

Network Traffic Feature Selection for Intrusion Detection Using Statistical and Correlation Analysis

Su Mon Tint Soe¹; Dr. Cho Me Me Maung²

¹University of Technology (Yatanarpon Cyber City); Department of Computer Engineering,
Pyin Oo Lwin, Myanmar

²Technological University (Mandalay); Department of Information Technology, Mandalay, Myanmar

Publication Date: 2026/10/13

Abstract: In the contemporary landscape of increasingly sophisticated network attacks, effectively selecting salient features from network traffic data has become a critical step in developing robust Intrusion Detection Systems (IDS). However, the efficacy of Machine Learning (ML) and Deep Learning (DL)-based IDS is often compromised by the high dimensionality of network traffic features, disparate data distributions, and the presence of irrelevant features. This research presents a Two-Stage Network Traffic Feature Extraction Framework for intrusion detection, evaluated on the CSE-CIC-IDS2018 Network Traffic Dataset. In the first stage, twenty-two (22) features—including Flow Duration, Packet Rate, Byte Rate, Packet Length Statistics, Inter-Arrival Time, and TCP Flag Information—were extracted to characterize the behavioral properties of network traffic. In the second stage, an additional sixteen (16) features encompassing Flow Information, Packet Statistics, Port Information, and Protocol-related Features were selected to facilitate the discrimination of network attacks. To assess the quality and significance of the extracted features, statistical and correlation analyses were conducted. The mean, standard deviation, minimum, maximum, and median values of each feature were computed to analyze network traffic characteristics. Furthermore, a correlation matrix and heatmaps were employed to examine inter-feature relationships, allowing for the identification of highly correlated feature pairs. The resulting feature extraction framework effectively captures the essential characteristics of network traffic data and serves as a foundational representation for the subsequent development of ML and DL-based intrusion detection models.

Keywords: Network Traffic Analysis, CSE-CIC-IDS2018, Feature Selection, Machine Learning, Correlation Analysis, Deep Learning, Intrusion Detection System.

How to Cite: Su Mon Tint Soe; Dr. Cho Me Me Maung (2026) Network Traffic Feature Selection for Intrusion Detection Using Statistical and Correlation Analysis. *International Journal of Innovative Science and Research Technology*, 11(9), 3393-3401. <https://doi.org/10.38124/ijisrt/26sep1445>

I. INTRODUCTION

The rapid proliferation of the Internet and digital communication technologies has led to the pervasive adoption of network systems across organizations, enterprises, and government institutions. However, as network utilization continues to expand, cyber threats have grown increasingly sophisticated. Denial of Service (DoS), Distributed Denial of Service (DDoS), Brute Force, Port Scanning, and other network attacks pose significant risks to the security of critical data and services within networked environments [1], [2].

To detect such cyber threats in real time, Intrusion Detection Systems (IDS) have become an essential technology in the domain of network security. Traditional IDS approaches rely on predefined signatures and rules to identify attacks, which limits their ability to detect emerging or previously unknown attack patterns [3]. Consequently, the

extraction of effective feature representations from network traffic data has emerged as a critical research focus for Machine Learning (ML) and Deep Learning (DL)-based IDS [4], [5].

Network traffic data encompasses a wide array of features that characterize network behavior, including Flow Duration, Packet Transmission Rate, Packet Length, Inter-Arrival Time, TCP Flag Information, and other relevant attributes [6]. However, the direct utilization of all available features can result in the inclusion of irrelevant features, high feature dimensionality, and increased computational cost [7]. Therefore, appropriate feature selection and feature extraction methods play a vital role at the foundational stage of IDS development.

This research presents a Two-Stage Network Traffic Feature Extraction Framework for network intrusion

detection, utilising the widely adopted CSE-CIC-IDS2018 Network Traffic Dataset [8]. In the first stage, twenty-two (22) features were extracted for behavioral anomaly analysis to characterize both normal and anomalous behaviors of network traffic. In the second stage, sixteen (16) features were selected for attack classification to identify the distinctive characteristics of various network attack types.

To evaluate the quality and significance of the extracted features, statistical and correlation analyses were employed. The mean, standard deviation, minimum, maximum, and median values of the features were computed to examine network traffic distributions, while inter-feature relationships were analyzed through a correlation matrix and heatmaps. Furthermore, feature importance analysis was intended to identify significant features for the subsequent development of intelligent intrusion detection models.

The primary objective of this research is not to directly construct a network intrusion detection model, but rather to establish a foundational network traffic feature extraction framework that enables the development of reliable and effective Machine Learning and Deep Learning models. The resulting feature analysis outcomes can serve as essential foundational information for the future advancement of AI-based Intrusion Detection Systems.

II. RELATED WORK

Network Intrusion Detection Systems (IDS) have been a prominent research topic in the field of network security for many years. Early IDS approaches primarily employed signature-based and rule-based detection methods, identifying attacks based on predefined attack patterns. While these methods proved effective at detecting known attacks, they exhibited limitations in identifying emerging or previously unknown attacks. Consequently, Machine Learning (ML) and Deep Learning (DL)-based IDS approaches, capable of automatically learning hidden patterns within network traffic data, have garnered increasing research attention.

In Machine Learning-based intrusion detection approaches, the extraction of suitable features from network traffic directly impacts model performance and constitutes a critical stage [9]. Studies have utilized flow-based features—such as Flow Duration, Packet Count, Byte Count, Packet Length Statistics, and Protocol Information—to characterize the behavioral properties of network traffic. However, the high dimensionality of features in network datasets can introduce irrelevant or redundant features, necessitating feature selection and feature analysis.

