Chapter 2
Advancing Cybersecurity: Leveraging Anomaly Detection for Proactive Threat Identification in Network and System Data

Dwijendra Nath Dwivedi
https://orcid.org/0000-0001-7662-415X
Krakow University of Economics, Poland

Ghanashyama Mahanty
https://orcid.org/0000-0002-6560-2825
Utkal University, India

Shafik Khashouf
University of Liverpool, UK

ABSTRACT

This chapter presents an innovative approach to cybersecurity by applying anomaly detection techniques to network and system data. The study uses a comprehensive dataset from simulated network environments to analyze various attack scenarios and evaluate classification algorithms. The approach uses an ensemble model to achieve superior detection accuracy and integrates feature importance analysis. The findings show that the proposed anomaly detection framework not only identifies known attack types but also detects novel threats, underscoring its potential as a pivotal tool in cybersecurity. This research paves the way for a new era in cybersecurity. These findings reveal that the proposed anomaly detection framework not only achieves high accuracy in identifying known attack types but also exhibits robustness in detecting novel threats, thereby underscoring its potential as a pivotal tool in the cybersecurity arsenal. This chapter advocates for a paradigm shift towards proactive threat identification, emphasizing the critical role of anomaly detection in fortifying network defenses against the ever-increasing sophistication of cyber-attacks.

DOI: 10.4018/979-8-3693-3597-0.ch002

INTRODUCTION

The advent of the digital era has brought about an unparalleled level of connectedness and ease, which has fundamentally altered the way in which we live, work, and interact with one another. On the other hand, this interconnection also creates an environment that is conducive to the development of cyber dangers, which is why enterprises, nations, and individuals alike should place a high priority on cybersecurity. The necessity for advanced and proactive security measures has grown more vital than it has ever been before as a result of the increasing sophistication of the strategies that hackers use. Using the ability of machine learning to identify and mitigate risks in network and system data, this study digs into the world of anomaly detection as a cornerstone for improving cybersecurity. Anomaly detection will be discussed as a cornerstone. Malware, phishing, and denial-of-service attacks are all examples of cyber threats. Other types of cyber threats include advanced persistent threats and zero-day exploits, which are more covert activities for hackers. Antivirus software that is based on signatures and firewalls are examples of traditional security methods that are effective at identifying known threats. However, these systems frequently fail to detect innovative or sophisticated attacks that do not match established patterns. An strategy that is more dynamic and flexible to threat detection and response is required because of this constraint, which highlights the requirement of such an approach.

The capacity to recognize unexpected patterns or behaviors that may indicate a potential security risk is made possible by anomaly detection, which emerges as a powerful solution to this difficulty. The focus of anomaly detection is on recognizing deviations from normal activity, which enables the detection of dangers that were not previously known. This is in contrast to conventional approaches, which depend on known threat signatures. In today's fast changing threat landscape, when new vulnerabilities and attack vectors are constantly appearing, this method is especially valuable because it allows for swift progress to be made. We investigate the use of a variety of machine learning models to analyze data from networks and systems in order to identify anomalies in this study. Our study makes use of a rich dataset that simulates a wide variety of network interactions, including both events that are not harmful and those that are malevolent. The rigorous exploratory data analysis that we perform allows us to obtain profound insights into the characteristics of the dataset, allowing us to recognize important qualities and patterns that are related with patterns of abnormal behavior. Our models are able to differentiate between regular and abnormal occurrences thanks to the utilization of a number of machine learning methods, such as Decision Trees, Random Forest, Gradient Boosting, and Support Vector Machines, amongst others. We ensure that a comprehensive review of each model's performance is carried out by utilizing evaluation criteria that include accuracy, precision, recall, and the F1 score.

In addition to this, we present an ensemble model that enhances detection capabilities by combining the favorable characteristics of individual classifiers. Not only does this ensemble technique improve overall accuracy, but it also provides support against deficiencies that are specific to the model. When it comes to giving security analysts and system administrators with relevant information, we also undertake feature importance analysis to determine which network attributes are most symptomatic of unusual behavior. This study contributes to the creation of security measures that are more proactive and adaptive by showing a comprehensive framework for utilizing anomaly detection in the field of cybersecurity. Our research demonstrates that machine learning has the potential to revolutionize cybersecurity operations by shifting from reactive to proactive threat identification and mitigation methodologies. Following this, we will outline our approach for data analysis and model creation, show our results and conclusions,