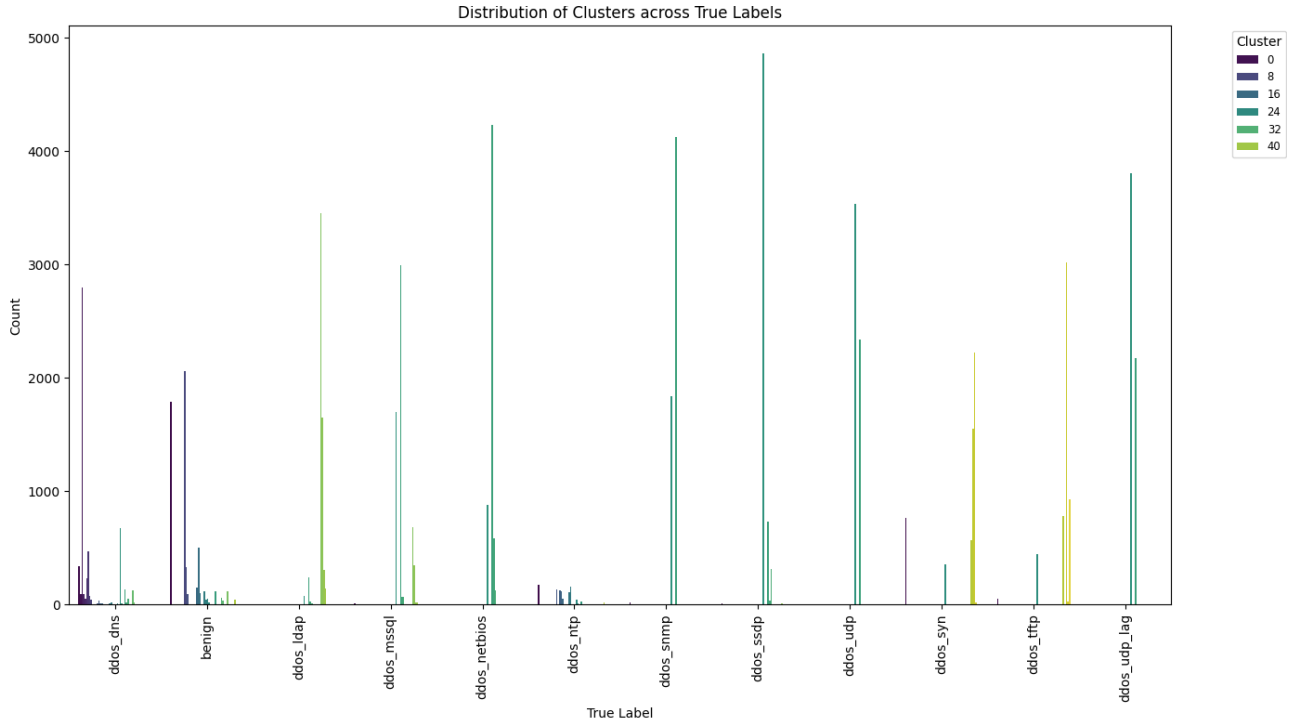
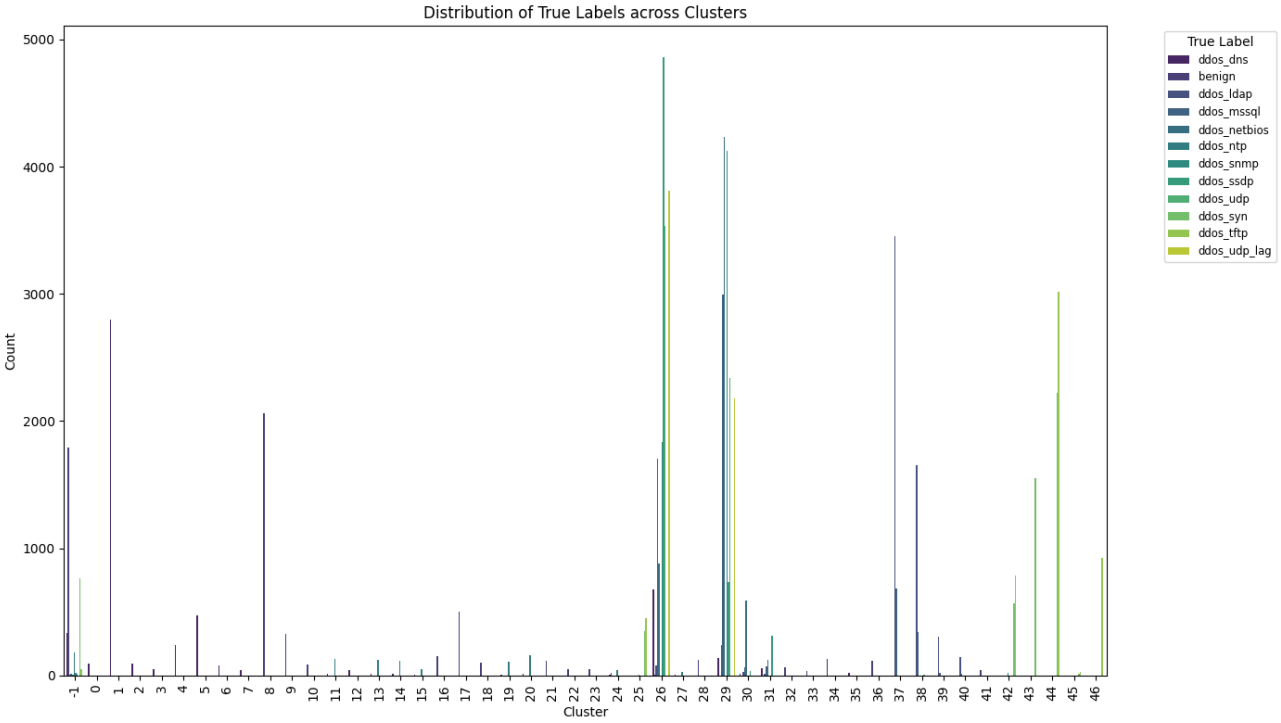
# DBscan

***Distribution of clusters in relation to the ground truth***

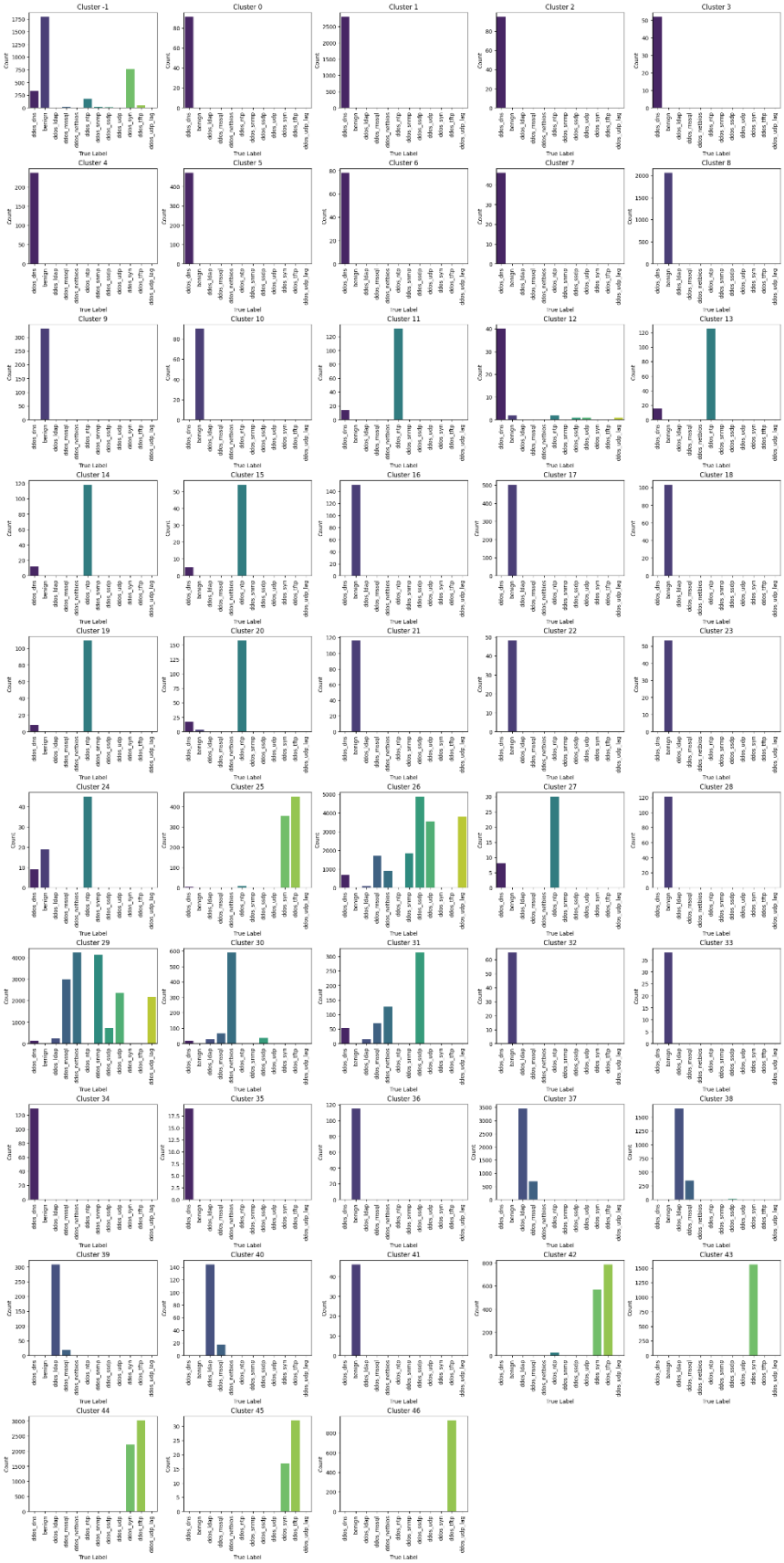
As already said, DBSCAN identifies clusters based on the density of points, with points in high-density regions forming clusters and points in low-density regions being labeled as noise. The eps parameter (0.77 in our case) sets the maximum distance between points in a cluster, and min_samples (30) sets the minimum number of points required to form a cluster. In DBSCAN, any point that does not belong to any cluster is labeled as -1. These points are considered noise because they are not within a dense region of points. By labeling and removing noise points, DBSCAN helps in forming more accurate and meaningful clusters. Noise points, if included in clusters, can distort the true structure of the data.
As already done with k-means, the figure below shows the distribution of various clusters in relation to the ground truth labels.



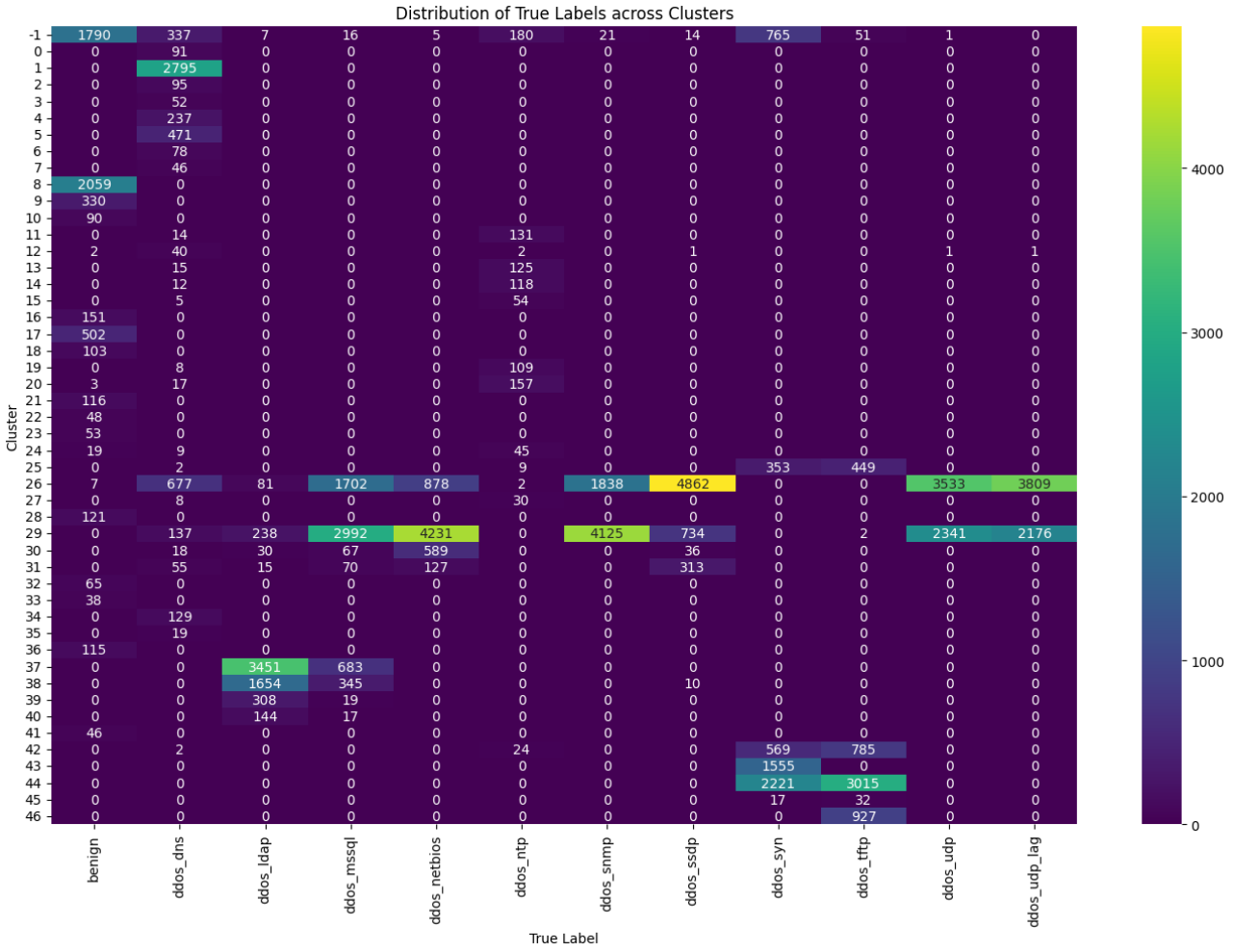

Mirroring the analysis with K-means, the labels composed of a mix of different clusters again are benign and ddos_dns, but with the addition of ddos_ntp, which, however, shows a significant decrease in the number of samples involved (y-axis) compared with the first two.




In the second figure, the one related to the inverse visualization, sizable clusters that contain only one "pure" label are: cluster 1, that contains only “ddos_dns” labels, cluster 8, that contains only “benign” ones and cluster 43, that contain only “ddos_syn”.


Cluster -1
Cluster 0
Cluster 1
Cluster 2
Cluster 3
Cluster 4
Cluster 5
Cluster 6
Cluster 7
Cluster 8
Cluster 9
Cluster 10
Cluster 11
Cluster 12
Cluster 13
Cluster 14
Cluster 15
Cluster 16
Cluster 17
Cluster 18
Cluster 19
Cluster 20
Cluster 21
Cluster 22
Cluster 23
Cluster 24
Cluster 25
Cluster 26
Cluster 27
Cluster 28
Cluster 29
Cluster 30
Cluster 31
Cluster 32
Cluster 33
Cluster 34
Cluster 35
Cluster 36
Cluster 37
Cluster 38
Cluster 39
Cluster 40
Cluster 41
Cluster 42
Cluster 43
Cluster 44
Cluster 45
Cluster 46
Count
True Label
ddos_dns
benign
ddos_ldap
ddos_mssql
ddos_netbios
ddos_ntp
ddos_snmp
ddos_ssdp
ddos_udp
ddos_syn
ddos_tftp
ddos_udp_lag

### *Contingency table*




The contingency table shows the distribution of true labels within each cluster, including the noise points. We proceed as we did with K-means.

*Pure (or nearly pure) clusters:*

| *Cluster* | 0 | 1 | 2 | 3 | 4 | 5 | 6 |
|---|---|---|---|---|---|---|---|
| *Label* | ddos_dns | ddos_dns | ddos_dns | ddos_dns | ddos_dns | ddos_dns | ddos_dns |

| *Cluster* | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 |
|---|---|---|---|---|---|---|---|---|---|
| *Label* | ddos_dns | benign | benign | benign | ddos_ntp | ddos_dns | ddos_ntp | ddos_ntp | ddos_ntp |

| *Cluster* | 16 | 17 | 18 | 19 |
|---|---|---|---|---|
| *Label* | benign | benign | benign | ddos_ntp |

| *Cluster* | 20 | 21 | 22 | 23 | 24 | 27 | 28 | 32 |
|---|---|---|---|---|---|---|---|---|
| *Label* | ddos_ntp | benign | benign | benign | ddos_ntp | ddos_ntp | benign | benign |

| *Cluster* | 33 | 34 | 35 | 36 | 41 | 43 | 46 |
|---|---|---|---|---|---|---|---|
| *Label* | benign | ddos_dns | ddos_dns | benign | benign | ddos_syn | ddos_tftp |

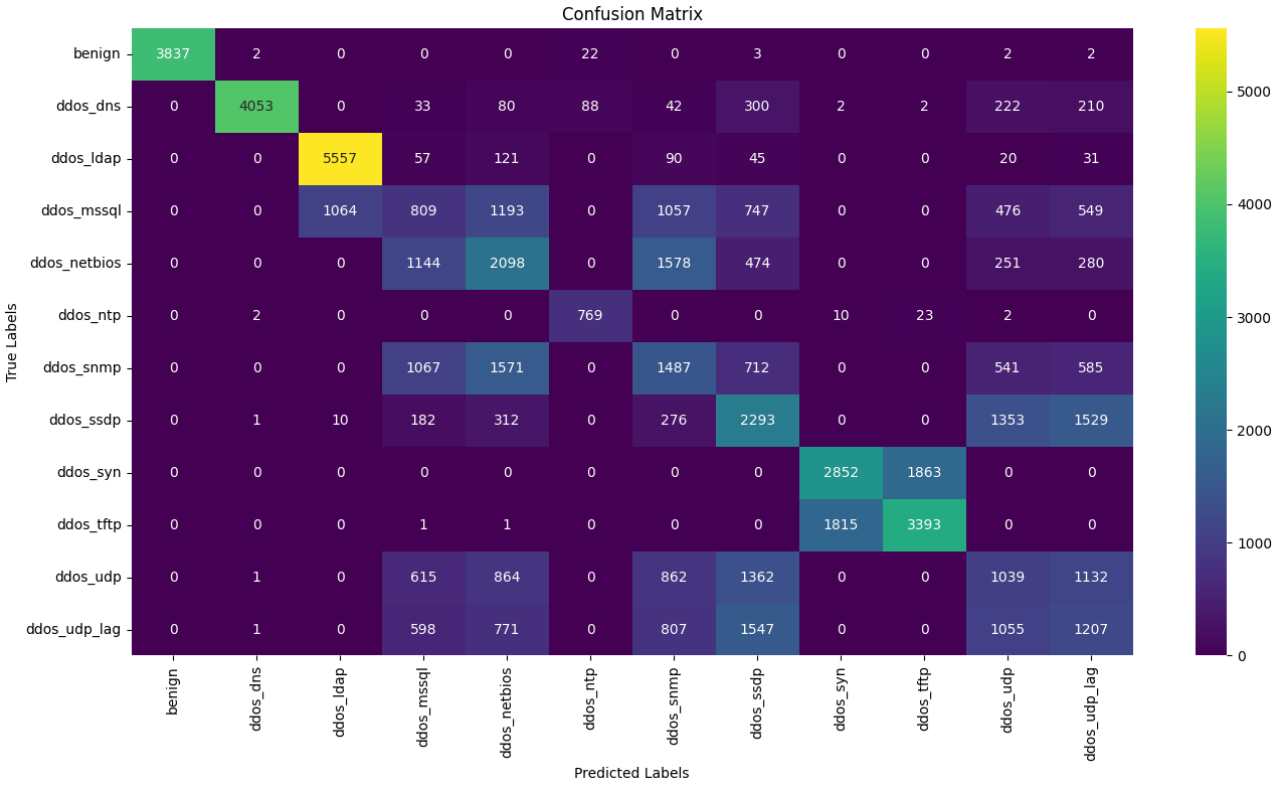
It is important to note that there are two clusters (26 and 29) that capture an exceptionally high number of samples, highlighting the model's difficulty in identifying well-defined clusters. Further in-depth analysis on this matter will be conducted subsequently.

Now, to allow the analysis to focus on well-defined clusters and have an interpretation of the results is easier and more reliable, the noise has been removed from the DataFrame.

Given the relative frequencies of true labels within each cluster, a probabilistic approach was adopted to assign labels to the clusters. This process involved several key steps:

-Calculating Relative Frequencies: A contingency table was created from the DataFrame to show the distribution of true labels within each cluster. The relative frequencies of these labels were then calculated, providing a distribution that indicated how likely each label was to appear in a given cluster.
-Filtering and Normalizing Probabilities: To ensure meaningful probabilistic assignment, only labels that represented at least 60% of the maximum frequency within a cluster were considered. These probabilities were then normalized to sum to one.
-Assigning Labels Probabilistically: For each cluster, labels were assigned based on the filtered and normalized probabilities. This method allowed for a nuanced and statistically informed assignment of labels, particularly useful for clusters containing a mix of true labels.
-Evaluation of Refined Clusters: To assess the performance of the probabilistic label assignment, the confusion matrix below was created to compare the true labels with the predicted labels assigned probabilistically. This matrix provided a detailed view of the clustering performance, highlighting areas of agreement and discrepancy between the true and assigned labels.

### *Cluster Mapping*




Here, it is evident that certain clusters, such as those labeled as benign, ddos_dns, and ddos_ldap, exhibit high purity with a significant majority of points correctly labeled. In contrast, other clusters have been assigned labels with varying degrees of accuracy, indicating a mix of true labels within those clusters:

Classification Report:

| | precision | recall | f1-score | support |
|---|---|---|---|---|
| benign | 1.00 | 0.99 | 1.00 | 3868 |
| ddos_dns | 1.00 | 0.81 | 0.89 | 5032 |
| ddos_ldap | 0.84 | 0.94 | 0.89 | 5921 |
| ddos_mssql | 0.18 | 0.14 | 0.16 | 5895 |
| ddos_netbios | 0.31 | 0.38 | 0.34 | 5825 |
| ddos_ntp | 0.87 | 0.95 | 0.91 | 806 |
| ddos_snmp | 0.25 | 0.26 | 0.25 | 5963 |
| ddos_ssdp | 0.30 | 0.38 | 0.34 | 5956 |
| ddos_syn | 0.61 | 0.61 | 0.61 | 4715 |
| ddos_tftp | 0.64 | 0.65 | 0.64 | 5210 |
| ddos_udp | 0.21 | 0.18 | 0.19 | 5875 |
| ddos_udp_lag | 0.22 | 0.20 | 0.21 | 5986 |
| accuracy | | | 0.48 | 61052 |
| macro avg | 0.54 | 0.54 | 0.54 | 61052 |
| weighted avg | 0.48 | 0.48 | 0.48 | 61052 |

Clusters to Label Mapping with Weights:

| Cluster | Labels with Weights |
|---|---|
| 0 | ddos_dns: 100.00% |
| 1 | ddos_dns: 100.00% |
| 10 | benign: 100.00% |
| 11 | ddos_ntp: 100.00% |
| 12 | ddos_dns: 100.00% |
| 13 | ddos_ntp: 100.00% |
| 14 | ddos_ntp: 100.00% |
| 15 | ddos_ntp: 100.00% |
| 16 | benign: 100.00% |
| 17 | benign: 100.00% |
| 18 | benign: 100.00% |
| 19 | ddos_ntp: 100.00% |
| 2 | ddos_dns: 100.00% |
| 20 | ddos_ntp: 100.00% |
| 21 | benign: 100.00% |
| 22 | benign: 100.00% |
| 23 | benign: 100.00% |
| 24 | ddos_ntp: 100.00% |
| 25 | ddos_syn: 44.01%, ddos_tftp: 55.99% |
| 26 | ddos_ssdp: 39.84%, ddos_udp: 28.95%, ddos_udp_lag: 31.21% |
| 27 | ddos_ntp: 100.00% |
| 28 | benign: 100.00% |
| 29 | ddos_mssql: 26.37%, ddos_netbios: 37.28%, ddos_snmp: 36.35% |
| 3 | ddos_dns: 100.00% |
| 30 | ddos_netbios: 100.00% |
| 31 | ddos_ssdp: 100.00% |
| 32 | benign: 100.00% |
| 33 | benign: 100.00% |
| 34 | ddos_dns: 100.00% |
| 35 | ddos_dns: 100.00% |
| 36 | benign: 100.00% |
| 37 | ddos_ldap: 100.00% |
| 38 | ddos_ldap: 100.00% |
| 39 | ddos_ldap: 100.00% |
| 4 | ddos_dns: 100.00% |
| 40 | ddos_ldap: 100.00% |
| 41 | benign: 100.00% |
| 42 | ddos_syn: 42.02%, ddos_tftp: 57.98% |
| 43 | ddos_syn: 100.00% |
| 44 | ddos_syn: 42.42%, ddos_tftp: 57.58% |
| 45 | ddos_tftp: 100.00% |
| 46 | ddos_tftp: 100.00% |
| 5 | ddos_dns: 100.00% |
| 6 | ddos_dns: 100.00% |
| 7 | ddos_dns: 100.00% |
| 8 | benign: 100.00% |
| 9 | benign: 100.00% |

This classification report provides a comprehensive performance evaluation, indicating high precision and recall for certain labels such as benign, ddos_dns, and ddos_ldap, while showing lower performance for labels such as ddos_mssql, ddos_snmp, and ddos_udp. Specifically, the benign label achieved near-perfect precision and recall, reflecting the model's strong ability to correctly identify benign instances with minimal false positives and negatives. The ddos_dns and ddos_ldap labels also performed well, with high precision and recall, suggesting that these types of attacks were distinctly recognized by the clustering model. Conversely, labels like ddos_mssql, ddos_snmp, and ddos_udp exhibited significantly lower performance, with both precision and recall values indicating higher rates of misclassification. This disparity suggests potential overlaps or confusion among these attack types within the feature space. The overall accuracy of the model was 48%, which, while moderate, highlights the challenges in clustering and classifying complex datasets with multiple attack types. The macro average F1-score of 0.54 and the weighted average F1-score of 0.48 reflect the model's varied performance across different labels, with the weighted average taking into account the support (number of true instances) for each label. In summary, DB-scan performed significantly worse than K-means.

