DAC-01-2581\_PRS2023\_ITS

**UNIVERSITAS NEGERI SEMARANG**

**DAC-01-0074**

**Statstronomers**

**DAC 2023**

**DAC 2023**

**DAC 2023**

**CHAPTER I: Introduction**

The rapid advancement of information technology and the widespread proliferation of telecommunications networks have fundamentally reshaped how we communicate, access information, and conduct global business [1]. Within the context of Indonesia's telecommunications sector, a significant challenge lies in the enduring threat of network attacks. These encompass a spectrum of deliberate actions aimed at undermining the security and functionality of computer networks. These attacks can manifest in various forms, such as Distributed Denial of Service (DDoS) attacks, infiltration by malware, or unauthorized access to sensitive data [2]. The consequences of these attacks may involve disruption of communication services, theft of customer data, or even compromise of the underlying network infrastructure.

However, the research delves deeply into the pivotal endeavor of identifying these network attacks. This intricate process necessitates the utilization of specialized tools and methodologies, frequently revolving around the implementation of Intrusion Detection Systems (IDS) or Network Intrusion Detection Systems (NIDS) [3], [4]. The inherent challenge lies in the creation of mechanisms capable of promptly recognizing and reacting to unauthorized activities within the network in real-time, thereby minimizing potential risks and damage.

Moreover, this research introduces an innovative approach to address these challenges – the utilization of multi-layer classification. This method aims to provide a more comprehensive understanding of network activities by starting with a binary classification, distinguishing between anomalies and normal behavior [2]. For detected anomalies, it goes a step further, utilized multiclass techniques to precisely identify the specific type of attack. By way of explanation, when an anomaly is initially identified as a Denial-of-Service (DoS) attack, a secondary classification process ensues to determine if it fits into subcategories like DoS, Neptune, or Smurf attacks. Similarly, for probe attacks, the system employs multiclass classification to determine if they can be categorized into subtypes such as Nmap, Portscan, Satan, or Ipsweep attacks [3], [5], [6].

This approach not only enhances our ability to detect network attacks but also provides a more detailed and nuanced view of the threat landscape. It empowers us to respond more effectively by tailoring our actions to specific attack types and subcategories, ultimately bolstering our network security and resilience.

This multi-layered approach holds several advantages, particularly in feature selection. It begins with the outermost layer, distinguishing anomalies from normal activities, and gradually delves into the specifics of attack types [7]. This process aids in identifying which attributes are most relevant for accurate classification, enhancing the overall effectiveness of the system [6]. In essence, the research sets out to contribute by not only detecting network attacks but also by providing a comprehensive understanding of the attack landscape, enabling more informed and targeted cybersecurity efforts in Indonesia's telecommunications sector.

**CHAPTER II: Theoretical Framework**

1. Intrusion Detection Systems

The Intrusion Detection System (IDS) stands as a highly efficient security reinforcement tool, crucial for detecting and safeguarding against cyber-attacks within any network or host [2]. Its fundamental role is to identify and respond to suspicious activities, serving as a proactive measure to protect the network from potential threats and reduce the economic losses that can result from security breaches [8]. This capability makes IDS an integral component in ensuring the security and resilience of digital infrastructures, reinforcing the defense mechanisms against a wide range of cyber threats.

The Intrusion Detection System (IDS) serves as a vital security measure against network attacks and can be classified based on its deployment location: Network-based IDS (NIDS) or Host-based IDS (HIDS) [4].

HIDS operates on a single device within the network, monitoring that device's activities for signs of suspicious behavior [4], [9]. However, HIDS can strain the resources of the host device and is better suited for protecting individual devices, making it less efficient for large-scale networks [4].

On the other hand, NIDS monitors the entire network and identifies potential threats to network devices. A typical NIDS operates in three key phases: monitoring, detection, and response [4]. During the monitoring phase, it collects statistical network features like packet counts and connections. These features are then used in the classification phase, where Machine Learning (ML) algorithms assess whether the observed characteristics indicate a potential network attack [4]. Based on the classification results, the system initiates suitable defensive actions during the response phase.

1. Machine Learning techniques for NIDS

Machine learning techniques play a pivotal role in fortifying the security of computer networks through the development of Network Intrusion Detection Systems (NIDS) within the cybersecurity domain [2]. Extensive research efforts have been dedicated to exploring various machine learning models, broadly categorized into traditional and advanced approaches, to enhance the effectiveness of NIDS [10].

In the realm of traditional machine learning algorithms, three prominent contenders have risen to prominence: K-Nearest Neighbors (KNN), Support Vector Machines (SVM), and Random Forest [10]. These algorithms are highly recognized for their proficiency in addressing the fundamental challenges of intrusion detection, excelling in classification tasks and feature selection [11]. Consequently, they serve as invaluable tools for constructing effective NIDSs.

