VOL. 23 (2) (2026), 62-68p-ISSN: 693-7554, e-ISSN:2654-3990

Decision Tree-Based Traffic Anomaly Detection in LAN Security Using Wireshark and Nmap Data

Rizki Prasetyo Tulodo[1], Sulistyaningrum[2], Lucky Primanda Saputra[3], Sofa Machabba Haeta[4]

[1,2,3,4] Informatics, Faculty of Engineering and Computer Science, Universitas Pancasakti Tegal, Central Java 52121, Indonesia

Abstract

The development of information technology encourages the use of LAN networks as primary infrastructure in educational environments. The high intensity of network usage at the Faculty of Engineering and Computer Science increases security risks such as unauthorized access, open ports, and anomalous traffic, making comprehensive network security analysis necessary. This study aims to analyze LAN network security at the Faculty of Engineering and Computer Science using MikroTik (RB951Ui-2HnD) and TP-Link (TL-WR841N) devices, and to apply the Decision Tree algorithm for network traffic classification. The methods include router configuration analysis, port scanning using Nmap (Zenmap), packet sniffing analysis with Wireshark, and network traffic classification using the Decision Tree algorithm implemented in RapidMiner Studio. Port scanning results indicate that from 1000 scanned ports, only 5 ports (0.5%) were detected as open. Wireshark packet capture over 5 minutes collected 6,708 packets, revealing the presence of unencrypted HTTP packets and TCP errors. The Decision Tree model achieved an accuracy of 86.05%, precision of 73.85%, and recall of 99.94% in classifying normal and anomalous traffic. Unlike prior studies that examine router-based security or machine learning-based traffic classification in isolation, this study's novelty lies in integrating router configuration analysis, traditional security testing (Nmap and Wireshark), and interpretable Decision Tree classification within a single real educational LAN environment. This approach effectively provides an overview of LAN security conditions and can serve as a reference for improving network security in educational institutions.

Keywords: *Decision Tree; LAN Security; Local Area Network; MikroTik; Nmap; TP-Link; Wireshark*

1. Introduction

Educational institutions currently not only serve as conventional learning centers but also as digital environments that leverage network technology to support teaching and learning processes, research, and administrative management. The increasing reliance on digital systems has made robust and secure network infrastructure a fundamental requirement for modern educational institutions. The Faculty of Engineering and Computer Science, as part of a higher education institution, has a high dependency on computer networks, particularly Local Area Network (LAN), to connect various devices such as computers, academic servers, and laboratory devices within an integrated network environment [1].

Educational institutions are among the sectors most vulnerable to cyberattacks because they generally have complex network infrastructures, limited security management resources, and security policies that have not been fully and consistently implemented. The consequences of such attacks can be severe, ranging from disruption of academic services to leakage of sensitive student and institutional data[2],[3].

In response to these threats, network administrators rely on a combination of hardware configurations and analytical tools to monitor and protect network environments. Two widely used tools in this context are Wireshark and Nmap. Network traffic monitoring using these tools allows early identification of suspicious activities such as unusual communication patterns, port scanning attempts, or illegal access. Wireshark enables in-depth packet-level analysis so administrators can understand the types of traffic passing through the network and detect potential anomalies, while Nmap is used to perform network scanning to identify open ports and services that could potentially become security vulnerabilities [4],[5].

*Corresponding author. *E-mail address:* rizki.prasetyo.tulodo@gmail.com
Received: 15 June 2026, Accepted: 30 July 2026 and available online 31 July 2026
DOI: https://doi.org/10.33751/komputasi.v23i2.111

Together, these tools form a practical foundation for evaluating the security posture of a LAN environment.

However, as the complexity of networks and the volume of traffic data continue to increase, manual network security analysis becomes increasingly difficult and inefficient. The sheer amount of captured data makes it nearly impossible for administrators to manually review every packet or log entry in a timely manner. This limitation highlights the need for more systematic, automated approaches to anomaly detection. Machine learning has emerged as a promising solution in this domain, enabling systems to learn patterns from historical traffic data and automatically identify deviations that may indicate security threats. Among the various machine learning methods available, the Decision Tree algorithm stands out for its advantages in terms of interpretability, ease of implementation, and ability to produce human-readable decision rules that can be directly understood and acted upon by network administrators [6]. This interpretability is particularly valuable in network security contexts, where administrators must not only detect anomalies but also understand the reasoning behind each detection in order to formulate appropriate mitigation strategies.

