Security Analytics: Challenges and Future Directions

Ganga Sharma*
Bhawana Tyagi**

Abstract

The frequency and type of cyber attacks are increasing day by day. However, well-known cyber security solutions are not able to cope with the increasing volume of data that is generated for providing security solutions. Therefore, current trend in research on cyber security is to apply Big Data Analytics (BDA) techniques to cyber security. This field, called security analytics (SA), can help network managers in the monitoring and surveillance of real-time network streams and real-time detection of malicious and/or suspicious patterns. Researchers believe that an SA system can assist in enhancing all traditional security mechanisms. Nonetheless, there are certain issues related to incorporating big data analytics to cyber security. This paper presents an analysis on the issues and challenges faced by Security Analytics, and further provides future directions in the field.

Keywords: cyber-security, big data, security analytics, big data analytics

I. Introduction

Big data analytics (BDA) is the large scale analysis and processing of information [1,14]. It uses advanced analytic and parallel techniques to process very large and diverse records that include different types of contents. BDA tools allow getting enormous benefits and valuable insights by dealing with any massive volume of mixed unstructured, semi-structured and structured data that is fast changing and difficult to process using conventional database techniques.

In recent years, BDA has gained popularity in the security community as it promises efficient processing and analysis of security-related data at large scale [3]. Corporate research is now focusing on Security Analytics, i.e., the application of Big Data Analytics techniques to cyber-security. Analytics can assist network managers particularly in the monitoring and surveillance of real-time network streams and real-time detection of both malicious and suspicious (outlying) patterns. Over the past ten years, enterprise security has gone incrementally more difficult as new and unanticipated threats/attacks surface. The existing security infrastructures collect, process and analyze terabytes of security data on monthly basis. This data is too large to be handled efficiently by the existing data storage architectures, algorithms, and query mechanisms. Therefore the application of Big data analytics (BDA) to security is the need of the hour.

This paper provides an overview of how big data analytics can help in enhancing the traditional cyber security mechanisms and thus provide a means for better security analysis. Rest of the paper is organized as follows: section 2 gives a brief overview of literature work, section 3 describes the basic BDA process, section 4 and 5 respectively provide the challenges and fututre directions in security analytics while section 6 concludes the paper.

II. Literature Review

Security analytics is a new technology and concept, therefore much research has not been conducted in this area. However, there are some significant contributions by several authors in this field. For e.g., Mahmood and Afzal[14] have presented a comprehensive survey on the state of the art of Security Analytics, i.e., its description, technology, trends, and tools. Gahi et al [1] highlight the benefits of Big Data Analytics and then provide a brief overview of challenges of security and privacy in big data environments itself. Further, they present some available protection techniques and propose some

Ganga Sharma*
Assistant Professor (IT Dept)
IITM Janakpuri

Bhawana Tyagi**
Assistant Professor (IT Dept)
IITM

possible tracks that enable security and privacy in a malicious big data context. Cybenko and Landwehr[7] stud-ied historical data from a variety of cyber- and national security domains in United state such as computer vulner-ability databases, offensive and defense, co-evolution of wormbots such as Conficker etc. They claim that security analytics can provide the ultimate solution for cyber-security. Cardenas et al[9]provide details of how the security analytics landscape is changing with the introduction and widespread use of new tools to leverage large quantities of structured and unstructured data. It also outlines some of the fundamental differences between security analytics and traditional analytic. Camargo et al[10] research on the use of big data analytics for security and analyze the perception of people for security. They found that big data can indeed provide a long-term solution for citizen's security, in particular cyber security.

III. Big Data And The Basic Bda Process

Big data is data whose complexity hinders it from being managed, queried and analyzed efficiently by the existing database architectures[4]. The "complexity" of big data is defined through 4V's: 1) volume – referring to terabytes, petabytes, or even exabytes (10006 bytes) of stored information, 2) variety – referring to the co-existence of unstructured, semi-structured and structured data, and 3) velocity – referring to the rapid pace at which big data is being generated and 4) veracity- to stress the importance of maintaining quality data within an organization.

The domain of Big Data Analytics (BDA) is concerned with the extraction of value from big data, i.e., insights which are nontrivial and previously unknown, implicit and potentially useful. These insights have a direct impact on deciding or manipulating the current business strategy [14]. The assumption is that patterns of usage, occurrences or behaviors exist in big data. BDA attempts to fit mathematical models on these patterns through different data mining techniques such as Predictive Analytics, Cluster Analysis, Association Rule Mining, and Prescriptive Analytics [13]. Insights from these techniques are typically represented on interactive dashboards and help corporations maintain the competitive edge, increase profits, and enhance their CRM.

Fig. 1 shows the basic stages of BDA process[14] . Initially, data to be analyzed is selected from real-time streams of big data and is pre-processed (i.e. cleaned). This is called ETL (Extract Transform Load). It can take up to 60% of the effort of BDA, e.g., catering for inconsistent, incomplete andmissing values, normalizing, discretizing and reducing data, ensuring statistical quality of data through boxplots, cluster analysis, normality testing etc., and understanding data through descriptive statistics (correlations, hypothesis testing, histograms etc.). Once data is cleaned, it is stored in BDA databases (cloud, mobile, network servers etc.) and analyzed with analytics. The results are then shown in interactive dashboards using computer visualization.