Additionally, the table displaying the weights of labels for each cluster shows a high degree of certainty for many clusters, such as those entirely composed of benign or ddos_dns labels.

For instance, clusters 10, 16, 17, 18, 21, 22, 23, 28, 32, 33, 36, 41, and 8 were all composed entirely of benign points, indicating a strong, uniform distribution and high confidence in the label assignment for these clusters. Similarly, clusters 0, 1, 2, 3, 4, 5, 6, 7, 12, and 34 were entirely composed of ddos_dns points, reflecting a clear and distinct clustering of this attack type.

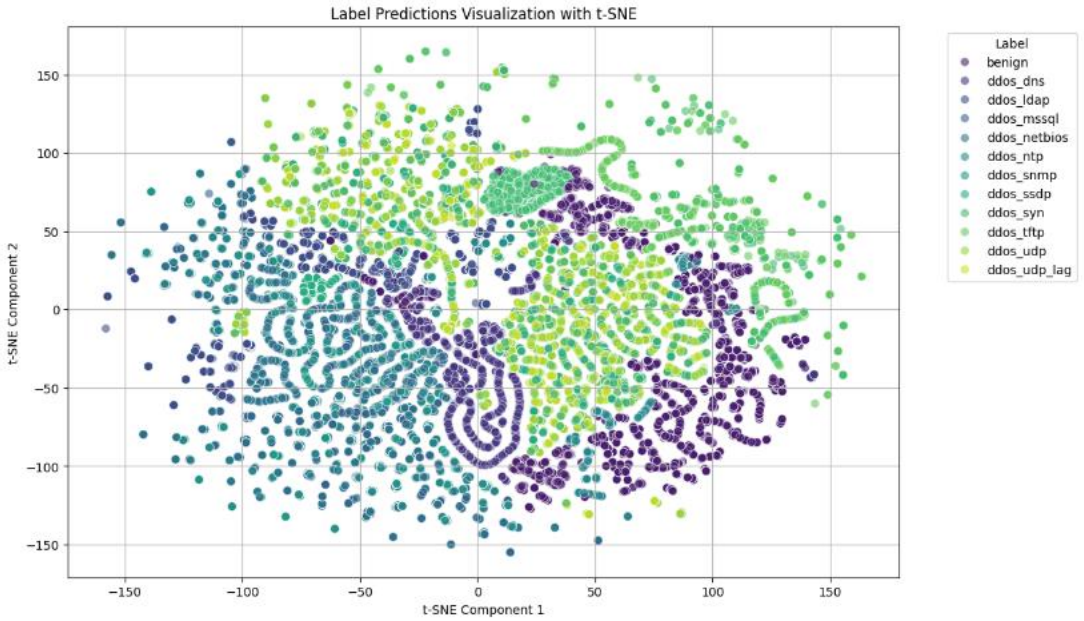
However, some clusters contained a mix of labels, reflecting the probabilistic nature of the assignment method and the inherent overlap in feature space among certain attack types. For example, clusters 25, 42, and 44 all contained a mix of ddos_syn and ddos_tftp labels, with weights of 44.01% and 55.99%, 42.02% and 57.98%, and 42.42% and 57.58%, respectively. The similarities observed across these clusters indicate the presence of common characteristics that can induce confusion in the clustering process. The nearly identical distributions of ddos_syn and ddos_tftp labels in these clusters suggest that these attack types share significant similarities in the dataset's feature space. This overlap could be due to several factors, including feature overlap, where the features used for clustering might not be distinctive enough to separate ddos_syn from ddos_tftp, as both attack types might exhibit similar behavior in terms of packet size, frequency, or other network characteristics. Additionally, similar traffic patterns might be followed by ddos_syn and ddos_tftp attacks, making it difficult for the clustering algorithm to differentiate between them. For instance, both might target the same network vulnerabilities or use comparable methods to overwhelm network resources.

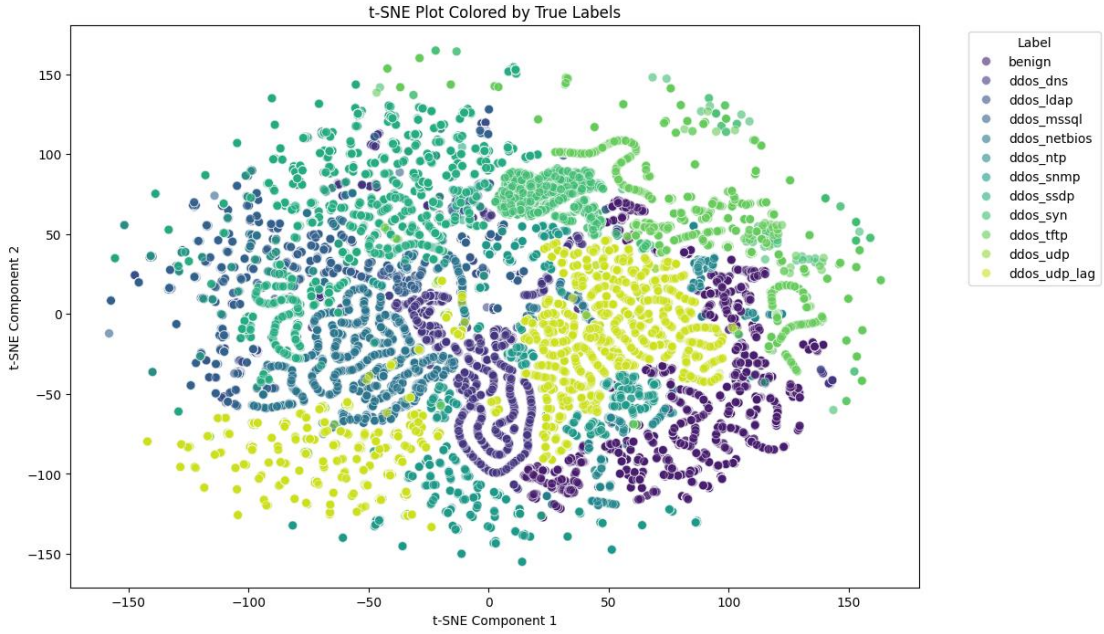
Cluster 26 displayed a more complex mix, with ddos_ssdp, ddos_udp, and ddos_udp_lag labels distributed at 39.84%, 28.95%, and 31.21%, respectively, while cluster 29 exhibited a mixed distribution with ddos_mssql, ddos_netbios, and ddos_snmp labels at 26.37%, 37.28%, and 36.35%, respectively.

This mapping of clusters to labels with their respective weights provides a view of the reliability and distribution of labels within the clusters. Clusters with uniform distributions demonstrate the model's effectiveness in distinctly identifying certain attack types or benign traffic. In contrast, mixed clusters reveal areas where the feature space might be shared among different labels.
Further analysis will be conducted to investigate these findings in more detail.

***TSNE***

To obtain a visual representation that contrasts the predicted label mapping with the actual labels, we applied the t-SNE technique to both of them:



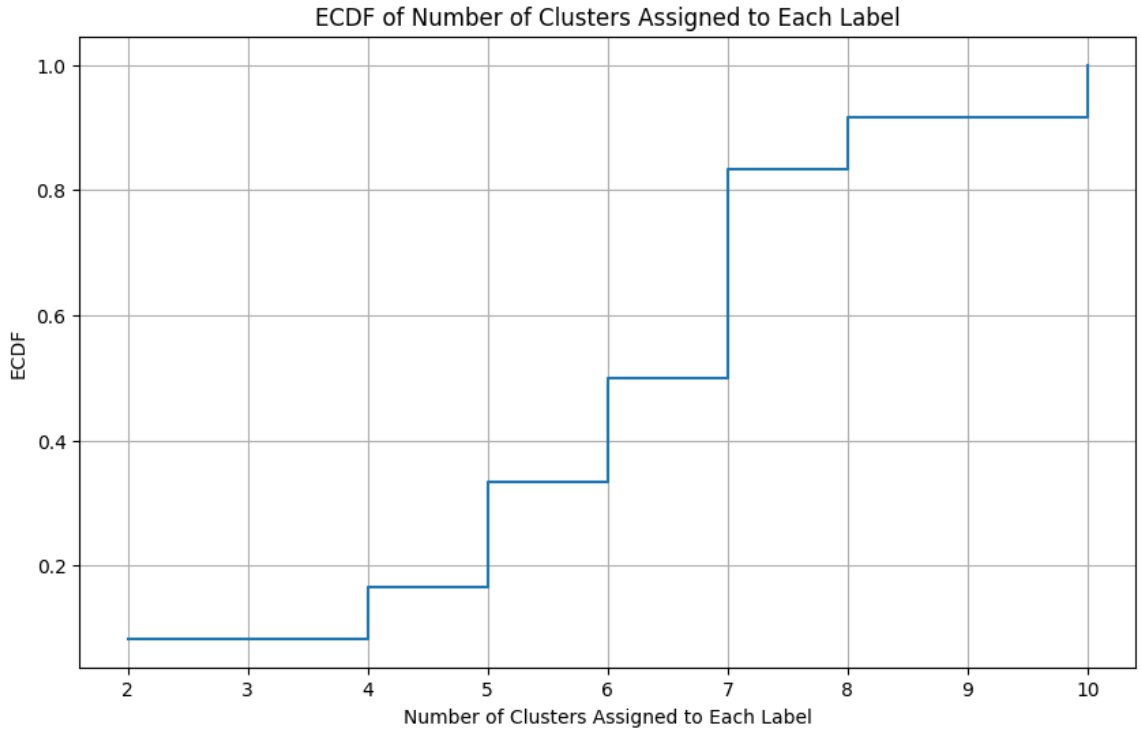

By comparing these plots, we can identify areas where DBSCAN performed well and where it encountered difficulties.
In both plots, benign traffic (colored purple) appears as distinct and relatively compact clusters. This indicates that DBSCAN was able to effectively identify benign traffic, which is characterized by its uniform behavior, making it easier to separate from other types of traffic, the clusters of ddos_dns and ddos_ldap are also well-defined in both the predicted and true label plots. DBSCAN successfully captured these attack types, which suggests that they have distinctive features that the algorithm can easily recognize.

Areas of Misclassification:

The t-SNE plot of predicted labels shows significant overlap between ddos_mssql ddos_snmp, and ddos_udp This overlap is less pronounced in the plot of true labels, indicating that DBSCAN struggled to differentiate between these types of attacks. This misclassification likely arises from similar features in these attacks, leading to confusion in the clustering process.

Clusters for ddos_syn and ddos_tftp exhibit substantial overlap in the predicted labels plot. The near-identical distribution in the predicted plot indicates common characteristics that might not be adequately differentiated by the current feature set.

***ECDF of number of cluster assigned to each label***

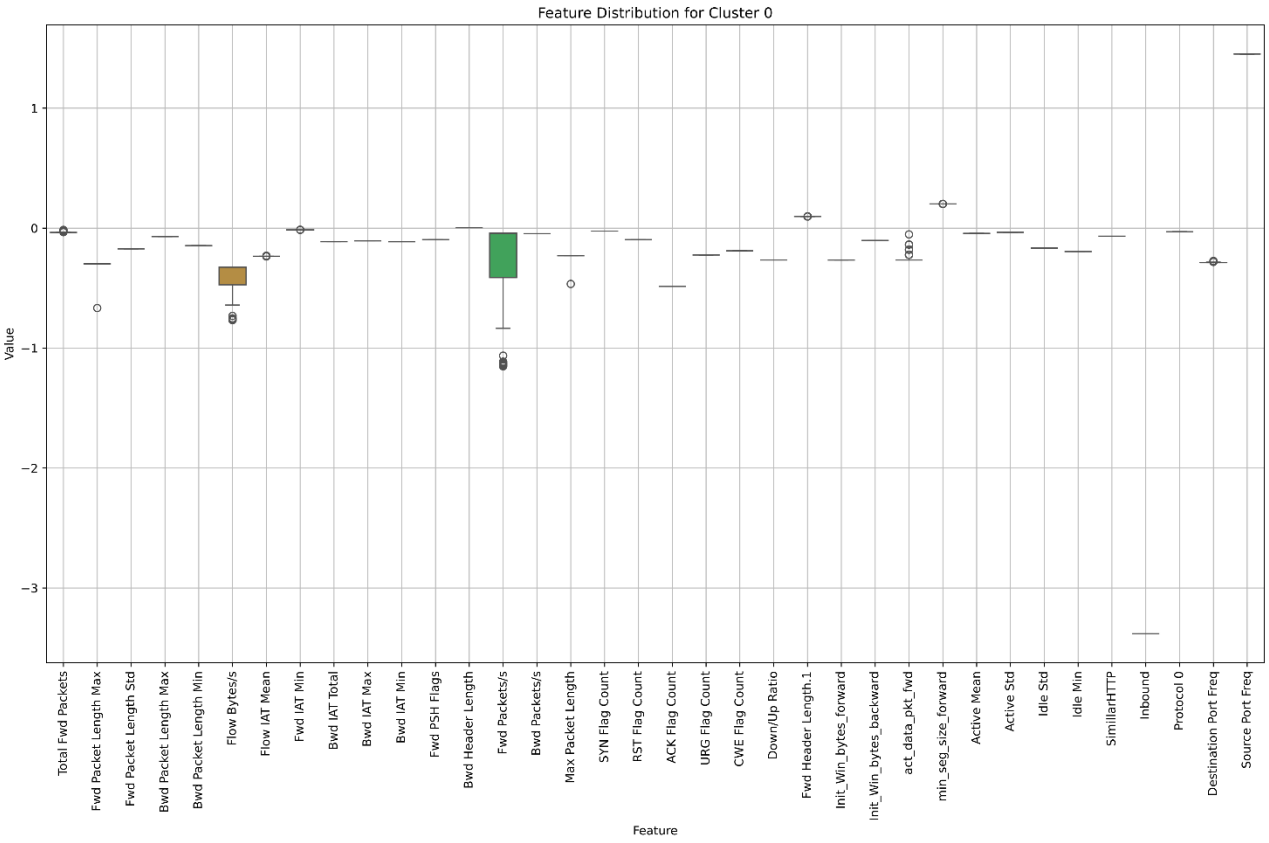


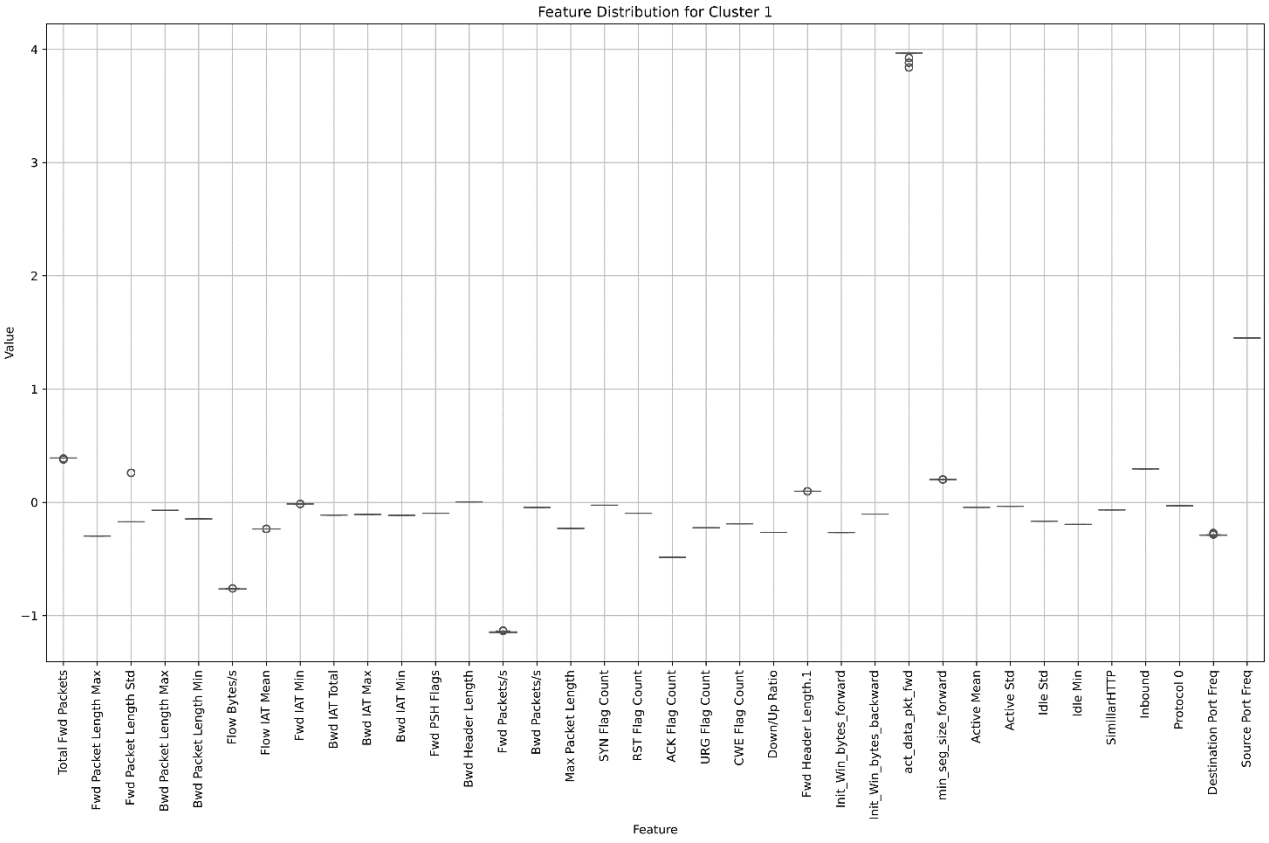

Approximately 20% of the labels are assigned to as few as 2 clusters, while around 60% of the labels are assigned to 6 or fewer clusters. This suggests that the majority of the labels are distributed among a moderate number of clusters.

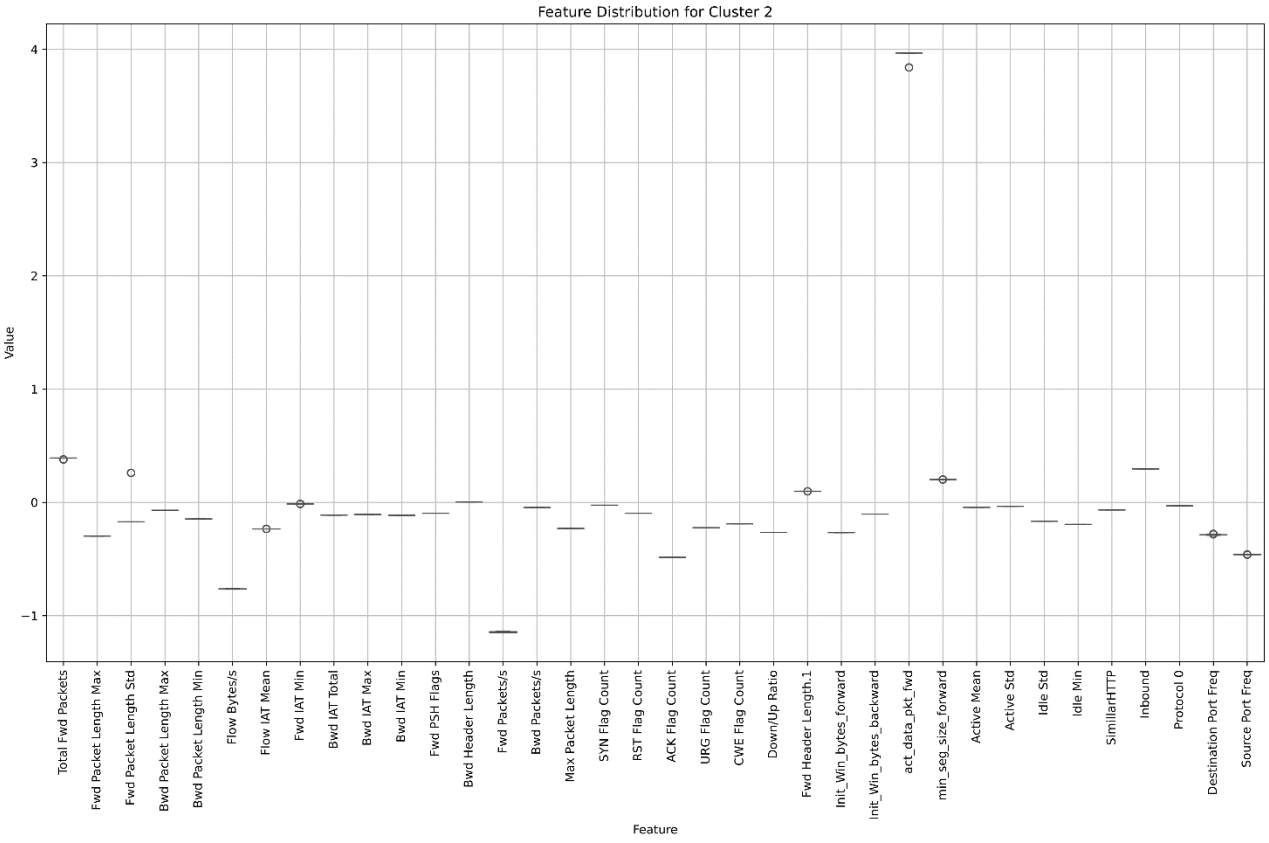
***Features importance***

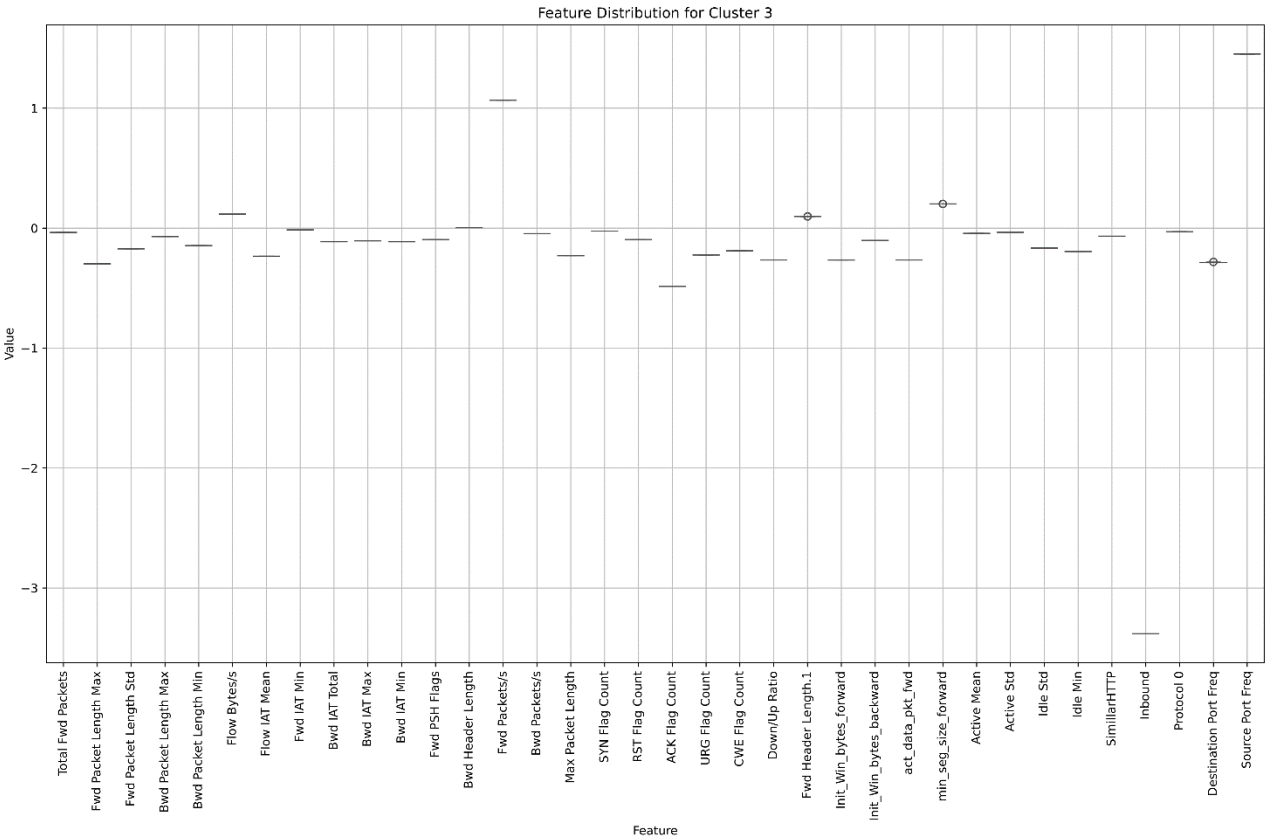
As done with k-means, the two methods used for feature analysis are multi-class classification and intra-cluster variable similarity