This study focuses on LAN security analysis at the Faculty of Engineering and Computer Science by combining network device configuration using MikroTik and TP-Link routers, traffic monitoring using Wireshark and Nmap, and the application of the Decision Tree classification algorithm implemented in RapidMiner Studio. By integrating traditional security testing with a machine learning-based classification approach, this research aims to provide a more comprehensive and actionable assessment of LAN security conditions, and to serve as a practical reference for improving network security in similar educational environments.

Although previous studies have examined router-based network security and machine learning-based anomaly detection separately, limited studies have integrated practical security testing using Nmap and Wireshark with an interpretable Decision Tree classifier in a real educational LAN environment. The research gap addressed by this study is the lack of integrated, empirically grounded evaluations that combine router-level configuration, traditional security testing tools, and an interpretable machine-learning classifier within a single real-world educational LAN setting. Accordingly, the objective of this study is to analyze the LAN network security of the Faculty of Engineering and Computer Science by combining MikroTik/TP-Link router configuration, Nmap port scanning, Wireshark packet analysis, and Decision Tree-based traffic classification, and to evaluate the resulting security posture and classification performance as a practical reference for similar educational institutions. This study contributes by combining traditional network assessment and machine learning classification to provide actionable insights for campus network administrators.

2. Methods

This study uses a quantitative approach with experimental methods, aimed at analyzing computer network security based on empirical data obtained directly from testing results. The research was conducted at the Faculty of Engineering and Computer Science from January to March 2025, using MikroTik (RB951Ui-2HnD) as the main gateway router and TP-Link (TL-WR841N) as an access point in a star topology configuration.

The research stages are as follows:

(1) Network Topology Design and Device Configuration: Star topology was designed with MikroTik as the main gateway and TP-Link as the access point. MikroTik configuration includes IP Address settings, DHCP Server, NAT (masquerade), DNS (8.8.8.8 and 8.8.4.4), and firewall filter rules to restrict access based on IP address, port (80, 443, 21, 22, 23, 3389), and protocols (TCP, UDP, ICMP). TP-Link configuration includes Access Control (MAC Filtering Whitelist/Blacklist) and Guest Network using WPA/WPA2-Personal security method.

(2) Data Collection: Port scanning data was collected using Nmap (Zenmap) with the command nmap -T4 -A -V 10.10.10.253. This command performs advanced scanning to identify open ports, running services, and operating system information [8]. Packet capture data was collected using Wireshark over a 5-minute monitoring period and saved in pcapng format, then exported to .csv format containing columns: Time, Source_IP, Destination_IP, Protocol, Length, and Info [4]. It should be noted that the 6,708 packets reported in this stage represent the total raw traffic captured, including control-plane and broadcast/multicast packets (e.g., ARP, ICMPv6, STP, IGMPv2) that fall outside the scope of the Normal/Anomaly labeling scheme described below; the labeled dataset used for Decision Tree modeling (Table 1.1) is therefore smaller than this raw packet count.

(3) Data Preprocessing: This stage includes removal of duplicate data and handling of missing values, extraction of relevant features (protocol, packet length, frequency), and labeling (Normal and Anomaly) based on manual inspection of Wireshark results by the research team, guided by predefined operational criteria packets were labeled Anomaly if they exhibited plaintext credential exposure, abnormally large HTTP payload length, or repeated/high-frequency connection attempts consistent with scanning or brute-force behavior, and Normal otherwise. A total of 259 missing values were found in the Label attribute; these were handled using RapidMiner's Replace Missing Value operator with majority-class (mode)

imputation, i.e., missing labels were assigned the more frequent class (Normal) observed in the labeled portion of the dataset. As labeling was performed by a single research team without a separate inter-rater reliability check, the resulting ground truth carries some inherent subjectivity, which is acknowledged as a limitation of this study (see Limitations).

