

Course Title: Research Methodology in Computer Science

Course Code: AI800

Weighting: 20%

Lecturer: Dr. Mohammed Anbar

Hand-out Date: 20 Oct 2024

Hand-in Date: 9 Nov 2024

Hand-in Method: Online Submission via E-Learning.

Assignment Report: literature Review in Botnet Detection in the Internet-of-Things Networks

Table of contents:

- Summary research paper's
- Critical Analysis:
- Connections
- Conflicts between the findings
- Synthesis Matrix for Literature Review:
- References

Summary research paper's:

Paper 1: Def-IDS: An Ensemble Defense Mechanism Against Adversarial Attacks for Deep Learning-based Network Intrusion Detection

This paper discusses the challenge of protecting network intrusion detection systems (NIDS) from adversarial attacks, which have been shown to reduce the accuracy of deep learning models by subtly manipulating input data. While adversarial vulnerabilities have been well-studied in image processing, their impact on NIDS has received less attention. To address this gap, the authors propose Def-IDS, a defense mechanism tailored for NIDS to counter both known and unknown adversarial attacks. Def-IDS uses a two-module training framework that combines multi-class generative adversarial networks (GANs) with multi-source adversarial retraining. The evaluation results demonstrated the effectiveness of Def-IDS using the CSE-CIC-IDS2018 benchmark dataset. (Wang et al., 2021)

Paper 2: SMOTE-DRNN: A Deep Learning Algorithm for Botnet Detection in the Internet-of-Things Networks

This paper addresses botnet detection in IoT networks using Deep Learning, focusing on the issue of imbalanced network traffic data. The authors propose a botnet detection algorithm that uses SMOTE (to generate additional samples for underrepresented classes) along with a Deep Recurrent Neural Network (DRNN) to learn hierarchical feature representations from balanced data. Implemented on the Bot-IoT dataset, the SMOTE-DRNN model outperformed the DRNN model, achieving high scores in Precision (99.50%), Recall (99.75%), F1-score (99.62%), AUC

(99.87%), GM (99.74%), and MCC (99.62%). This approach demonstrates improved detection performance in imbalanced data compared to other machine learning and deep learning models.(Popoola et al., 2021)

Paper 3: Botnet Detection on IoT using Ensemble Learning Technique

This study proposes an ensemble learning model for botnet detection in IoT networks by combining Artificial Neural Networks (ANN) and Decision Trees (DT) to improve accuracy and efficiency. The model achieved 100% accuracy in detecting botnets within 11.36 seconds using only 20 features, outperforming previous studies. However, the model's limitations include testing only in a lab environment, which restricts its access to diverse IoT devices and data types. Future work could enhance detection speed, test with IPv6, and develop methods to reduce dataset requirements while maintaining accuracy, addressing issues of data scarcity in machine learning for IoT security.(Rezaei, 2021)

Paper 4: A Collaborative Approach to Early Detection of IoT Botnet

This paper introduces a collaborative machine learning model for early detection of IoT botnets, moving beyond traditional full-time series data-dependent methods. The model, which achieved 99.37% accuracy on a dataset of 5,023 IoT botnet and 3,888 benign samples, allows for faster detection by relying on a diverse set of features rather than complete execution data. The model's effectiveness was demonstrated in a V-Sandbox

environment, showcasing its capability for automated, real-time botnet detection. Future work will focus on improving V-Sandbox's efficiency to further reduce detection times.(Nguyen et al., 2022)

Paper 5: Intrusion Detection System for IoT Botnet Attacks Using Deep Learning

This study presents a Deep Neural Network (DNN)-based intrusion detection system for identifying DDoS botnet attacks in IoT networks. Recognizing the challenges posed by IoT's resource constraints and network heterogeneity, the model was developed using a realistic dataset generated at UNSW Canberra's Cyber Range Lab, combining normal and attack traffic data. The proposed DNN demonstrated superior detection accuracy and precision over existing systems, achieving an accuracy of 0.94. The research underscores IoT's rapid growth and the increased risk of cyberattacks, highlighting the need for adaptable, low-cost, and efficient security solutions. Future work aims to extend detection capabilities to other IoT devices and enhance the model's effectiveness in recognizing emerging botnet threats.(Jithu P et al., 2021)

Paper 6: ELBA-IoT: An Ensemble Learning Model for Botnet Attack Detection in IoT Networks