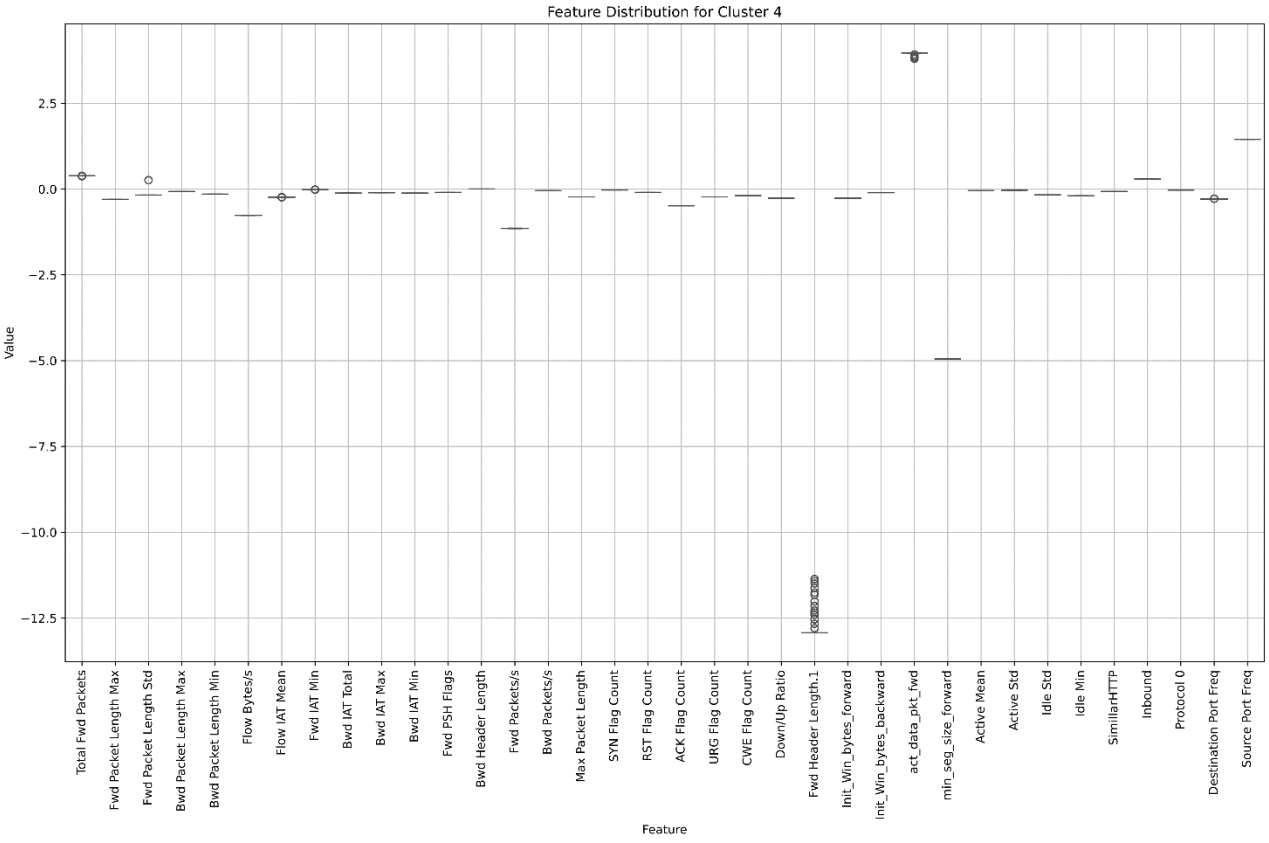
1. The intra-cluster variable similarity analysis using box plots for each cluster provides a detailed visualization of how different features vary within each cluster. This helps in identifying the most important features that contribute to the clustering process, understanding the distinct characteristics of each cluster, and pinpointing areas where the clustering algorithm might need improvement.

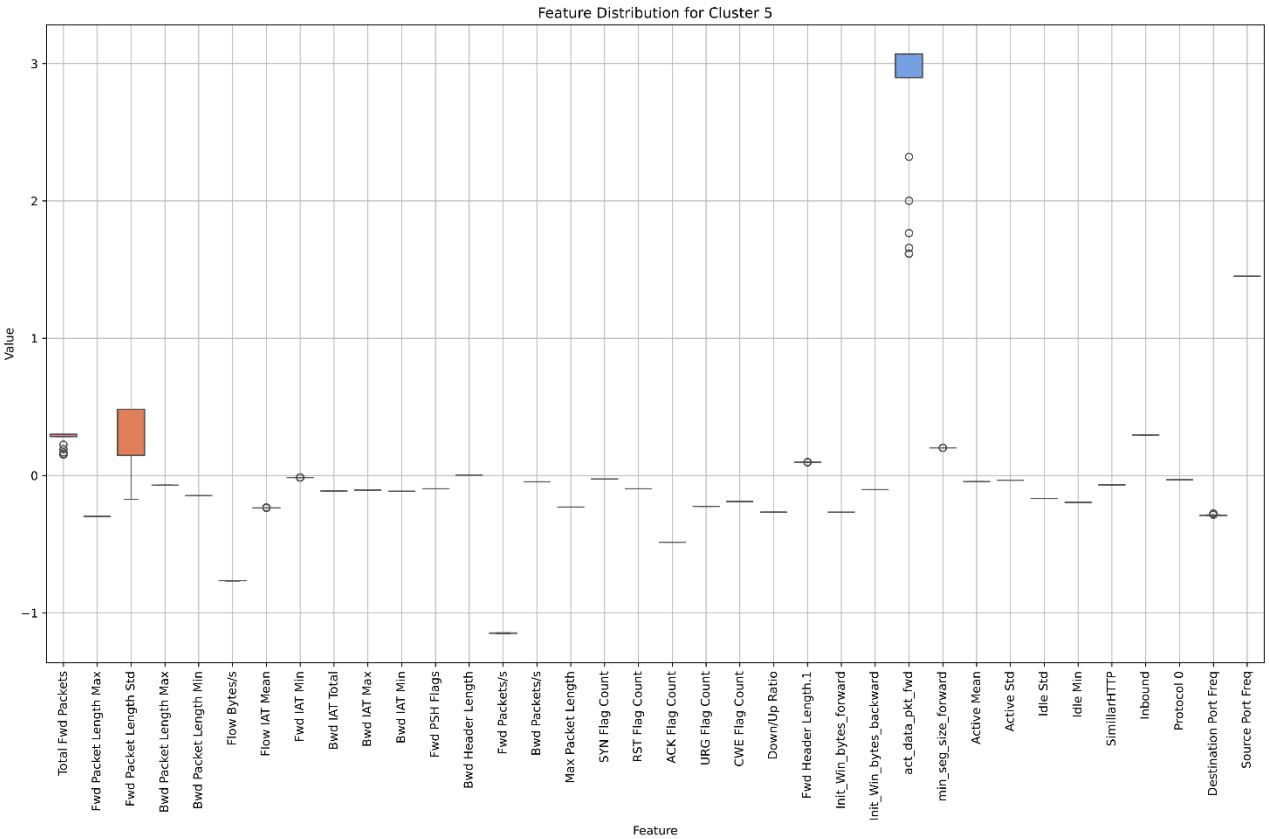

Feature Distribution for Cluster 0
Value
Feature
Total Fwd Packets
Fwd Packet Length Max
Fwd Packet Length Std
Bwd Packet Length Max
Bwd Packet Length Min
Flow Bytes/s
Flow IAT Mean
Fwd IAT Min
Bwd IAT Total
Bwd IAT Max
Bwd IAT Min
Fwd PSH Flags
Bwd Header Length
Fwd Packets/s
Bwd Packets/s
Max Packet Length
SYN Flag Count
RST Flag Count
ACK Flag Count
URG Flag Count
CWE Flag Count
Down/Up Ratio
Fwd Header Length.1
Init_Win_bytes_forward
Init_Win_bytes_backward
act_data_pkt_fwd
min_seg_size_forward
Active Mean
Active Std
Idle Std
Idle Min
SimillarHTTP
Inbound
Protocol 0
Destination Port Freq
Source Port Freq

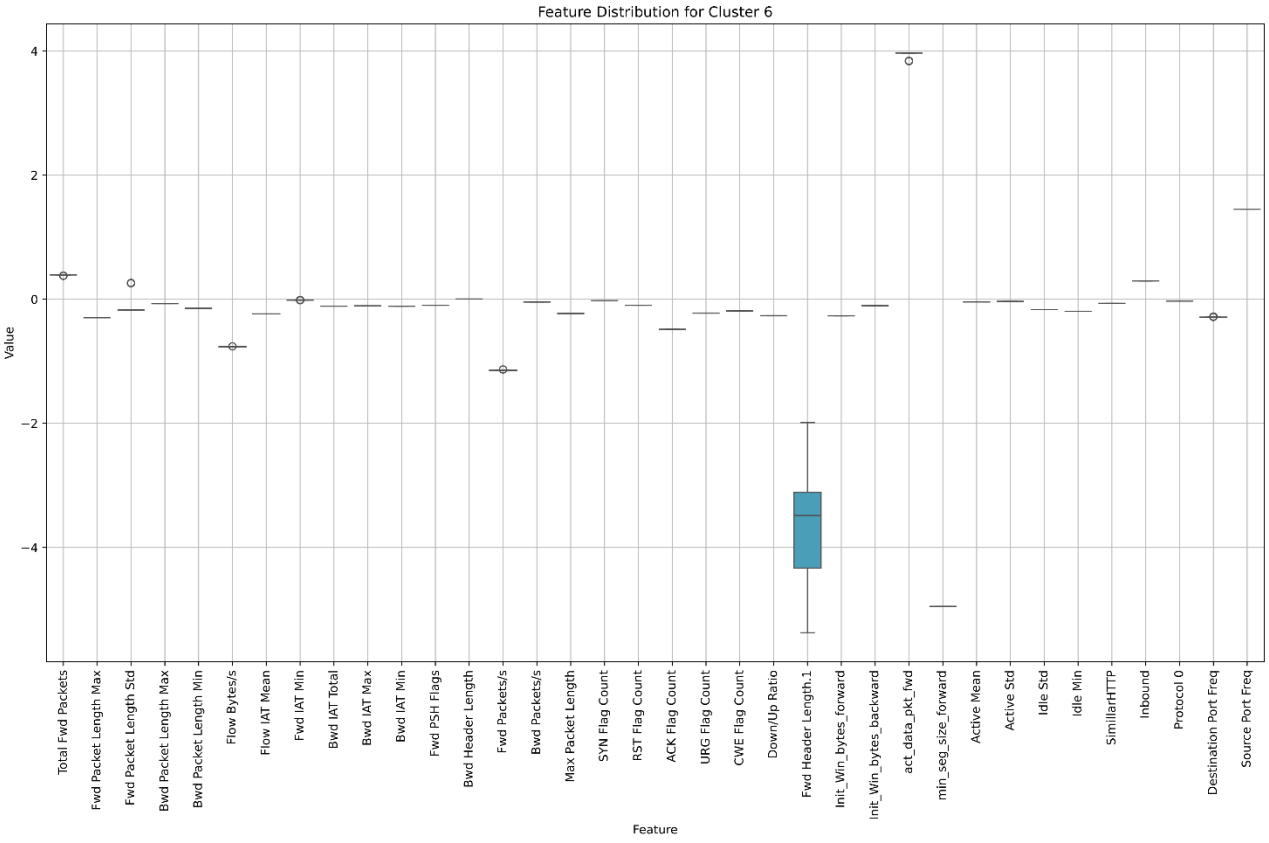

Feature Distribution for Cluster 1
Value
Feature
Total Fwd Packets
Fwd Packet Length Max
Fwd Packet Length Std
Bwd Packet Length Max
Bwd Packet Length Min
Flow Bytes/s
Flow IAT Mean
Fwd IAT Min
Bwd IAT Total
Bwd IAT Max
Bwd IAT Min
Fwd PSH Flags
Bwd Header Length
Fwd Packets/s
Bwd Packets/s
Max Packet Length
SYN Flag Count
RST Flag Count
ACK Flag Count
URG Flag Count
CWE Flag Count
Down/Up Ratio
Fwd Header Length.1
Init_Win_bytes_forward
Init_Win_bytes_backward
act_data_pkt_fwd
min_seg_size_forward
Active Mean
Active Std
Idle Std
Idle Min
SimillarHTTP
Inbound
Protocol 0
Destination Port Freq
Source Port Freq

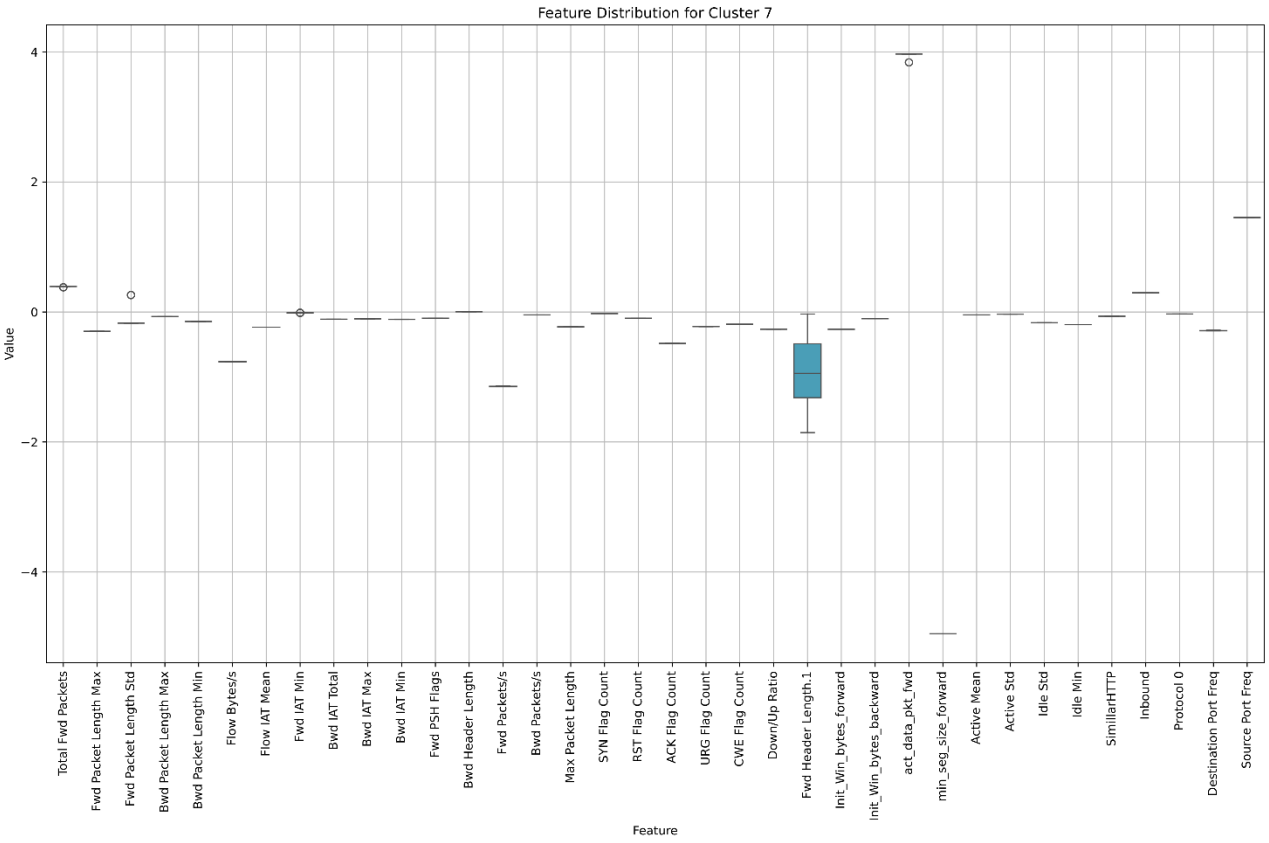

Feature Distribution for Cluster 2
Value
Feature
Total Fwd Packets
Fwd Packet Length Max
Fwd Packet Length Std
Bwd Packet Length Max
Bwd Packet Length Min
Flow Bytes/s
Flow IAT Mean
Fwd IAT Min
Bwd IAT Total
Bwd IAT Max
Bwd IAT Min
Fwd PSH Flags
Bwd Header Length
Fwd Packets/s
Bwd Packets/s
Max Packet Length
SYN Flag Count
RST Flag Count
ACK Flag Count
URG Flag Count
CWE Flag Count
Down/Up Ratio
Fwd Header Length.1
Init_Win_bytes_forward
Init_Win_bytes_backward
act_data_pkt_fwd
min_seg_size_forward
Active Mean
Active Std
Idle Std
Idle Min
SimillarHTTP
Inbound
Protocol 0
Destination Port Freq
Source Port Freq

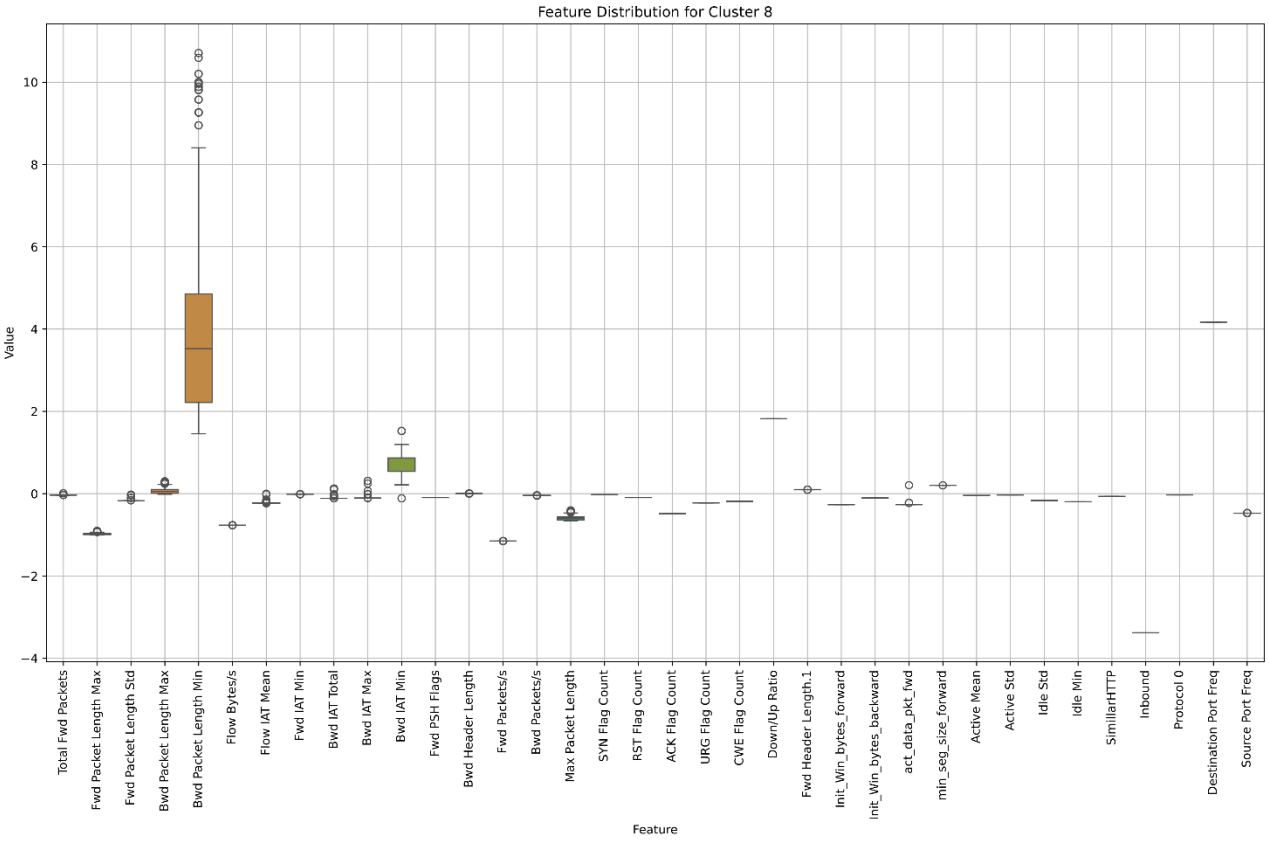

Feature Distribution for Cluster 3
Value
Feature
Total Fwd Packets
Fwd Packet Length Max
Fwd Packet Length Std
Bwd Packet Length Max
Bwd Packet Length Min
Flow Bytes/s
Flow IAT Mean
Fwd IAT Min
Bwd IAT Total
Bwd IAT Max
Bwd IAT Min
Fwd PSH Flags
Bwd Header Length
Fwd Packets/s
Bwd Packets/s
Max Packet Length
SYN Flag Count
RST Flag Count
ACK Flag Count
URG Flag Count
CWE Flag Count
Down/Up Ratio
Fwd Header Length.1
Init_Win_bytes_forward
Init_Win_bytes_backward
act_data_pkt_fwd
min_seg_size_forward
Active Mean
Active Std
Idle Std
Idle Min
SimillarHTTP
Inbound
Protocol 0
Destination Port Freq
Source Port Freq

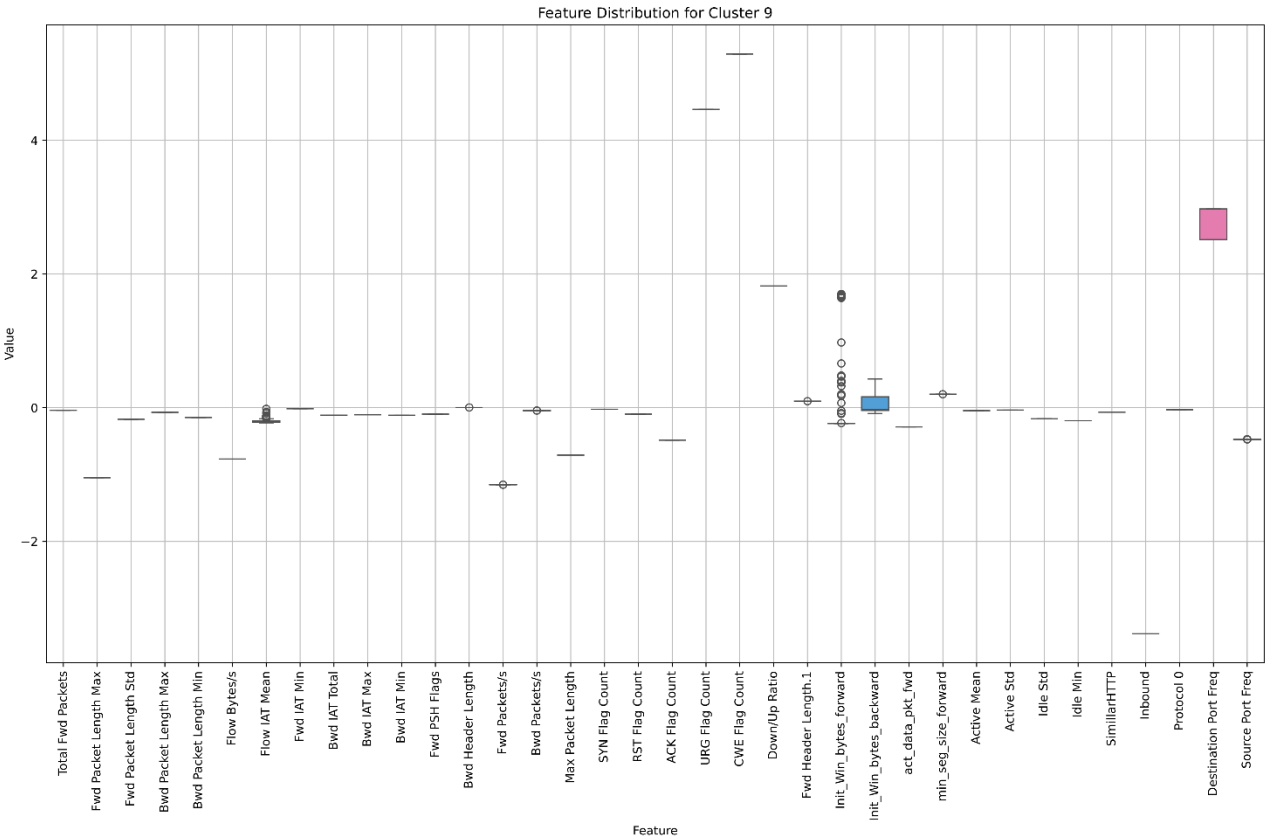

Feature Distribution for Cluster 4
Value
Feature
Total Fwd Packets
Fwd Packet Length Max
Fwd Packet Length Std
Bwd Packet Length Max
Bwd Packet Length Min
Flow Bytes/s
Flow IAT Mean
Fwd IAT Min
Bwd IAT Total
Bwd IAT Max
Bwd IAT Min
Fwd PSH Flags
Bwd Header Length
Fwd Packets/s
Bwd Packets/s
Max Packet Length
SYN Flag Count
RST Flag Count
ACK Flag Count
URG Flag Count
CWE Flag Count
Down/Up Ratio
Fwd Header Length.1
Init_Win_bytes_forward
Init_Win_bytes_backward
act_data_pkt_fwd
min_seg_size_forward
Active Mean
Active Std
Idle Std
Idle Min
SimillarHTTP
Inbound
Protocol 0
Destination Port Freq
Source Port Freq