In the study by L. Dhanabal and S. Shantharajah, the NSL-KDD dataset was employed for network intrusion detection, examining the significance of traffic features and comparing the classification performance of Machine Learning algorithms [10]. Their work demonstrated that the appropriate selection of network traffic features is essential in constructing an Intrusion Detection System.

Furthermore, owing to the increasing complexity of modern network environments, IDS research has continued using more contemporary datasets. The CSE-CIC-IDS2018 dataset, developed by the Canadian Institute for Cybersecurity, has been widely adopted in ML and DL-based IDS research, as it incorporates realistic network traffic, benign traffic, and diverse attack scenarios. This dataset includes flow-based features generated by CICFlowMeter, enabling detailed examination of network traffic behavior [11].

Recent studies have employed Deep Learning methods—including Convolutional Neural Networks (CNN), Recurrent Neural Networks (RNN), Long Short-Term Memory (LSTM), and hybrid CNN-LSTM models—to detect complex patterns within network traffic [12]. Although Deep Learning models can reduce the need for manual feature engineering, data preprocessing and feature extraction stages remain essential for obtaining high-quality input feature representations.

Feature correlation analysis and statistical analysis are important methods for understanding the distribution and relationships of network traffic features. By examining the mean, standard deviation, minimum, maximum, and correlation coefficients of features, it is possible to determine which features are significant in characterizing network behavior and attack characteristics. Moreover, identifying highly correlated features enables the reduction of redundant information and enhances the efficiency of subsequently developed detection models.

However, the majority of previous studies have primarily focused on classification performance by directly utilizing features, with limited separate analysis of network traffic behavior analysis and attack-specific feature representation. Therefore, this research presents a Two-Stage Network Traffic Feature Extraction Framework based on the CSE-CIC-IDS2018 dataset, in which twenty-two (22) features are separately extracted for behavioral anomaly analysis and sixteen (16) features for attack classification analysis, with feature characteristics examined in detail through statistical and correlation analyses.

III. METHODOLOGY

This research proposes a Two-stage Network Traffic Feature Extraction Framework to construct foundational input representations for network intrusion detection. The primary objective of the proposed framework is to effectively extract significant traffic characteristics from raw network traffic data for intrusion detection, and to examine the quality and relationships of the features through statistical and correlation analyses.

The dataset employed in this study is the CSE-CIC-IDS2018 Network Traffic Dataset, which comprises normal (benign) traffic and diverse network attack traffic. Data preprocessing, feature extraction, and feature analysis stages were performed based on the flow-based network traffic information contained within the dataset.

➤ Overall Flow of the Proposed Framework

Define According to the overall workflow described in figure 1, this study processes network traffic data from the CSE-CIC-IDS2018 dataset in several stages to extract feature representations suitable for intrusion detection. First, the raw network traffic CSV files in the dataset are collected and read for data preprocessing. At this stage, data cleaning, missing value handling, duplicate header removal, data type conversion, and infinite value removal are carried out to prepare a clean network traffic dataset that can be further analyzed.

After preprocessing, the resulting network traffic features are divided into two stages based on the research objectives: Stage-1 Behavioral Feature Extraction and Stage-2 Attack Feature Extraction. In Stage-1, 22 features are extracted to study normal and abnormal behaviors in network traffic, and these are used for behavior anomaly analysis. In Stage-2, 16 features are selected to study the specific traffic characteristics of network attacks, and these are used for attack traffic analysis.

Based on the features obtained from these two stages, the feature dataset is further analyzed. First, statistical feature analysis is performed by calculating the mean, standard deviation, minimum, maximum, and median values of each feature. Then, a correlation matrix and heatmaps are used to study the relationships between features and to identify highly correlated features. This helps in understanding the relationships among features and any redundant information they may share.

Finally, feature importance analysis is conducted to identify the features that contribute more significantly to describing network traffic behavior and attack characteristics. The final feature representation obtained from these processes serves as a foundational feature dataset intended for use as input features when building machine learning and deep learning-based intrusion detection models in future work.

