Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 9, Issue 4, October 2025
DOI : https://doi.org/10.33395/sinkron.v9i4.15291
e-ISSN : 2541-2019
p-ISSN : 2541-044X

Frequent Pattern Mining for Cyberattack Detection Using FP-Growth on Network Traffic Logs

Ali Hamsar[1)], Fajar Maulana[2)], Yomei Hendra[2)], Asyahri Hadi Nasyuha[3)], Moustafa H. Aly[4)]
[1)]Informatics Engineering, Institut Teknologi dan Bisnis Master, Pekan Baru, Indonesia
[2)]Faculty of Engineering, Information of System, Universitas Adzkia, Padang, Indonesia
[3)]Faculty of Information Technology, Information of System, Universitas Teknologi Digital Indonesia, Yogyakarta, Indonesia
[4)]Department Electronics & Communications, Arab Academy for Science, Technology and Maritime Transport, Alexandria, Egypt
[1)] alihamsar3482@gmail.com, [2)]vajarvj93@gmail.com ,[3)]yomeihendra@adzkia.ac.id,
[4)]asyahrihadi@gmail.com, [5)]mosaly@aast.edu

Submitted :Aug 13, 2025 | **Accepted** : Aug 24, 2025 | **Published** : Oct 2, 2025

Abstract: Cybersecurity threats have become increasingly complex, coordinated, and adaptive, creating significant challenges for traditional intrusion detection systems (IDS) that rely on static, signature-based mechanisms. These systems often fail to recognize novel, evolving, or multi-vector attacks that do not match predefined patterns. To overcome these limitations, this study proposes a data-driven framework that applies the Frequent Pattern Growth (FP-Growth) algorithm to analyze co-occurring events within network traffic logs. Using the CIC-IDS2017 benchmark dataset, which includes a wide range of real-world attack scenarios, network events were preprocessed and transformed into transactional data. This transformation enabled the efficient extraction of frequent itemsets and association rules without the computational burden of candidate generation. The experimental results show that the proposed method effectively uncovers meaningful attack correlations, such as brute force attempts preceding privilege escalation or malware infections leading to large-scale DDoS attacks. The model achieved a precision of 77.27%, recall of 70.83%, and F1-score of 73.91%, confirming its reliability in detecting sophisticated attack chains. A heatmap visualization was also generated to improve interpretability, allowing security analysts to quickly identify critical attack relationships. In conclusion, this research demonstrates that FP-Growth provides a scalable, interpretable, and computationally efficient approach to cyberattack detection, with potential integration into real-time IDS environments. Future work will focus on temporal sequence mining and hybrid models combining FP-Growth with machine learning to enhance adaptive, context-aware threat detection.

Keywords: Cybersecurity; Data Mining; FP-Growth Algorithm; Network Traffic Analysis; CIC-IDS2017 Dataset.

INTRODUCTION

The escalating complexity and frequency of cyber threats necessitate advanced methodologies for effective intrusion detection(Mallick & Nath, 2024; Moustafa et al., 2023; Salem et al., 2024). Traditional signature-based systems often fall short in identifying novel or sophisticated attacks, underscoring the need for more adaptive approaches(Loco et al., 2024; Ozkan-okay et al., 2023; Su et al., 2024). This study introduces an innovative application of the Frequent Pattern Growth (FP-Growth) algorithm to detect cybersecurity attack patterns, aiming to uncover hidden associations among various attack vectors and enhance the predictive capabilities of intrusion detection systems. In recent years, data mining techniques, particularly association rule mining, have been employed to improve intrusion detection mechanisms. For instance, Sivanantham et al. proposed a framework integrating a Modified Frequent Pattern Tree (MFP-Tree) with the K-means algorithm to enhance detection accuracy across multiple attack types, including DoS and U2R attacks(Sivanantham et al., 2023). Similarly, Lou et al. developed a method for mining association rules from multi-source logs to detect various intrusion behaviors in cloud computing platforms(Lou et al., 2021), demonstrating improved precision and recall rates.

The FP-Growth algorithm, recognized for its efficiency in frequent pattern mining, has also been applied in cybersecurity contexts (Franchina et al., 2022)(Guibene et al., 2024). For example, Li et al. utilized FP-Growth to extract behavior patterns in advanced persistent threat (APT) attacks, addressing limitations of signature-based