In contrast, recent research endeavors have been focused on advanced machine learning techniques, including Multilayer Perceptrons (MLP), Autoencoders, Gradient Boosting, CatBoost, and XGBoost [10]. These cutting-edge models stand out due to their exceptional ability to identify intricate and subtle patterns within network data [11]. As a result, they make substantial contributions to the development of NIDSs with enhanced capabilities, enabling the detection of even the most sophisticated and rapidly evolving network threats. These advanced techniques not only bolster NIDSs' accuracy but also enhance their adaptability to the rapidly changing threat landscape, ensuring the security and resilience of computer networks [11].

The integration of machine learning techniques, spanning both traditional and advanced approaches, is instrumental in steering the progression of NIDSs. These techniques drive the continuous evolution of NIDS, equipping them with the versatility and agility required to tackle the multifaceted challenges posed by the ever-evolving threat landscape within the cybersecurity domain. NIDS, serving as the first line of defense in safeguarding computer networks, heavily relies on these machine learning approaches to enhance its detection and response capabilities [2]. Traditional algorithms ensure that NIDS can accurately distinguish between normal network behavior and potentially malicious activities [10], while advanced methods empower NIDSs to detect even the most sophisticated and rapidly evolving threats [11]. In essence, this integration empowers NIDSs to remain at the forefront of network security, contributing significantly to the security and resilience of computer networks.

**CHAPTER III: Analytical Steps**

1. Data Preparation and Data Preprocessing

Data preparation and preprocessing serve as the initial steps in the analytical process. This phase involves collecting and organizing the data required for the analysis. Raw data is cleaned, transformed, and made ready for further exploration and modeling. It encompasses tasks such as data cleaning to handle missing or erroneous values, data transformation to ensure uniform formats, and data integration to combine information from multiple sources. Proper data preparation and preprocessing lay the foundation for accurate and meaningful insights in subsequent stages.

1. Data Exploration

Data exploration is a critical step to gain a comprehensive understanding of the dataset. It involves the use of descriptive statistics, data visualization, and exploratory data analysis techniques. The primary objective is to uncover patterns, relationships, and potential outliers within the data. Exploratory data analysis helps in identifying trends, correlations, and insights that inform subsequent modeling decisions. It plays a crucial role in feature selection, which is pivotal for the effectiveness of machine learning models.

1. Mechine Learning Model Selection

Machine learning model selection is a pivotal decision in the analytical process. It entails choosing the most suitable algorithms and models based on the nature of the problem, the dataset, and the desired outcomes. This step involves experimenting with various machine learning techniques, including traditional ones like K-Nearest Neighbors (KNN), Support Vector Machines (SVM), and Random Forest, as well as advanced models like Multilayer Perceptrons (MLP), Autoencoders, Gradient Boosting, CatBoost, and XGBoost. The goal is to identify the models that exhibit the highest predictive accuracy and are well-aligned with the specific requirements of the intrusion detection task.

1. Multi-Layer Classification

Multi-layer classification is a distinctive approach employed in this research. It involves a hierarchical classification process that starts with a binary classification to distinguish between anomalies and normal network behavior. For detected anomalies, it delves deeper into multiclass classification techniques to precisely identify the specific type of network attack. This multilayered approach allows for a nuanced understanding of network activities and facilitates more targeted responses to different attack types and subcategories.

1. Evaluation and Response

The final phase involves the evaluation of the selected machine learning models' performance. Various evaluation metrics are employed to assess their accuracy, precision, recall, F1-score, and other relevant measures. Additionally, the research emphasizes the importance of response mechanisms. When a network attack is detected, the system must initiate appropriate defensive actions promptly to mitigate potential risks and minimize damage. The effectiveness of the response mechanisms is a critical aspect of ensuring network security and resilience.

**CHAPTER IV: Analysis of Results**

(font 12) Lorem ipsum dolor sit amet, consectetur adipiscing elit. Nullam eget urna nec arcu interdum ullamcorper [3]. Sed rhoncus sapien at justo vehicula, id interdum nisi ullamcorper. Quisque non tellus in neque vulputate dictum sit amet id justo. Sed et volutpat ipsum. Sed vel nunc eu metus blandit dictum. Vestibulum ut tincidunt ligula, in facilisis tortor [4]. Suspendisse ultricies semper justo, eget ultrices dolor pellentesque nec. In hac habitasse platea dictumst. Nullam vel magna urna. Sed tristique, arcu a bibendum lacinia, urna sapien iaculis libero, ut venenatis ligula purus id massa.. These results can serve as a reference to determine in the clustering which provinces have high and low stunting rates.

