IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.102 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 13, Issue 2, February 2024

DOI: 10.17148/IJARCCE.2024.13208

Data Collection Technologies Using Network Security

B. Ashwini[1], R. Ranjani[2], M. Ezhilrani[3]

PG Student, Department of Computer Science, R.B. Gothi Jain College for Women, Chennai-600057, India.[1,2]

Research Scholar, Department of Information Technology, Bharathiar University, Coimbatore-641016, India.[3]

Abstract: According to Security threats and economic loss caused by network attacks, invasion, and vulnerabilities have motivated intensive studies on network security. Normally, the data collected in a network system can return or can be used to detect security threats. Examining and investigate security-related data can help detect network attacks and invasion, Hence making it for possible to further measure the security level of the whole network system. Obviously, the first step in detecting network attacks and invasion is to collect security-related data. Following we provide that requirements and objectives for security-related data collection and present a taxonomy of data collection technologies. In this paper we discuss network security-related data collection, requirement, objectives, technologies, future research trends.

Keywords: security, data collection, CIA Triangle.

I. INTRODUCTION

With the express development of network and communication technologies, there has been increasing amount of attention on the security of network systems. Network security is usually reflected by to the point data generated, originated or Extracted from the network system. By studying the data related to network security events, the security of the network system can be quantified and measured. We refer to the data that indicates security threats and shows abnormality with regard to security, safety, privacy and trust as ***network security-related data***, in short security-related data. Obviously, the first step to detect network attacks and intrusions is to collect the security-related data. However, there are a lot of challenges in collecting the security-related data in the current era of big data and the next generation of network systems (in short 5G). In the context of big data, the amount of data shared, originated, produced in the network is enormous. The security-related data has 5V (i.e., Volume, Variety, Value, Velocity and Veracity) characteristics, which pose tremendous difficulties in collecting these data.

In this paper, we that provide a concluded review on network data collection technologies and compare existing works according to a number of functional and security objectives for the purpose of high quality security- related data collection. We then discuss open issues, challenges and future research trends in this field. Specifically, our contributions can be summarized as follows:

➢ We propose a number of requirements and objectives in terms of collecting the security- related data in a large scale heterogeneous network system, which can be used to evaluate existing related works;

➢ We comprehensively review the methods, mechanisms and technologies for collecting network data by applying the proposed requirements and objectives to evaluate their performance towards high quality network security-related data collection;

II. NETWORK SECURITY RELATED DATA COLLECTION

Purposes and Application Scenarios:

Collecting network data is becoming increasingly important, particularly with the flourishing of big data and Internet of Things (IoT). The purposes of collecting network data mainly include 1) intrusion detection, 2) network management, 3) traffic accounting, 4) network forensics, and 5) malware detection.

1) Intrusion Detection:
Intrusion detection refers to the behavior of monitoring and detecting nasty activities or policy abuse in a network or system. The most used scenarios of network security-related data collection are IDS and other network security systems or security devices that detect network attacks and intrusions.

DOI: 10.17148/IJARCCE.2024.13208

The data collection module of IDS is responsible for monitoring the host state, the network data and user behaviors. Network data here include the parameters of network activities, the number of network connections, the number of packets, the content of packages, etc.

2) Network Management:

Network management includes network opinion, fault detection, network configuration and design. A Network Management System (NMS) troubleshoots network fault, performs network configuration, and monitors quality of service. An efficient NMS focuses on collecting real-time network data in order to monitor relevant resources and network performance, thus bond the efficiency and robustness of the network system. Network data collection has long been the central component of NMS or Network Protocol Analyzer.

III. REQUIREMENTS OF SECURITY-RELATED DATA COLLECTION

1) Functional Requirements (FR)

The functional requirements (FR) are those that must be implemented in order to collect network security-related data. We list the requirements based on the current literature as below.

➢ FR1: Must be able to collect required security- related data in different situations and contexts.
➢ FR2: Must be able to store collected data in a storage medium.
➢ FR3: Must be able to know where and when to collect security-related data.
➢ FR4: Must be able to load information about which data should be collected.
➢ FR5: Must be able to export data to other systems or create external database.
➢ FR6: Must have the ability to manage and control data.
➢ FR7:Must be efficient and stable when collecting data.
➢ FR8: Must be flexible and scalable in data collection.
➢ FR9: Must not cost too much computation resources, storage resources or other resources in collecting data in some scenarios.
➢ FR10: Must be automatic and adaptive, with a certain degree of intelligence and learning ability in order to adapt to the changes of a network environment.
➢ FR11: Must not ruin the original network system.
➢ FR12: Must be universal and generic and be able to support a variety of application scenarios.
➢ FR13: Must not produce new data that might affect the accuracy of collected data.

2) Security Requirements (SR)

The security requirements (SR) are those that deal with quality and security issues when collecting network security-related data.

