|Jurnal Jaringan Komputer dan Keamanan |JJKK|Vol. 6| No. 1|Feb|Year 2025|
This is an Open Access article, published by Institute of Information Technology and Social Science (IITSS), Indonesia

Analysis of Mikrotik Server Log Data Mining for Attack Pattern Analysis with Clustering at Bekangdam II Sriwijaya

M.R. Kareliansyah[1*], Heri Suroyo [2]

Abstract

Research conducted at Bekangdam II Sriwijaya aims to analyze the pattern of attacks carried out by crackers using data mining of logs on Mikrotik. The method employed in this research is the descriptive method. The research directly involves fieldwork because it requires a lengthy analysis. The benefits derived from conducting this research include identifying the techniques launched by crackers against the existing network at Bekangdam II Sriwijaya. Developments in the world of technology have reached the Industrial Era 4.0. In this era, telecommunication networks have undergone many changes, both in wired and wireless networks. The increasing rate of this technological development is inseparable from the growing cybercrime activity. The results of this study indicate that data mining of logs on the Bekangdam II Sriwijaya Server can be analyzed, with the analysis revealing an IP that frequently accesses the server continuously, specifically 103.39.9.81, utilizing the Telnet medium 205 times in a day. The clustering process using K-means resulted in 4 clusters: the first cluster (1) is shown in blue, the second cluster (2) is purple, the third cluster (3) is green, and the fourth cluster is red. The yellow star color indicates the cluster's center point or its centroid point. Based on the access media, the Cluster Center Initialization can be grouped according to the following table. The medium most frequently used to launch this attack is Telnet, with a frequency of 535 times.

Keywords

Mikrotikm, Server log data mining, Attack pattern

Article History
Received 21 November 2024
Accepted 18 February 2025

How to Cite
Kareliansyah, M.R, & Suroyo, H. (2025). Analysis of Mikrotik Server Log Data Mining for Attack Pattern Analysis with Clustering at Bekangdam II Sriwijaya. ***Jurnal Jaringan Komputer dan Keamanan,*** *6*(1), 22-29.

[1*] Universitas Bina Darma, Indonesia, Corresponding email: kareliansyah@gmail.com
[2] Universitas Bina Darma, Indonesia, email: herisuroyo@binadarma.ac.id

Introduction

The rapid development of information technology continues to evolve in line with human needs, offering various conveniences that simplify daily activities and professional tasks (Mansoori et al., 2012). Technological innovation has transformed how people communicate, retrieve information, and interact across vast distances. Telecommunication networks, supported by the internet, enable individuals to connect instantaneously without geographical boundaries. Through the internet, users access expansive knowledge repositories and digital services using unique network identifiers known as IP Addresses, which allow devices to communicate in a standardized and interconnected ecosystem.

As digital connectivity expands, all software and hardware components become integrated into a unified global network. This interconnected environment greatly facilitates information management, enabling users to access diverse data sources with ease. However, this openness also introduces significant risks. Information stored on websites or servers cannot be fully guaranteed secure, as vulnerabilities on the server side may be exploited by unauthorized individuals. In practice, information is categorized into public (open) and private, requiring different levels of protection based on its sensitivity and potential impact if breached.

Given these challenges, robust security systems must be established to minimize vulnerabilities. Weak security mechanisms on applications or servers create opportunities for attackers to conduct cybercrimes such as data breaches, unauthorized access, and denial-of-service attacks. As emphasized by Hartawan & Desnanjaya (2018), both application-level and server-level security must be strengthened to ensure the confidentiality, integrity, and availability of information. This requires continuous monitoring, regular risk assessments, and the deployment of advanced protection techniques capable of identifying attack patterns at an early stage.

One crucial aspect of securing server infrastructure is network monitoring. Server computers, including commonly used network devices such as Mikrotik routers, maintain real-time service activity logs that record all incoming and outgoing network traffic, login attempts, bandwidth usage patterns, and potential intrusions. These logs serve as valuable forensic data for analyzing suspicious activities and identifying abnormal traffic behavior. Because servers are frequent targets of penetration attempts, continuous log-based monitoring becomes essential for early detection of cyberattacks.