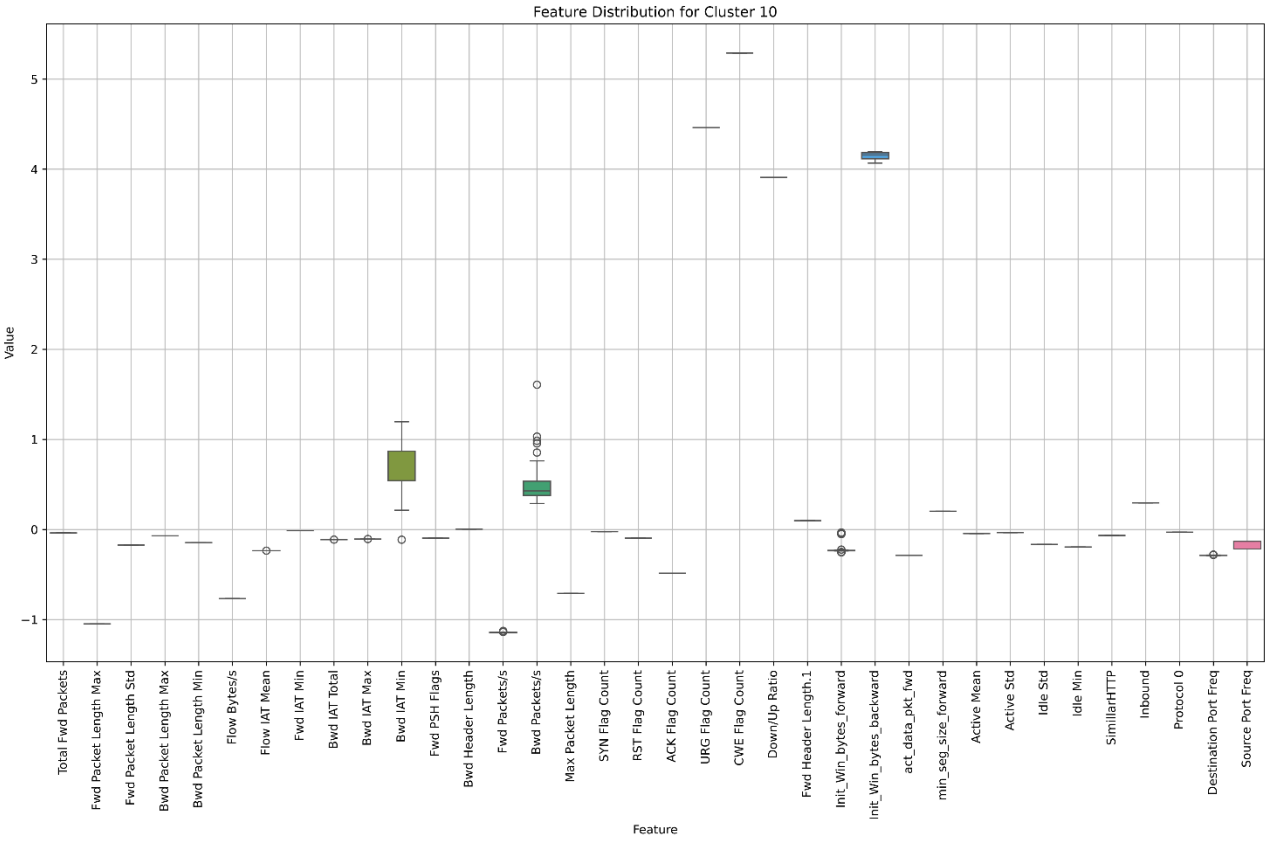

Feature Distribution for Cluster 5
Value
Feature
Total Fwd Packets
Fwd Packet Length Max
Fwd Packet Length Std
Bwd Packet Length Max
Bwd Packet Length Min
Flow Bytes/s
Flow IAT Mean
Fwd IAT Min
Bwd IAT Total
Bwd IAT Max
Bwd IAT Min
Fwd PSH Flags
Bwd Header Length
Fwd Packets/s
Bwd Packets/s
Max Packet Length
SYN Flag Count
RST Flag Count
ACK Flag Count
URG Flag Count
CWE Flag Count
Down/Up Ratio
Fwd Header Length.1
Init_Win_bytes_forward
Init_Win_bytes_backward
act_data_pkt_fwd
min_seg_size_forward
Active Mean
Active Std
Idle Std
Idle Min
SimillarHTTP
Inbound
Protocol 0
Destination Port Freq
Source Port Freq

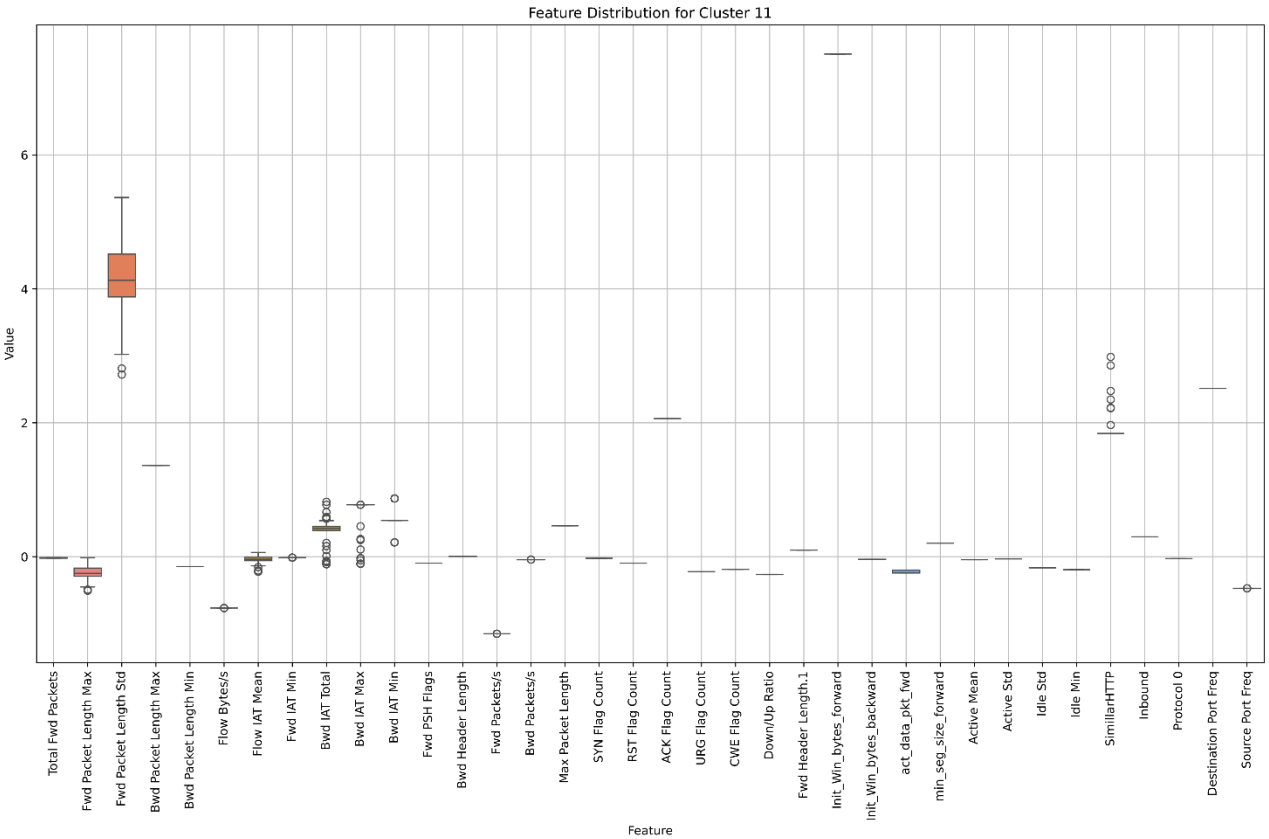

Feature Distribution for Cluster 6
Value
Feature
Total Fwd Packets
Fwd Packet Length Max
Fwd Packet Length Std
Bwd Packet Length Max
Bwd Packet Length Min
Flow Bytes/s
Flow IAT Mean
Fwd IAT Min
Bwd IAT Total
Bwd IAT Max
Bwd IAT Min
Fwd PSH Flags
Bwd Header Length
Fwd Packets/s
Bwd Packets/s
Max Packet Length
SYN Flag Count
RST Flag Count
ACK Flag Count
URG Flag Count
CWE Flag Count
Down/Up Ratio
Fwd Header Length.1
Init_Win_bytes_forward
Init_Win_bytes_backward
act_data_pkt_fwd
min_seg_size_forward
Active Mean
Active Std
Idle Std
Idle Min
SimillarHTTP
Inbound
Protocol 0
Destination Port Freq
Source Port Freq

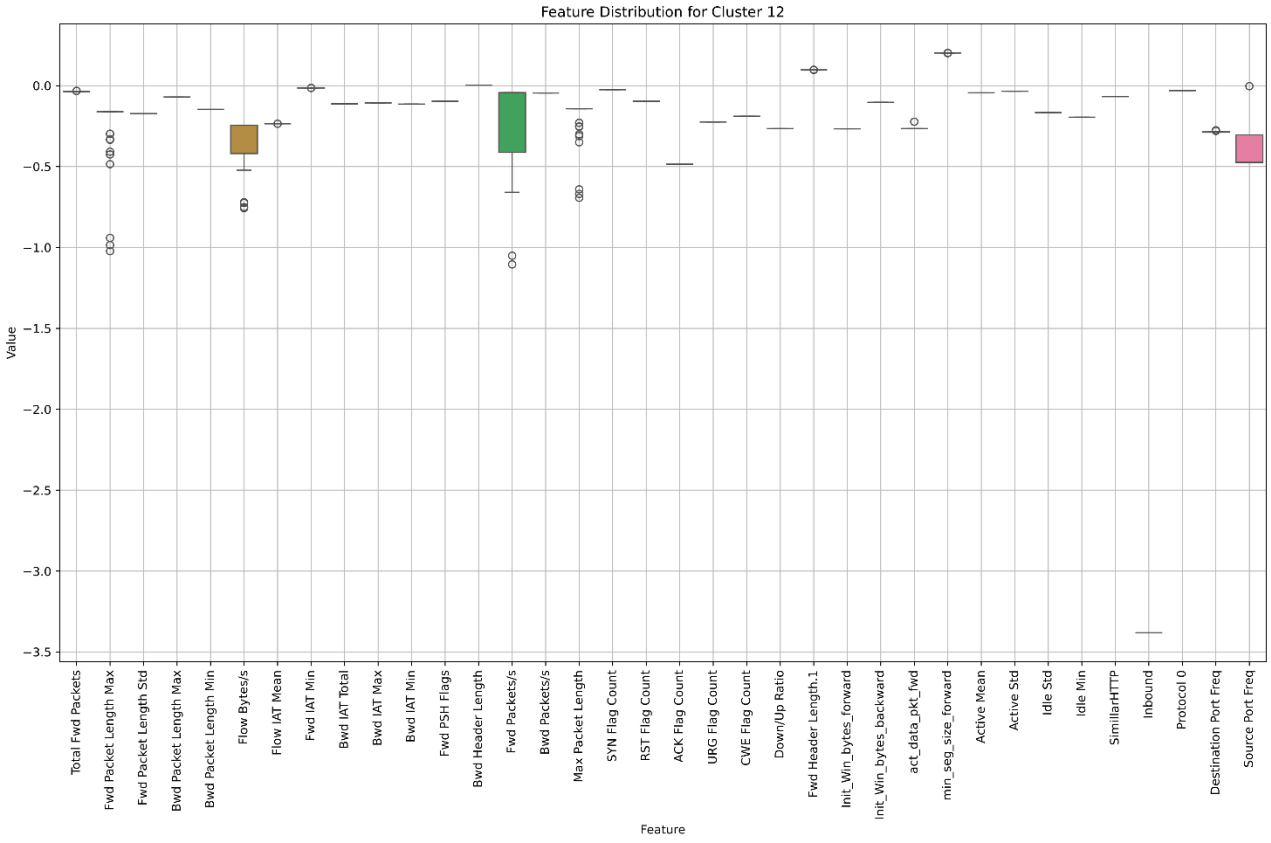

Feature Distribution for Cluster 7
Value
Feature
Total Fwd Packets
Fwd Packet Length Max
Fwd Packet Length Std
Bwd Packet Length Max
Bwd Packet Length Min
Flow Bytes/s
Flow IAT Mean
Fwd IAT Min
Bwd IAT Total
Bwd IAT Max
Bwd IAT Min
Fwd PSH Flags
Bwd Header Length
Fwd Packets/s
Bwd Packets/s
Max Packet Length
SYN Flag Count
RST Flag Count
ACK Flag Count
URG Flag Count
CWE Flag Count
Down/Up Ratio
Fwd Header Length.1
Init_Win_bytes_forward
Init_Win_bytes_backward
act_data_pkt_fwd
min_seg_size_forward
Active Mean
Active Std
Idle Std
Idle Min
SimillarHTTP
Inbound
Protocol 0
Destination Port Freq
Source Port Freq

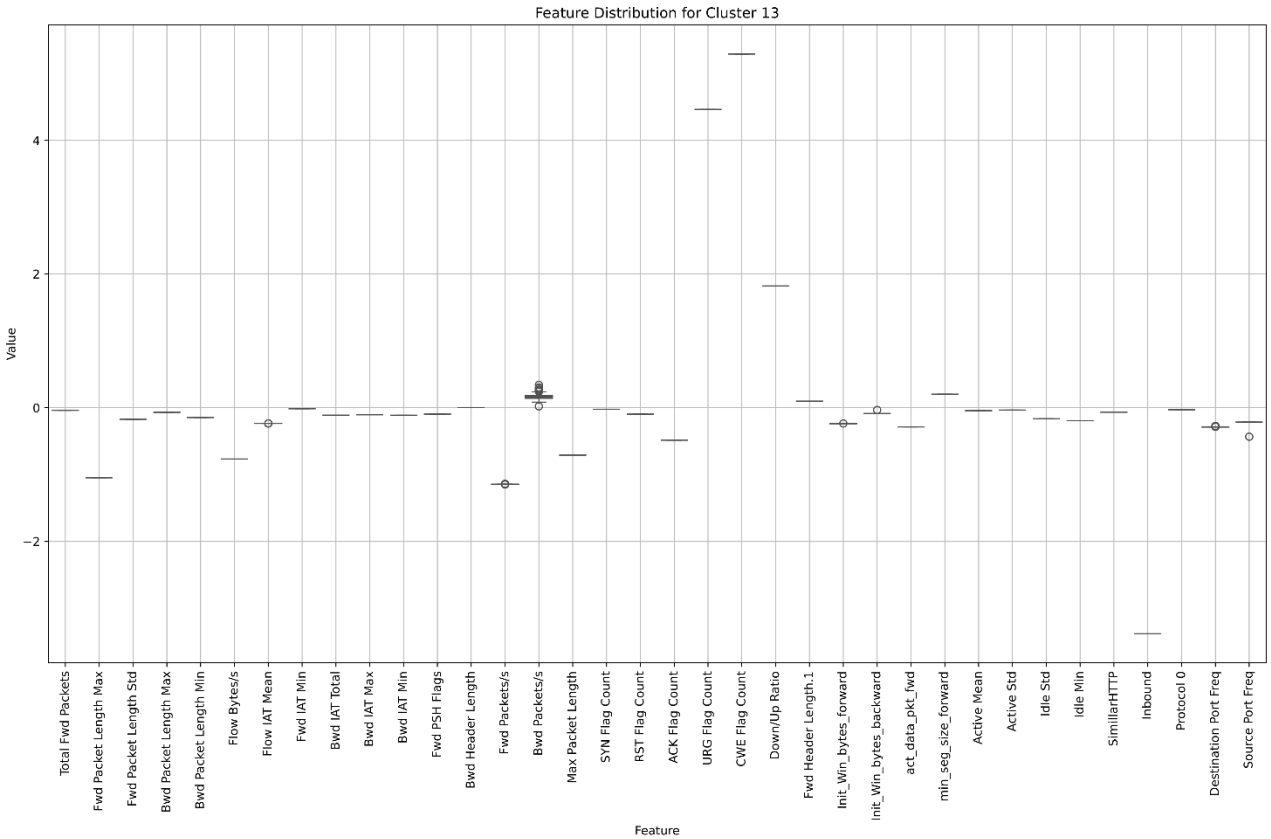

Feature Distribution for Cluster 8
Value
Feature
Total Fwd Packets
Fwd Packet Length Max
Fwd Packet Length Std
Bwd Packet Length Max
Bwd Packet Length Min
Flow Bytes/s
Flow IAT Mean
Fwd IAT Min
Bwd IAT Total
Bwd IAT Max
Bwd IAT Min
Fwd PSH Flags
Bwd Header Length
Fwd Packets/s
Bwd Packets/s
Max Packet Length
SYN Flag Count
RST Flag Count
ACK Flag Count
URG Flag Count
CWE Flag Count
Down/Up Ratio
Fwd Header Length.1
Init_Win_bytes_forward
Init_Win_bytes_backward
act_data_pkt_fwd
min_seg_size_forward
Active Mean
Active Std
Idle Std
Idle Min
SimillarHTTP
Inbound
Protocol 0
Destination Port Freq
Source Port Freq

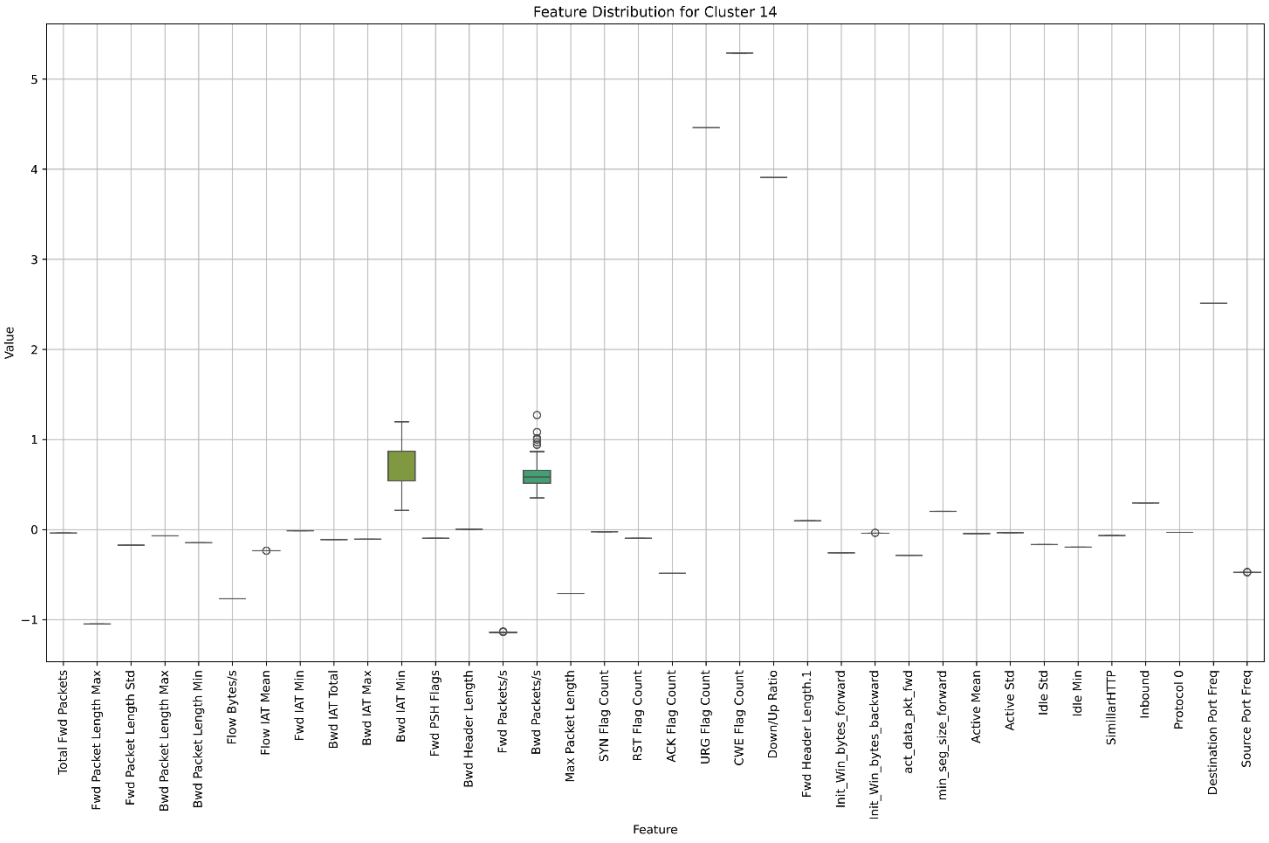

Feature Distribution for Cluster 9
Value
Feature
Total Fwd Packets
Fwd Packet Length Max
Fwd Packet Length Std
Bwd Packet Length Max
Bwd Packet Length Min
Flow Bytes/s
Flow IAT Mean
Fwd IAT Min
Bwd IAT Total
Bwd IAT Max
Bwd IAT Min
Fwd PSH Flags
Bwd Header Length
Fwd Packets/s
Bwd Packets/s
Max Packet Length
SYN Flag Count
RST Flag Count
ACK Flag Count
URG Flag Count
CWE Flag Count
Down/Up Ratio
Fwd Header Length.1
Init_Win_bytes_forward
Init_Win_bytes_backward
act_data_pkt_fwd
min_seg_size_forward
Active Mean
Active Std
Idle Std
Idle Min
SimillarHTTP
Inbound
Protocol 0
Destination Port Freq
Source Port Freq

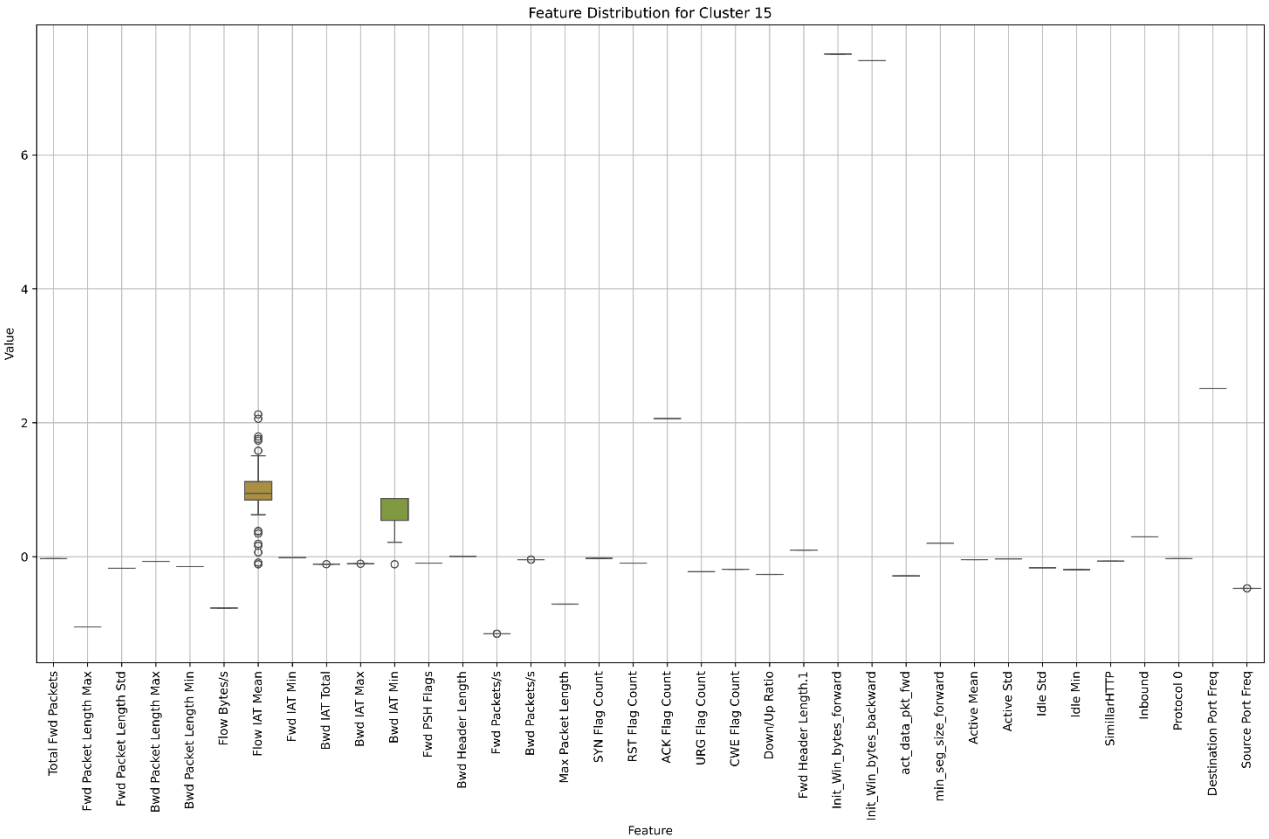

Feature Distribution for Cluster 10
Value
Feature
Total Fwd Packets
Fwd Packet Length Max
Fwd Packet Length Std
Bwd Packet Length Max
Bwd Packet Length Min
Flow Bytes/s
Flow IAT Mean
Fwd IAT Min
Bwd IAT Total
Bwd IAT Max
Bwd IAT Min
Fwd PSH Flags
Bwd Header Length
Fwd Packets/s
Bwd Packets/s
Max Packet Length
SYN Flag Count
RST Flag Count
ACK Flag Count
URG Flag Count
CWE Flag Count
Down/Up Ratio
Fwd Header Length.1
Init_Win_bytes_forward
Init_Win_bytes_backward
act_data_pkt_fwd
min_seg_size_forward
Active Mean
Active Std
Idle Std
Idle Min
SimillarHTTP
Inbound
Protocol 0
Destination Port Freq
Source Port Freq