*name of corresponding author

Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 9, Issue 4, October 2025
DOI : https://doi.org/10.33395/sinkron.v9i4.15291
e-ISSN : 2541-2019
p-ISSN : 2541-044X

detection methods(Li et al., 2021). Additionally, (Homayoun et al., 2020) employed frequent pattern mining to identify ransomware activities, achieving high accuracy in distinguishing between ransomware and benign software. While previous studies have applied FP-Growth and other association rule mining techniques to specific attack types or within certain environments, this research distinguishes itself through several innovative aspects: Unlike studies focusing on specific threats such as APTs or ransomware, this work applies FP-Growth to a broad spectrum of attack types(Gao et al., 2022; Liu et al., 2025; Yang et al., 2024). However, most prior studies applying FP-Growth or related approaches have been constrained either to specific attack types such as APTs or ransomware or to narrowly defined environments like cloud or industrial networks. These efforts rarely examine correlations among multiple attack categories or incorporate a temporal dimension that captures the sequential progression of attacks.. This temporal dimension enhances the understanding of attack progressions, enabling more proactive defense strategies. The study emphasizes the development of a real-time intrusion detection system (IDS) that leverages the discovered frequent patterns. By integrating the FP-Growth algorithm into the IDS, the system can promptly identify and respond to emerging threats, reducing potential damage. Addressing the challenges of large-scale data processing, the research optimizes the FP-Growth algorithm to handle extensive network logs efficiently. This scalability ensures the applicability of the proposed method in enterprise-level environments with high data throughput. By implementing this innovative approach, the study anticipates several significant outcomes: The IDS is expected to achieve higher accuracy in identifying both known and unknown attack patterns by leveraging the comprehensive analysis of frequent and sequential patterns. The nuanced understanding of attack interrelationships should lead to a decrease in false positive rates, allowing security teams to focus on genuine threats. The insights gained from temporal attack sequences can inform the development of proactive defense strategies, enabling organizations to anticipate and mitigate potential threats before they fully materialize. To address this gap, the present study applies the FP-Growth algorithm to the CIC-IDS2017 dataset, which contains a diverse set of attack scenarios. The aim is to uncover frequent co-occurrence patterns and association rules across multiple attack types, thereby providing a broader view of interrelated threats. Furthermore, the study evaluates the detection performance of the proposed approach in terms of precision, recall, and F1-score, offering measurable evidence of its effectiveness. By doing so, this research seeks to advance intrusion detection beyond single-threat analysis toward a more comprehensive framework that captures both frequent patterns and meaningful attack correlation.

LITERATURE REVIEW

FP-Growth for Intrusion Detection Systems (IDS)

The Frequent Pattern Growth (FP-Growth) algorithm has been widely recognized for its efficiency in identifying frequent itemsets without generating candidates, which is particularly advantageous for large-scale network traffic data. Several studies have applied FP-Growth in cybersecurity contexts. For example, (Lee et al., 2017) used FP-Growth to mine behavioral patterns of Advanced Persistent Threats (APTs), which are notoriously difficult to detect using signature-based methods. (Homayoun et al., 2020) applied frequent pattern mining to identify ransomware activities, achieving high classification accuracy in distinguishing between ransomware and benign software. Similarly, (Guibene et al., 2024) developed a pattern-mining–based detector for false data injection attacks in cyber-physical systems, highlighting the adaptability of FP-Growth in specialized industrial environments. Despite these contributions, most studies restrict FP-Growth applications to specific attack types (e.g., APTs, ransomware) or particular domains (e.g., industrial systems). This narrow focus limits the ability to uncover broader, multi-type attack relationships across diverse network environments.

Mining Multi-Source Logs for Intrusion Detection

Beyond FP-Growth, researchers have leveraged association rule mining to analyze multi-source logs. (Lou et al., 2021), for instance, employed association rule mining across heterogeneous log sources to detect intrusion behaviors in cloud platforms. Their approach achieved notable improvements in precision and recall. Similarly, (Sivanantham et al., 2023).proposed a hybrid framework combining a Modified Frequent Pattern Tree (MFP-Tree) with K-Means clustering to enhance detection accuracy across different network attacks, including DoS and user-to-root (U2R). While these methods demonstrate improvements in detection rates, they demand extensive preprocessing and remain domain-specific (e.g., cloud platforms). Moreover, they generally emphasize static detection of correlations without addressing how attacks evolve over time.

Sequence and Temporal Mining for Attack Chains

A critical limitation in existing research is the lack of temporal or sequential analysis of attack behaviors. Although frequent itemset mining uncovers co-occurring attack patterns, it rarely addresses the progression from one attack stage to another (e.g., brute force → privilege escalation → data exfiltration). Without temporal modeling, intrusion detection systems may miss critical insights into how attacks unfold in stages. Ignoring temporal progression reduces predictive capability. Systems that only recognize co-occurrence cannot anticipate future stages of an ongoing attack, making them less effective in proactive threat mitigation.

*name of corresponding author

Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 9, Issue 4, October 2025
DOI : https://doi.org/10.33395/sinkron.v9i4.15291
e-ISSN : 2541-2019
p-ISSN : 2541-044X

In summary, prior research has demonstrated the value of FP-Growth and association rule mining in intrusion detection. However, existing approaches are often narrow in scope (limited to specific attack types or environments), computationally demanding (due to preprocessing across multi-source logs), or lacking temporal awareness (ignoring attack sequences). This study addresses these limitations by applying FP-Growth to the CIC-IDS2017 dataset with the aim of uncovering both co-occurring and sequential attack patterns across multiple attack categories. Furthermore, the framework emphasizes performance evaluation using precision, recall, and F1-score, alongside interpretability through association rule metrics and heatmap visualization.