To support this monitoring process, analytical methods such as K-Means Clustering can be employed. According to Zulfadhilah et al. (2016), K-Means is effective for grouping and classifying attack patterns contained within Mikrotik log data. By clustering similar data entries, administrators can quickly identify anomalies or unusual traffic groups that may indicate malicious activities. The insights obtained from these clusters allow for more targeted and proactive network defense strategies.

Various clustering techniques—including Fuzzy K-Means, DBSCAN, Hierarchical Clustering, and K-Means Clustering—are widely used in previous research for pattern recognition and anomaly detection in network traffic. Among these, K-Means remains one of the most popular due to its simplicity, computational efficiency, and ability to generate clear grouping structures. Studies that apply K-Means to server log analysis demonstrate its effectiveness in detecting inconsistencies and mapping attack patterns based on log-generated data. The resulting cluster information provides valuable insights into the types, frequencies, and behaviors of attacks, thereby supporting administrators in strengthening network security and optimizing intrusion-response strategies.

Methodology

The researcher decided to conduct this study using the descriptive research method. According to experts, Hidayat (2010) states that the descriptive method is a broader type of research in the use of acquired data. The term 'broader' implies a focus on a lengthy analysis from start to finish, so the results obtained are detailed and specific.

In conducting research using the descriptive method, every researcher must be required to have a strong commitment to the research being conducted, both theoretically and practically, because when going directly into the field, every essential piece of information must be collected. This research method requires a long and strong analysis.

Data Collection Methodology

The data collection methods used in this research involve several techniques:

a. Literature Method Data collection using this method involves gathering various written sources related to the research implementation by reading, studying, and recording important matters connected to the issue being discussed to obtain a detailed overview that can support the thesis preparation.
b. Observation Method This is a data collection method by conducting a direct review or observation of the object being studied. Data is collected directly on the research object by recording matters related to the research and the thesis title, thereby obtaining complete and accurate data (Pratama et al., 2019).

Data Analysis Methodology

The research development method used in this study is the Clustering Technique using the RapidMiner application. By employing the K-means Clustering method and assisted by the RapidMiner application, data can be processed based on existing columns. This research will yield several data groups corresponding to frequency or quantity based on the percentages used, such as High, Low, Medium, and Sufficient.

The following are the stages in using the Clustering method:

a. Preprocessing In this stage, several parameters will be added to meet the requirements for grouping the risk levels of a data, in this case, Mikrotik server log data at Bekangdam II Sriwijaya. The parameters that will be used include Priority, Port, and Frequency. The following is an explanation of the parameters provided (Haris et al., 2020):

Parameter	Explanation
Priority	Each log data will be determined based on the order of each attack data that must be considered. Priority will be divided into three: **Low** (1 and 2), **Medium** (3), and

	High (4).[62]
Port	A value will be assigned to the acquired log data to determine which ports are used for access. Ports will be divided into three: **Well-Known Port** (4), **Registered Port** (3), and **Dynamic / Private Port** (1-2).[63]
Frequency	Frequency will determine the number of attacks that have been sent. The frequency parameter values assigned are: **1-2** (Low), **3** (Medium), and **4** (Highest frequency, High).[64]

b. Clustering After grouping the data through the preprocessing stage, the next stage is clustering, which involves grouping the log data obtained from the preprocessing stage. The resulting data will have characteristics based on the mentioned parameters using the K-Means Clustering method.

c. Risk Assessment After the preprocessing stage, the next stage is the Risk Assessment stage. In this stage, the result of each clustering will yield three final centroid center points. These three centroid points will result in the average frequency present in each cluster. Each cluster will have its own centroid, and no two will be the same.