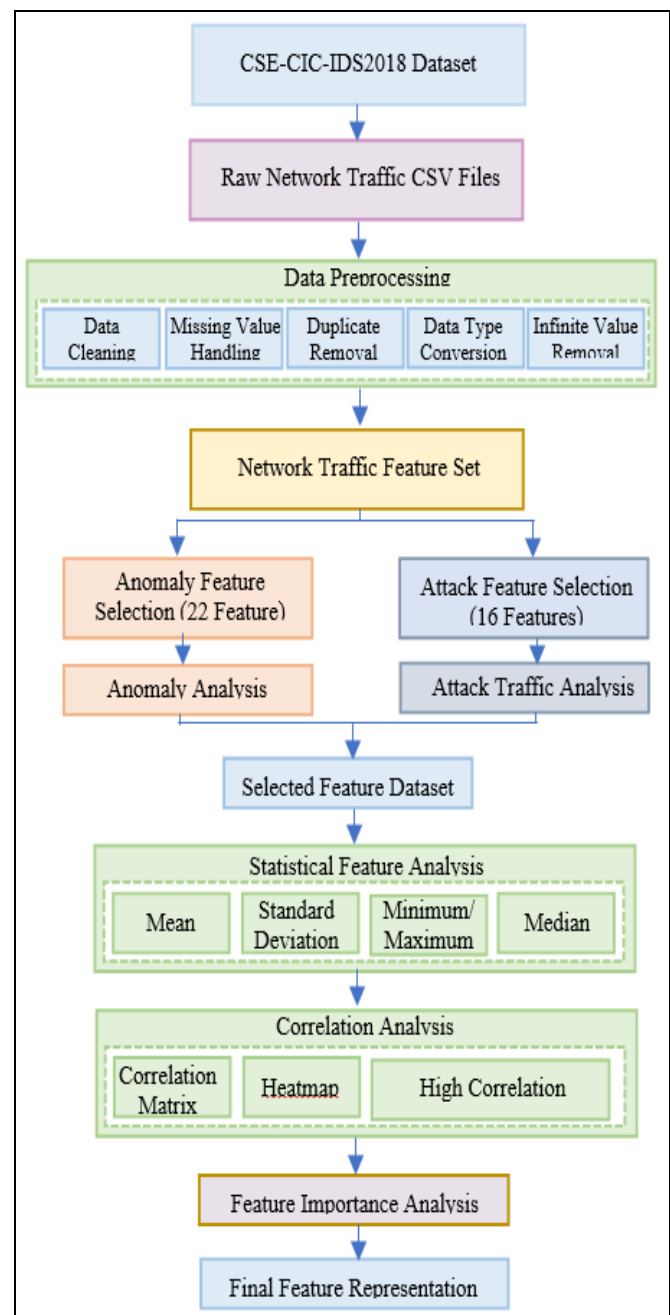


Fig 1 Overall Workflow of the Proposed Network Traffic Feature Selection Framework

➤ Dataset Preprocessing

The CSE-CIC-IDS2018 dataset contains flow-based network traffic records generated by CICFlowMeter. Since each dataset includes a large number of features and may vary in data quality, data preprocessing was carried out before feature extraction. First, all CSV files were systematically read, and the column names in the datasets were converted into a standard format. Duplicate header rows, missing values, and infinite values (Inf, -Inf) were then removed.

In addition, since the network traffic features are numerical data, values stored in string format were converted into numeric format. This step was performed to obtain a clean feature dataset suitable for machine learning and statistical analysis.

➤ *Two-Stage Feature Extraction*

In this study, the features obtained from network traffic data are extracted in two stages based on the research objective and the type of analysis: Stage-1 Behavioral Anomaly Feature Extraction and Stage-2 Attack Classification Feature Extraction. This separation allows the general behavior of network traffic and the specific characteristics of attack traffic to be studied independently.

In the first stage, 22 features were selected to identify normal and abnormal behaviors in network traffic. These features include flow duration, flow packets per second, flow bytes per second, total forward packets, total backward packets, packet length statistics, flow inter-arrival time, forward/backward packet rate, TCP flag information, and active/idle time statistics. Since these features represent the duration, traffic volume, packet transmission rate, packet size, and communication timing patterns of network communication, they are used to analyze normal and abnormal behaviors in network traffic.

In the second stage, 16 features were selected to study the specific characteristics of network attack traffic. These features include flow duration, packet rate, byte rate, destination port, packet size information, forward/backward packet statistics, and TCP SYN, ACK, and RST flag counts, as well as flow inter-arrival time. Because these features can describe the traffic volume, packet transmission behavior, targeted network service, and connection characteristics of attack traffic, they are used to study the distinct characteristics of different attack traffic types.

The feature sets obtained from these two stages were used for subsequent statistical analysis, correlation analysis, and feature importance analysis. They are also prepared as feature representations that can be used to build machine learning and deep learning-based intrusion detection models in future work.

➤ *Feature Statistical Analysis*

Statistical analysis was performed to study the distribution and numerical characteristics of the feature datasets obtained from the feature extraction stage in detail. For each feature, the mean, standard deviation, minimum, maximum, and median values were calculated. The mean represents the average value of the corresponding feature, while the median indicates the central value of the values in the dataset. The minimum and maximum values describe the lowest and highest ranges of a feature, and the standard deviation shows how closely the values are clustered around the mean or how widely they are spread. Therefore, by examining these statistical parameters together, the value range, distribution, and variation of each network traffic feature can be understood. This provides basic information for assessing suitable features for subsequent feature analysis and intrusion detection models.

➤ *Feature Correlation Analysis*

Pearson correlation analysis was used to study the interrelationships and the level of association among the features obtained after feature extraction. The correlation

values between features were calculated and presented using a correlation matrix, and a correlation heatmap was used for visualization to make the relationship patterns among features clearer. A correlation coefficient close to +1 indicates a strong positive relationship between two features, a value close to 0 indicates a weak linear relationship, and a value close to -1 indicates a strong negative relationship. In addition, by separately examining feature pairs with high correlation values, redundant features that may provide the same or overlapping information were identified. This analysis provides basic information that can be used for feature selection and feature optimization in future work.

➤ *Feature Importance Analysis*

Feature importance analysis was carried out to study which features obtained from the feature extraction stage are more important in describing network traffic behavior and attack characteristics. In this analysis, the impact and contribution level of each feature were assessed, and the features that are important for distinguishing and describing network traffic patterns were identified.

The results obtained from the feature importance analysis make it possible to understand the importance level of the features and help reduce unnecessary or less effective features. In addition, selecting and using features with high importance allows the input feature set of machine learning and deep learning-based intrusion detection models to be prepared more effectively in future work.

In this study, feature importance analysis was used to identify the important features among flow-based features, packet-based features, and network communication characteristics. These features provide information that can be used as a basis for feature selection when building AI-based intrusion detection systems in the future.