This study proposes ELBA-IoT, an ensemble learning-based detection model designed for lightweight, anomaly-based detection of botnet attacks in IoT networks. Targeting IoT constraints in computation, storage, and communication, ELBA-IoT profiles normal and malicious behavior in network traffic using three decision-tree-based techniques: AdaBoost, RUSBoost, and bagging. Using the N-BaIoT-2021 dataset, the model

achieved high detection accuracy (99.6%) with low inference overhead (40 seconds), outperforming existing solutions. The authors suggest deploying ELBA-IoT on physical IoT devices (e.g., Raspberry Pi) for real-time detection and exploring resource concerns like energy use, memory, and communication in IoT. Future work includes creating new datasets to test the model across various IoT applications and peak times. (Abu Al-Haija & Al-Dala'ien, 2022)

Critical Analysis:

The reviewed studies offer valuable insights into botnet and intrusion detection in IoT networks, highlighting various strengths and addressing different aspects of IoT security challenges. Together, these papers explore IoT security using innovative approaches like ensemble learning, deep learning, and anomaly-based detection models, achieving effective detection accuracy and robust performance despite resource constraints.

Notably, the Def-IDS model tackles adversarial attacks by using multi-class GANs and adversarial retraining—an area overlooked in other IoT botnet detection studies. This resilience against adversarial manipulation makes Def-IDS a critical contribution to the field, as traditional models are more vulnerable to these types of attacks.

Similarly, SMOTE-DRNN excels in addressing imbalanced network traffic data through data augmentation, improving detection across minority classes—a crucial aspect in IoT network traffic analysis. This approach offers a distinct solution to class imbalance, a limitation not directly addressed by ensemble models like ELBA-IoT, which focus more on model architecture than data balancing strategies.

From a practical perspective, models like ELBA-IoT and the Collaborative Approach for Early Detection of IoT Botnet stand out for their lightweight, low-latency structures and high detection accuracy suited for real-time applications. These models demonstrate strong potential for real-world deployment within actual IoT networks due to their efficient processing times. However, the lab-based testing environments and limited dataset diversity in some studies (such as the Ensemble Learning Technique for Botnet Detection on IoT) restrict their generalizability to real-world IoT settings, where diverse network conditions are the norm.

Despite these contributions, there are some notable shortcomings across the studies. Although many models show high detection accuracy, they lack thorough testing under real-world conditions. Few studies consider the computational burden of deploying these models on constrained IoT devices. Additionally, models like Def-IDS and DNN-based Intrusion Detection System for IoT Botnet Attacks Using Deep Learning have limited real-time applicability due to higher resource demands, highlighting a trade-off between robustness and efficiency that requires further exploration.

Connections:

All six studies share a common goal of enhancing botnet and intrusion detection in IoT networks through machine learning and deep learning methods. The papers leverage different ML/DL techniques to address specific challenges, such as data imbalance, resource limitations, and the complexity of real-time detection in heterogeneous IoT environments. Across the board, there's a trend favoring ensemble and collaborative models (like ELBA-IoT and Def-IDS) and data augmentation methods (like SMOTE in SMOTE-DRNN) to increase model robustness and adaptability across various IoT scenarios.

Conflicts Between Findings:

1.Dataset Requirements: While Def-IDS and ELBA-IoT utilize ensemble learning to detect anomalous traffic adaptively, Def-IDS specifically focuses on adversarial robustness. This creates a gap in other models that prioritize accuracy without defenses against adversarial attacks. The absence of adversarial safeguards in other models may limit their effectiveness in real-world IoT environments, where adversarial threats are becoming increasingly common.

2.Data Completeness and Imbalance: SMOTE-DRNN addresses data imbalance through oversampling, enhancing its performance with underrepresented classes. In contrast, ELBA-IoT does not explicitly tackle imbalance, instead relying on a carefully selected decision-tree ensemble, achieving high detection accuracy with low computational requirements. This divergence illustrates a split between approaches that focus on data manipulation (e.g., SMOTE-DRNN) and those that emphasize model architecture (e.g., ELBA-IoT) for reliable detection in imbalanced datasets.

3.Computation vs. Real-Time Efficiency: Both SMOTE-DRNN and ELBA-IoT models exhibit high detection accuracy and low inference time; however, ELBA-IoT is specifically designed for real-time, lightweight deployment, suggesting practical advantages over models like Def-IDS and DNN-based IDS that rely on more complex architectures. The contrast in computational demands emphasizes the trade-off between model complexity and real-time efficiency across these studies.