To calculate the risk value of each attack carried out on the server at Bekangdam II Sriwijaya, the risk value calculation formula based on the determined parameters will be used. The formula for determining the result can be seen in the following:

$$P = \{1 - 4\}$$
$$D = \{1 - 5\}$$
$$F = \{1 - 5\}$$
$$MaxRA = 10$$
$$RA = \frac{P\ x\ D\ x\ F}{X}$$

Every risk value result obtained will be further categorized based on the attack risk categories, as shown in the following table (Haris et al., 2020):

Risk Value	Category
1-2	Low [81]
2	Medium
3	High

d. **Evaluation** The obtained results must first undergo an evaluation process to determine whether the results are considered optimal or not. If the results obtained are optimal, then the research on the attack risk value is correct. If not, further research on the obtained results is necessary.

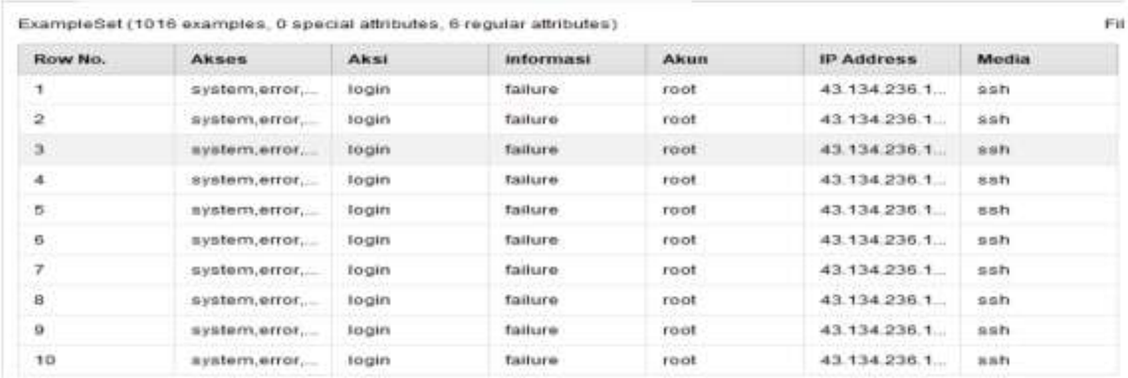
Results

This stage will describe the results of the data and text mining analysis process based on the data obtained from mining the log data on the Bekangdam II Sriwijaya server. The following are the stages for describing the results of the data and text mining analysis process based on the data obtained.

Discussion

Data Preprocessing

The data obtained from Mikrotik log mining is still in text form and has not been filtered based on columns. Therefore, the data must be grouped into a table so it can be processed in the RapidMiner application. The following image is the result of Data Preprocessing using the RapidMiner application :



ExampleSet (1016 examples, 0 special attributes, 6 regular attributes) Fil

Row No.	Akses	Aksi	Informasi	Akun	IP Address	Media
1	system,error,...	login	failure	root	43.134.236.1...	ssh
2	system,error,...	login	failure	root	43.134.236.1...	ssh
3	system,error,...	login	failure	root	43.134.236.1...	ssh
4	system,error,...	login	failure	root	43.134.236.1...	ssh
5	system,error,...	login	failure	root	43.134.236.1...	ssh
6	system,error,...	login	failure	root	43.134.236.1...	ssh
7	system,error,...	login	failure	root	43.134.236.1...	ssh
8	system,error,...	login	failure	root	43.134.236.1...	ssh
9	system,error,...	login	failure	root	43.134.236.1...	ssh
10	system,error,...	login	failure	root	43.134.236.1...	ssh

Figure 3.1 Preprocessing Results in RapidMiner

Clustering

From the data preprocessing using RapidMiner Studio, the researcher can view the population of a data based on a graph by selecting the statistics menu. Based on the IP addresses attempting to access the Mikrotik server at Bekangdam II Sriwijaya, the following graph is generated:

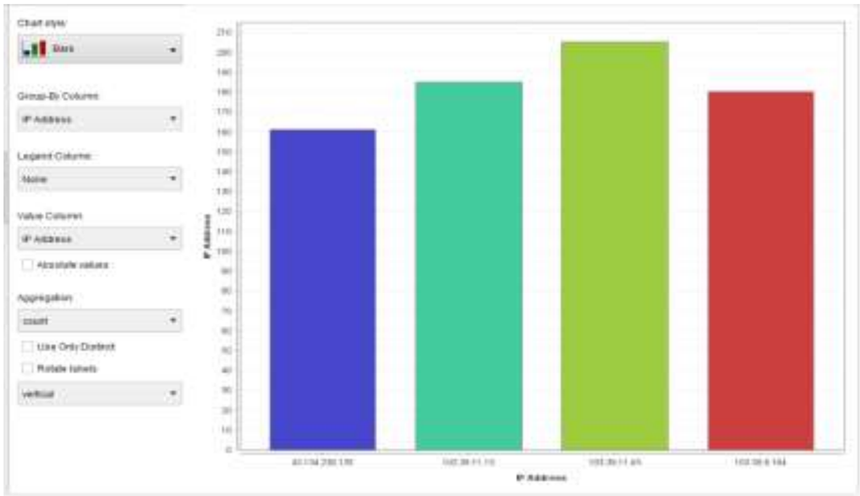


Chart style: Bars
Group-By Column: IP Address
Legend Column: None
Value Column: IP Address
□ Absolute values
Aggregation: count
□ Use Only Distinct
□ Rotate labels
vertical

Figure 3.2 Bar Chart Depicting the Processed Data in RapidMiner

From the graphical data, it is clearly visible that several IP Addresses are attempting to log in to the Mikrotik Server using random Usernames and Passwords. This attack technique is called a Brute Force Attack. A Brute Force Attack is an act by an individual to penetrate a server by continuously attempting random logins against a number of words, usually those inputted by default by the server administrator. The individual can access the system if one of those random words successfully performs authentication on the server.

Risk Assessment

Based on the Bar Graph data, the IPs that frequently attempt authentication against the Bekangdam II Sriwijaya Server are:

Table 2.3 Initial Assignment of Cluster Centroids According to IP Address

Attack Technique	IP Address	Count/Frequency	Cluster Center Initialization
Brute Force Attack	103.39.9.81	205	1
Brute Force Attack	103.39.9.77	188	2
Brute Force Attack	103.39.11.13	180	3
Brute Force Attack	43.134.236.135	162	4

Conclusion and Recommendations

Based on the preceding chapters and the results of the clustering analysis conducted on log data mining from the Bekangdam II Sriwijaya Server, several key conclusions can be drawn. The analysis revealed that the IP address 103.39.9.81 frequently performed continuous access attempts, utilizing the Telnet protocol 205 times in a single day. Such repeated access behavior is typically associated with Brute Force Attack techniques, which involve the use of random usernames and passwords to repeatedly attempt authentication.

The clustering process using the K-means algorithm produced four distinct clusters: Cluster 1 (blue), Cluster 2 (purple), Cluster 3 (green), and Cluster 4 (red). The yellow star symbol represents the centroid or center point of each cluster. Based on the access medium, the initialization of cluster centers can be categorized accordingly.

In terms of attack frequency by medium, Telne ranked first with 535 occurrences, followed by SSH with 161 occurrences, and again SSH with 35 occurrences, indicating multiple SSH-based attempts from different sources. The log data processed via RapidMiner confirmed that the dominant technique used by the most active IP addresses was **Brute Force Attack.

The top three IP addresses involved in these authentication attempts were:

103.39.9.81 with **205 attempts
103.39.9.77 with **188 attempt
103.39.11.13 with **180 attempts

It is also evident that the quality of data preprocessing particularly data cleaning significantly influences the accuracy and clarity of the analytical output.

The researcher proposes the following recommendations based on the findings of this study:

This research is expected to serve as a reference for future studies, encouraging other researchers to conduct more refined and comprehensive investigations in the field of network security and log data analysis. Continuous server monitoring is essential. It is recommended to implement mechanisms that **block IP addresses** exhibiting persistent authentication attempts, as such behavior can overload the server and disrupt internet traffic for client systems. Given the clear indication of Brute Force Attacks, network administrators are advised not to use default router passwords. Instead, they should configure **strong, unpredictable passwords to mitigate the risk of unauthorized access.