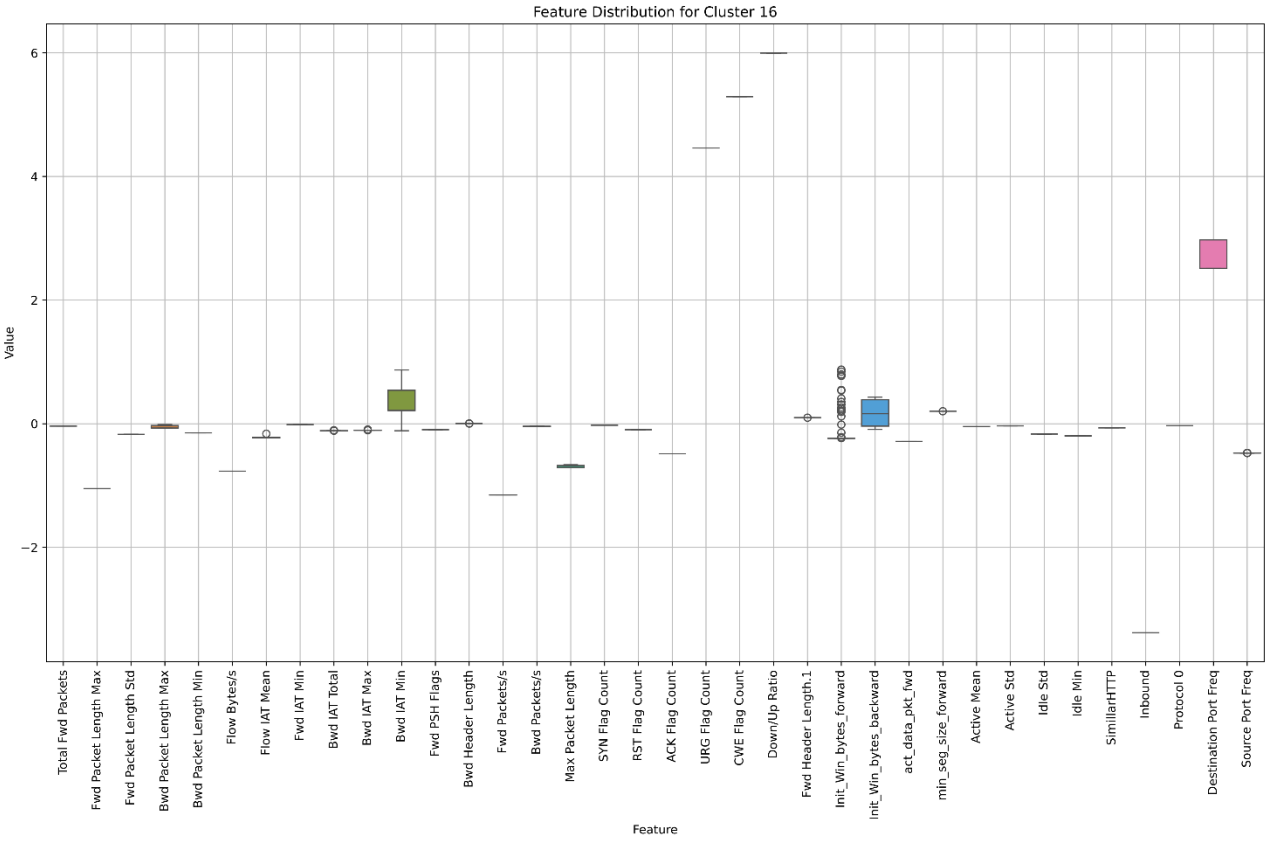

Feature Distribution for Cluster 11
Value
Feature
Total Fwd Packets
Fwd Packet Length Max
Fwd Packet Length Std
Bwd Packet Length Max
Bwd Packet Length Min
Flow Bytes/s
Flow IAT Mean
Fwd IAT Min
Bwd IAT Total
Bwd IAT Max
Bwd IAT Min
Fwd PSH Flags
Bwd Header Length
Fwd Packets/s
Bwd Packets/s
Max Packet Length
SYN Flag Count
RST Flag Count
ACK Flag Count
URG Flag Count
CWE Flag Count
Down/Up Ratio
Fwd Header Length.1
Init_Win_bytes_forward
Init_Win_bytes_backward
act_data_pkt_fwd
min_seg_size_forward
Active Mean
Active Std
Idle Std
Idle Min
SimillarHTTP
Inbound
Protocol 0
Destination Port Freq
Source Port Freq

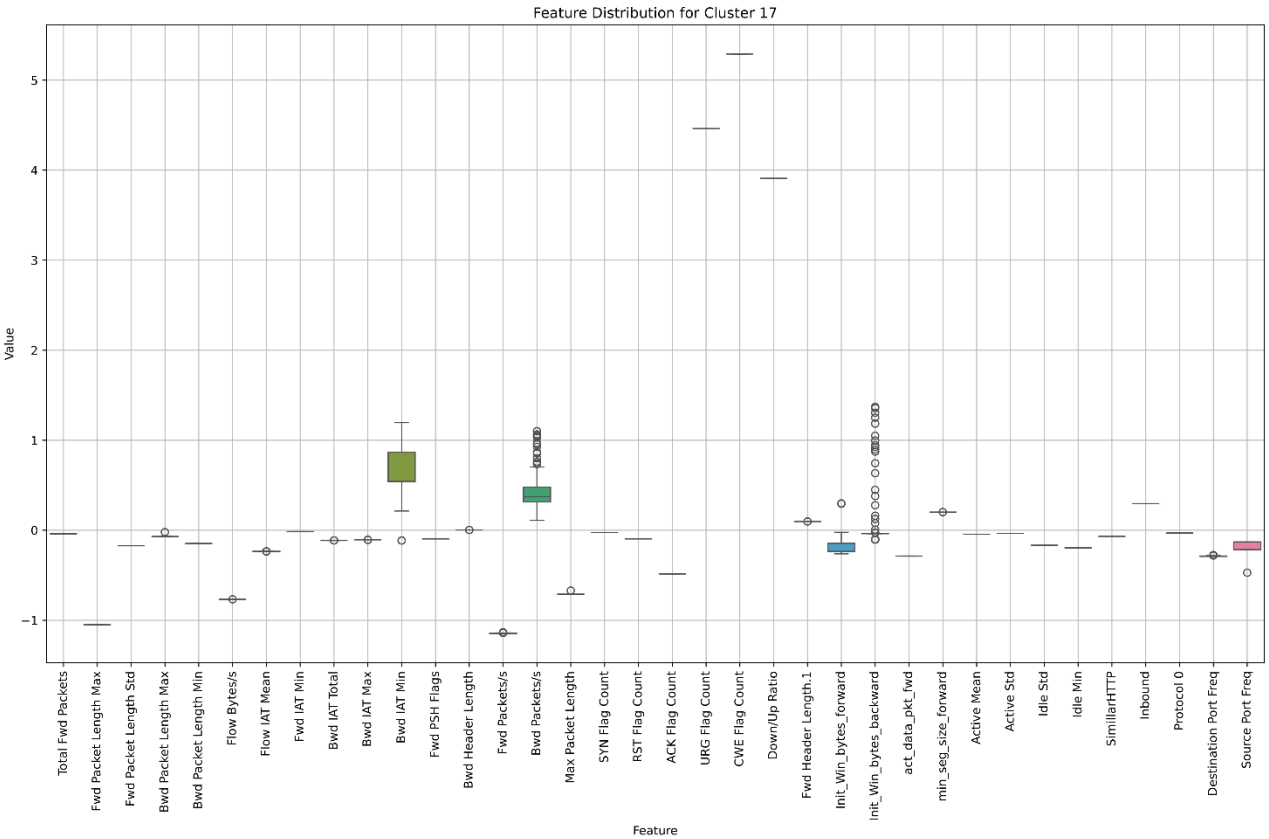

Feature Distribution for Cluster 12
Value
0.0
−0.5
−1.0
−1.5
−2.0
−2.5
−3.0
−3.5
Total Fwd Packets
Fwd Packet Length Max
Fwd Packet Length Std
Bwd Packet Length Max
Bwd Packet Length Min
Flow Bytes/s
Flow IAT Mean
Fwd IAT Min
Bwd IAT Total
Bwd IAT Max
Bwd IAT Min
Fwd PSH Flags
Bwd Header Length
Fwd Packets/s
Bwd Packets/s
Max Packet Length
SYN Flag Count
RST Flag Count
ACK Flag Count
URG Flag Count
CWE Flag Count
Down/Up Ratio
Fwd Header Length.1
Init_Win_bytes_forward
Init_Win_bytes_backward
act_data_pkt_fwd
min_seg_size_forward
Active Mean
Active Std
Idle Std
Idle Min
SimillarHTTP
Inbound
Protocol 0
Destination Port Freq
Source Port Freq
Feature

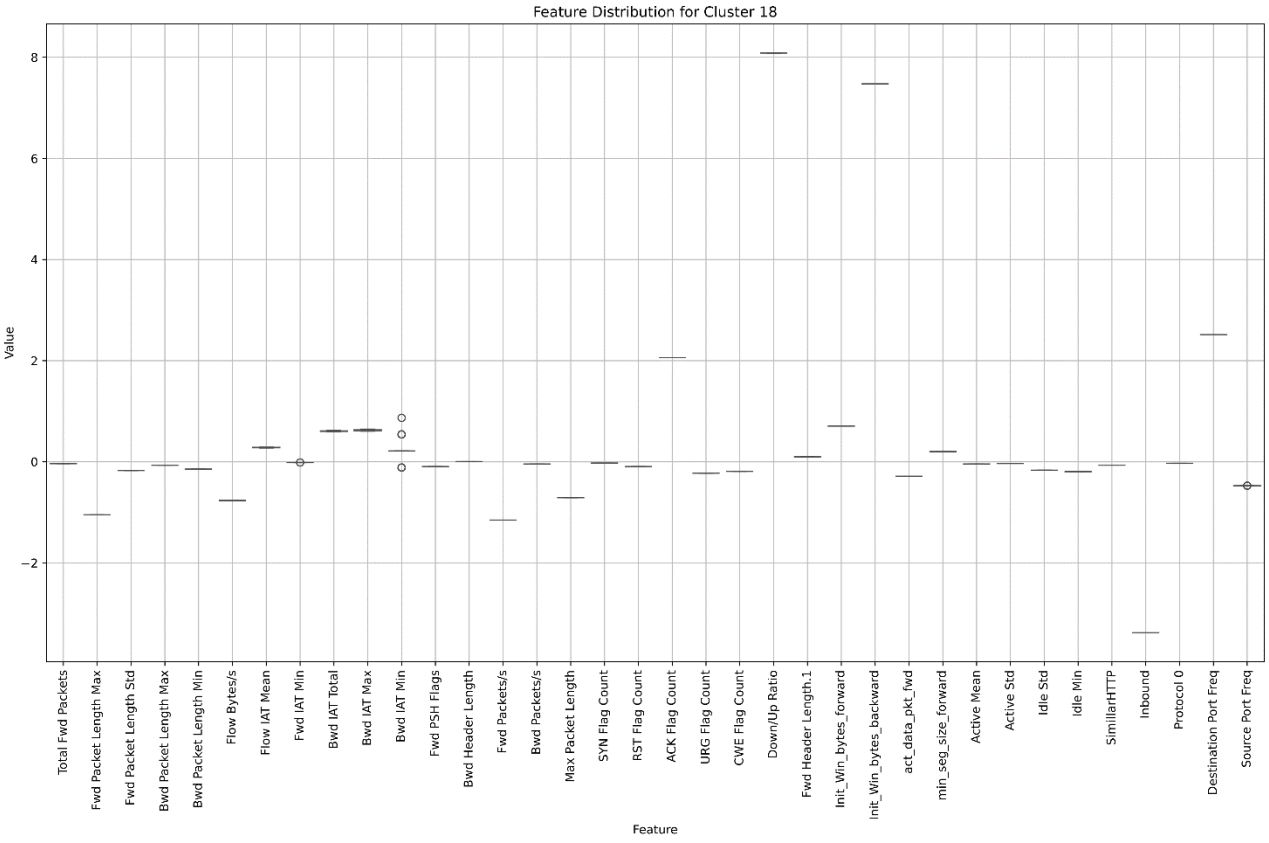

Feature Distribution for Cluster 13
Value
4
2
0
−2
Total Fwd Packets
Fwd Packet Length Max
Fwd Packet Length Std
Bwd Packet Length Max
Bwd Packet Length Min
Flow Bytes/s
Flow IAT Mean
Fwd IAT Min
Bwd IAT Total
Bwd IAT Max
Bwd IAT Min
Fwd PSH Flags
Bwd Header Length
Fwd Packets/s
Bwd Packets/s
Max Packet Length
SYN Flag Count
RST Flag Count
ACK Flag Count
URG Flag Count
CWE Flag Count
Down/Up Ratio
Fwd Header Length.1
Init_Win_bytes_forward
Init_Win_bytes_backward
act_data_pkt_fwd
min_seg_size_forward
Active Mean
Active Std
Idle Std
Idle Min
SimillarHTTP
Inbound
Protocol 0
Destination Port Freq
Source Port Freq
Feature

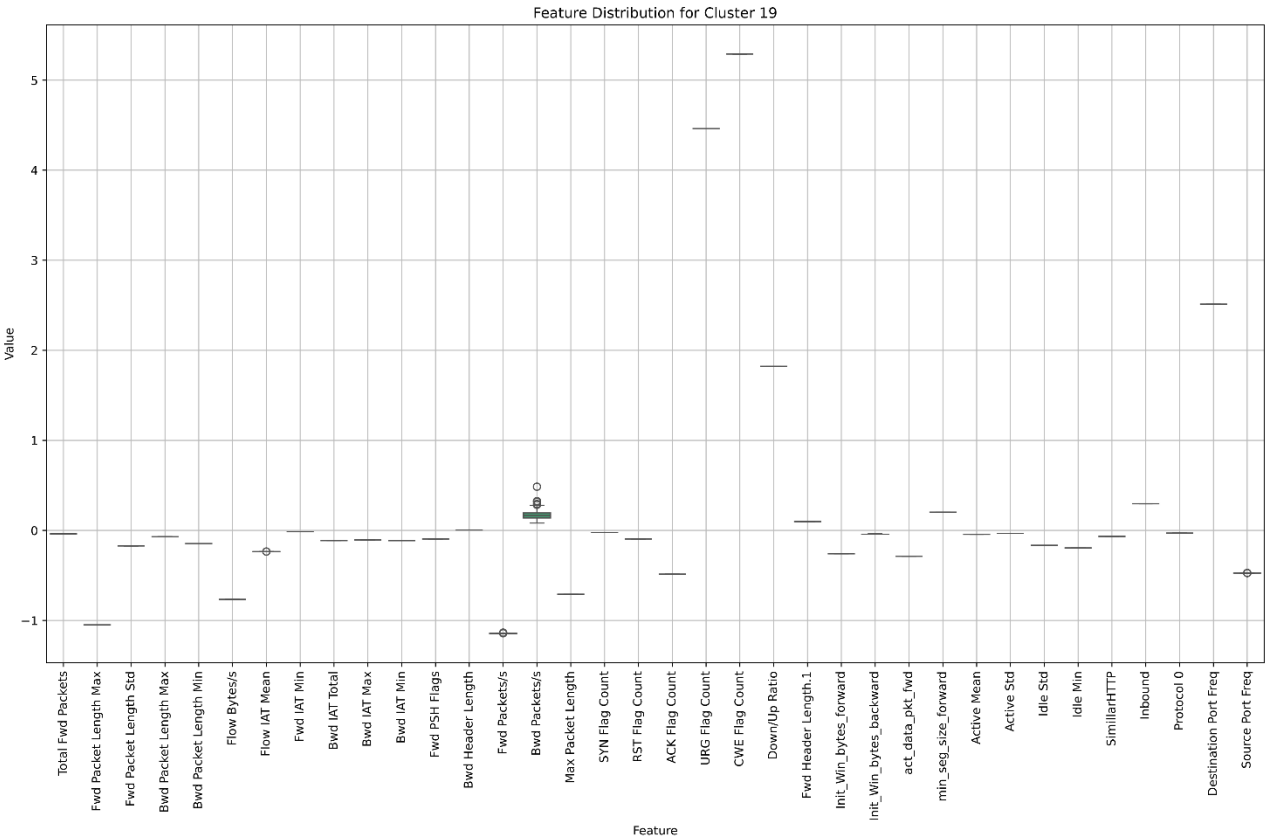

Feature Distribution for Cluster 14
Value
Feature
Total Fwd Packets
Fwd Packet Length Max
Fwd Packet Length Std
Bwd Packet Length Max
Bwd Packet Length Min
Flow Bytes/s
Flow IAT Mean
Fwd IAT Min
Bwd IAT Total
Bwd IAT Max
Bwd IAT Min
Fwd PSH Flags
Bwd Header Length
Fwd Packets/s
Bwd Packets/s
Max Packet Length
SYN Flag Count
RST Flag Count
ACK Flag Count
URG Flag Count
CWE Flag Count
Down/Up Ratio
Fwd Header Length.1
Init_Win_bytes_forward
Init_Win_bytes_backward
act_data_pkt_fwd
min_seg_size_forward
Active Mean
Active Std
Idle Std
Idle Min
SimillarHTTP
Inbound
Protocol 0
Destination Port Freq
Source Port Freq

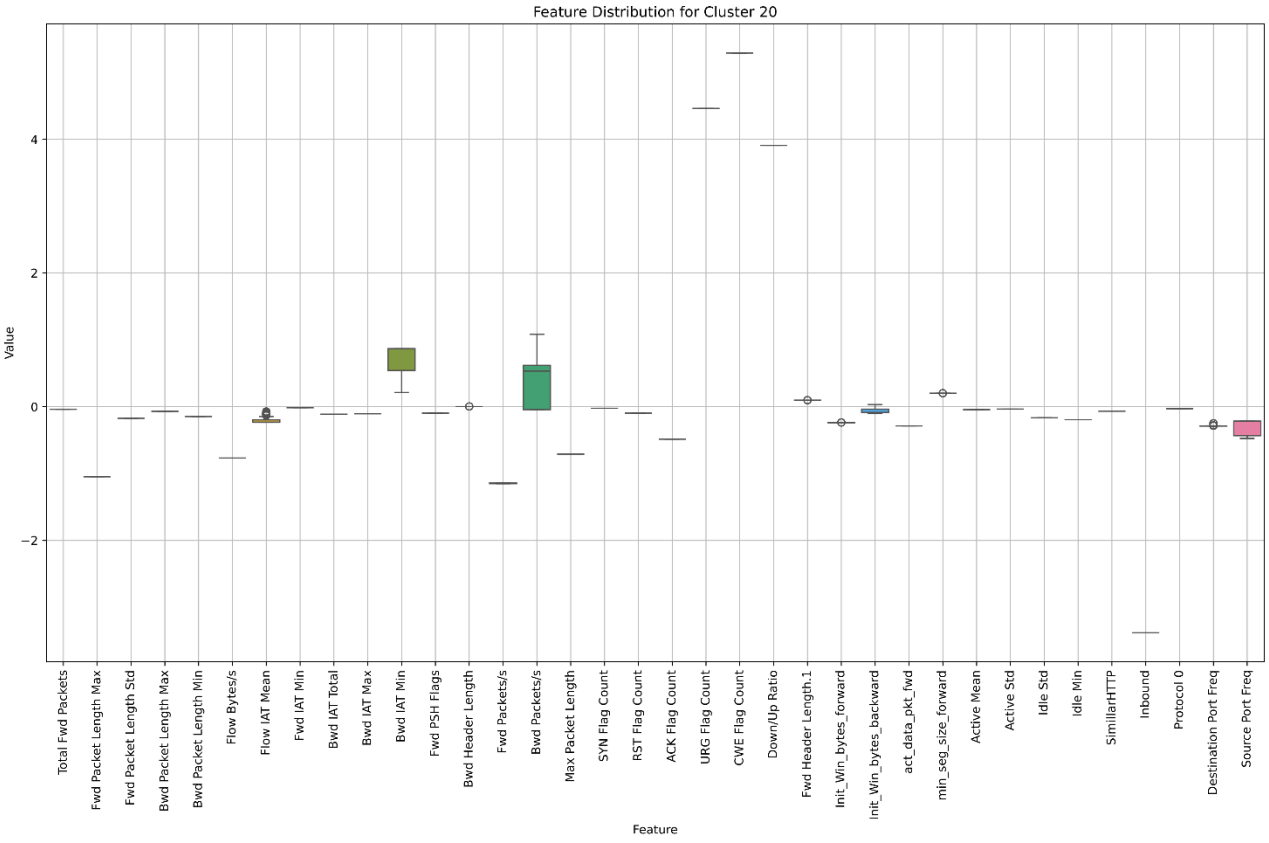

Feature Distribution for Cluster 15
Value
Feature
Total Fwd Packets
Fwd Packet Length Max
Fwd Packet Length Std
Bwd Packet Length Max
Bwd Packet Length Min
Flow Bytes/s
Flow IAT Mean
Fwd IAT Min
Bwd IAT Total
Bwd IAT Max
Bwd IAT Min
Fwd PSH Flags
Bwd Header Length
Fwd Packets/s
Bwd Packets/s
Max Packet Length
SYN Flag Count
RST Flag Count
ACK Flag Count
URG Flag Count
CWE Flag Count
Down/Up Ratio
Fwd Header Length.1
Init_Win_bytes_forward
Init_Win_bytes_backward
act_data_pkt_fwd
min_seg_size_forward
Active Mean
Active Std
Idle Std
Idle Min
SimillarHTTP
Inbound
Protocol 0
Destination Port Freq
Source Port Freq