A graph with different colored bars

Description automatically generated

**(Font 12) Figure 4.1.** Barchart diagram

Nullam eget urna nec arcu interdum ullamcorper. Sed rhoncus sapien at justo vehicula, id interdum nisi ullamcorper. Quisque non tellus in neque vulputate dictum sit amet id justo. Sed et volutpat ipsum. Sed vel nunc eu metus blandit dictum.

Based on **Attachment 1,** it can be observed for provinces from the highest to the lowest average stunting rates. These results can serve as a reference to determine in the clustering which provinces have high and low stunting rates.

**CHAPTER V: Conclusion and Recommendation**

(Font 12) The conclusion should be concise and clear

# REFERENCES

[1] A. Li and S. Yi, “Intelligent Intrusion Detection Method of Industrial Internet of Things Based on CNN-BiLSTM,” *Security and Communication Networks*, vol. 2022, 2022, doi: 10.1155/2022/5448647.

[2] P. Toupas, D. Chamou, K. M. Giannoutakis, A. Drosou, and D. Tzovaras, “An intrusion detection system for multi-class classification based on deep neural networks,” in *Proceedings - 18th IEEE International Conference on Machine Learning and Applications, ICMLA 2019*, Institute of Electrical and Electronics Engineers Inc., Dec. 2019, pp. 1253–1258. doi: 10.1109/ICMLA.2019.00206.

[3] S. Mohammadi, H. Mirvaziri, M. Ghazizadeh-Ahsaee, and H. Karimipour, “Cyber intrusion detection by combined feature selection algorithm,” *Journal of Information Security and Applications*, vol. 44, pp. 80–88, Feb. 2019, doi: 10.1016/j.jisa.2018.11.007.

[4] K. He, D. D. Kim, and M. R. Asghar, “Adversarial Machine Learning for Network Intrusion Detection Systems: A Comprehensive Survey,” *IEEE Communications Surveys and Tutorials*, vol. 25, no. 1, pp. 538–566, 2023, doi: 10.1109/COMST.2022.3233793.

[5] C. Ambedkar and V. Kishore Babu, “Detection of Probe Attacks Using Machine Learning Techniques,” 2015. [Online]. Available: www.arcjournals.org

[6] S. Norwahidayah, Noraniah, N. Farahah, A. Amirah, N. Liyana, and N. Suhana, “Performances of Artificial Neural Network (ANN) and Particle Swarm Optimization (PSO) Using KDD Cup ’99 Dataset in Intrusion Detection System (IDS),” in *Journal of Physics: Conference Series*, IOP Publishing Ltd, Jun. 2021. doi: 10.1088/1742-6596/1874/1/012061.

[7] A. Halbouni, T. S. Gunawan, M. H. Habaebi, M. Halbouni, M. Kartiwi, and R. Ahmad, “CNN-LSTM: Hybrid Deep Neural Network for Network Intrusion Detection System,” *IEEE Access*, vol. 10, pp. 99837–99849, 2022, doi: 10.1109/ACCESS.2022.3206425.

[8] Y. Zhang, X. Chen, L. Jin, X. Wang, and D. Guo, “Network Intrusion Detection: Based on Deep Hierarchical Network and Original Flow Data,” *IEEE Access*, vol. 7, pp. 37004–37016, 2019, doi: 10.1109/ACCESS.2019.2905041.

[9] A. Khraisat, I. Gondal, P. Vamplew, and J. Kamruzzaman, “Survey of intrusion detection systems: techniques, datasets and challenges,” *Cybersecurity*, vol. 2, no. 1, Dec. 2019, doi: 10.1186/s42400-019-0038-7.

[10] I. H. Sarker, “Machine Learning: Algorithms, Real-World Applications and Research Directions,” *SN Computer Science*, vol. 2, no. 3. Springer, May 01, 2021. doi: 10.1007/s42979-021-00592-x.

[11] P. Nancy, S. Muthurajkumar, S. Ganapathy, S. V. N. Santhosh Kumar, M. Selvi, and K. Arputharaj, “Intrusion detection using dynamic feature selection and fuzzy temporal decision tree classification for wireless sensor networks,” *IET Communications*, vol. 14, no. 5, pp. 888–895, Mar. 2020, doi: 10.1049/iet-com.2019.0172.

**ATTACHMENT**

Attachment 1: Diagram

A black background with white text

Description automatically generated

Attachment 2: Merging Logic

|  |
| --- |
| Load binary prediction save as Y |
| if Y == 1:  1. hapus kolom Y  2. modul prediksi multi-class attack  3. new\_Y  if new\_Y == 1:  1. hapus kolom new\_Y  2. model prediksi toa  3. new\_Y2 (ouput: dos, smurf, neptune)  elif new\_Y == 2:  1. hapus kolom new\_Y  2. model prediksi toa  3. new\_Y2 (ouput: nmap, ipsweep, portsweep, satan)  else:  0 -> normal  else -> normal |

Attachment 3:

Attachment 4:

**SYNTAX**