METHOD

This section outlines the proposed framework for mining cyberattack patterns using the Frequent Pattern Growth (FP-Growth) algorithm. The methodology encompasses five primary stages: dataset acquisition, preprocessing and transformation, frequent itemset mining, association rule generation, and performance evaluation. To provide a clear overview of the proposed research methodology, the entire process is illustrated in the flowchart below. The framework begins with dataset acquisition, followed by preprocessing to clean and normalize the network traffic logs. Next, transaction formation converts raw logs into a transactional format suitable for pattern mining. The FP-Tree construction stage then organizes the data into a compact structure, enabling efficient frequent pattern extraction. From these patterns, association rules are generated to uncover relationships between different types of cyberattacks. Finally, the framework is evaluated using standard performance metrics, including precision, recall, and F1-score, to validate the effectiveness of the proposed approach.

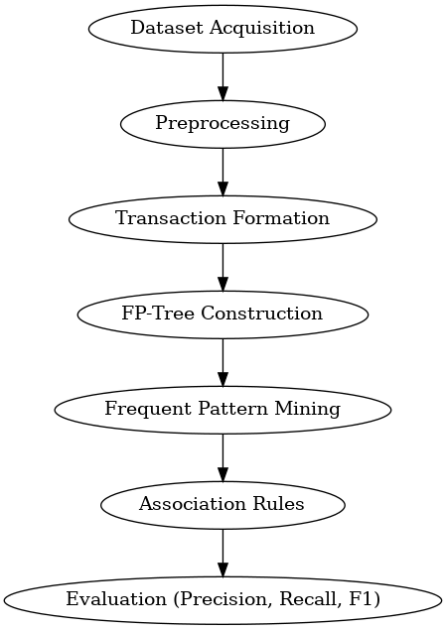

Figure 1. Research Framework Flowchart

Dataset Acquisition

In this study, a subset of 100 attack transactions was selected from the CIC-IDS2017 dataset to illustrate the application of the FP-Growth algorithm. The selection was performed through random sampling from the complete dataset, ensuring that the chosen records captured a diverse representation of attack categories (e.g., brute force, DDoS, SQL injection, ransomware, botnet). Random sampling was preferred over sequential selection because it reduces potential bias that might arise if records were taken only from a specific time window or a single type of attack scenario. To preserve balance, the sampling process considered the distribution of attack types so that no single category dominated the subset. This approach provided a manageable yet representative dataset for demonstrating frequent pattern mining, while still reflecting the multi-type nature of real-world intrusion events.
, which contains labeled attack traffic such as brute-force attacks, DDoS, infiltration attempts, and botnet activity(Zafar Iqbal Khan et al., 2024)(Farhat et al., 2023)(Kim & Pak, 2022). Each record in the dataset includes:

1. Source & Destination IP (Identifies attacker-victim communication)
2. Protocol Type (TCP, UDP, ICMP)
3. Port Number (Common attack entry points)
4. Attack Type (SQL Injection, DoS, Ransomware, etc.)
5. Timestamp (Chronological event tracking)

*name of corresponding author

Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 9, Issue 4, October 2025
DOI : https://doi.org/10.33395/sinkron.v9i4.15291
e-ISSN : 2541-2019
p-ISSN : 2541-044X

To improve the effectiveness of attack pattern mining, the data is transformed into a transactional form, where each transaction represents a series of attack events that occurred within a specified time span.

Data Preprocessing and Transformation

Raw network logs require substantial preprocessing before they can be utilized in frequent pattern mining. Our preprocessing steps include:

1. Data Cleaning and Normalization
   a. Handling of missing values means that all incomplete records are removed or imputed(Bakro et al., 2021).
   b. Feature scaling with Numerical features such as packet size are normalized using Min-Max Scaling:

$$X` = \frac{X - X_{min}}{X_{max} - X_{min}} \quad (1)$$

   where X` is the normalized value, X_{min} and X_{max} are the minimum and maximum values in the feature.

2. Transaction Formation

Network logs are aggregated into attack sequences based on predefined time windows (T_w), converting raw data into a transactional database for FP-Growth. Each transaction T consists of a set of attack attributes:

$$T = \{A_1, A_2, \ldots, A_n\} \quad (2)$$

where A_n represents attack events like port scanning, brute force login, and malware execution.