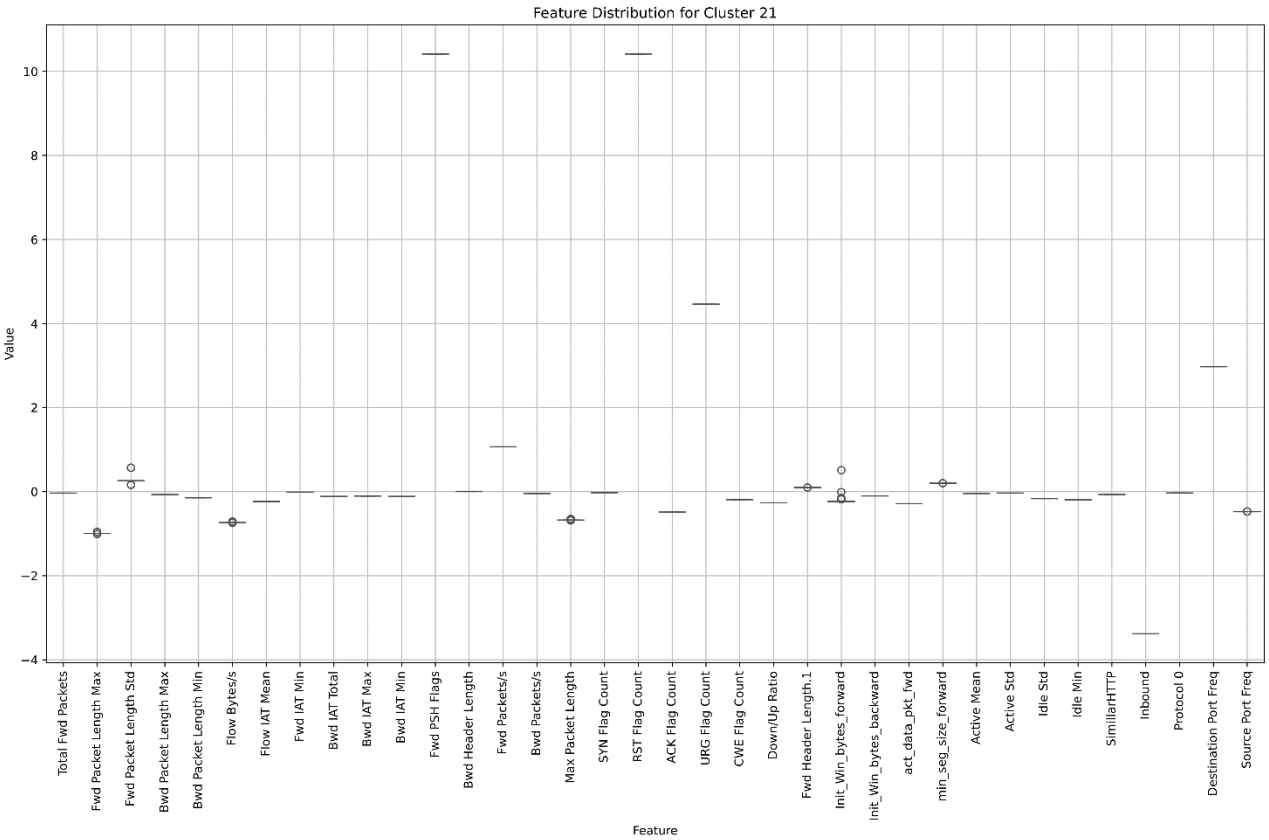

Feature Distribution for Cluster 16
Value
Feature
Total Fwd Packets
Fwd Packet Length Max
Fwd Packet Length Std
Bwd Packet Length Max
Bwd Packet Length Min
Flow Bytes/s
Flow IAT Mean
Fwd IAT Min
Bwd IAT Total
Bwd IAT Max
Bwd IAT Min
Fwd PSH Flags
Bwd Header Length
Fwd Packets/s
Bwd Packets/s
Max Packet Length
SYN Flag Count
RST Flag Count
ACK Flag Count
URG Flag Count
CWE Flag Count
Down/Up Ratio
Fwd Header Length.1
Init_Win_bytes_forward
Init_Win_bytes_backward
act_data_pkt_fwd
min_seg_size_forward
Active Mean
Active Std
Idle Std
Idle Min
SimillarHTTP
Inbound
Protocol 0
Destination Port Freq
Source Port Freq

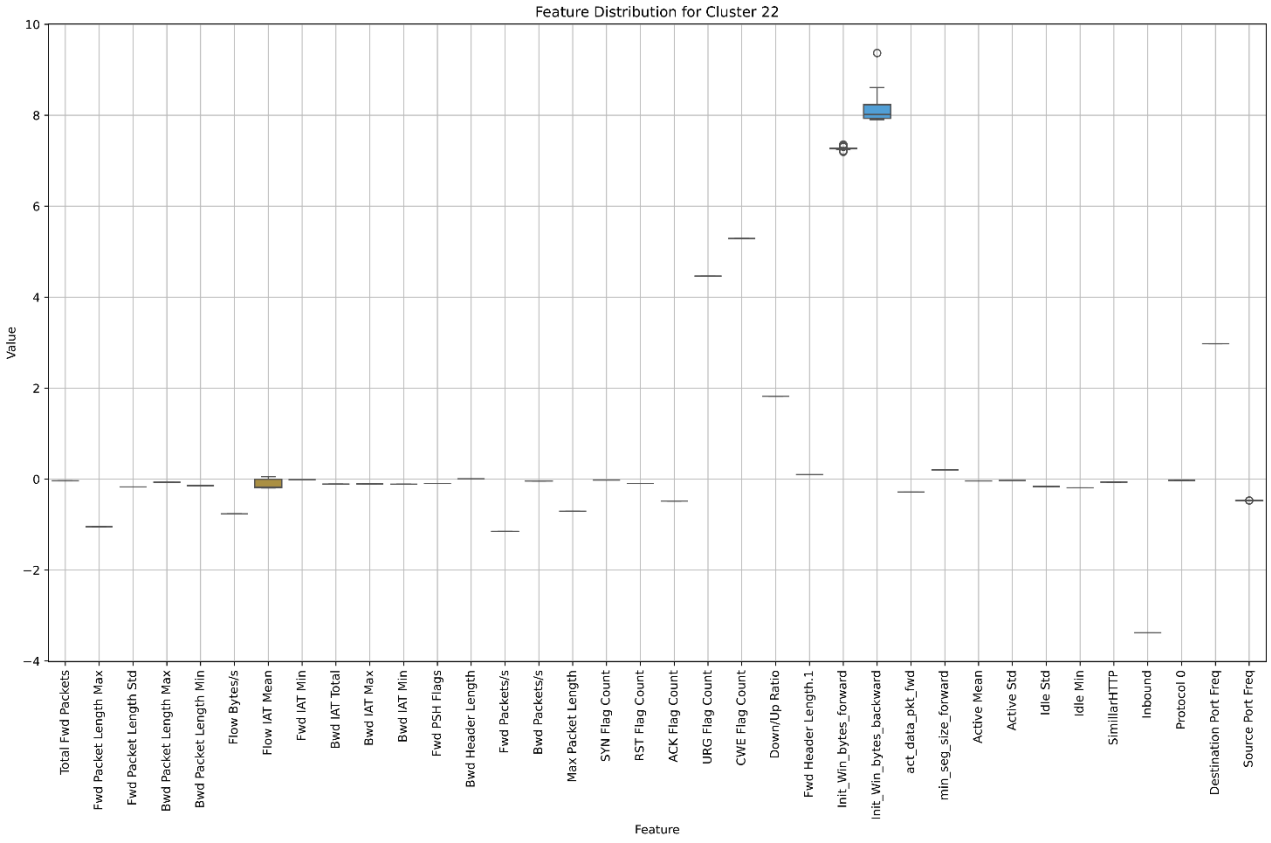

Feature Distribution for Cluster 17
Value
Feature
Total Fwd Packets
Fwd Packet Length Max
Fwd Packet Length Std
Bwd Packet Length Max
Bwd Packet Length Min
Flow Bytes/s
Flow IAT Mean
Fwd IAT Min
Bwd IAT Total
Bwd IAT Max
Bwd IAT Min
Fwd PSH Flags
Bwd Header Length
Fwd Packets/s
Bwd Packets/s
Max Packet Length
SYN Flag Count
RST Flag Count
ACK Flag Count
URG Flag Count
CWE Flag Count
Down/Up Ratio
Fwd Header Length.1
Init_Win_bytes_forward
Init_Win_bytes_backward
act_data_pkt_fwd
min_seg_size_forward
Active Mean
Active Std
Idle Std
Idle Min
SimillarHTTP
Inbound
Protocol 0
Destination Port Freq
Source Port Freq

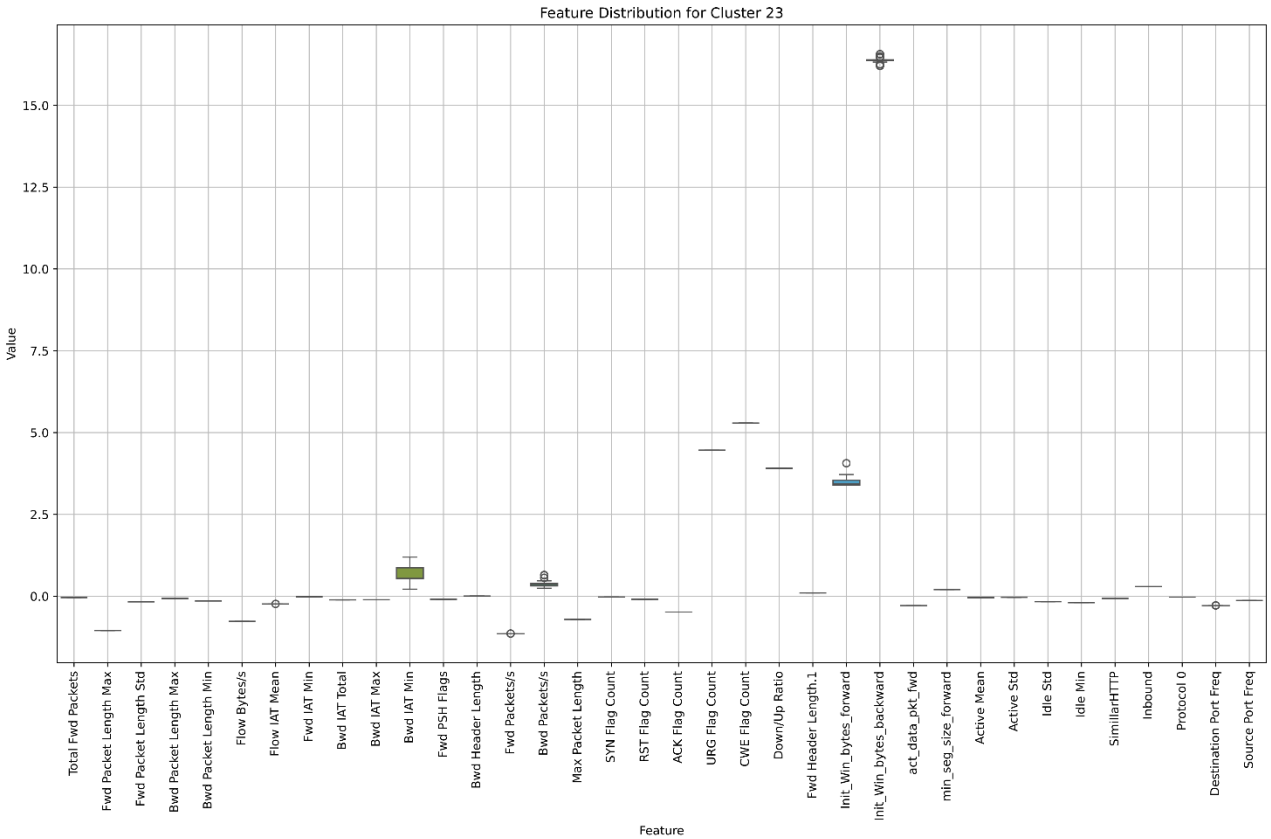

Feature Distribution for Cluster 18
Value
Feature
Total Fwd Packets
Fwd Packet Length Max
Fwd Packet Length Std
Bwd Packet Length Max
Bwd Packet Length Min
Flow Bytes/s
Flow IAT Mean
Fwd IAT Min
Bwd IAT Total
Bwd IAT Max
Bwd IAT Min
Fwd PSH Flags
Bwd Header Length
Fwd Packets/s
Bwd Packets/s
Max Packet Length
SYN Flag Count
RST Flag Count
ACK Flag Count
URG Flag Count
CWE Flag Count
Down/Up Ratio
Fwd Header Length.1
Init_Win_bytes_forward
Init_Win_bytes_backward
act_data_pkt_fwd
min_seg_size_forward
Active Mean
Active Std
Idle Std
Idle Min
SimillarHTTP
Inbound
Protocol 0
Destination Port Freq
Source Port Freq

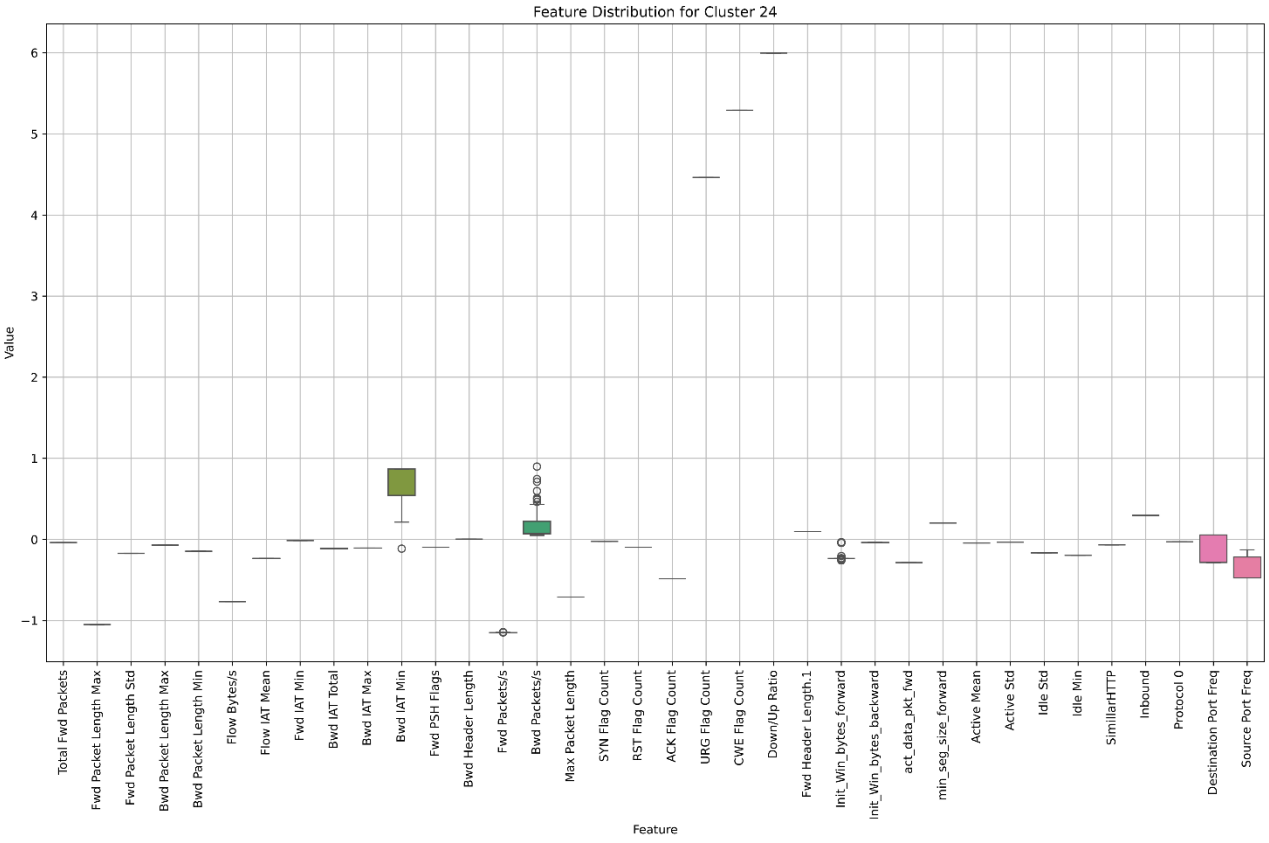

Feature Distribution for Cluster 19
Value
Feature
Total Fwd Packets
Fwd Packet Length Max
Fwd Packet Length Std
Bwd Packet Length Max
Bwd Packet Length Min
Flow Bytes/s
Flow IAT Mean
Fwd IAT Min
Bwd IAT Total
Bwd IAT Max
Bwd IAT Min
Fwd PSH Flags
Bwd Header Length
Fwd Packets/s
Bwd Packets/s
Max Packet Length
SYN Flag Count
RST Flag Count
ACK Flag Count
URG Flag Count
CWE Flag Count
Down/Up Ratio
Fwd Header Length.1
Init_Win_bytes_forward
Init_Win_bytes_backward
act_data_pkt_fwd
min_seg_size_forward
Active Mean
Active Std
Idle Std
Idle Min
SimillarHTTP
Inbound
Protocol 0
Destination Port Freq
Source Port Freq

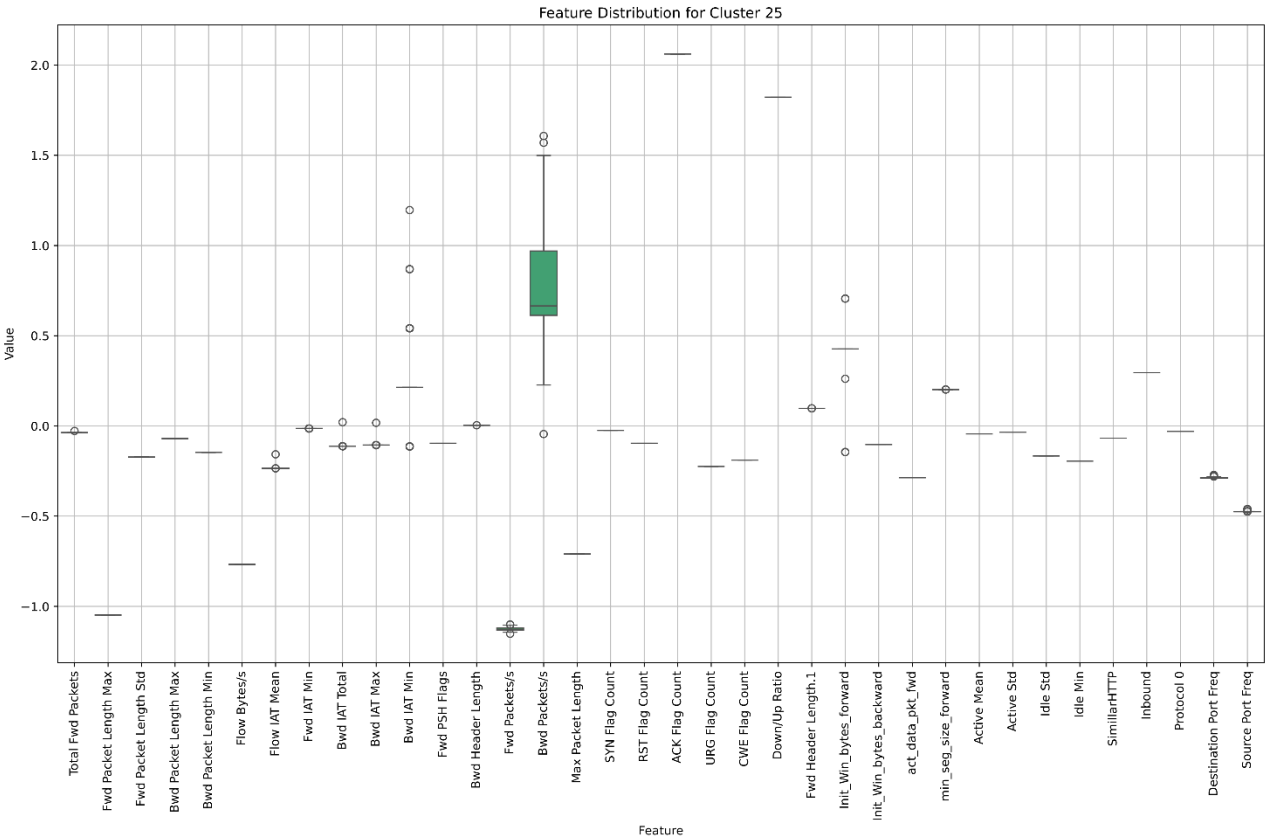

Feature Distribution for Cluster 20
Value
Feature
Total Fwd Packets
Fwd Packet Length Max
Fwd Packet Length Std
Bwd Packet Length Max
Bwd Packet Length Min
Flow Bytes/s
Flow IAT Mean
Fwd IAT Min
Bwd IAT Total
Bwd IAT Max
Bwd IAT Min
Fwd PSH Flags
Bwd Header Length
Fwd Packets/s
Bwd Packets/s
Max Packet Length
SYN Flag Count
RST Flag Count
ACK Flag Count
URG Flag Count
CWE Flag Count
Down/Up Ratio
Fwd Header Length.1
Init_Win_bytes_forward
Init_Win_bytes_backward
act_data_pkt_fwd
min_seg_size_forward
Active Mean
Active Std
Idle Std
Idle Min
SimillarHTTP
Inbound
Protocol 0
Destination Port Freq
Source Port Freq

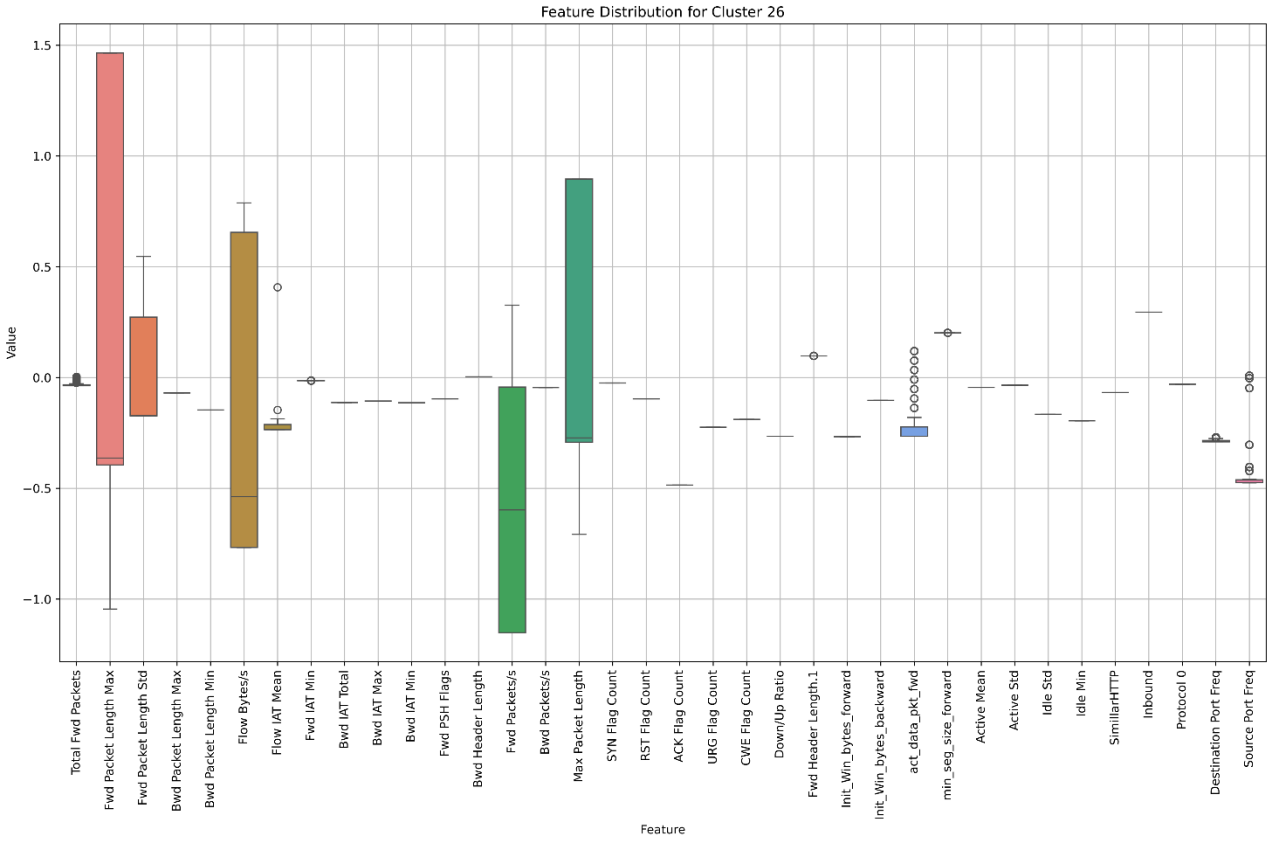

Feature Distribution for Cluster 21
Value
Feature
Total Fwd Packets
Fwd Packet Length Max
Fwd Packet Length Std
Bwd Packet Length Max
Bwd Packet Length Min
Flow Bytes/s
Flow IAT Mean
Fwd IAT Min
Bwd IAT Total
Bwd IAT Max
Bwd IAT Min
Fwd PSH Flags
Bwd Header Length
Fwd Packets/s
Bwd Packets/s
Max Packet Length
SYN Flag Count
RST Flag Count
ACK Flag Count
URG Flag Count
CWE Flag Count
Down/Up Ratio
Fwd Header Length.1
Init_Win_bytes_forward
Init_Win_bytes_backward
act_data_pkt_fwd
min_seg_size_forward
Active Mean
Active Std
Idle Std
Idle Min
SimillarHTTP
Inbound
Protocol 0
Destination Port Freq
Source Port Freq