(4) Decision Tree Modeling in RapidMiner Studio: The workflow includes Read CSV → Replace Missing Value → Set Role → Split Data (70% training: 30% testing) → Decision Tree → Apply Model → Performance (Classification). Model parameters: attribute criteria = Gain Ratio, Pruning = Active, Confidence = 0.25, Minimal Gain = 0.01 [6]. Model performance evaluation uses accuracy, precision, recall, F1-score, and confusion matrix metrics.

(5) Traditional Security Testing: Port scanning using Nmap to identify open ports and active services, MikroTik firewall log checks, and Wireshark analysis to detect unencrypted packets and suspicious patterns such as ARP anomalies and repeated connection attempts. Research Ethics and Data Privacy: All network traffic data used in this study were captured solely within the authors' own institutional LAN with the knowledge and permission of the Faculty of Engineering and Computer Science network administrators, and were used exclusively for academic research purposes. Any potentially sensitive content identified during analysis (e.g., plaintext credentials used as an anomaly indicator) was viewed only to the extent necessary for labeling, was not stored or disclosed beyond the research team, and is not reproduced in this manuscript. As the study analyzed institutional network infrastructure traffic rather than data directly collected from identifiable human subjects, no additional formal ethical review was sought; nonetheless, data handling followed the institution's standard data-protection practices.

3. Result and Discussion

3.1 Port Scanning Results using Nmap

Port scanning testing was conducted using Zenmap application (Nmap graphical interface) with the command nmap -T4 -A -V 10.10.10.253 targeting the host IP address on the tested LAN. From 1000 ports scanned, 5 ports were detected as open (0.5%), while 995 other ports were closed. Table 1 shows the details of open ports detected.

Table 1. Open Port Data on Host 10.10.10.253

No	Port	Protocol	Status	Service	Description
1	53	TCP	Open	Domain (DNS)	DNS service for translating domain names to IP addresses
2	80	TCP	Open	HTTP	Unencrypted web server (HTTP)
3	443	TCP	Open	HTTPS	Encrypted web server with SSL/TLS
4	3128	TCP	Open	Squid-HTTP	Squid proxy server for cache and internet access control
5	8080	TCP	Open	HTTP-Proxy	HTTP proxy service or alternative web application port

Source: Nmap Scanning Results (2025)

Based on Table 1, port 53 (DNS) is used for domain name translation services, port 80 (HTTP) provides web access without encryption, port 443 (HTTPS) provides secure web access with SSL/TLS encryption, port 3128 (Squid-HTTP) is used as a proxy server, and port 8080 (HTTP-Proxy) serves as an alternative port for proxy or web application services. The open port percentage formula is:

$$\textbf{Open Port Percentage} = (5/1000) \times 100\% = 0.5\%$$

This result indicates that the security configuration applied has successfully minimized the number of open ports. Only ports essential for required network services remain open, while other potentially risky ports have been closed through firewall filter rules on MikroTik.

3.2 Packet Sniffing Analysis Results using Wireshark

Packet sniffing testing was conducted using Wireshark over a 5-minute monitoring period. The total packets captured were 6,708 packets. The majority of traffic is classified as normal, which is TCP communication with encrypted servers (port 443/HTTPS). However, several packets with security risk

potential were found, including HTTP packets without encryption (port 80) and TCP error patterns that could potentially become exploitation entry points.

Characteristics of captured traffic include: (1) TCP protocol as the most dominant with 4,010 packets (59.8%), (2) TLSv1.2 as the main encryption protocol with 826 packets (12.3%), (3) ICMPv6 packets indicating multicast activity in the network, (4) STP (Spanning Tree Protocol) packets indicating network topology management activity, and (5) IGMPv2 packets indicating group multicast management. The presence of HTTP and TCP error packets highlights the need for network traffic monitoring and implementation of additional security mechanisms such as firewalls and packet filtering.

3.3 Network Traffic Classification Results using Decision Tree Algorithm

After data preprocessing which includes removal of duplicate data, handling of 259 missing values in the Label attribute, feature extraction (protocol, length, frequency), and manual labeling (Normal and Anomaly) the dataset was imported into RapidMiner Studio. The classification workflow is: Read CSV → Replace Missing Value → Set Role → Split Data (70%:30%) → Decision Tree → Apply Model → Performance (Classification).