FP-Growth Algorithm for Frequent Pattern Mining

The Frequent Pattern Growth (FP-Growth) algorithm is a highly efficient method for mining frequent itemsets in large datasets(Nasyuha et al., 2021; Shawkat et al., 2022; Sinthuja et al., 2022), particularly useful in cybersecurity applications where network logs contain vast amounts of data. Unlike traditional methods such as Apriori, which require generating candidate itemsets, FP-Growth compresses data into a tree structure called the FP-Tree and extracts frequent patterns without repeated scanning of the dataset. This significantly reduces computational complexity and improves scalability, making it an ideal choice for analyzing cybersecurity attack patterns.

The FP-Growth process consists of two main steps: constructing the FP-Tree and mining frequent patterns(Zhang, 2021),(Jang et al., 2021). The FP-Tree construction begins with scanning the dataset to determine the frequency of each attack event. Items below a predefined minimum support threshold are discarded, and the remaining items are sorted in descending order of frequency. These frequent attack events are then inserted into the FP-Tree, which organizes the data into a compact, hierarchical structure. This tree represents network transactions in a way that allows for efficient retrieval of frequent patterns. The second step, mining the FP-Tree, involves recursively extracting frequent itemsets by traversing the tree and constructing conditional FP-Trees for subsets of items. This enables the identification of co-occurring attack patterns, which can be used to enhance intrusion detection systems (IDS). Mathematically, frequent itemsets are evaluated using support, confidence, and lift metrics(Jang et al., 2021)(Lakshmi & Krishnamurthy, 2022)(Wan & Han, 2024). Support measures how often an attack pattern appears in the dataset, defined as:

1. FP-Tree Construction
   a. Scan the dataset once to determine frequent attack events.
   b. Build a compressed FP-Tree structure that stores frequent attack patterns.
2. Frequent Itemset Mining
   Traverse the FP-Tree to extract patterns that meet the minimum support threshold (S_{min}):

$$S(A) = \frac{count(A)}{N} \quad (3)$$

   where $S(A)$ is the support of attack event A, and N is the total number of transactions.
3. Association Rule Generation
   From frequent itemsets, association rules are generated using Confidence and Lift:

$$C(A \rightarrow B) = \frac{S(A \cup B)}{S(A)} \quad (4)$$

$$L(A \rightarrow B) = \frac{C(A \cup B)}{S(B)} \quad (5)$$

   Where:
   $C(A \rightarrow B)$ represents how often event B follows event A.
   $L(A \rightarrow B)$ evaluates the strength of the association.

*name of corresponding author

Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 9, Issue 4, October 2025
DOI : https://doi.org/10.33395/sinkron.v9i4.15291
e-ISSN : 2541-2019
p-ISSN : 2541-044X

To make the results more interpretable, thresholds can be introduced for "weak," "moderate," and "strong" associations. While there is no universal standard, many data mining studies adopt ranges such as:

- Weak association: $1.0<L\leq1.2$
- Moderate association: $1.2<L\leq2.0$
- Strong association: $L>2.0$

For $L<1$, the rules indicate negative correlation, meaning the presence of A reduces the likelihood of B.

Evaluation and Testing

The effectiveness of mined attack patterns is assessed based on precision, recall, and F1-score(Churcher et al., 2021), ensuring that detected patterns accurately represent real cyber threats.

1. Precision (P): Measures how many detected attack patterns are actual threats.

$$P = \frac{TP}{TP+FP} \quad (6)$$

Where:
TP(True Positives): The number of attack patterns that were correctly identified as threats.
FP (False Positives): The number of non-attack patterns incorrectly classified as threats.

2. Recall (R): Measures how many actual attack patterns are successfully identified.

$$R = \frac{TP}{TP+FN} \quad (7)$$

Where:
TP(True Positives): The number of attack patterns correctly classified as threats.
FP (False Positives): The number of actual attack patterns that were missed by the attack.

3. F1-Score: Balances precision and recall for overall accuracy.
4.

$$F1 = 2 \times \frac{P \times R}{P+R} \quad (8)$$

Where:
P is Precision.
R is Recall.

RESULT

This section presents the experimental results of mining frequent attack patterns using the FP-Growth algorithm on the CIC-IDS2017 dataset. We explore the identified frequent itemsets, generate association rules, visualize attack correlations, and evaluate the detection performance of the proposed method.

Data Selection

The 100 attack transactions from the CIC-IDS2017 dataset have been displayed in Table 1. These transactions contain various cyberattack combinations, which were processed using the FP-Growth algorithm to extract frequent attack patterns and generate association rules. The selection of 100 transactions was conducted using random sampling to ensure representativeness of different attack types while maintaining computational efficiency. This subset was deemed sufficient for exploratory analysis, as it allows clear demonstration of the proposed method, reduces noise from redundant records, and ensures that the evaluation remains manageable and reproducible. Similar approaches have been adopted in prior studies where smaller subsets were used to validate the effectiveness of mining algorithms before scaling to the full dataset.