References

Anonymous. (n.d.). *Data Mining Analysis of Mikrotik Server Logs for Attack Pattern Detection Using Clustering at Bekangdam II Sriwijaya*.

Davies, P. (2004). *Database Systems* (3rd ed.).

Ditanaya, T. H. (2016). *Design and Development of a Syslog and Cassandra-Based Log Server System for Network Monitoring at ITS*.

Fatmawati, K., & Windarto, A. P. (2018). Data mining: Application of RapidMiner with K-Means Clustering on Dengue Fever Outbreak Areas by Province. *Computer Engineering, Science and System Journal, 3*(2), 173. https://doi.org/10.24114/cess.v3i2.9661

Han, J., & Kamber, M. (2001). *Data Mining: Concepts and Techniques*. Academic Press.

Haris, G., Wibawa, P., Made, I. G., Sasmita, A., & Raharja, I. M. S. (2020). Honeypot Log Data Analysis Using K-Means Clustering Method. *Jurnal Ilmiah Merpati Universitas Udayana, 8*(1), 13–21.

Hartawan, I. N. B., & Desnanjaya, I. G. M. N. (2018). Performance Analysis of Zigbee Protocol Indoors and Outdoors as a Data Communication Medium in Wireless Sensor Networks. *Jurnal RESISTOR (Rekayasa Sistem Komputer), 1*(2), 65–72. https://doi.org/10.31598/jurnalresistor.v1i2.320

Mansoori, M., Zakaria, O., & Gani, A. (2012). Enhancing Intrusion Deception System Exposure Through Hybrid Honeypot Implementation. *International Arab Journal of Information Technology, 9*(5).

Nofitri, R., & Irawati, N. (2019). Profit Data Analysis Using RapidMiner Software. *JURTEKSI (Journal of Technology and Information Systems), 5*(2), 199–204. https://doi.org/10.33330/jurteksi.v5i2.365

Pramudiono, P. (2007). *Introduction to Data Mining: Extracting Knowledge Gems from the Data Mountain*.

Pratama, Y., Ependi, U., & Suroyo, H. (2019). Optimization of Wireless Network Performance Using the Hierarchical Token Bucket (Case Study: Muhammadiyah University of Palembang). *Journal, 1*(1), 49–59.

Rahmat, C. T. I. B., Gafar, A. A., Fajriani, N., Ramdani, U., Uyun, F. R., Purnamasari, P. Y., & Ransi, N. (2017). Implementation of K-Means Clustering in RapidMiner for Accident-Prone Area Analysis. *Proceedings of the National Seminar on Applied Quantitative Research*, April, 58–60.

Sadikin, R. (2012). *Cryptography for Network Security*. Yogyakarta.

Santosa, B. (2007). *Data Mining: Techniques for Utilizing Data for Business Purposes*. Graha Ilmu.

Sofana, I. (2017). *Computer Network Book Based on Mikrotik*. Informatika.

Suyanto. (2017). *Data Mining for Data Classification and Clustering*. Bandung: Informatika Publishing.

International Telecommunication Union. (1991). *ITU Proceedings*, 987–990. https://doi.org/10.18356/cbabbce2-en

Wikipedia. (2021). Komando Daerah Militer II/Sriwijaya. Retrieved from https://id.wikipedia.org/wiki/Komando_Daerah_Militer_II/Sriwijaya

Zulfadhilah, M., Riadi, I., & Prayudi, Y. (2016). Log Classification Using K-Means Clustering to Identify Internet User Behaviors. *International Journal of Computer Applications, 154*(3).

Zulfadhilah, M., Riadi, I., & Prayudi, Y. (2016). Log Classification Using K-Means Clustering to Identify Internet User Behaviors. *International Journal of Computer Applications, 154*(3), 34–39. https://doi.org/10.5120/ijca2016912076