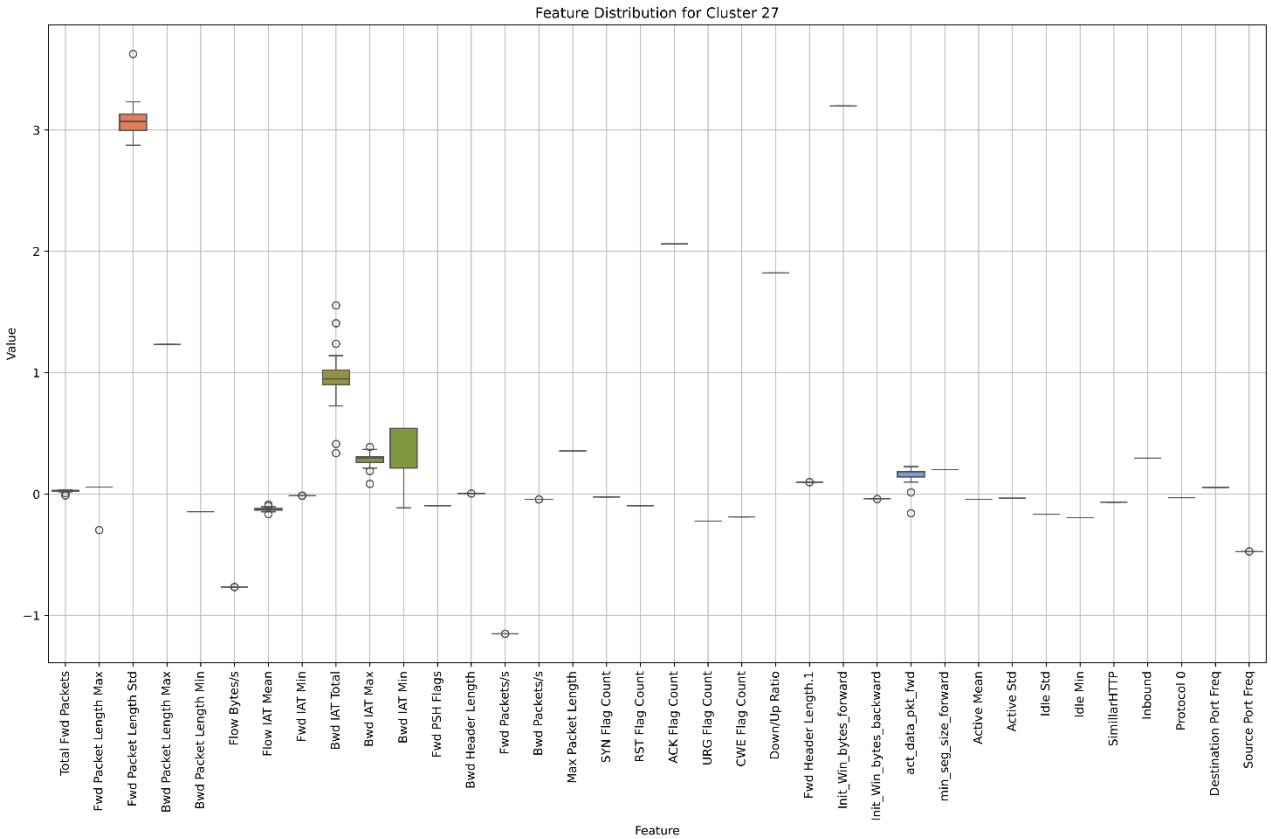

Feature Distribution for Cluster 22
Value
Feature
Total Fwd Packets
Fwd Packet Length Max
Fwd Packet Length Std
Bwd Packet Length Max
Bwd Packet Length Min
Flow Bytes/s
Flow IAT Mean
Fwd IAT Min
Bwd IAT Total
Bwd IAT Max
Bwd IAT Min
Fwd PSH Flags
Bwd Header Length
Fwd Packets/s
Bwd Packets/s
Max Packet Length
SYN Flag Count
RST Flag Count
ACK Flag Count
URG Flag Count
CWE Flag Count
Down/Up Ratio
Fwd Header Length.1
Init_Win_bytes_forward
Init_Win_bytes_backward
act_data_pkt_fwd
min_seg_size_forward
Active Mean
Active Std
Idle Std
Idle Min
SimillarHTTP
Inbound
Protocol 0
Destination Port Freq
Source Port Freq

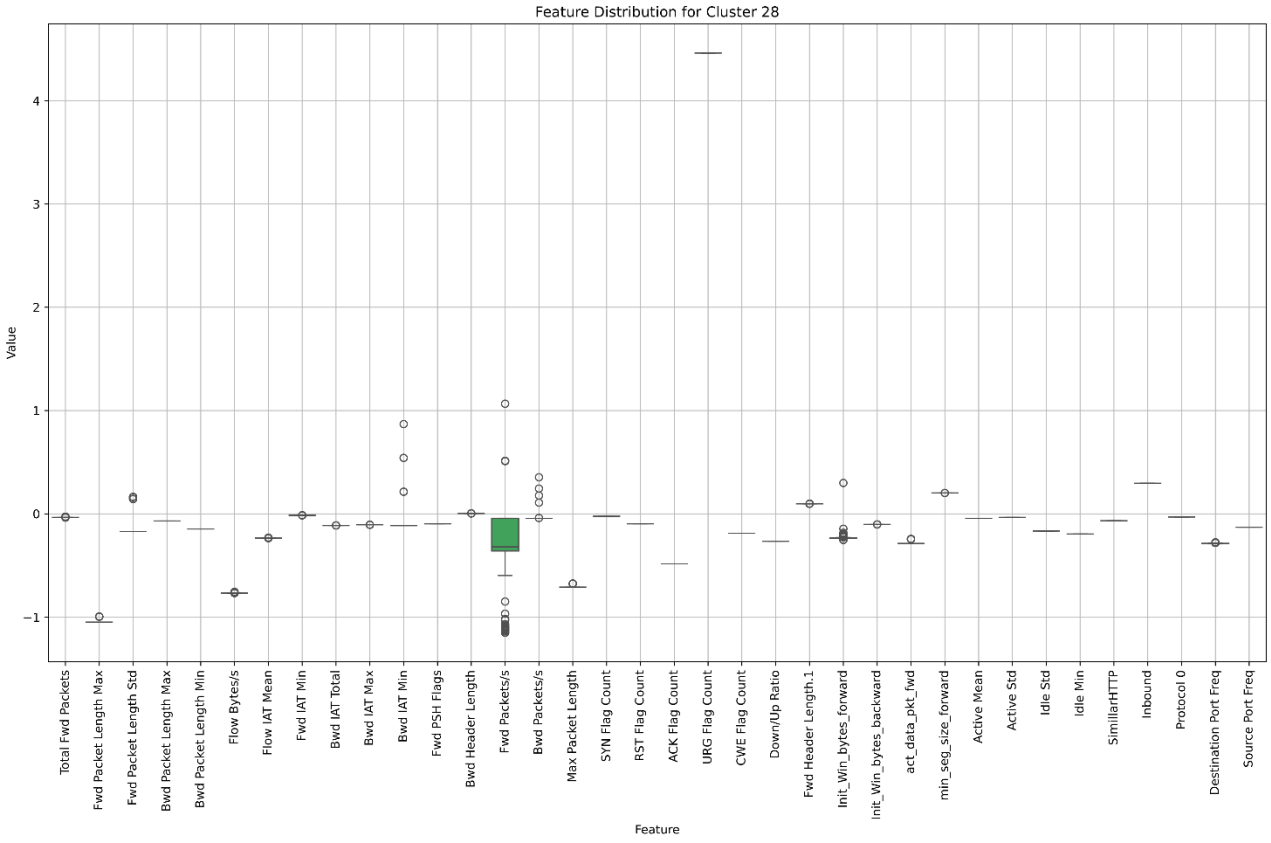

Feature Distribution for Cluster 23
Value
Feature
Total Fwd Packets
Fwd Packet Length Max
Fwd Packet Length Std
Bwd Packet Length Max
Bwd Packet Length Min
Flow Bytes/s
Flow IAT Mean
Fwd IAT Min
Bwd IAT Total
Bwd IAT Max
Bwd IAT Min
Fwd PSH Flags
Bwd Header Length
Fwd Packets/s
Bwd Packets/s
Max Packet Length
SYN Flag Count
RST Flag Count
ACK Flag Count
URG Flag Count
CWE Flag Count
Down/Up Ratio
Fwd Header Length.1
Init_Win_bytes_forward
Init_Win_bytes_backward
act_data_pkt_fwd
min_seg_size_forward
Active Mean
Active Std
Idle Std
Idle Min
SimillarHTTP
Inbound
Protocol 0
Destination Port Freq
Source Port Freq

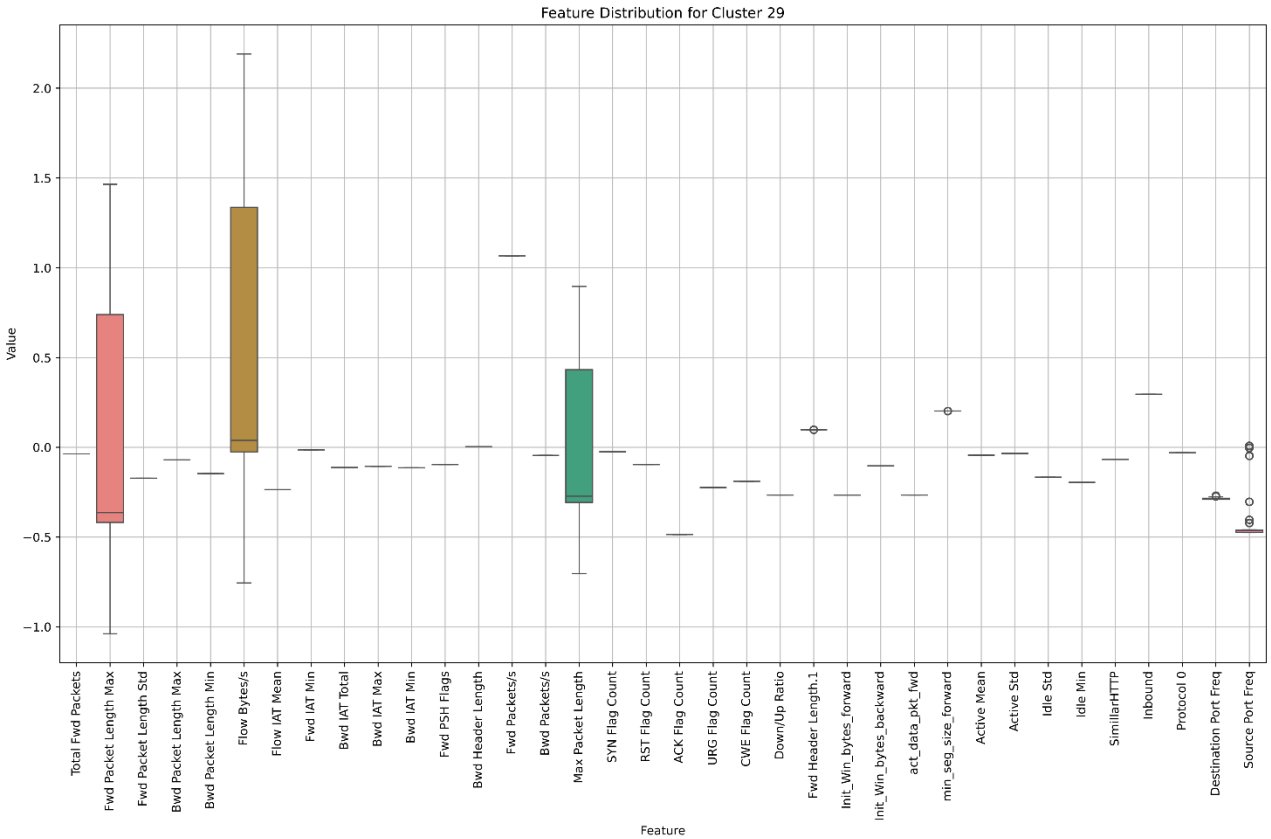

Feature Distribution for Cluster 24
Value
Feature
Total Fwd Packets
Fwd Packet Length Max
Fwd Packet Length Std
Bwd Packet Length Max
Bwd Packet Length Min
Flow Bytes/s
Flow IAT Mean
Fwd IAT Min
Bwd IAT Total
Bwd IAT Max
Bwd IAT Min
Fwd PSH Flags
Bwd Header Length
Fwd Packets/s
Bwd Packets/s
Max Packet Length
SYN Flag Count
RST Flag Count
ACK Flag Count
URG Flag Count
CWE Flag Count
Down/Up Ratio
Fwd Header Length.1
Init_Win_bytes_forward
Init_Win_bytes_backward
act_data_pkt_fwd
min_seg_size_forward
Active Mean
Active Std
Idle Std
Idle Min
SimillarHTTP
Inbound
Protocol 0
Destination Port Freq
Source Port Freq

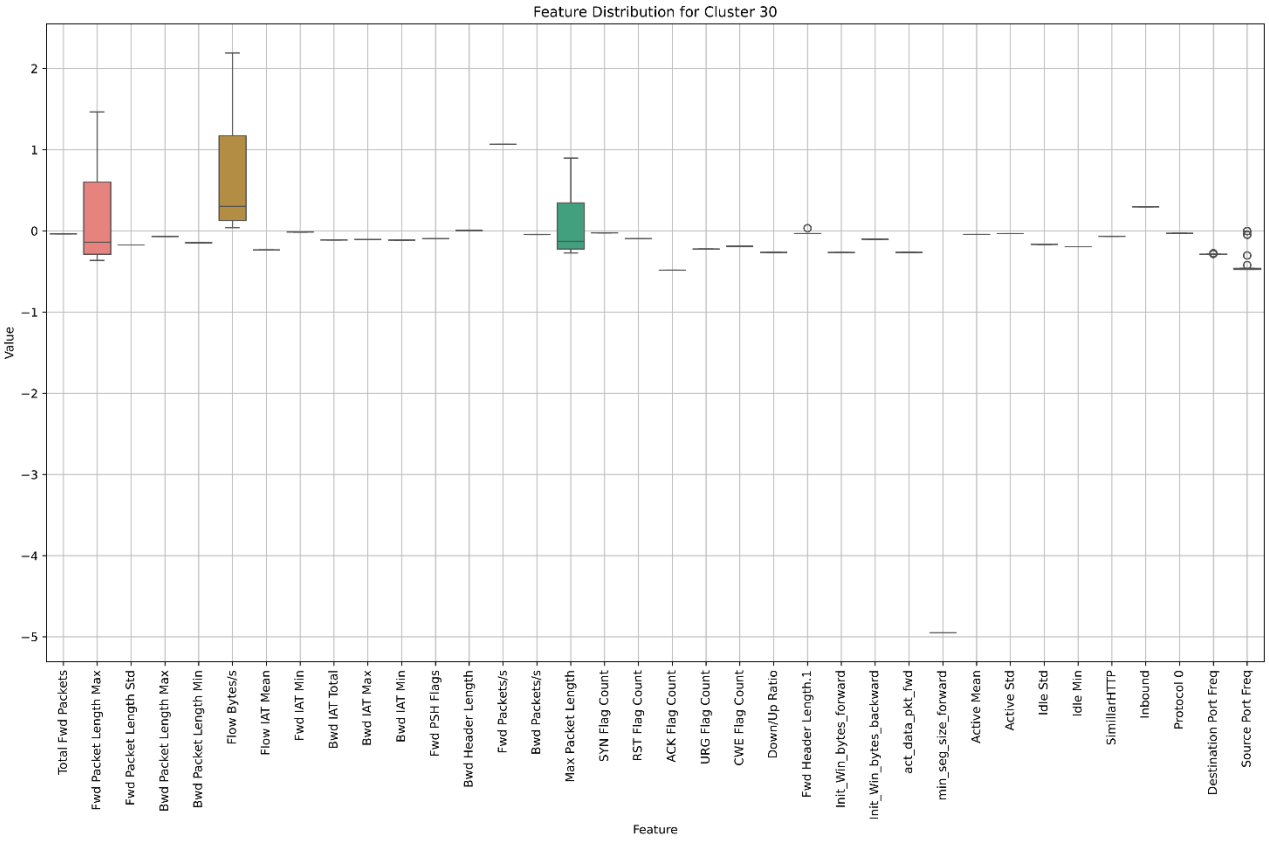

Feature Distribution for Cluster 25
Value
Feature
Total Fwd Packets
Fwd Packet Length Max
Fwd Packet Length Std
Bwd Packet Length Max
Bwd Packet Length Min
Flow Bytes/s
Flow IAT Mean
Fwd IAT Min
Bwd IAT Total
Bwd IAT Max
Bwd IAT Min
Fwd PSH Flags
Bwd Header Length
Fwd Packets/s
Bwd Packets/s
Max Packet Length
SYN Flag Count
RST Flag Count
ACK Flag Count
URG Flag Count
CWE Flag Count
Down/Up Ratio
Fwd Header Length.1
Init_Win_bytes_forward
Init_Win_bytes_backward
act_data_pkt_fwd
min_seg_size_forward
Active Mean
Active Std
Idle Std
Idle Min
SimillarHTTP
Inbound
Protocol 0
Destination Port Freq
Source Port Freq

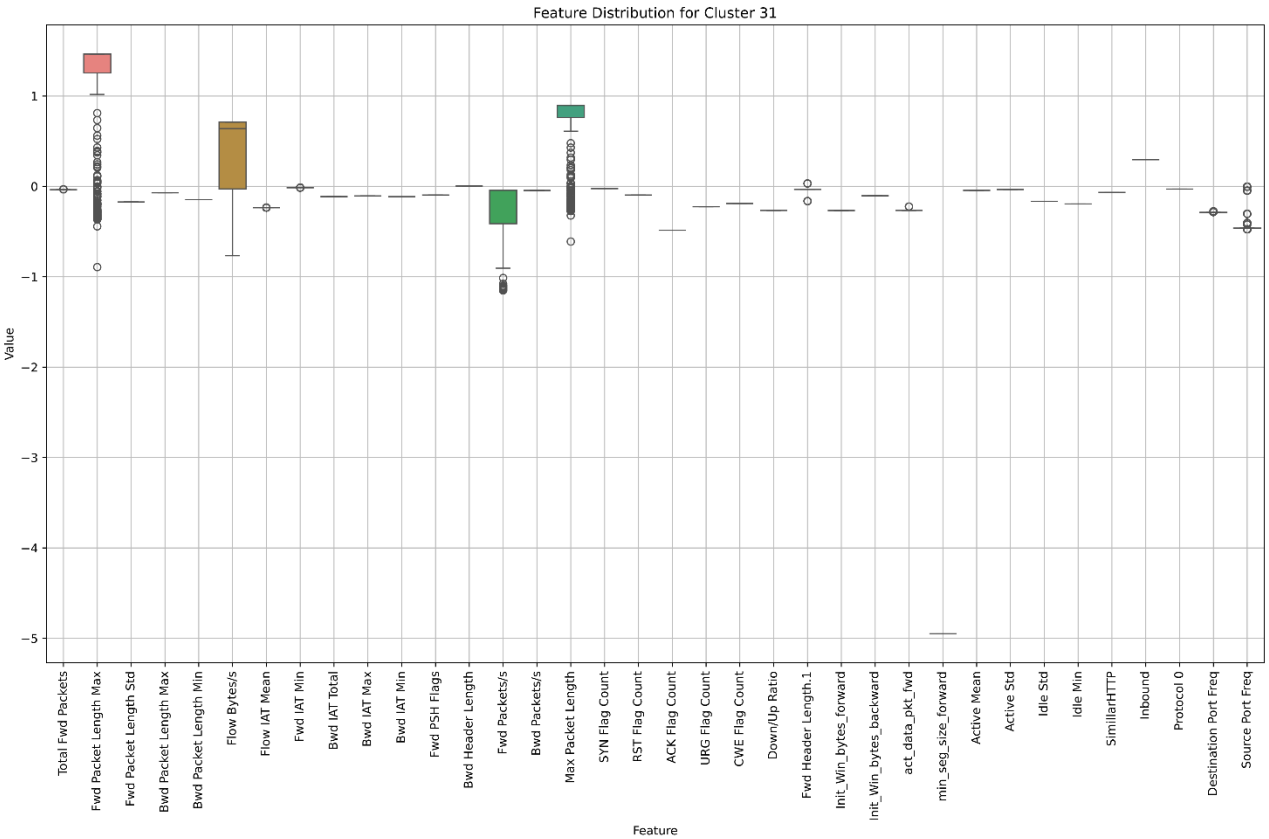

Feature Distribution for Cluster 26
Value
Feature
Total Fwd Packets
Fwd Packet Length Max
Fwd Packet Length Std
Bwd Packet Length Max
Bwd Packet Length Min
Flow Bytes/s
Flow IAT Mean
Fwd IAT Min
Bwd IAT Total
Bwd IAT Max
Bwd IAT Min
Fwd PSH Flags
Bwd Header Length
Fwd Packets/s
Bwd Packets/s
Max Packet Length
SYN Flag Count
RST Flag Count
ACK Flag Count
URG Flag Count
CWE Flag Count
Down/Up Ratio
Fwd Header Length.1
Init_Win_bytes_forward
Init_Win_bytes_backward
act_data_pkt_fwd
min_seg_size_forward
Active Mean
Active Std
Idle Std
Idle Min
SimillarHTTP
Inbound
Protocol 0
Destination Port Freq
Source Port Freq

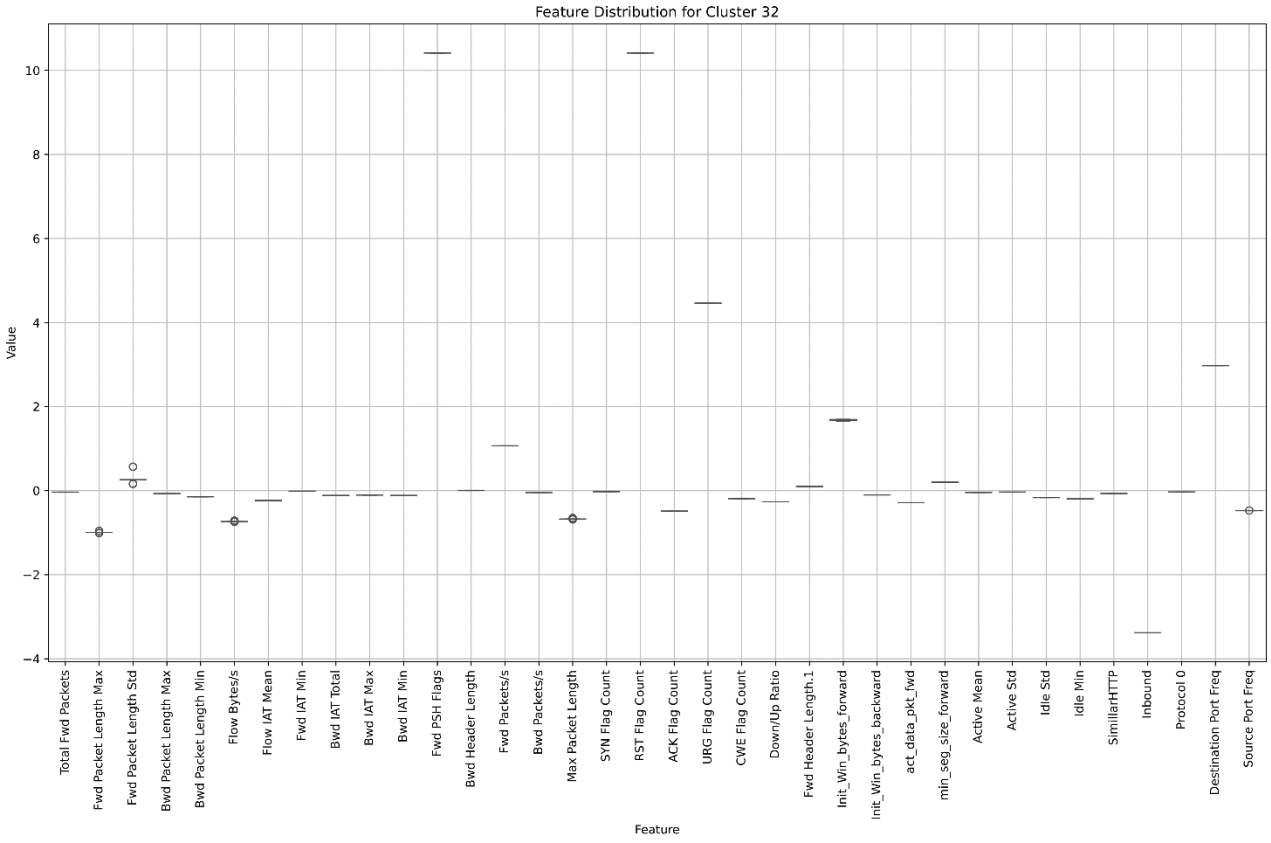

Feature Distribution for Cluster 27
Value
Feature
Total Fwd Packets
Fwd Packet Length Max
Fwd Packet Length Std
Bwd Packet Length Max
Bwd Packet Length Min
Flow Bytes/s
Flow IAT Mean
Fwd IAT Min
Bwd IAT Total
Bwd IAT Max
Bwd IAT Min
Fwd PSH Flags
Bwd Header Length
Fwd Packets/s
Bwd Packets/s
Max Packet Length
SYN Flag Count
RST Flag Count
ACK Flag Count
URG Flag Count
CWE Flag Count
Down/Up Ratio
Fwd Header Length.1
Init_Win_bytes_forward
Init_Win_bytes_backward
act_data_pkt_fwd
min_seg_size_forward
Active Mean
Active Std
Idle Std
Idle Min
SimillarHTTP
Inbound
Protocol 0
Destination Port Freq
Source Port Freq