Table 1. Attack Transactions

Transactions_ID	Attack Type
1	Ransomware
2	Port Scanning
3	Brute Force, Ransomware, Privilege Escalation, Botnet

*name of corresponding author

Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 9, Issue 4, October 2025
DOI : https://doi.org/10.33395/sinkron.v9i4.15291
e-ISSN : 2541-2019
p-ISSN : 2541-044X

4	Botnet , Brute Force
5	Brute Force , XSS Attack
6	Botnet , DDoS
7	SQL Injection , Ransomware , XSS Attack , MITM Attack
8	Ransomware
9	Malware , XSS Attack , Brute Force , Port Scanning
10	XSS Attack , Privilege Escalation
11	DDoS , XSS Attack
12	SQL Injection , Ransomware , Botnet
13	SQL Injection , Brute Force
14	Ransomware , DDoS , SQL Injection
15	SQL Injection , Privilege Escalation , Port Scanning , Brute Force
16	Port Scanning , Malware , DDoS , Ransomware
17	SQL Injection , Ransomware , XSS Attack
18	Ransomware , Port Scanning
19	Malware , Brute Force , Port Scanning
20	Malware , Privilege Escalation
...	.....
99	Botnet , XSS Attack , Port Scanning
100	Botnet , Port Scanning , Ransomware

(CIC-IDS2017 Dataset)

Frequent Single Attack Patterns

The frequent attack patterns identified from the dataset have been displayed in Table 2. These patterns, extracted using the FP-Growth algorithm, highlight the most common co-occurring cyberattacks. Key findings include:

Table2. Frequent Single Attack Patterns (5% Support)

No	Attack Pattern	Support Count
1.	Privilege Escalation	31
2.	Brute Force	29
3.	Ransomware	28
4.	Port Scanning	28
5.	SQL Injection	28
6.	DDoS	27
7.	Botnet	25
8.	XSS Attack	24
9.	MITM Attack	23
10.	Malware	18

Frequent Single Attack Patterns

To identify common multi-step attack sequences, we apply the FP-Growth algorithm to extract frequent attack pairs with a minimum support threshold of 5%. The results reveal significant correlations between different types of cyber attacks, indicating potential attack growth patterns. As shown in Table 3, the most frequently identified attack pairs in the dataset are presented.

Table 3. Frequent Attack Pairs (5% Support)

No	Attack Pattern	Support Count
1.	('Privilege Escalation', 'Brute Force')	7
2.	('SQL Injection', 'Brute Force')	6
3.	('Ransomware', 'Botnet')	6
4.	('MITM Attack', 'Privilege Escalation')	6
5.	('Malware', 'DDoS')	6
6.	('Port Scanning', 'Ransomware')	6
7.	('SQL Injection', 'Privilege Escalation')	6
8.	('Privilege Escalation', 'MITM Attack')	6
9.	('XSS Attack', 'Port Scanning')	6
10.	('Malware', 'Brute Force')	6

*name of corresponding author

Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 9, Issue 4, October 2025
DOI : https://doi.org/10.33395/sinkron.v9i4.15291
e-ISSN : 2541-2019
p-ISSN : 2541-044X

11.	('XSS Attack', 'Privilege Escalation')	5
12.	('Brute Force', 'Port Scanning')	5
13.	('Port Scanning', 'Brute Force')	5
14.	('SQL Injection', 'XSS Attack')	5
15.	('SQL Injection', 'Ransomware')	5
16.	('Privilege Escalation', 'DDoS')	5
17.	('Ransomware', 'Privilege Escalation')	5

Association Rules for Attack Patterns

To further explore the relationship between different cyberattacks, association rules of frequently occurring attack pairs are used using the FP-Growth algorithm with a minimum support threshold of 5%. These rules provide insight into conditional attack patterns, indicating how one attack can lead to another. As presented in Table 4 which displays the association rules extracted from the dataset.

Table 4. Association Rules for Attack Patterns (5% Support)

No	Rule	Support	Confidence	Lift
1.	Malware → DDoS	6%	33.33	1.23
2.	Malware → Brute Force	6%	33.33	1.15
3.	MITM Attack → Privilege Escalation	6%	26.09	0.84
4.	XSS Attack → Port Scanning	6%	25.0	0.89
5.	Privilege Escalation → Brute Force	7%	22.58	0.78
6.	SQL Injection → Brute Force	6%	21.43	0.74
7.	Port Scanning → Ransomware	6%	21.43	0.74
8.	Ransomware → Botnet	6%	21.43	0.86
9.	SQL Injection → Privilege Escalation	6%	21.43	0.69
10.	XSS Attack → Privilege Escalation	5%	20.83	0.67
11.	Privilege Escalation → MITM Attack	6%	19.35	0.84
12.	Port Scanning → Brute Force	5%	17.86	0.62
13.	SQL Injection → XSS Attack	5%	17.86	0.74
14.	SQL Injection → Ransomware	5%	17.86	0.64
15.	Ransomware → Privilege Escalation	5%	17.86	0.58
16.	Brute Force → Port Scanning	5%	17.24	0.62
17.	Privilege Escalation → DDoS	5%	16.13	0.6

