An In-Depth Analysis of the Potential and Risks Linked to Big Data in Enhancing the Efficacy of Cybersecurity Safeguards

Drishti Arora

Any Graphics Pvt. Ltd., Greater Noida
Uttar Pradesh, India

ABSTRACT

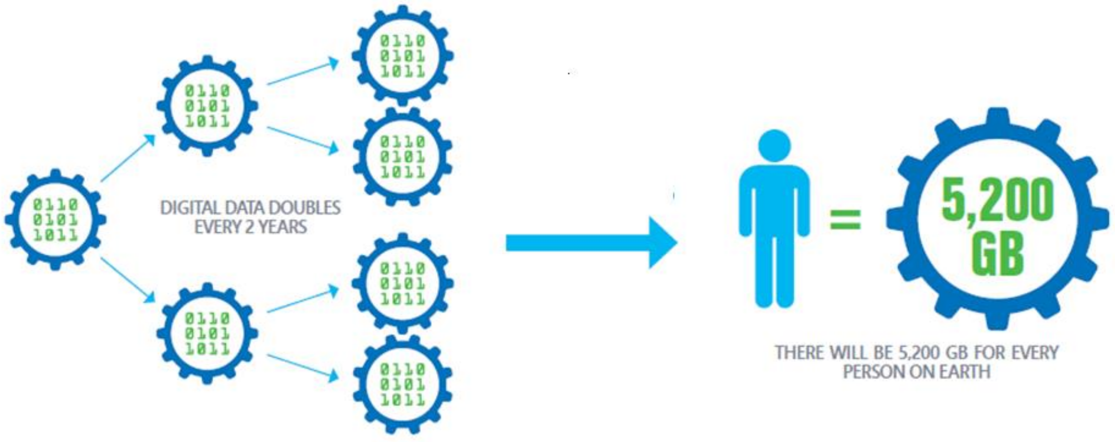
This research paper presents an examination of the challenges and possibilities that arise from the intersection of Big Data and cybersecurity. The exponential growth of digital data—increasing by a factor of 30 and doubling every two years—presents a significant landscape. By 2025, the projected data volume will exceed 40 trillion gigabytes, amounting to 5200 gigabytes for every individual on Earth. This proliferation of data also signifies an opportunity for organized crime to exploit.

In today's 24/7 online lifestyle, there are vast opportunities to connect globally, yet this connectivity also opens doors for cybercriminals. These nefarious actors leverage Big Data to gain insights into infected machines, breached databases, and compromised information systems. By discerning trends, failures, and successes, cybercriminals enhance the efficacy of their subsequent attacks.

The foundational three Vs of Big Data—Volume, Velocity, and Variety—offer a framework to integrate security considerations proactively within Big Data architectures. This approach aims to address security concerns by design, safeguarding against cyber threats.

INTRODUCTION

Big Data commonly denotes the increasing prevalence of large, intricate datasets. It encompasses not only the sheer volume of data but also its variety and the speed at which it is generated, linked, and modified. This surge is attributed to the expansive realm of sensors, information technology services, and interconnected devices, all contributing to the exponential growth of data. The digital data landscape is expanding rapidly, with information created, replicated, or consumed growing by a factor of 30 and doubling every two years. By 2020, projections indicate an astounding 40 trillion gigabytes of digital data, equivalent to 5200 gigabytes per person on Earth.


DIGITAL DATA DOUBLES EVERY 2 YEARS
0110 0101 1011
=
5,200 GB
THERE WILL BE 5,200 GB FOR EVERY PERSON ON EARTH

As data production and storage escalate, so does the readiness of organized crime to exploit it. An illustrative instance of this occurred during the 2008 Mumbai attacks in India, where the assailants utilized cyber space, strategically harnessing big data for their nefarious plans.

STATEMENT OF PROBLEM

In today's world, the proliferation of data—both in terms of volume and variety—has reached unprecedented levels and continues to accelerate. Safeguarding the information of individuals and organizations against online threats has become an urgent imperative. The pervasive 24/7 online lifestyle presents vast opportunities for connectivity, yet also creates avenues for cybercriminals to exploit.

HYPOTHESIS OF THE STUDY

A. This study aims to investigate the risk perspectives associated with Big Data and Cybersecurity, offering insights into enhanced data protection strategies for organizations.

B. Furthermore, it seeks to explore how Big Data can be leveraged to combat cybercrimes, aiding in their prediction and prevention."

METHODOLOGY

The methodology employed in this study is based on the analysis of secondary data gathered from reputable articles in journals, books, and reputable websites.