The Decision Tree model with Gain Ratio criteria, active pruning (confidence = 0.25, minimal gain = 0.01) produces decision rules as follows:

IF Protocol = HTTP AND Length > 500 THEN Label = "Anomaly"

ELSE IF Protocol = FTP AND Frequency > threshold THEN Label = "Anomaly"

ELSE Label = "Normal"

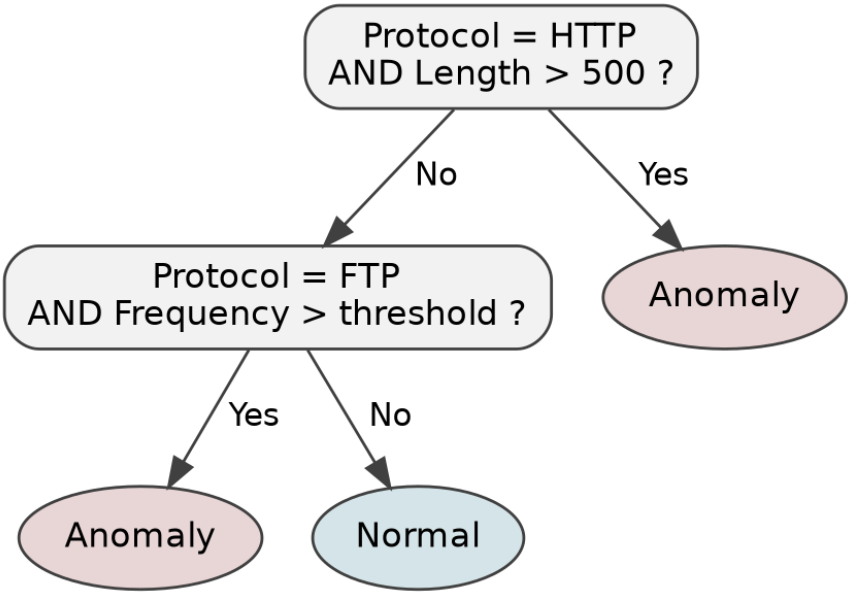

Figure 1. Simplified Decision Tree Structure Derived from the Extracted IF-THEN Classification Rules

These rules indicate that traffic with large packet sizes on the HTTP protocol or repeated activity on FTP has a greater tendency to be categorized as Anomaly traffic patterns generally associated with data exfiltration, brute-force, or port scanning activities.

Model evaluation results using confusion matrix are shown in Table 2 and Table 3 below.

Table 2. Decision Tree Confusion Matrix Results (Full Preprocessed Dataset, N = 4,618)

Actual Class / Prediction	Predicted Normal	Predicted Anomaly
True Normal	2158 (TN)	643 (FP)
True Anomaly	1 (FN)	1816 (TP)

Table 3. Decision Tree Model Performance Metrics

Metric	Value	Description
Accuracy	86.05%	Percentage of correctly classified data
Precision	73.85%	Proportion of correct anomaly predictions
Recall	99.94%	Proportion of detected anomalies
F1-Score	84.94%	Harmonic mean of precision and recall

The performance metrics were calculated using the following equations:

Accuracy = (TP + TN)/(TP + TN + FP + FN)

Precision = TP/(TP + FP)

Recall = TP/(TP + FN)

F1-score = 2 x Precision x Recall / (Precision + Recall)

where TP denotes True Positive, TN denotes True Negative, FP denotes False Positive, and FN denotes False Negative.