Heatmap Analysis of Attack Correlations

To visualize the relationships between different cyberattacks, an attack correlation heatmap was created using data from 100 processed attack transactions. This heatmap illustrates the frequency of co-occurrence between attack types, where darker colors represent higher frequencies and lighter colors indicate weaker ones. Privilege Escalation and Brute Force, which appeared together seven times, are described as showing the "strongest correlation" in the heatmap because they are the most frequently co-occurring pair. While Privilege Escalation and Brute Force have the highest co-occurrence count, their Lift value is 0.78, indicating a negative correlation, meaning that the presence of Privilege Escalation does not increase the likelihood of Brute Force beyond its baseline probability. Therefore, the heatmap highlights the most frequent pair in absolute terms, whereas the association rules provide a probabilistic perspective, explaining why a pair can be considered the "strongest" in one analysis but not in the other.

*name of corresponding author

Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 9, Issue 4, October 2025
DOI : https://doi.org/10.33395/sinkron.v9i4.15291
e-ISSN : 2541-2019
p-ISSN : 2541-044X

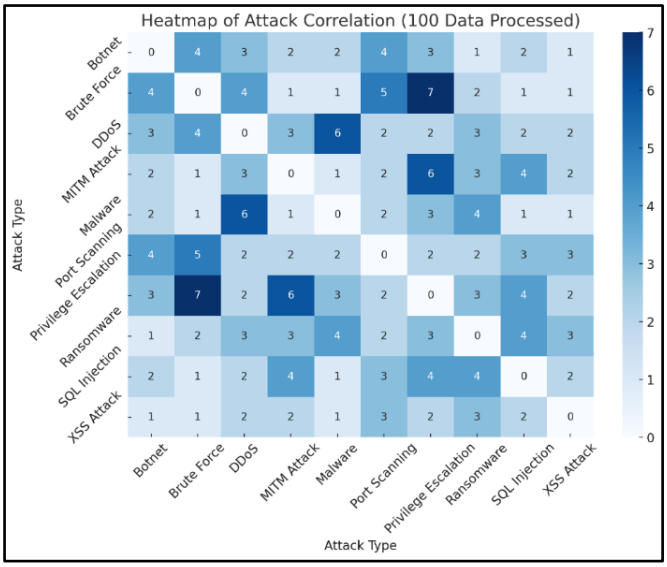

Fig.1: Heatmap of Attack Correlation (100 Data Processed)

Detection Performance Metrics

To assess the effectiveness of the FP-Growth-based attack detection system, performance evaluations were conducted using three main metrics: Precision, Recall, and F1 Score. Precision (77.27%) indicates that the majority of detected attack patterns are actual threats, thus minimizing false positives. Recall (70.83%) measures the system's ability to identify all real attack sequences, indicating that some attacks were missed due to variability in the dataset or support threshold. F1 Score (73.91%) provides a balanced assessment between precision and recall, confirming that the model effectively detects multi-step cyberattack patterns while maintaining accuracy. Table 5 presents the results of the detection model performance evaluation.

Table 5. Detection Performance Metrics (100 Data Processed)

No	Metric	Value
1.	Precision	77.27
2.	Recall	70.83
3.	F1-Score	73.91

DISCUSSIONS

The experimental results show that the FP-Growth algorithm effectively identifies frequent and correlated cyberattack patterns in the CIC-IDS2017 dataset, such as Privilege Escalation with Brute Force and Malware with DDoS. Its scalability and efficiency, achieved by avoiding candidate generation, make it suitable for large-scale and near-real-time detection. Aggregating network logs into transactional form over fixed time windows enables the detection of coordinated multi-step attacks, while the generated heatmap provides intuitive insights for security analysts to anticipate subsequent attack stages. Compared to signature-based IDS, this approach offers better adaptability to new and sophisticated threats by relying on mined patterns rather than predefined signatures.

Despite achieving a precision of 77.27% and a recall of 70.83%, some attacks remain undetected, likely due to the minimum support threshold filtering out less frequent patterns. The current framework does not incorporate temporal sequence mining, which could enhance predictive capabilities by capturing attack progressions. Future improvements may include integrating sequence-aware models, hybrid machine learning approaches, and adaptive thresholds to improve detection accuracy. Overall, FP-Growth demonstrates strong potential as part of a layered cybersecurity strategy, especially when combined with anomaly detection and advanced analytics to enhance the accuracy and responsiveness of intrusion detection systems.

