Chapter 13
Decision Trees Unleashed: Simplifying IoT Malware Detection With Advanced AI Techniques

Derek Mohammed
Capitol Technology University, USA & Saint Leo University, USA

Marwan Omar
Capitol Technology University, USA & Illinois Institute of Technology, USA

ABSTRACT

This chapter presents an in-depth study on the application of decision tree-based classifiers for the detection of malware in internet of things (IoT) environments. With the burgeoning expansion of IoT devices, the threat landscape has grown increasingly complex, making traditional security measures insufficient. This study proposes an innovative approach using decision tree algorithms to address the growing concern of IoT malware. The research methodology encompasses a comprehensive analysis of IoT vulnerabilities, focusing on malware threats and the development of a decision tree-based classifier. The classifier is empirically validated using the MaleVis dataset, a rich source of real-world IoT malware data. Performance metrics such as precision, recall, specificity, F1-score, accuracy, and processing time are meticulously evaluated to determine the efficacy of the model.

INTRODUCTION

The concept of intrusion in the context of computer security refers to attempts to breach security protocols by compromising a system's integrity. In response to this pervasive threat, a variety of tools and techniques, including advanced detection systems, have been developed to fortify networks and systems against such intrusions. Intrusion detection, as delineated in extant literature (Chiba, 2019; Irshad, 2020; Omar, 2022; Irshad, 2019; Chaudry, 2020), entails the classification of data activity into normative or intrusive categories to pinpoint undesirable activities. An intrusion detection system (IDS) primarily functions to detect and thwart intrusion attempts, originating either externally or internally, within a monitored network. Predominantly, IDS utilizes two detection methodologies: misuse detection, which

DOI: 10.4018/979-8-3693-1906-2.ch013

Copyright © 2024, IGI Global. Copying or distributing in print or electronic forms without written permission of IGI Global is prohibited.

identifies intrusions using known attack signatures, and anomaly detection, which hinges on deviations from established normal behavior models (Guezzaz, 2021; Omar, 2023).

Despite the efficacy of IDSs, they are not devoid of limitations, particularly in areas such as real-time detection, alarm generation, and data accuracy, which occasionally culminate in less than optimal detection results. This ongoing challenge underscores the continued relevance and dynamism of intrusion detection as a research field. Recent explorations in this domain have centered around the integration of machine learning (ML) methodologies to bolster intrusion detection capabilities and reinforce overall computer security. Several scholarly investigations (Fernandes, 2019; Kheraisat, 2019; Omar, 2021) have focused on the application of ML techniques to improve data quality and training, thereby augmenting the performance of intrusion detection systems. Decision trees, a prevalent ML tool for classification, test individual features independently and assign classifications following each branch split, as corroborated by studies employing algorithms like ID3 and C4.5 (Cavos, 2019; Jeong, 2016).

However, the structured nature of data collection often poses a challenge, necessitating preprocessing for unstructured data before analysis. The selection of relevant features is critical in minimizing computational costs associated with modeling and enhancing the predictive model's performance (Masdarri, 2020; Alazzam, 2020).

This paper presents a novel methodology that employs decision trees for network intrusion detection, particularly emphasizing decision-making accuracy. Feature engineering techniques were utilized to refine data quality. The study makes two notable contributions: firstly, it enhances data quality through the implementation of the entropy decision method. Secondly, it develops a classifier model using decision tree algorithms for effective network intrusion detection. For detailed insights, readers are referred to the subsequent sections.

Significance of the Study: This research concentrates on the application of decision trees to detect IoT malware, addressing the increasing vulnerability of IoT devices to such attacks. The proliferation and autonomous operation of IoT devices have rendered them susceptible to malicious exploitation. This study positions itself as a crucial intervention in understanding and countering IoT malware threats.

The experimental outcomes, utilizing the real-world dataset MaleVis, attest to the method's efficacy. With precision and recall rates of 97.23% and 95.89% respectively, and additional metrics like a specificity of 96.58%, an F1-score of 96.40%, an accuracy of 96.43%, and an average processing time per malware classification of 789 ms, the proposed methodology demonstrates superior performance over existing solutions.

The significance of this study is manifold:

1. Enhanced Detection of IoT Malware: The study introduces a more effective decision tree-based approach for detecting and classifying IoT malware.
2. Simplification of Analysis: The decision tree framework offers a straightforward and interpretable means for analyzing IoT malware, facilitating quicker response and mitigation strategies.
3. Validation in a Real-World Context: Using the publicly available MaleVis dataset, the study's findings are substantiated in a practical setting, enhancing their credibility and applicability.
4. Contributions to Cybersecurity: By leveraging machine learning techniques with a focus on IoT malware detection, this research contributes significantly to the broader field of cybersecurity, highlighting the potential of decision trees in enhancing intrusion detection systems and strengthening IoT security.

17 more pages are available in the full version of this document, which may be purchased using the "Add to Cart" button on the publisher's webpage:
www.igi-global.com/chapter/decision-trees-unleashed/336894

Related Content

Developing More Effective and Adaptive U.S. Governmental Healthcare Leaders
Amalisha Sabie Aridi (2022). *International Journal of Health Systems and Translational Medicine (pp. 1-25).*
www.irma-international.org/article/developing-more-effective-and-adaptive-us-governmental-healthcare-leaders/314579

A survey of unsupervised learning in medical image registration
(2022). *International Journal of Health Systems and Translational Medicine (pp. 0-0).*
www.irma-international.org/article//282677

EarLocalizer: A Deep-Learning-Based Ear Localization Model for Side Face Images in the Wild
Aman Kamboj, Rajneesh Raniand Aditya Nigam (2019). *Design and Implementation of Healthcare Biometric Systems (pp. 137-159).*
www.irma-international.org/chapter/earlocalizer/219958

Digital Medicine: The Quality Standpoint
Anastasius Moumtzoglou (2017). *Design, Development, and Integration of Reliable Electronic Healthcare Platforms (pp. 179-195).*
www.irma-international.org/chapter/digital-medicine/169550

Advancing IoT Security Posture K-Means Clustering for Malware Detection
Ali Dayouband Marwan Omar (2024). *Innovations, Securities, and Case Studies Across Healthcare, Business, and Technology (pp. 221-239).*
www.irma-international.org/chapter/advancing-iot-security-posture-k-means-clustering-for-malware-detection/336893