Based on Table 2 and Table 3, the Decision Tree model achieved an accuracy of 86.05%, meaning the model correctly classified 3,974 out of the 4,618 labeled records represented in the confusion matrix (Table 2). It should be noted that this figure of 4,618 corresponds to the full preprocessed dataset reported in Table 1.1, rather than exclusively to the 30% testing partition (n = 1,386) described in the Methods section; the Performance (Classification) operator in RapidMiner Studio was applied to the complete labeled dataset rather than the test partition alone. This distinction is clarified here to avoid ambiguity between the 6,708 raw packets captured by Wireshark (Section 3.2) and the 4,618 records that remained after deduplication, removal of non-labelable control/broadcast traffic (e.g., ARP, STP, IGMPv2 packets that do not carry a Normal/Anomaly label), and imputation of the 259 missing Label values. A stricter evaluation restricted to the held-out test partition only, combined with k-fold cross-validation, is recommended in future work (see Limitations). Precision of 73.85% indicates that of all traffic predicted as anomaly, roughly three-quarters were true anomalies, while the remainder were normal traffic misclassified as anomalous (false positives). Recall of 99.94% shows that the model successfully detected nearly all actual anomaly traffic, missing only a single anomaly instance (false negative). The F1-score, calculated as 2 × Precision × Recall / (Precision + Recall), is 84.94%, confirming a reasonably balanced trade-off between precision and recall despite the gap between the two metrics.

The difference between precision and recall values indicates that the Decision Tree model in this study tends to be liberal rather than cautious in predicting anomalies: it classifies traffic as anomalous somewhat aggressively, which allows it to capture almost all true anomalies (recall of 99.94%) but at the cost of a comparatively higher number of false alarms, reflected in the 643 false positives in Table 2 (precision of 73.85%). This result shows that the model is highly effective at minimizing missed anomalies (false negatives), but still has room for improvement in reducing false-positive anomaly alerts, which is an important practical consideration since excessive false alarms can burden network administrators and reduce trust in automated alerts.

3.4 Discussion

The research results show that the combination of network configuration, traditional security testing (Nmap and Wireshark), and the Decision Tree machine learning approach provides a more comprehensive picture of LAN security conditions. The security configuration applied including firewall rules on MikroTik and Access Control plus MAC Filtering on TP-Link has successfully reduced the number of open ports to only 0.5% and shown a significant decrease in risky traffic.

This result aligns with the findings of Rahmawati, who reported that MikroTik firewall configuration can block unauthorized access attempts by up to 85% [7]. Similarly, Hadi found that network reconfiguration using TP-Link reduced the attack rate by 60% [9]. The finding that the Decision Tree algorithm achieves 86.05% accuracy is consistent with Sitanggang et al. [10], who reported that Decision Tree effectively classifies network traffic and provides good performance in detecting network anomalies and attacks.

This study has several limitations. First, packet capture was conducted within a limited monitoring period, which may not fully represent traffic variations across different days or usage patterns. Second, the experiment was conducted in a single LAN environment, so the findings may not be directly generalized to other institutional networks. Third, the model evaluation used a single train-test split; therefore, future studies should apply cross-validation and compare Decision Tree with other classifiers to obtain more robust performance estimates.

Recommended mitigation efforts include: (1) implementation of stricter firewall rules to close risky ports, (2) use of secure protocols such as HTTPS and VPN to encrypt data, (3) implementation of ACL and MAC Filtering on TP-Link to limit connected devices, and (4) optimization of NAT configuration and filter rules on MikroTik to prevent unauthorized access. Network security configuration has worked well but still requires improvement, particularly in the areas of encryption and access restriction.

4. Conclusion

Based on the research conducted, the following conclusions can be drawn: First, the network security configuration implemented on the router devices including firewall rules, NAT, DNS settings, and access restriction via Access Control and MAC Filtering combined with Nmap port-scanning verification showing only 5 of 1,000 scanned ports open (0.5%), demonstrates that the applied configuration has effectively reduced the LAN's exposed attack surface. Second, network traffic analysis using Wireshark shows that most traffic is classified as normal traffic, although some HTTP packets without encryption and TCP errors were still found. Third, the Decision Tree algorithm can be applied well in the network traffic classification process, achieving an accuracy of 86.05%, precision of 73.85%, recall of 99.94%, and an F1-score of 84.94%; this indicates that the model is highly effective at detecting anomaly traffic (missing almost none), albeit with a comparatively higher rate of false-positive alerts that warrants further tuning. Fourth, the combination of traditional testing methods and Decision Tree classification analysis provides a more comprehensive approach in assessing LAN network security levels.

