BIG DATA ANALYTICS IN CYBER SECURITY

Jui Roy

Department of Mathematics, University Institute of Sciences, Chandigarh University, Gharuan, Mohali, Punjab-140413, India

---***---

ABSTRACT: In the field of defence, big data analytics refers to the ability to collect large volumes of digital data in order to analyse, visualise, and draw conclusions that can help predict and prevent cyber attacks. It improves our cyber defence posture when combined with security technologies. They allow businesses to spot activity trends that indicate network threats. This research paper focuses on We'll look at how Big Data will help with information security. In this paper, I focus to how Big data can improve information security.

KEYWORD: Big data, Threat detection, Cyber security, Data base security.

INTRODUCTION:

The term "Big Data" refers to data sets that are extremely large or complex, so that conventional data set processing application software is insufficient or incapable of dealing with them. The most significant distinction between traditional and big data is in terms of scale, velocity, and variance. Volume denotes the volume of data generated; velocity denotes the rate at which the data is generated; and variance denotes the different types of structured and unstructured data. Big data is becoming a hot topic for research in almost every field these days, particularly cyber security. [3] Social networking sites and smart devices are the primary sources of this info. At this stage, data is being produced. This massive number of malware infections in only one economic field demonstrates the magnitude of the threat to the global economy each year. As a result, it is clear that the cyber-security of IT systems, company networks, and web applications may be insufficient to deal with the rapid evolution of cyber attacks.

BIG DATA ANALYTICS:

According to Gartner, “big data” tends to refer to “high-volume, high-velocity, and high-variety information assets that necessitate cost-effective, creative information processing for increased insight and decision making.” There are numerous meanings for big data available, but To me, Gartner's description made the most sense. Visualization is one notable absence in this description. The importance of visualisation to me stems from the fact that data processing often outpaces our ability to extract meaning from it.[1] This is due to the scarcity of actionable information. Most big data analytics methods, such as machine learning, statistics, predictive analysis, behavioural analysis, and so on, have been around for a long time. These techniques have traditionally been used on structured data sets ranging from a few MBs to a few GBs in size. They can now accommodate even larger quantities of both structured and unstructured data, up to petabytes. The following factors are behind this rapid change: The amount of data (Number) - Size: the volume of datasets is a crucial factor, as it determines how much data is produced. The amount of data (Number) - Size: the size of datasets is important because it decides how much data is processed. Complexity (structure, behaviour, and permutations of datasets)-Velocity (rate of data generation and transmission)-Complexity structure, behaviour.

THE-CYBER ATTACK LANDSCAPE:

According to a security report (Internet Security Threat Report, 2016), more than 430 million new pieces of malware were discovered in 2015, and what was even more remarkable about this is that this finding came as no surprise to the researchers. The security report goes on to explain that targeted, sophisticated and persistent attacks against government organisations and businesses of all sizes are on the rise and pose a serious threat to national security and economy. More than 430 million new pieces of malware were found in 2015, according to a security study (Internet Security Threat Report, 2016), and what's more surprising is that the researchers were not surprised. The security report goes on to say that targeted, complex, and persistent attacks on government agencies and companies of all sizes are a growing problem.[6] Cyber-attacks take many types, ranging from simple viruses to extremely advanced malware like cyber-weapons.

According to Virvilis and Gritzalis (2013), an APT has the following characteristics: They are usually targeted at specific and high-value targets, and thus for specific operating systems or platforms; they usually have an initial attack vector, such as malicious office documents or removable drives; and they are equipped with a list of evasion techniques to get around anti-virus software, and intrusion detection systems (IDS) implement command and control mechanisms; encryption of network traffic is one of their evasive techniques; they use stolen but authentic digital certificates to trick the targeted systems into assuming they are safe. Crime gangs, state-sponsored hackers, hacktivists, and lone wolf hackers are examples of external attackers. The attackers' method of attack includes the use of highly advanced malware that is difficult to detect even with similarly sophisticated security systems. For example, it was claimed that the malware used in the Sony attack may have gotten past most of the network defences.[5]

THREAT DETECTION WITH BDA:

The big data threat categories are three types : surveillance, disclosure, and discrimination.

Surveillance refers to the sensation of being observed as a result of the gathering, aggregation, and/or utilisation of one's data. The sensation of being watched could be an issue in and of itself, equivalent to mental distress. It could also be an issue because such feelings can influence how people behave if they begin to second-guess what they do, read, or look for.

Data that has been disclosed outside of the context in which it was acquired. One potential source of information exposure is a nosy employee who searches a corporate database for individuals he knows. An identity thief who successfully hacks into a database is another possibility. In this sense, insecurity issues are transparency issues.

Discrimination is defined as treating someone unfairly based on information gathered about them. Discrimination threatens people in a variety of ways. The most obvious example would be attempting to forecast membership in a protected group, such as race or religion, and then discriminating based on that. Some people may also object to any discrimination based on a protected characteristic.[4]

Antivirus applications, network IDS/IPS, host IDS/IPS, network device incidents, logging, FIM and white listing, and SIEM are the traditional categories for detecting and preventing cyber-attacks. While these systems are beneficial in many respects, they are not without flaws. against today's stealthy cyber-attacks is proving to be largely ineffective. This is due to the fact that, in addition to running independently of one another, these systems produce a large volume of data that is difficult and time consuming to analyse without the appropriate tool, making it possible to miss key cyber-attack events. This implies that these disparate networks can be made more effective with the proper implementation of the right tool that can sift through data much faster.[2]

CONCLUSION:

The aim of Big Data analytics for security is to obtain real-time actionable intelligence. In three ways, Big Data can have a significant effect on your current company. It will assist you in the following ways: They allow data to be stored in a single "window" of events, but they require a lot of computational power to analyse these "windows." And solutions based on the actor model were less resource-intensive on the machine, but more memory-intensive due to the need to repeat data for each entity. As a result, solutions based on MapReduce modifications took a middle ground role.

REFERENCES:

1. Wang, L. and Jones, R., 2020. Big data analytics in cyber security: network traffic and attacks. Journal of Computer Information Systems,
2. Kantarcioglu, M. and Xi, B., 2016, October. Adversarial data mining: Big data meets cyber security. In Proceedings of the 2016 ACM SIGSAC conference on computer and communications security (pp. 1866-1867).
3. Angin, P., Bhargava, B. and Ranchal, R., 2019. Big data analytics for cyber security.
4. Mahmood, T. and Afzal, U., 2013, December. Security analytics: Big data analytics for cybersecurity: A review of trends, techniques and tools. In 2013 2nd national conference on Information assurance (ncia) (pp. 129-134). IEEE.
5. Mahmood, T. and Afzal, U., 2013, December. Security analytics: Big data analytics for cybersecurity: A review of trends, techniques and tools. In 2013 2nd national conference on Information assurance (ncia) (pp. 129-134). IEEE.
6. Alguliyev, R. and Imamverdiyev, Y., 2014, October. Big data: Big promises for information security. In 2014 IEEE 8th International Conference on Application of Information and Communication Technologies (AICT) (pp. 1-4). IEEE.