IV. RESULTS AND DISCUSSION

In this study, the proposed two-stage network traffic feature extraction framework was implemented using the network traffic flows in the CSE-CIC-IDS2018 dataset. Before feature extraction, data preprocessing steps were carried out to improve the quality of the dataset. These steps included duplicate header removal, missing value handling, infinite value removal, and numeric data conversion. After the preprocessing steps were completed, a total of 15,727,888 network traffic flow samples were obtained.

From the resulting network traffic data, features were extracted in two stages for two different analysis objectives. In the first stage, behavioral anomaly feature extraction, 22 features were selected to describe normal and abnormal behaviors in network traffic. These features include flow duration, flow packets per second, flow bytes per second, forward/backward packet count, packet length statistics, flow inter-arrival time, and TCP flag information. These features represent network communication behavior, traffic volume, and packet transmission patterns.

In the second stage, attack classification feature extraction, 16 features were selected to identify the specific characteristics of attack traffic. These features include destination port, packet statistics, flow rate, packet size information, and TCP connection state information, and they were used to describe the communication characteristics and network access behavior of attack traffic.

In studying the traffic class distribution of the CSE-CIC-IDS2018 dataset, it was found that the dataset contains a total of 15 traffic classes as shown in figure 2. When analyzing the traffic distribution within the dataset, benign traffic was found to be the largest class, with 13,390,249 flows. Among the attack traffic, DDoS attack-HOIC with 686,012 flows, DDoS attacks-LOIC-HTTP with 576,191 flows, DoS attacks-Hulk with 461,912 flows, and Bot with 286,191 flows were the most prevalent attack categories.

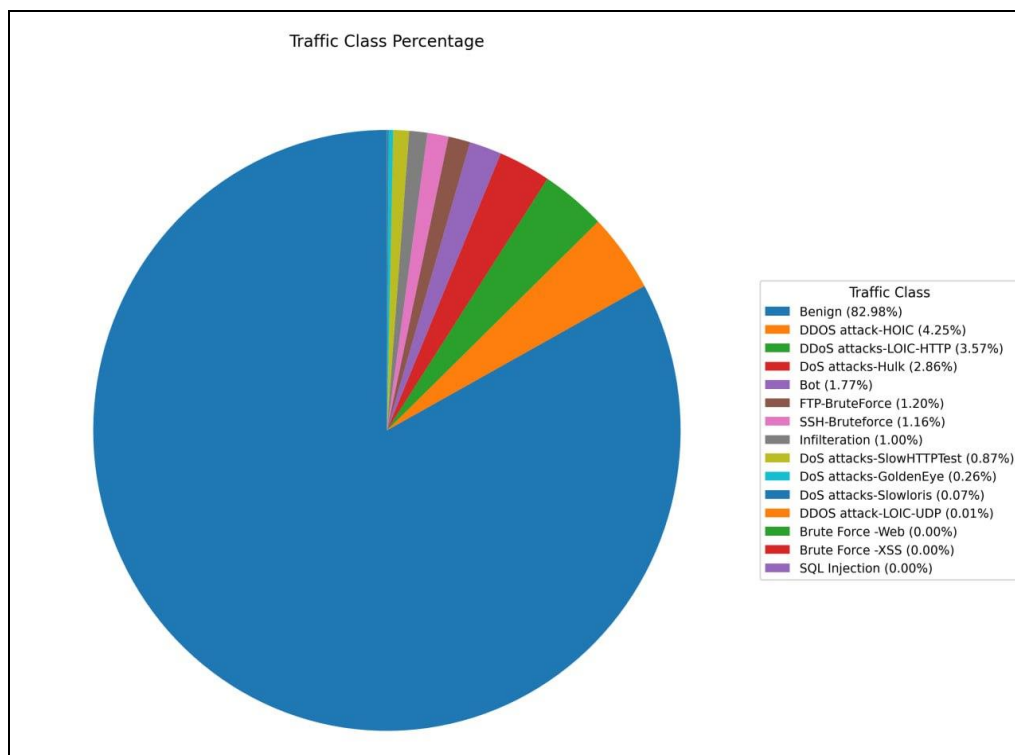


Fig 2 Traffic Class Distribution of the CSE-CIC-IDS2018 Dataset

These results show that the distribution of normal traffic and attack traffic in the CSE-CIC-IDS2018 dataset is imbalanced. In particular, since the number of benign traffic flows is significantly larger than that of most attack traffic, this indicates that the dataset has a class imbalance problem. Therefore, sampling strategies and class balancing techniques need to be considered when building machine learning and deep learning models in future work.

The statistical characteristics of the 22 features extracted for behavioral anomaly analysis were studied using mean, standard deviation as shown in table 1. The results showed that flow-based features have higher value ranges compared to other features.

The Flow_Duration feature has a mean value of 11,883,905.42 and a standard deviation of 4.95×10^8 , and high variation was also observed in the Flow_Pkts_s and Flow_Byts_s features. This result indicates that the connection duration, packet transmission rate, and data transfer volume of network traffic flows vary depending on the traffic type.

In addition, the high standard deviation values in flow duration, flow IAT, and idle time-related features reflect the variation in network communication timing patterns. Packet length and packet size-related features can also describe traffic volume and data transmission behavior, and they were found to be important features for behavioral anomaly analysis.