IV. Challenges in Security Analytics

The big data is a recent technology and has been widely adopted to provide solutions to organsational decision making[11]. One of the most important area to benefit from the advancements in big data analytics is cyber security. This area is now being stated as security analytics. An important goal for security analytics is to enable organisations to identify unknown indicators of attack, and uncover things like when compromised credentials are being used to bypass defenses[2]. However, handling unstructured data and combing it with structured data to arrive at an accurate assessment is one of the big challenges in security analytics.

In the past, information security was really based on event correlation designed for monitoring and detecting known attack patterns[9]. This model alone is no longer adequate as multidimensional cyber-attacks are dynamic and can use different tactics and techniques to find their way into and out of an organization. In addition, the traditional set of security devices is designed and optimized to look for particular aspects of attacks: a network perspective, an attack perspective, a malware perspective, a host perspective, a web traffic perspective, etc[12]. These different technologies see isolated aspects of an attack and lack the bigger picture.

1. Cyber-attacks are extremely difficult to distinguish or investigate, because until all the event data is combined, it's extremely hard to determine what an attacker is trying to accomplish[6,8].

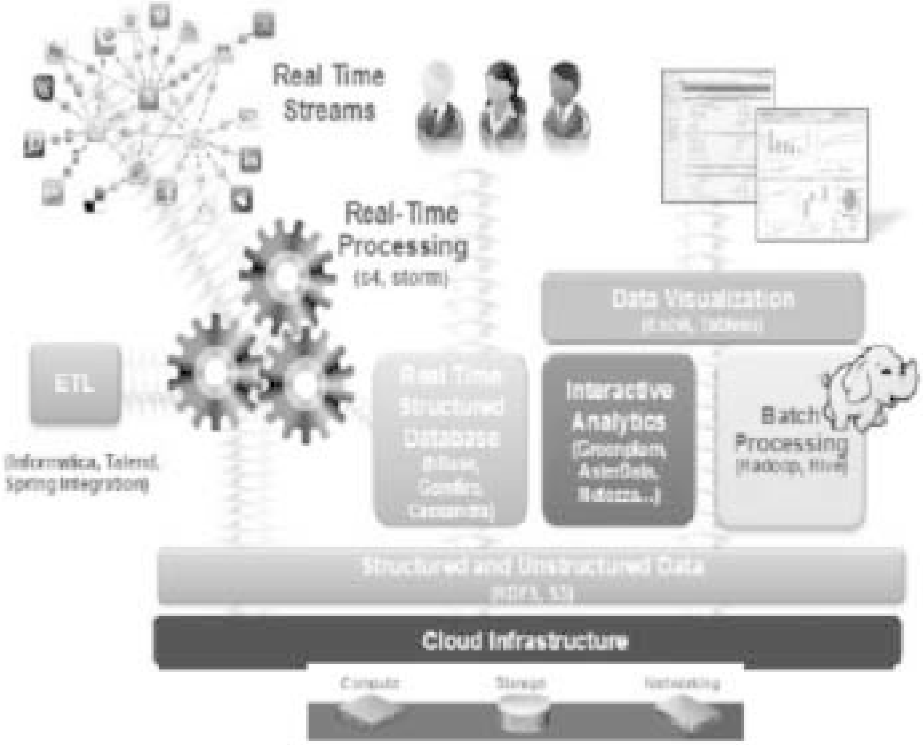

Fig1. Basic BDA process[14]

Addressing new types of cyber-threats requires a commitment to data collection and processing as well as much greater diligence on security data analytics.

2. The main idea behind big data is to extract useful insights by performing specific computations. However, it is important to secure and protect these computations to avoid any risk or attempt to change or skew the extracted results. It is also important to protect the systems from any attempt to spy on the nature or the number of performed computations.

3. In an open context, large volume of content collected through big data is not always a good metric for the quality of extracted results. Therefore, it may not always be possible to achieve good threat detection and prevention.

4. Since cyber-attacks can be multidimensional can happen over long periods of time, historical analysis must also be incorporated so that analysts can perform root cause analysis and attack scoping to determine the breadth of a compromise or data breach.

5. While original data formats should be preserved, security analysts must also have the ability to tag, index, enrich, and query any data element or group of data elements together to get a broader perspective for threat detection/response. Otherwise, security data will remain a black hole if it can't be easily queried and understood by security professionals .

6. Systems must provide a simple interface and search-based access to broaden and simplify access to data. This will empower security analysts to investigate threats and gain valuable experience. Systems should also allow for straightforward ways to create dashboards and reports to streamline security operations.

