On Soft Computing: Special Issue On Soft Computing Models For Big Data, July 2015, Volume: 05, Issue: 04 1035

[7]. John Gantz, David Reinsel, -Digital Universe in 2020, IDC IView Report, December 2012.

[8]. O' Brien, S. (2016, May 05). Challenges to Cyber Security & How Big Data Analytics Can Help. Retrieved March 13, 2017, from http://datameer-wp-production-origin.datameer.com/company/datameerblog/challenges-to-cyber-security-and-how-big-data-analytics-can-help

[9]. Robert Eastman, -Big Data and Predictive Analytics: On the Cyber security Front Line, IDC Whitepaper, February 2015

[10]. Shen Yin, Okyay Kaynak,Big Data for Modern Industry: Challenges and Trends, Vol. 103, No. 2, February 2015, proceedings of the IEEE

[11]. Stephen Kaisler et.al, -Big Data: Issues and Challenges Moving Forward, IEEE Computer Society Intl Conf in Hawaii Jun13

[12]. Z. Spalevic, Cyber security as a global challenge today, Singidunum Journal of Applied Sciences, 2014, pp. 687 -692.

[13]. Solving Cyber Security Challenges using Big Data,Prajakta Joglekar, Nitin Pise, International Journal of Computer Applications (0975–8887)Volume 154–No.4, November 2016