Table 1 Statistical Analysis Results of Behavioral Anomaly Features

No	Feature	Mean	Std
1	Flow_Duration	11883905.42	495475116.39
2	Flow_Pkts_s	52296.98	264725.04
3	Flow_Byts_s	257034.89	3667840.99
4	Tot_Fwd_Pkts	23.66	1525.64
5	Tot_Bwd_Pkts	6.35	164.51
6	TotLen_Fwd_Pkts	978.73	62353.20

7	TotLen_Bwd_Pkts	4759.01	235160.22
8	Pkt_Len_Mean	78.04	103.86
9	Pkt_Len_Std	122.15	162.77
10	Pkt_Size_Avg	90.00	108.37
11	Flow_IAT_Mean	3355189.24	223608721.03
12	Flow_IAT_Std	1283143.50	339114493.94
13	Fwd_IAT_Mean	3683026.57	223620563.58
14	Bwd_IAT_Mean	829090.71	4352891.13
15	Fwd_Pkts_s	36890.13	213937.31
16	Bwd_Pkts_s	15406.85	92686.16
17	SYN_Flag_Cnt	0.04	0.20
18	ACK_Flag_Cnt	0.33	0.47
19	RST_Flag_Cnt	0.19	0.39
20	Active_Mean	173903.46	2513315.47
21	Idle_Mean	5046729.31	264097629.99
22	Idle_Std	288350.15	169689954.83

The 16 features selected for attack classification were further studied through statistical analysis as shown in table 2. The results showed that flow duration, flow packets per second, flow bytes per second, and packet size-related features can describe the different communication characteristics among attack traffic types.

In particular, the destination port (Dst_Port) feature was found to be an important feature for attack traffic characterization, as it can represent information about the

network service target and access patterns. Destination port information can help in analyzing attack behaviors directed at different network services.

Pearson correlation analysis was used to study the relationships among the extracted features. According to the correlation analysis results for the behavioral feature set, high correlations were found among flow timing features, packet size features, and packet count features that is shown in figure 3.

Table 2 Statistical Analysis Results of Attack Classification Features

No	Feature	Mean	Std
1	Flow_Duration	11883905.42	495475116.39
2	Flow_Pkts_s	52296.98	264725.04
3	Flow_Byts_s	257034.89	3667840.99
4	Tot_Fwd_Pkts	23.66	1525.64
5	Tot_Bwd_Pkts	6.35	164.51
6	Pkt_Size_Avg	90.00	108.37
7	Pkt_Len_Var	41413.53	213584.69
8	Fwd_Pkt_Len_Mean	50.57	60.58
9	TotLen_Bwd_Pkts	4759.01	235160.22
10	Active_Mean	173903.46	2513315.47
11	Dst_Port	8917.96	18696.00
12	SYN_Flag_Cnt	0.04	0.20
13	ACK_Flag_Cnt	0.33	0.47
14	RST_Flag_Cnt	0.19	0.39
15	Flow_IAT_Mean	3355189.24	223608721.03
16	Flow_IAT_Std	1283143.50	339114493.94

In particular, Flow_IAT_Mean and Fwd_IAT_Mean were found to have a very strong positive relationship, with a correlation value of 0.999963. In addition, Tot_Bwd_Pkts and TotLen_Bwd_Pkts had a correlation value of 0.993508, indicating that backward packet count and data length are closely related to each other.

The correlation value of 0.992933 between Pkt_Len_Mean and Pkt_Size_Avg suggests that there may be overlapping information among packet size-related features. These correlation analysis results can help reduce redundant features during feature selection in future work.

When correlation analysis was performed for the attack classification feature set, it was found that there were fewer high correlation feature pairs compared to the behavioral feature set. Among the high correlation pairs found, Tot_Bwd_Pkts and TotLen_Bwd_Pkts had a correlation value of 0.993508, while Flow_Duration and Flow_IAT_Std had a negative correlation of -0.715130 that were shown in figure 5.

This result indicates that the features selected for attack classification can represent different network characteristics and have better feature independence.

In the final stage, Random Forest-based feature importance analysis was used to study the importance of the extracted features. In the behavioral anomaly feature analysis, Fwd_Pkts_s, Flow_Pkts_s, Tot_Bwd_Pkts,

Tot_Fwd_Pkts, Bwd_Pkts_s, Pkt_Len_Mean, and Pkt_Size_Avg were found to be important features that are shown in figure 4.