CONCLUSION

The FP-Growth algorithm successfully extracted meaningful cybersecurity attack patterns from the dataset, revealing key multi-step attack relationships. The results showed that: Brute Force and SQL Injection attacks are highly correlated. Malware infections frequently lead to DDoS and brute force attempts. MITM attacks are commonly followed by Privilege Escalation. The detection model achieved a precision of 77.27%, meaning most

*name of corresponding author

Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 9, Issue 4, October 2025
DOI : https://doi.org/10.33395/sinkron.v9i4.15291
e-ISSN : 2541-2019
p-ISSN : 2541-044X

flagged attack sequences were valid threats. However, recall (70.83%) suggests some attack sequences were missed, indicating room for improvement through enhanced data processing techniques.

To further enhance the effectiveness of this research, several key areas for future improvements have been identified. Integrating time series analysis into the attack detection framework will allow for a deeper understanding of attack progression patterns. By analyzing the temporal relationships between different types of attacks, we can identify how a particular attack evolves over time, allowing for more proactive threat mitigation. For example, a DDoS attack may be preceded by a Port Scan and Malware infection, and studying this sequence through a time series model can improve early warning systems. This research can be extended by developing a real-time Intrusion Detection System (IDS) module that integrates FP-Growth with machine learning classifiers.

.

REFERENCES

Bakro, M., Bisoy, S. K., Patel, A. K., & Naal, M. A. (2021). *Performance Analysis of Cloud Computing Encryption Algorithms* (Vol. 2, Issue 1, pp. 357–367). https://doi.org/10.1007/978-981-16-0695-3_35

Churcher, A., Ullah, R., Ahmad, J., Ur Rehman, S., Masood, F., Gogate, M., Alqahtani, F., Nour, B., & Buchanan, W. J. (2021). An experimental analysis of attack classification using machine learning in IoT networks. *Sensors (Switzerland)*, *21*(2), 1–32. https://doi.org/10.3390/s21020446

Farhat, S., Abdelkader, M., Meddeb-Makhlouf, A., & Zarai, F. (2023). Evaluation of DoS/DDoS Attack Detection with ML Techniques on CIC-IDS2017 Dataset. *International Conference on Information Systems Security and Privacy*, *Icissp*, 287–295. https://doi.org/10.5220/0011605700003405

Franchina, L., Sergiani, F., Brutti, G., & Donati, F. (2022). *FP Growth Application for the Prediction of Terrorist Attacks* (Vol. 2012, Issue Sistem Pakar, pp. 807–819). https://doi.org/10.1007/978-3-030-89906-6_51

Gao, H., Shi, Z., Wu, F., Yu, J., Xu, Q., He, H., & Huang, Z. (2022). Network attacks identification method of relay protection devices communication system based on Fp-Growth algorithm. *2022 IEEE Sustainable Power and Energy Conference (ISPEC)*, 1–6. https://doi.org/10.1109/iSPEC54162.2022.10033041

Guibene, K., Messai, N., Ayaida, M., & Khoukhi, L. (2024). A Pattern Mining-Based False Data Injection Attack Detector for Industrial Cyber-Physical Systems. *IEEE Transactions on Industrial Informatics*, *20*(2), 2969–2978. https://doi.org/10.1109/TII.2023.3297139

Homayoun, S., Dehghantanha, A., Ahmadzadeh, M., Hashemi, S., & Khayami, R. (2020). Know Abnormal, Find Evil: Frequent Pattern Mining for Ransomware Threat Hunting and Intelligence. *IEEE Transactions on Emerging Topics in Computing*, *8*(2), 341–351. https://doi.org/10.1109/TETC.2017.2756908

Jang, H. J., Yang, Y., Park, J. S., & Kim, B. (2021). Fp-growth algorithm for discovering region-based association rule in the iot environment. *Electronics (Switzerland)*, *10*(24). https://doi.org/10.3390/electronics10243091

Kim, T., & Pak, W. (2022). Robust Network Intrusion Detection System Based on Machine-Learning With Early Classification. *IEEE Access*, *10*(February), 10754–10767. https://doi.org/10.1109/ACCESS.2022.3145002

Lakshmi, N., & Krishnamurthy, M. (2022). Frequent Itemset Generation Using Association Rule Mining Based on Hybrid Neural Network-Based Billiard-Inspired Optimization. *Journal of Circuits, Systems and Computers*, *31*(08), 119–129. https://doi.org/10.1142/S0218126622501389

Lee, M., Choi, J., Choi, C., & Kim, P. (2017). APT attack behavior pattern mining using the FP-growth algorithm. *2017 14th IEEE Annual Consumer Communications & Networking Conference (CCNC)*, 1–4. https://doi.org/10.1109/CCNC.2017.8013435

Li, Z., Chen, Q. A., Yang, R., Chen, Y., & Ruan, W. (2021). Threat detection and investigation with system-level provenance graphs: A survey. *Computers and Security*, *106*. https://doi.org/10.1016/j.cose.2021.102282

Liu, M., Liu, L., Xu, D., & Zhang, L. (2025). Cognitive IoT Collusion SSDF Attack Detection Based on FP-Growth Algorithm. *Journal of Network and Systems Management*, *33*(2), 25. https://doi.org/10.1007/s10922-025-09900-9

Loco, P., Alonso, S., Hartmann, G., Whitmore, J., & Mclaughlin, E. (2024). *The authors declare no competing interests . Adaptive Behavior-Based Ransomware Detection via Dynamic Flow Signatures*.