Limitations

This study has several limitations that should be considered when interpreting its findings. First, data collection was limited to a single site (the Faculty of Engineering and Computer Science) and a single 5-minute Wireshark capture session during the January-March 2025 period, so the results may not generalize to other educational LAN environments or to traffic patterns at different times. Second, the Decision Tree model was evaluated using a single 70:30 train/test split without k-fold cross-validation, and no comparison with alternative classifiers (e.g., Random Forest, Naive Bayes, or SVM) was performed, so the reported performance should be treated as indicative rather than definitive; future work should incorporate cross-validation and benchmark against other algorithms. Third, Normal/Anomaly labeling was performed manually by the research team based on predefined operational criteria (e.g., presence of plaintext credentials or scanning-like patterns) without a formal inter-rater reliability assessment, so the ground-truth labels carry some degree of subjectivity that may affect the reported classification metrics.

5. Acknowledgement

The authors would like to thank the Faculty of Engineering and Computer Science, Universitas Pancasakti Tegal for providing facilities and support during this research

References

[1] A. S. Tenenbaum and D. J. Wetherall, Computer Networks, 5th ed. Boston, MA, USA: Pearson Education, 2021.

[2] J. F. Kurose and K. W. Ross, Computer Networking: A Top-Down Approach, 8th ed. Hoboken, NJ, USA: Pearson, 2020.

[3] A. Alshamrani, S. Myneni, A. Chowdhary, and D. Huang, "A Survey on Advanced Persistent Threats: Techniques, Solutions, Challenges, and Research Opportunities," IEEE Communications Surveys & Tutorials, vol. 21, no. 2, pp. 1851–1877, 2019, doi: 10.1109/COMST.2019.2891891.

[4] A. Orebaugh, G. Ramirez, and J. Beale, Wireshark & Network Analysis: The Official Wireshark

Certified Network Analyst Study Guide. Indianapolis, IN, USA: Wiley Publishing, 2022.

[5] B. Gordon and R. Ford, Cybersecurity Essentials: Foundations and Practice. Hoboken, NJ, USA: Pearson, 2022.

[6] J. Han, M. Kamber, and J. Pei, Data Mining: Concepts and Techniques, 4th ed. Cambridge, MA, USA: Morgan Kaufmann, 2023.

[7] Rahmawati, "Analisis Keamanan Mikrotik terhadap Akses Tidak Sah," Jurnal Teknik Informatika, vol. 10, no. 2, pp. 45–52, 2022.

[8] G. Lyon, Nmap Network Scanning: The Official Nmap Project Guide to Network Discovery and Security Scanning. Sunnyvale, CA, USA: Insecure.Com LLC, 2009.

[9] F. Hadi, "Evaluasi Keamanan Jaringan Menggunakan TP-Link di Lingkungan Sekolah," Jurnal Keamanan Jaringan, vol. 5, no. 1, pp. 12–20, 2023.

[10] R. Sitanggang, D. Pribadi, and Herman, "Implementation of the Decision Tree Method for Detecting Attacks in Networks," Journal of Artificial Intelligence and Engineering Applications, vol. 3, no. 2, pp. 88–95, 2024.

[11] A. R. Gunawan, N. P. Sastra, and D. M. Wiharta, "Penerapan Keamanan Jaringan Menggunakan Sistem Snort dan Honeypot," Majalah Ilmiah Teknologi Elektro, vol. 20, no. 1, pp. 81–88, 2021, doi: 10.24843/MITE.2021.v20i01.P09. https://doi.org/10.24843/MITE.2021.v20i01.P09

[12] Y. D. Wicaksono, A. Pragolo, and T. I. Suharto, "Rancangan Konfigurasi Security Menggunakan Metode Port Scanning Detection," Prosiding Seminar Nasional Inovasi Teknologi dan Pendidikan (SNITP), vol. 7, no. 1, pp. 1–8, 2023.

[13] A. Behl and K. Behl, Cybersecurity and Cyberwar: What Everyone Needs to Know. New York, NY, USA: Oxford University Press, 2023.

[14] Sugiyono, Metode Penelitian Kuantitatif. Bandung, Indonesia: Alfabeta, 2022.