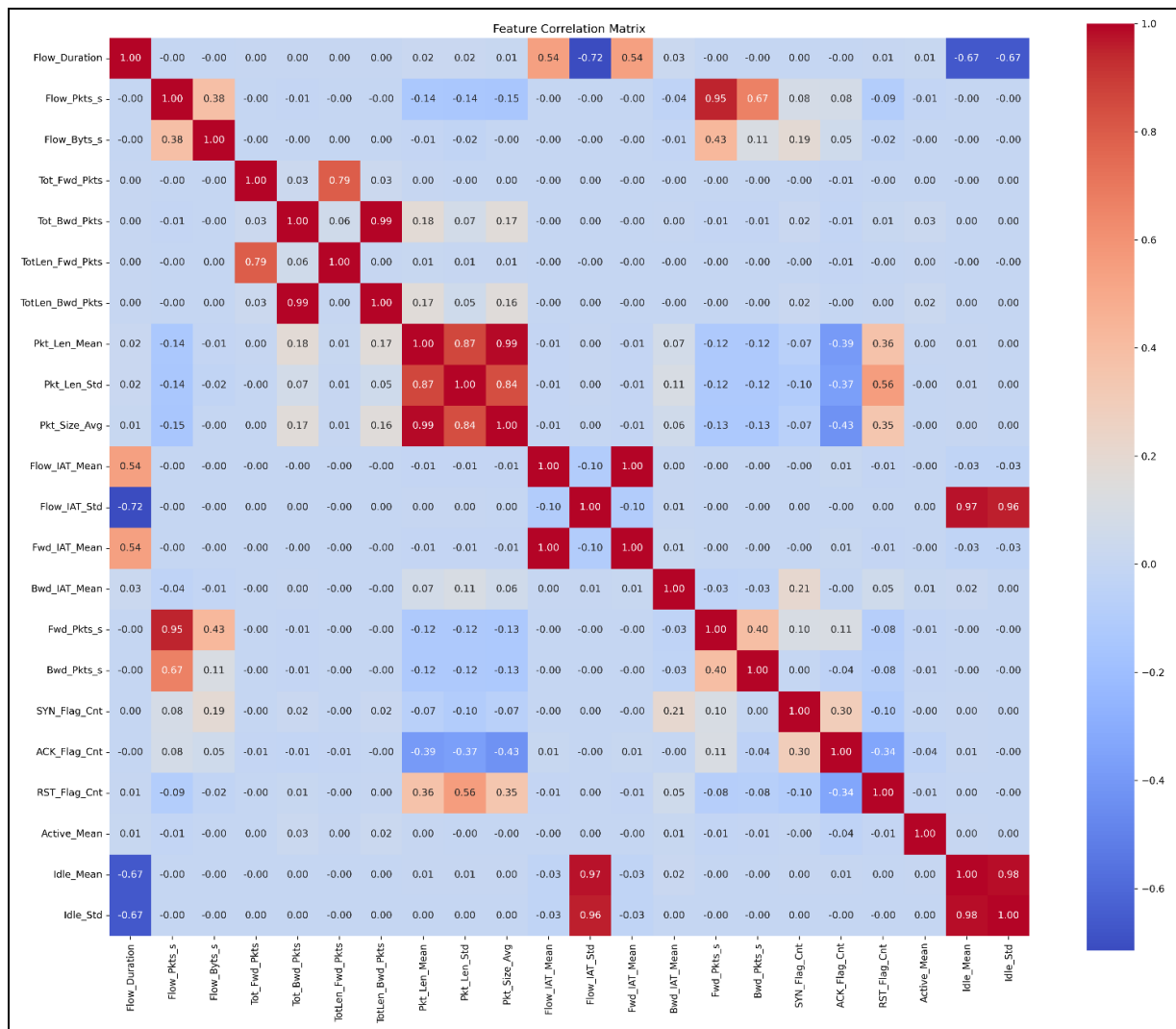


Fig 3 Correlation Heatmap of Behavioral Anomaly Features

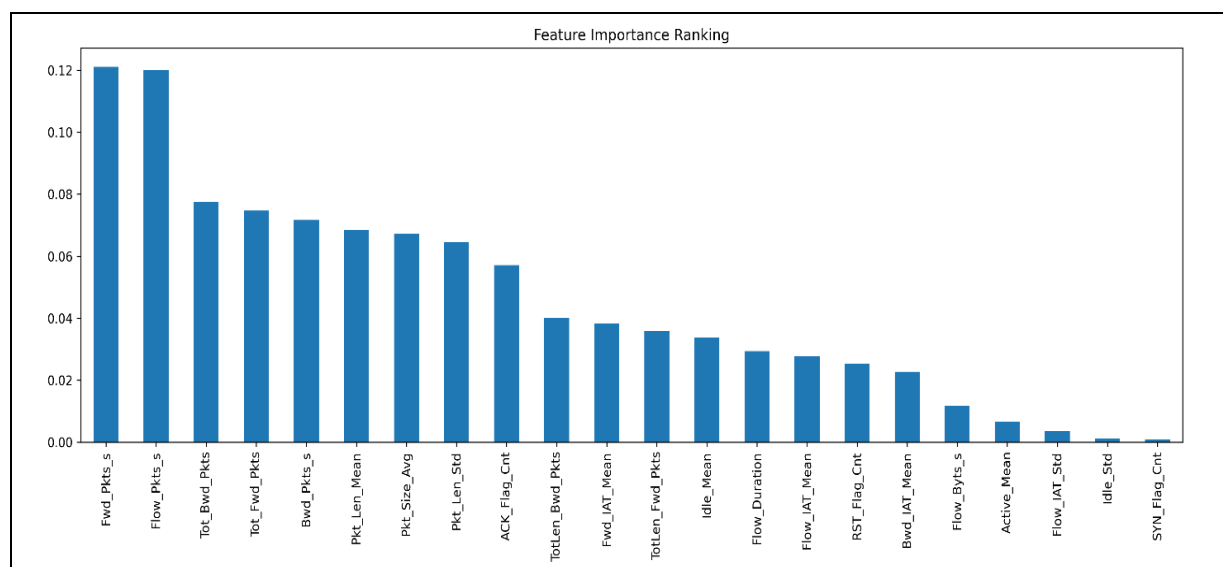


Fig 4 Importance Ranking of Behavioral Anomaly Features

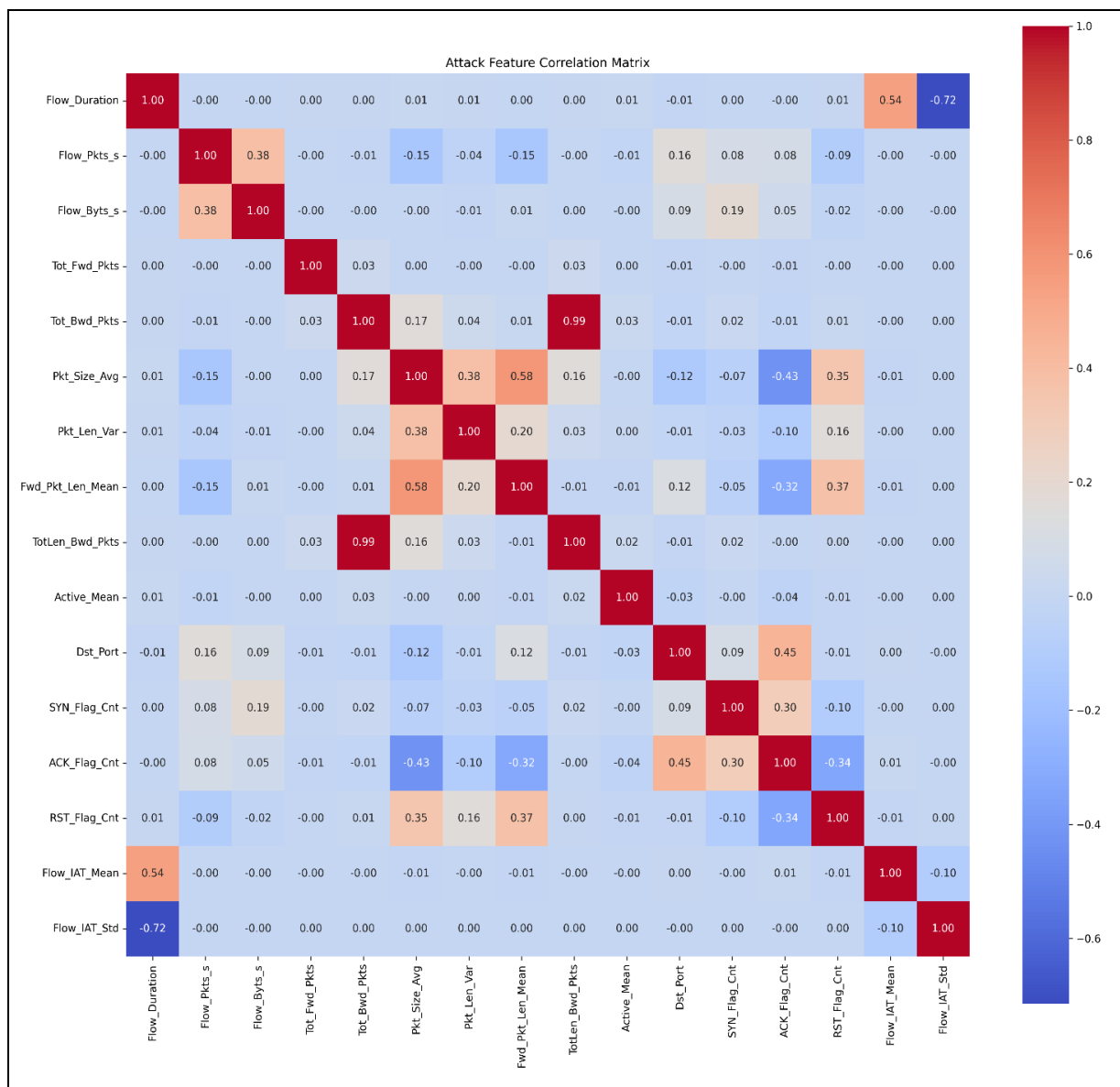


Fig 5 Correlation Heatmap of Attack Classification Features