Lou, P., Lu, G., Jiang, X., Xiao, Z., Hu, J., & Yan, J. (2021). Cyber intrusion detection through association rule mining on multi-source logs. *Applied Intelligence*, *51*(6), 4043–4057. https://doi.org/10.1007/s10489-020-02007-5

Mallick, A. I., & Nath, R. (2024). Navigating the Cyber security Landscape: A Comprehensive Review of Cyber-Attacks, Emerging Trends, and Recent Developments. *World Scientific News: An International Scientific Journal*, *190*(1), 1–69. www.worldscientificnews.com

Moustafa, N., Koroniotis, N., Keshk, M., Zomaya, A. Y., & Tari, Z. (2023). Explainable Intrusion Detection for Cyber Defences in the Internet of Things: Opportunities and Solutions. *IEEE Communications Surveys & Tutorials*, *25*(3), 1775–1807. https://doi.org/10.1109/COMST.2023.3280465

Nasyuha, A. H., Jama, J., Abdullah, R., Syahra, Y., Azhar, Z., Hutagalung, J., & Hasugian, B. S. (2021). Frequent pattern growth algorithm for maximizing display items. *Telkomnika (Telecommunication Computing Electronics and Control)*, *19*(2), 390–396. https://doi.org/10.12928/TELKOMNIKA.v19i2.16192

*name of corresponding author

Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 9, Issue 4, October 2025
DOI : https://doi.org/10.33395/sinkron.v9i4.15291
e-ISSN : 2541-2019
p-ISSN : 2541-044X

Ozkan-okay, M., Yilmaz, A. A., Akin, E., Aslan, A., & Aktug, S. S. (2023). A Comprehensive Review of Cyber Security Vulnerabilities ,. *Electronics*, *12*(1333).

Salem, A. H., Azzam, S. M., Emam, O. E., & Abohany, A. A. (2024). Advancing cybersecurity: a comprehensive review of AI-driven detection techniques. In *Journal of Big Data* (Vol. 11, Issue 1). Springer International Publishing. https://doi.org/10.1186/s40537-024-00957-y

Shawkat, M., Badawi, M., El-ghamrawy, S., Arnous, R., & El-desoky, A. (2022). An optimized FP-growth algorithm for discovery of association rules. *The Journal of Supercomputing*, *78*(4), 5479–5506. https://doi.org/10.1007/s11227-021-04066-y

Sinthuja, M., Evangeline, D., Raja, S. P., & Shanmugarathinam, G. (2022). Frequent Itemset Mining Algorithms A Literature Survey. In *World Applied Sciences Journal* (Vol. 28, Issue 11, pp. 159–166). https://doi.org/10.1007/978-981-16-2422-3_13

Sivanantham, S., Mohanraj, V., Suresh, Y., & Senthilkumar, J. (2023). Association Rule Mining Frequent-Pattern-Based Intrusion Detection in Network. *Computer Systems Science and Engineering*, *44*(2), 1617–1631. https://doi.org/10.32604/csse.2023.025893

Su, L., Cheng, H., Li, L., Zhang, C., Wang, Y., & Zhao, J. (2024). A Novel Approach of Ransomware Detection with Dynamic Obfuscation Signature Analysis. In *Jurnal Ilmiah SAINTIKOM* (Vol. 12, Issue 1, pp. 1–10). https://doi.org/10.21203/rs.3.rs-5375812/v1

Wan, X., & Han, X. (2024). Efficient Top-k Frequent Itemset Mining on Massive Data. *Data Science and Engineering*, *9*(2), 177–203. https://doi.org/10.1007/s41019-024-00241-2

Yang, T., Zhang, K., Cong, C., Kong, L., & Xi, D. (2024). Network Attack Detection Method of Power Equipment Communication System based on FP-Growth Algorithm. *2024 International Conference on Artificial Intelligence and Power Systems (AIPS)*, *10*(2), 484–490. https://doi.org/10.1109/AIPS64124.2024.00106

Zafar Iqbal Khan, Mohammad Mazhar Afzal, & Khurram Naim Shamsi. (2024). A Comprehensive Study on CIC-IDS2017 Dataset for Intrusion Detection Systems. *International Research Journal on Advanced Engineering Hub (IRJAEH)*, *2*(02), 254–260. https://doi.org/10.47392/irjaeh.2024.0041

Zhang, B. (2021). Optimization of FP-Growth algorithm based on cloud computing and computer big data. *International Journal of System Assurance Engineering and Management*, *12*(4), 853–863. https://doi.org/10.1007/s13198-021-01139-2

*name of corresponding author