V. Future Directions

It is no longer a matter of if, but when, attackers will break into your network. They'll use zero-day attacks, stolen access credentials, infected mobile devices, a vulnerable business partner, or other tactics. Security success is not just about keeping threats out of your network. Instead it's about quickly responding to and thwarting an attack when it happens[4,5]. According to a very reputed organization providing security solutions "Organizations are failing at early breach detection, with more than 92 percent of breaches undetected by the breached organization." It is clear that we need to play a far more active role in protecting our organizations[8]. We need to constantly monitor what is going on within our infrastructure and have an established, cyclical means of responding before attacks wreak havoc on our networks and reputations. Therefore, some of the primary requirements for the security analytics solution are:

1. Secure sensitive data entering Big database systems and then provide control access to Protected data by monitoring which applications and which users gets access to which original data.
2. Protection of sensitive data that maintains usable, realistic values for accurate analytics and modeling on data in its encrypted form.
3. Assure global regulatory compliance. Securely capture, analyze and store data from global sources, and ensure compliance with international data security, residency and privacy regulations. Address compliance comprehensively, not system-by-system.
4. Optimize performance and scalability.
5. Integrate data security, with quick implementation and an efficient, low-maintenance solution that should scale up. Leverage IT investments by integrating with the existing IT environment and extending current controls and processes into Big Databases.
6. As far as possible provide block layer encryption, which will improve security but also enable big data clusters to scale and perform[7,8].
7. Leverage security tools or third-party products. Tools may include SSL/TLS for secure communication, Kerberos for node authentication, transparent encryption for data-at-rest[13].

VI. Conclusion

Security analytics is the new technical foundation of an informed, reliable detection and response strategy for cyber attacks. Mature security organizations recognize this and are leading with building their security analytics capabilities today. A security analytics system combines and integrates the traditional ways of cyber threat detection to provide security analysts a platform with both enterprise-scale detection and investigative capabilities. It will not only help identify events that are happening now, but will also assess the state of security within the enterprise in order to predict what may occur in the future and enable more proactive security decisions.

References

1. Gahi, Y., Guennoun, M., & Mouftah, H. T. (2016, June). Big Data Analytics: Security and privacy challenges. In *Computers and Communication (ISCC), 2016 IEEE Symposium on* (pp. 952-957). IEEE.
2. Verma, R., Kantarcioglu, M., Marchette, D., Leiss, E., & Solorio, T. (2015). Security analytics: essential data analytics knowledge for cybersecurity professionals and students. *IEEE Security & Privacy, 13*(6), 60-65.
3. Oltsik, J. (2013). The Big Data Security Analytics Era Is Here. White Paper, Retrieved *from https://www.emc.com/collateral/analyst-reports/security-analytics-esg-ar.pdf* on on 30th December, 2016
4. Shackleford D. (2013). SANS Security Analytics Survey, WhitePaper, SANS Institute InfoSec Reading Room. Downloaded on 3^{0th} December, 2016.
5. Gawron, M., Cheng, F., & Meinel, C. (2015, August). Automatic detection of vulnerabilities for advanced security analytics. In *Network Operations and Management Symposium (APNOMS), 2015 17th Asia-Pacific* (pp. 471-474). IEEE.
6. Gantsou, D. (2015, August). On the use of security analytics for attack detection in vehicular ad hoc networks. In *Cyber Security of Smart Cities, Industrial Control System and Communications (SSIC), 2015 International Conference on* (pp. 1-6). IEEE.

7. Cybenko, G., & Landwehr, C. E. (2012). Security analytics and measurements. *IEEE Security & Privacy, 10*(3), 5-8.

8. Cheng, F., Azodi, A., Jaeger, D., & Meinel, C. (2013, December). Multi-core Supported High Performance Security Analytics. In *Dependable, Autonomic and Secure Computing (DASC), 2013 IEEE 11th International Conference on* (pp. 621-626). IEEE.

9. Cardenas, A. A., Manadhata, P. K., & Rajan, S. P. (2013). Big data analytics for security. *IEEE Security & Privacy, 11*(6), 74-76.

10. Camargo, J. E., Torres, C. A., Martínez, O. H., & Gómez, F. A. (2016, September). A big data analytics system to analyze citizens' perception of security. In *Smart Cities Conference (ISC2), 2016 IEEE International* (pp. 1-5). IEEE.

11. Alsuhibany, S. A. (2016, November). A space-and-time efficient technique for big data security analytics. In *Information Technology (Big Data Analysis)(KACSTIT), Saudi International Conference on* (pp. 1-6). IEEE.

12. Rao, S., Suma, S. N., & Sunitha, M. (2015, May). Security Solutions for Big Data Analytics in Healthcare. In *Advances in Computing and Communication Engineering (ICACCE), 2015 Second International Conference on* (pp. 510-514). IEEE.

13. Marchetti, M., Pierazzi, F., Guido, A., & Colajanni, M. (2016, May). Countering Advanced Persistent Threats through security intelligence and big data analytics. In *Cyber Conflict (CyCon), 2016 8th International Conference on* (pp. 243-261). IEEE.

14. T. Mahmood and U. Afzal, "Security Analytics: Big Data Analytics for cyber-security: A review of trends, techniques and tools," 2nd National Conference on Information Assurance (NCIA), 2013