A. Challenges and Opportunities of Big Data:

In the age of big data, the first line of defense against cybercrime is awareness. Recent surveys indicate that while most cyber security professionals recognize the significance of big data, the concept itself is not always clearly understood.

- Organizations are advised to integrate customized processes and technical solutions tailored to their specific risks and needs for the collection, processing, storage, analysis, and sharing of data.

- The integration of big data analytics into a robust infrastructure to provide and develop security solutions is crucial. Equally important is the employment of skilled IT professionals to implement these solutions effectively.

- As the need for advanced data analysis grows, bolstering cyber security teams with highly skilled data scientists and analytics experts may become increasingly essential.

- Future technological investments should prioritize flexible, analytics-based solutions that can adapt to evolving business requirements and security threats.

B. Challenges and Opportunities of Cybersecurity:

Robert Eastman [11] categorizes the majority of current cybersecurity threats into the following broad categories:

a) Advanced Persistent Threats (APT):

- APTs are continuous, discreet hacking operations often orchestrated by humans, targeting specific entities such as organizations or nations for business or political motives.

b) Insider Data Theft:

- Insider threats pose a malicious risk to organizations, originating from individuals within the institution itself, including employees, contractors, or business associates. Data theft aims to compromise privacy or gain access to confidential information.

c) Distributed Denial of Service (DDoS):

- A denial-of-service attack attempts to render a network resource inaccessible to its intended users by overwhelming the host's services. In a DDoS attack, the sources are often thousands of unique IP addresses.

d) Trojan Attacks:

- Trojans are malicious computer programs disguised as useful or benign applications to deceive victims into installing them. Trojans are commonly spread through social engineering tactics.

e) Phishing:

- Phishing involves attempts to acquire sensitive information such as usernames, passwords, and credit card details by impersonating trustworthy entities.

f) External Software Introduction including Malware:

- Malware encompasses any software designed to disrupt computer operations, gather sensitive information, or gain unauthorized access to private systems. It may also display unwanted advertising.

g) SQL Injection:

- SQL injection is a code injection technique targeting data-driven applications. Malicious SQL statements are inserted into input fields for execution, potentially resulting in database destruction.

h) Zero-day Attacks:

- Zero-day vulnerabilities refer to security flaws in software unknown to the vendor. Hackers exploit these vulnerabilities before the vendor can address them, leading to zero-day attacks.

i) URL Redirection or Parameter Tampering:

- Web parameter tampering involves the manipulation of parameters exchanged between clients and servers to modify application data, such as user credentials, permissions, or item prices and quantities. This data is typically stored in cookies, hidden form fields, or URL query strings.

The threat actors associated with these categories can be classified as insiders, opportunists, and accidental users.

DATA ANALYSIS METHODS

1. Apache Spark

- Apache Spark is a high-speed engine designed for large-scale data processing. This open-source cluster-computing framework is invaluable for cyber security professionals in analyzing data and addressing key questions:

  - Which internal servers within the company are attempting to connect to servers based internationally?

  - Have the access patterns of users to internal resources changed over time?

  - Which users are displaying unusual behavior patterns, such as connecting through non-standard ports?

Apache Spark-powered solutions for big data discovery enable the detection of anomalies and outliers within extensive datasets. Visualization techniques prove especially beneficial when analyzing petabytes of data.

2. Fortscale Services

- Fortscale presents a robust big data solution against Advanced Persistent Threat (APT) attacks. APT attacks often unfold over extended periods while the victim organization remains unaware of the intrusion. According to Fortscale, big data analysis stands as an effective method for APT detection. Leveraging the Cloudera Hadoop distribution, Fortscale scrutinizes network traffic data to identify any potential invasions.

3. IBM Security QRadar

- This tool harnesses the power of big data capabilities to stay ahead of advanced threats and proactively prevent attacks. It assists in uncovering hidden relationships within vast amounts of security data, employing analytics to condense billions of security events into a manageable set of prioritized incidents. Key features of its Big Data solution include:

- Real-time correlation and anomaly detection of diverse security data.

- High-speed querying of security intelligence data.

- Flexible big data analytics spanning both structured and unstructured data.

- A graphical front-end tool for visualizing and exploring big data.

CONCLUSION

Rather than relying on outdated and traditional cyber security methods, the utilization of big data with behavioural analytics presents the most promising opportunity for enhancing information security. The three V's of big data—Volume, Velocity, and Variety—can form the foundation of a framework that integrates security concerns from the ground up, utilizing big data architecture to its fullest potential.

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