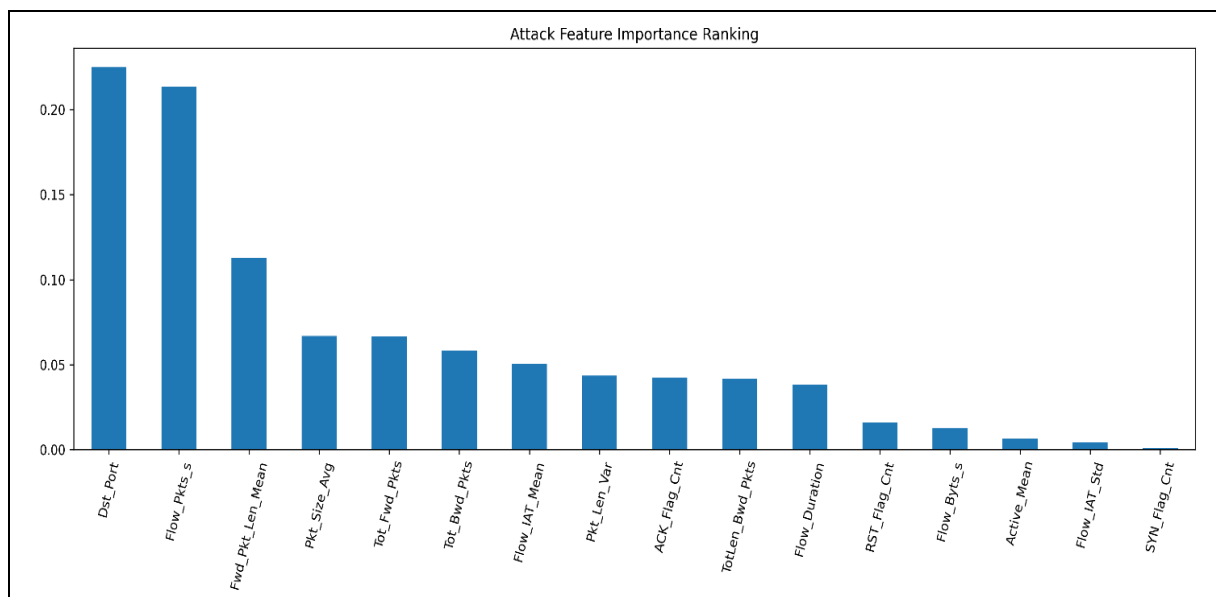


Fig 6 Importance Ranking of Attack Classification Features

Since these features can describe packet transmission rate, flow volume, and packet size characteristics, they were found to be important for network traffic behavior analysis.