4.Dataset Limitations: The Ensemble Learning Technique for Botnet Detection on IoT and DNN-based Intrusion Detection System for IoT Botnet Attacks highlight the importance of realistic testing environments while also revealing a limitation in dataset diversity. The lab-based datasets used in these studies, while controlled, don’t fully capture the complexity of real-world IoT networks, potentially impacting generalizability. Conversely, ELBA-IoT and Def-IDS utilize datasets more representative of various IoT applications, suggesting better real-world applicability, though models tested under highly controlled conditions may still risk overfitting.

Synthesis Matrix for Literature Review:

Study	Authors	Year	Research design	Key Findings/ Results	Methodology	Limitations
Def-IDS: An Ensemble Defense Mechanism Against Adversarial Attacks for Deep Learning-based Network Intrusion Detection	Wang et al.	2021	Experimental	Developed Def-IDS, which improves NIDS resilience to adversarial attacks by using multi-class GANs and adversarial retraining	Multi-module GANs + Adversarial retraining	Tested only on specific datasets; limited focus on types of adversarial attacks applicable to real-world IoT

SMOTE-DRNN: A Deep Learning Algorithm for Botnet Detection in the Internet-of-Things Networks	Popoola et al.	2021	Experimental	SMOTE-DRNN model addresses data imbalance in IoT traffic and achieves high accuracy (e.g., Precision: 99.5%)	SMOTE + Deep Recurrent Neural Network	Limited generalizability; tested in lab-based settings with a single dataset
Using Ensemble Learning Technique for Detecting Botnet on IoT	Rezaei	2021	Experimental	Combined ANN and DT in an ensemble, achieving 100% accuracy with only 20 features and in 11.36 seconds	Artificial Neural Networks (ANN) + Decision Tree (DT)	Tested only in controlled environments; no real-world deployment
A Collaborative Approach to Early Detection of IoT Botnet	Nguyen et al	2022	Experimental	Collaborative ML model detects IoT botnet early, achieving 99.37% accuracy while reducing dependency on full-time series data	Collaborative ML model; feature diversity rather than time-series data	Relies on simulated environments; lacks testing in varied real-world IoT networks
Intrusion Detection System for IoT Botnet Attacks Using Deep Learning	Jithu P et al.	2021	Experimental	Proposed a DNN-based IDS for DDoS detection in IoT, achieving 94% accuracy	Deep Neural Network (DNN)	Limited to DDoS attacks and lab-generated datasets; generalizability to varied botnet types is untested
ELBA-IoT: An Ensemble Learning Model for Botnet Attack Detection in IoT Networks	Abu Al-Haija & Al-Dala'ien	2022	Experiment al	ELBA-IoT using AdaBoost, RUSBoost, and Bagging techniques achieves 99.6% accuracy with low inference time	Ensemble model (AdaBoost, RUSBoost, Bagging)	Requires further testing on physical IoT devices; limited exploration of resource efficiency

References:

Wang, J., Pan, J., AlQerm, I., & Liu, Y. (2021, July). Def-ids: An ensemble defense mechanism against adversarial attacks for deep learning-based network intrusion detection. In 2021 International Conference on Computer Communications and Networks (ICCCN) (pp. 1-9). IEEE.

Popoola, S. I., Adebisi, B., Ande, R., Hammoudeh, M., Anoh, K., & Atayero, A. A. (2021). smote-drnn: A deep learning algorithm for botnet detection in the internet-of-things networks. *Sensors*, 21(9), 2985.

Rezaei, A. (2021). Using ensemble learning technique for detecting botnet on IoT. *SN Computer Science*, 2(3), 148.

Nguyen, G. L., Dumba, B., Ngo, Q. D., Le, H. V., & Nguyen, T. N. (2022). A collaborative approach to early detection of IoT Botnet. *Computers & Electrical Engineering*, 97, 107525.

Shareena, J., Ramdas, A., & AP, H. (2021). Intrusion detection system for iot botnet attacks using deep learning. *SN Computer Science*, 2(3), 1-8.

Abu Al-Haija, Q., & Al-Dala'ien, M. A. (2022). ELBA-IoT: An ensemble learning model for botnet attack detection in IoT networks. *Journal of Sensor and Actuator Networks*, 11(1), 18.