➢ SR1: Must be able to prevent data loss and ensure data truth in collecting data (data integrity, veracity and availability).
➢ SR2: Must protect user privacy in collecting data.
➢ SR3: Must ensure the security of collected data and be able to prevent data leakage.
➢ SR4: Must be able to verify the integrity and accuracy of the collected data.
➢ SR5: Must have access control capability that can authenticate users who want to access data and enable the access for eligible users.

Objectives of Security-Related Data Collection

1) Functional Objectives

➢ Applicability(APP)

Based on FR1 and FR4, we propose an objective named Applicability. Applicability in this paper refers to that a proposed network data collection technique or mechanism can be planted into a real network environment to collect network security-related data.

➢ Adaptability (ADA)

Based on FR10,we propose an objective named work ability. Adaptability refers to that a collection mechanism can adjust to different network contexts and situations. For example, the collected content can be chosen according to network context variation or the collecting frequency can be adjusted based on network data variation

DOI: 10.17148/IJARCCE.2024.13208

➢ Scalability (SCA)

BasedonFR8,we propose an objective named Scalability. Scalability refers to the quality of being scalable for a network security-related data collection technology. A elastic data collection technology should be based on a scalable architecture to support data collection of different types and various volumes of data.

IV. OPEN ISSUES AND FUTURE RESEARCH TRENDS

A. Open Issues

1) Heterogeneity - 5G era is coming, thus it is critical to implement a proper data collection mechanism in a large-scale heterogeneous network system. There are still two unsolved issues in terms of collecting network security-related data.

2) Adaptability

The data needed to be collected is different (e.g., security-related data)for different application requirements. In order to solve the problem of the validity of data collection, we need to implement specific types of data identification in mass data

3) Scalability with Light Complexity

The rule-based data collection method has high scalability, but it suffers from the rule conflict problem. If we find an efficient solution to solve this problem, scalability can be supported for network security-data collection.

V. CONCLUSION

Studying network security-related data collection is essential for the detection of network attacks and intrusions, thus contributing to ensure the security of a whole network system. In this paper, we introduced the concept of security-related data collection, specified its requirements and defined its objectives regarding both functionalities and security. Furthermore, we presented a taxonomy and classification of data collection technologies. With regard to data collection technologies, we mainly reviewed data collection nodes, data collection tools and specific data collection mechanisms and hopefully, they can also benefit other researchers and practitioners in this field

REFERENCES

[1]. D. L. Donoho, High-Dimensional Data Analysis: The Curses and Blessings of Dimensionality. Lecture on August 8, 2000, to the American Mathematical Society "Math Challenges of the 21st Century". Available from http://www-stat.stanford.edu/~donoho/.

[2]. M. Turk, A. Pentland, Eigenfaces for Recognition, J. Cognitive Neuroscience, 3-1 (1991) 71-96.

[3]. R. Bellmann, Adaptive Control Processes: A Guided Tour. Princeton University Press, 1961.

[4]. A. Barron, Universal Approximation Bounds for Superpositions of a Sigmoidal Function, IEEE Tr. On Information Theory, 8-3 (1993) 930-945.

[5]. D. W. Scott, J. R. Thompson, Probability density estimation in higher dimensions. In: J.E. Gentle (ed.), Computer Science and Statistics: Proceedings of the Fifteenth Symposium on the Interface, Amsterdam, New York, Oxford, North Holland-Elsevier Science Publishers, 1983, pp. 173-179.

[6]. P. Comon, J.-L. Voz, M. Verleysen, Estimation of performance bounds in supervised classification, *European Symposium on Artificial Neural Networks*, Brussels (Belgium), April 1994, pp. 37-42.

[7]. Chen, C.-M., Lee, H.-M., and Hwang, C.-W. A hierarchical neural network document classifier with linguistic feature selection. Applied Intelligence, 23(3):277–294, Dec. 2005.

[8]. Chen, L., Tokuda, N., and Nagai, A. A new differential LSI space-based probabilistic document classifier. Information Processing Letters, 88(5):203–212, 2003

[9]. Díaz, I., Ranilla, J., Montañes, E., Fernández, J., and Combarro, E. Improving performance of text categorization by combining filtering and support vector machines. Journal of the American Society for Information Science and Technology (JASIST), 55(7):579â˘A ¸S–592, May 2004

[10]. Drucker, H., Wu, D., and Vapnik, V. Support vector machines for spam categorization. IEEE Transactions on Neural Networks, 10(5):1048–1054, 1999.

[11]. Deng, Z.-H., Tang, S., Yang, D., Zhang, M., Li, L.-Y., and Xie, K.-Q. A comparative study on feature weight in text categorization. In Proc. of the Advanced Web Technologies and Applications, 6th Asia-Pacific Web Conference, APWeb 2004, pages 588–597, Hangzhou, China, Apr. 14–17, 2004. Springer-Verlag New York, Inc.

[12]. Napoleon D. and Praneesh M. “Detection of Brain Tumor using Kernel Induced Possiblistic C-Means Clustering”, volume no.3, issue no.9, pp 436-438, 2013