In the attack classification feature analysis, *Dst_Port* was found to be the highest-ranked feature, with an importance value of 0.224968. In addition, *Flow_Pkts_s*, *Fwd_Pkt_Len_Mean*, *Pkt_Size_Avg*, *Tot_Fwd_Pkts*, and *Tot_Bwd_Pkts* were also important features in describing attack traffic characteristics that were shown in figure 6.

These feature importance analysis results can be used as basic information for selecting suitable features for machine learning and deep learning-based intrusion detection models to be built in future work.

V. CONCLUSION

This study presented a two-stage network traffic feature extraction framework to build a foundational feature representation for network intrusion detection. The proposed framework is intended to effectively identify the general behavior of network traffic and attack-related characteristics based on the CSE-CIC-IDS2018 dataset.

In the proposed framework, feature extraction was performed in two stages: behavioral anomaly analysis and attack classification analysis. In the first stage, 22 features, including flow-based, packet-based, and timing-based features, were extracted to represent network traffic behavior. In the second stage, 16 features, including destination port, packet statistics, flow rate, and TCP flag information, were selected to describe the distinctive characteristics of attack traffic.

Statistical analysis, correlation analysis, and feature importance analysis were carried out to study the quality and importance of the extracted features. The statistical analysis results showed that features such as flow duration, flow rate, and inter-arrival time can effectively describe the variation in network traffic behavior. In addition, the correlation analysis made it possible to identify relationships and redundant information among features, providing important information for feature selection and optimization.

According to the results obtained from the feature importance analysis, flow-based features, packet transmission features, and features describing network communication characteristics were found to be important for network traffic analysis. Therefore, the proposed feature extraction framework can extract relevant feature representations from large-scale network traffic datasets and can be used as a solid preprocessing framework for building AI-based intrusion detection models in future work.

REFERENCES

- [1]. S. Z. M. Shaqadan, M. Al-Zewairi, and M. Alawida, "A survey of cyber security threats and defense mechanisms," *IEEE Access*, vol. 10, pp. 123456–123478, 2022.
- [2]. R. Sommer and V. Paxson, "Outside the closed world: On using machine learning for network intrusion detection," in *Proc. IEEE Symposium on Security and Privacy*, Oakland, CA, USA, 2010, pp. 305–316.
- [3]. S. Axelsson, "Intrusion detection systems: A survey and taxonomy," Technical Report No. 99-15, Department of Computer Engineering, Chalmers University of Technology, Sweden, 2000.
- [4]. A. Khraisat, I. Gondal, P. Vamplew, and J. Kamruzzaman, "Survey of intrusion detection systems: Techniques, datasets and challenges," *Cybersecurity*, vol. 2, no. 1, pp. 1–22, 2019.
- [5]. Y. Xin, L. Kong, Z. Liu, Y. Chen, Y. Li, H. Zhu, M. Gao, H. Hou, and C. Wang, "Machine learning and deep learning methods for cybersecurity," *IEEE Access*, vol. 6, pp. 35365–35381, 2018.
- [6]. A. Sharafaldin, H. Gharib, and A. A. Ghorbani, "Toward generating a new intrusion detection dataset and intrusion traffic characterization," in *Proc. 4th International Conference on Information Systems Security and Privacy (ICISSP)*, Portugal, 2018, pp. 108–116.
- [7]. A. L. Buczak and E. Guven, "A survey of data mining and machine learning methods for cyber security intrusion detection," *IEEE Communications Surveys & Tutorials*, vol. 18, no. 2, pp. 1153–1176, 2016.
- [8]. I. Sharafaldin, A. H. Lashkari, and A. A. Ghorbani, "Toward generating a new intrusion detection dataset and intrusion traffic characterization," in *Proc. International Conference on Information Systems Security and Privacy (ICISSP)*, 2018.
- [9]. M. Tavallaei, E. Bagheri, W. Lu, and A. A. Ghorbani, "A detailed analysis of the KDD CUP 99 data set," in *Proc. IEEE Symposium on Computational Intelligence for Security and Defense Applications*, Ottawa, Canada, 2009, pp. 1–6.
- [10]. L. Dhanabal and S. P. Shantharajah, "A study on NSL-KDD dataset for intrusion detection system based on classification algorithms," *International Journal of Advanced Research in Computer and Communication Engineering*, vol. 4, no. 6, pp. 446–452, 2015.
- [11]. A. Sharafaldin, A. H. Lashkari, and A. A. Ghorbani, "Developing realistic distributed denial of service (DDoS) attack dataset and taxonomy," in *Proc. IEEE International Conference on Information Systems Security and Privacy*, 2019.
- [12]. R. Vinayakumar, K. P. Soman, P. Poornachandran, and S. Sachin Kumar, "Deep learning approaches for cyber security intrusion detection systems," in *Proc. IEEE International Conference on Advances in Computing, Communications and Informatics (ICACCI)*, 2018, pp. 2626–2632.