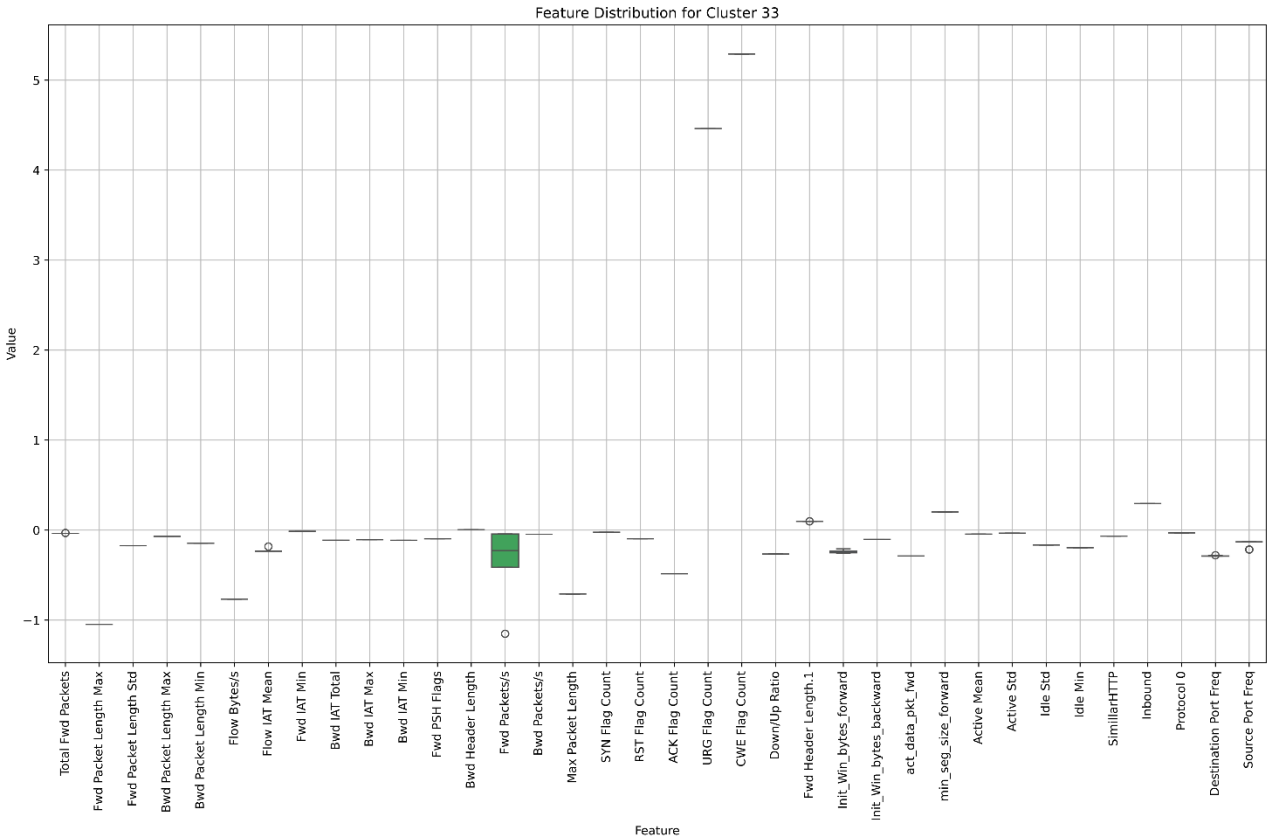

Feature Distribution for Cluster 28
Value
Feature
Total Fwd Packets
Fwd Packet Length Max
Fwd Packet Length Std
Bwd Packet Length Max
Bwd Packet Length Min
Flow Bytes/s
Flow IAT Mean
Fwd IAT Min
Bwd IAT Total
Bwd IAT Max
Bwd IAT Min
Fwd PSH Flags
Bwd Header Length
Fwd Packets/s
Bwd Packets/s
Max Packet Length
SYN Flag Count
RST Flag Count
ACK Flag Count
URG Flag Count
CWE Flag Count
Down/Up Ratio
Fwd Header Length.1
Init_Win_bytes_forward
Init_Win_bytes_backward
act_data_pkt_fwd
min_seg_size_forward
Active Mean
Active Std
Idle Std
Idle Min
SimillarHTTP
Inbound
Protocol 0
Destination Port Freq
Source Port Freq

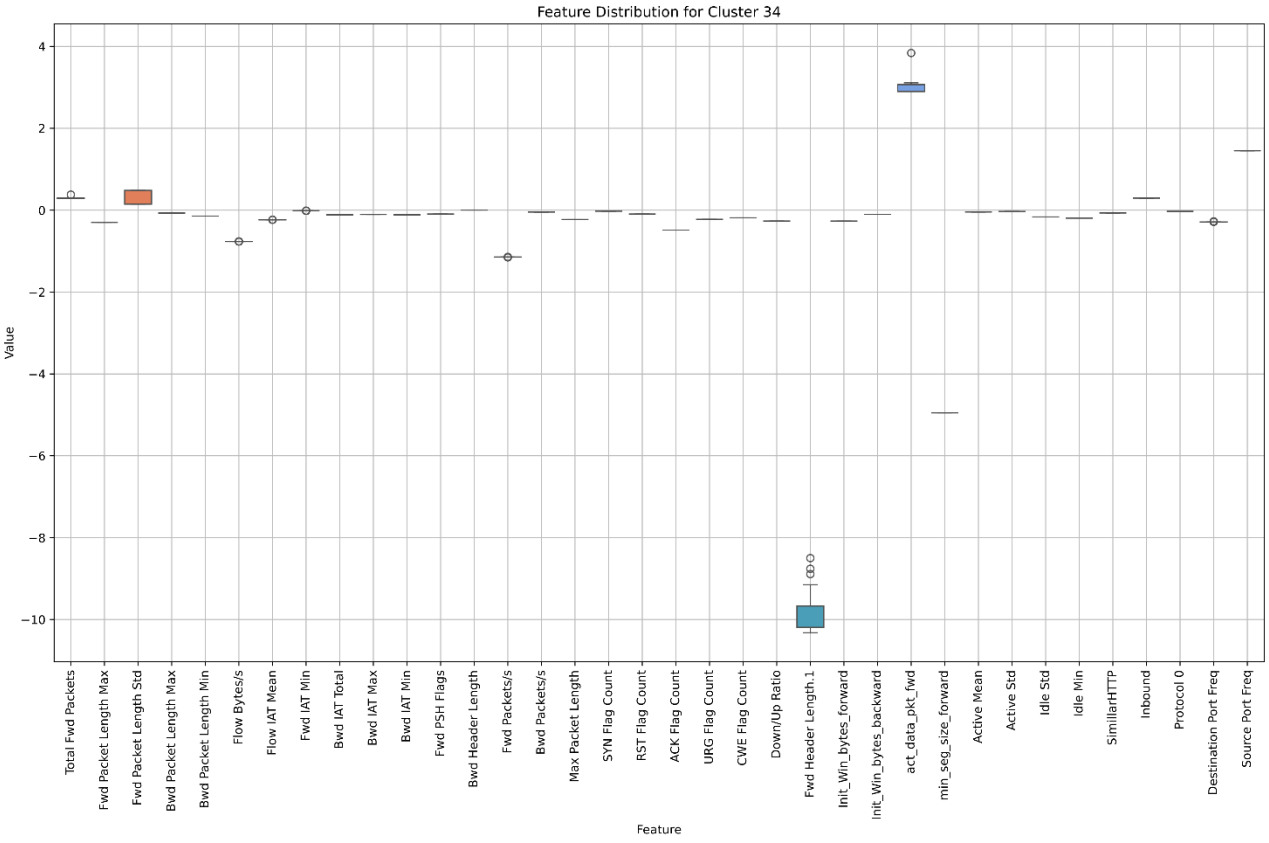

Feature Distribution for Cluster 29
Value
Feature
Total Fwd Packets
Fwd Packet Length Max
Fwd Packet Length Std
Bwd Packet Length Max
Bwd Packet Length Min
Flow Bytes/s
Flow IAT Mean
Fwd IAT Min
Bwd IAT Total
Bwd IAT Max
Bwd IAT Min
Fwd PSH Flags
Bwd Header Length
Fwd Packets/s
Bwd Packets/s
Max Packet Length
SYN Flag Count
RST Flag Count
ACK Flag Count
URG Flag Count
CWE Flag Count
Down/Up Ratio
Fwd Header Length.1
Init_Win_bytes_forward
Init_Win_bytes_backward
act_data_pkt_fwd
min_seg_size_forward
Active Mean
Active Std
Idle Std
Idle Min
SimillarHTTP
Inbound
Protocol 0
Destination Port Freq
Source Port Freq

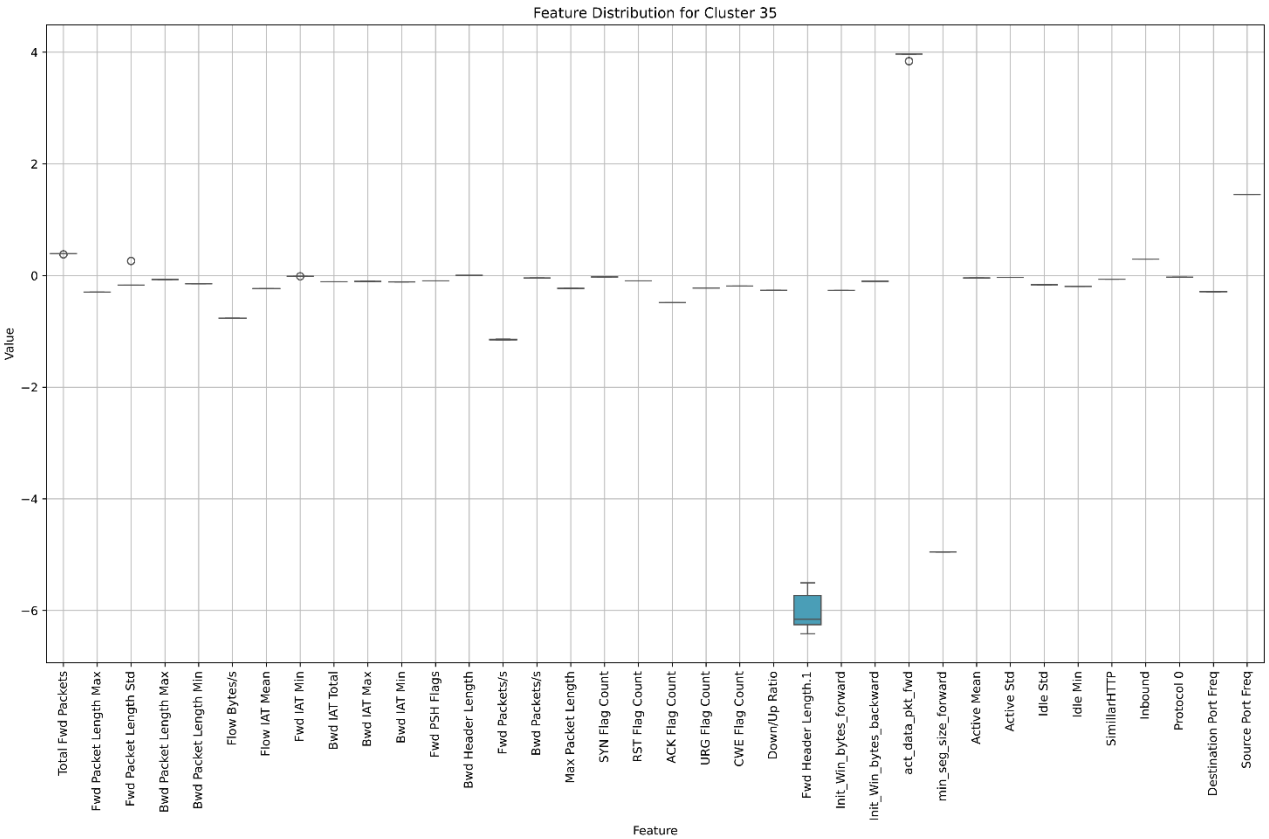

Feature Distribution for Cluster 30
Value
Feature
Total Fwd Packets
Fwd Packet Length Max
Fwd Packet Length Std
Bwd Packet Length Max
Bwd Packet Length Min
Flow Bytes/s
Flow IAT Mean
Fwd IAT Min
Bwd IAT Total
Bwd IAT Max
Bwd IAT Min
Fwd PSH Flags
Bwd Header Length
Fwd Packets/s
Bwd Packets/s
Max Packet Length
SYN Flag Count
RST Flag Count
ACK Flag Count
URG Flag Count
CWE Flag Count
Down/Up Ratio
Fwd Header Length.1
Init_Win_bytes_forward
Init_Win_bytes_backward
act_data_pkt_fwd
min_seg_size_forward
Active Mean
Active Std
Idle Std
Idle Min
SimillarHTTP
Inbound
Protocol 0
Destination Port Freq
Source Port Freq

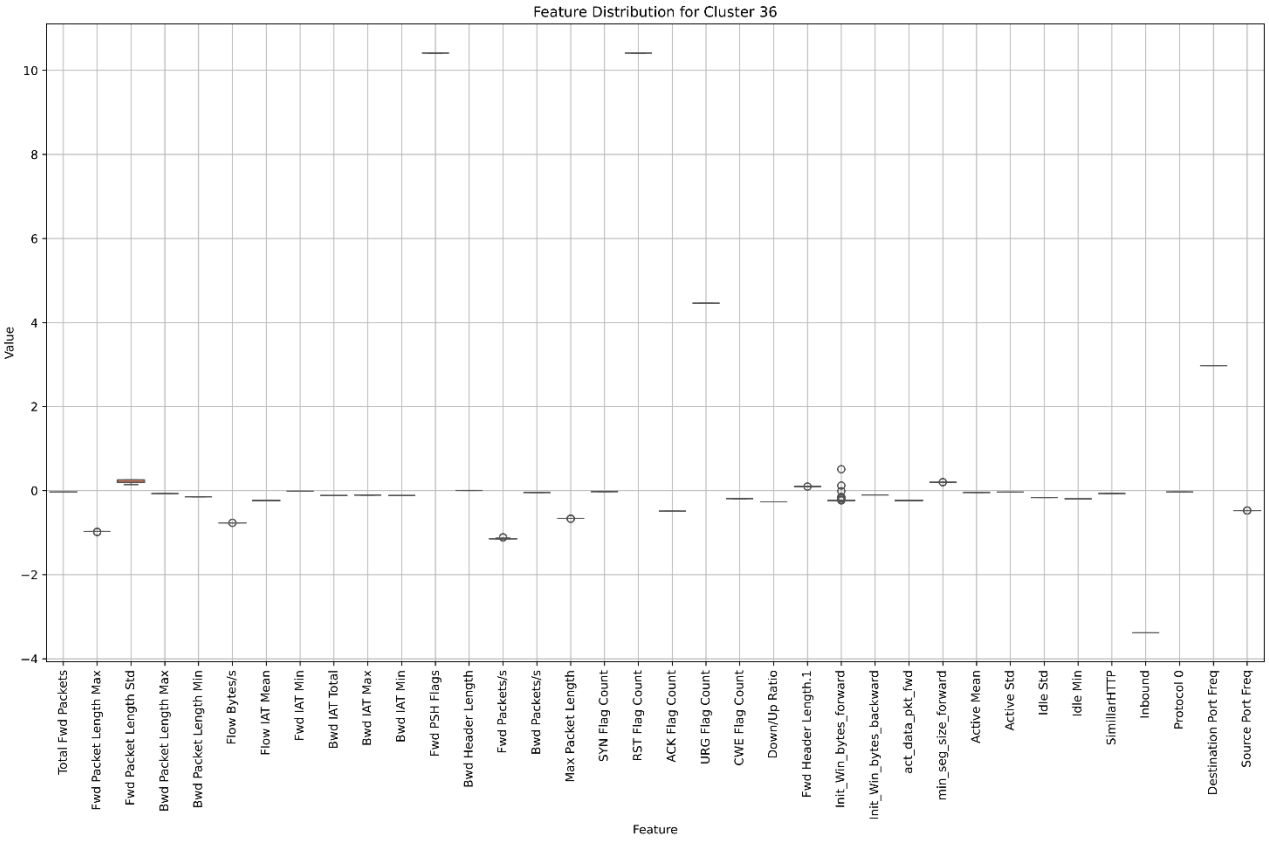

Feature Distribution for Cluster 31
Value
Feature
Total Fwd Packets
Fwd Packet Length Max
Fwd Packet Length Std
Bwd Packet Length Max
Bwd Packet Length Min
Flow Bytes/s
Flow IAT Mean
Fwd IAT Min
Bwd IAT Total
Bwd IAT Max
Bwd IAT Min
Fwd PSH Flags
Bwd Header Length
Fwd Packets/s
Bwd Packets/s
Max Packet Length
SYN Flag Count
RST Flag Count
ACK Flag Count
URG Flag Count
CWE Flag Count
Down/Up Ratio
Fwd Header Length.1
Init_Win_bytes_forward
Init_Win_bytes_backward
act_data_pkt_fwd
min_seg_size_forward
Active Mean
Active Std
Idle Std
Idle Min
SimillarHTTP
Inbound
Protocol 0
Destination Port Freq
Source Port Freq

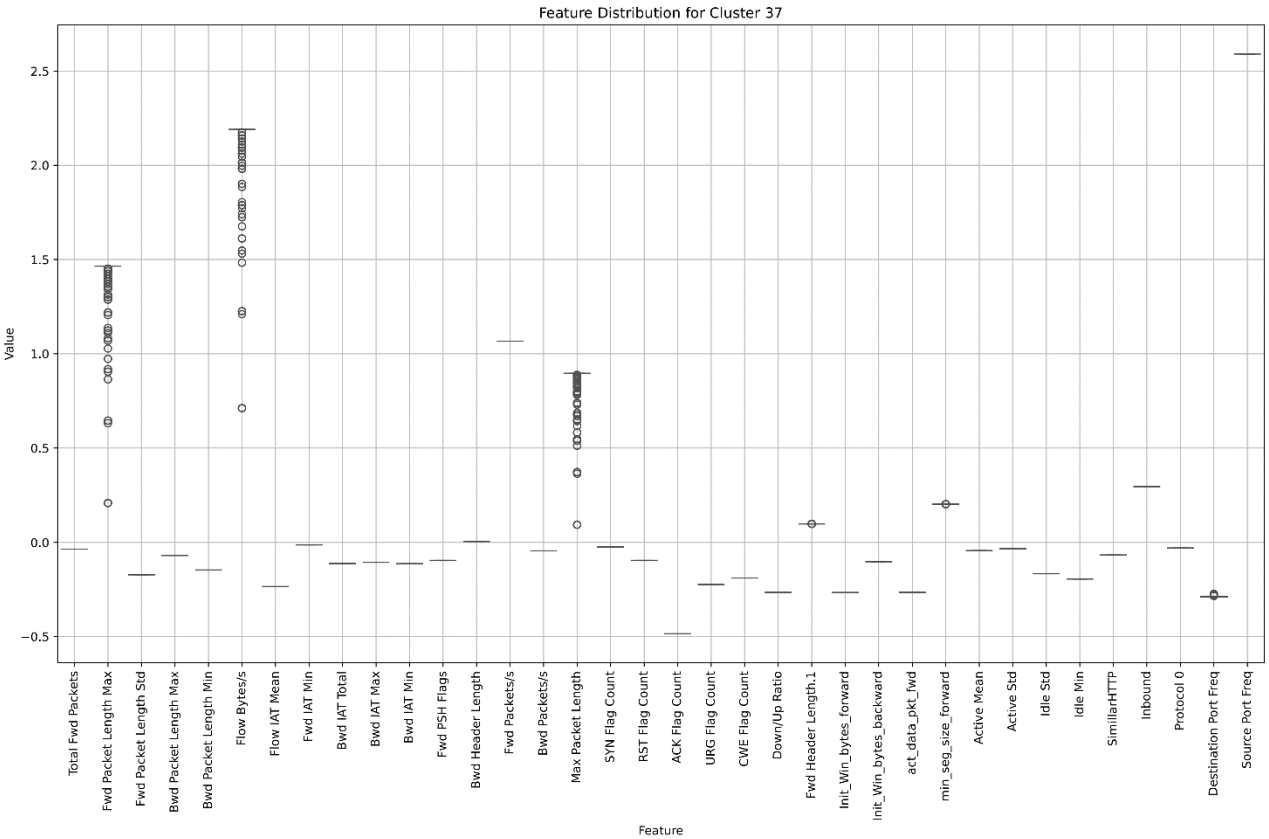

Feature Distribution for Cluster 32
Value
Feature
Total Fwd Packets
Fwd Packet Length Max
Fwd Packet Length Std
Bwd Packet Length Max
Bwd Packet Length Min
Flow Bytes/s
Flow IAT Mean
Fwd IAT Min
Bwd IAT Total
Bwd IAT Max
Bwd IAT Min
Fwd PSH Flags
Bwd Header Length
Fwd Packets/s
Bwd Packets/s
Max Packet Length
SYN Flag Count
RST Flag Count
ACK Flag Count
URG Flag Count
CWE Flag Count
Down/Up Ratio
Fwd Header Length.1
Init_Win_bytes_forward
Init_Win_bytes_backward
act_data_pkt_fwd
min_seg_size_forward
Active Mean
Active Std
Idle Std
Idle Min
SimillarHTTP
Inbound
Protocol 0
Destination Port Freq
Source Port Freq

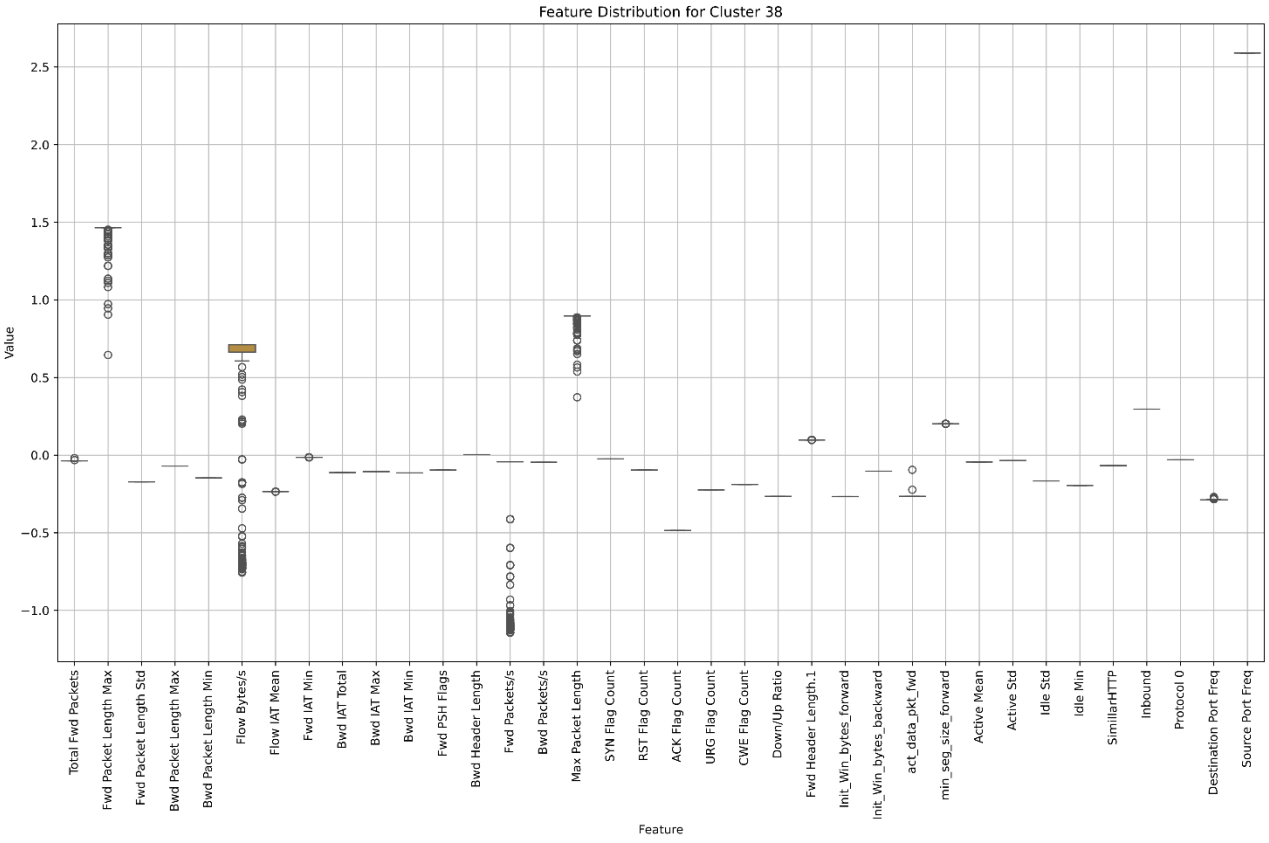

Feature Distribution for Cluster 33
Value
Feature
Total Fwd Packets
Fwd Packet Length Max
Fwd Packet Length Std
Bwd Packet Length Max
Bwd Packet Length Min
Flow Bytes/s
Flow IAT Mean
Fwd IAT Min
Bwd IAT Total
Bwd IAT Max
Bwd IAT Min
Fwd PSH Flags
Bwd Header Length
Fwd Packets/s
Bwd Packets/s
Max Packet Length
SYN Flag Count
RST Flag Count
ACK Flag Count
URG Flag Count
CWE Flag Count
Down/Up Ratio
Fwd Header Length.1
Init_Win_bytes_forward
Init_Win_bytes_backward
act_data_pkt_fwd
min_seg_size_forward
Active Mean
Active Std
Idle Std
Idle Min
SimillarHTTP
Inbound
Protocol 0
Destination Port Freq
Source Port Freq

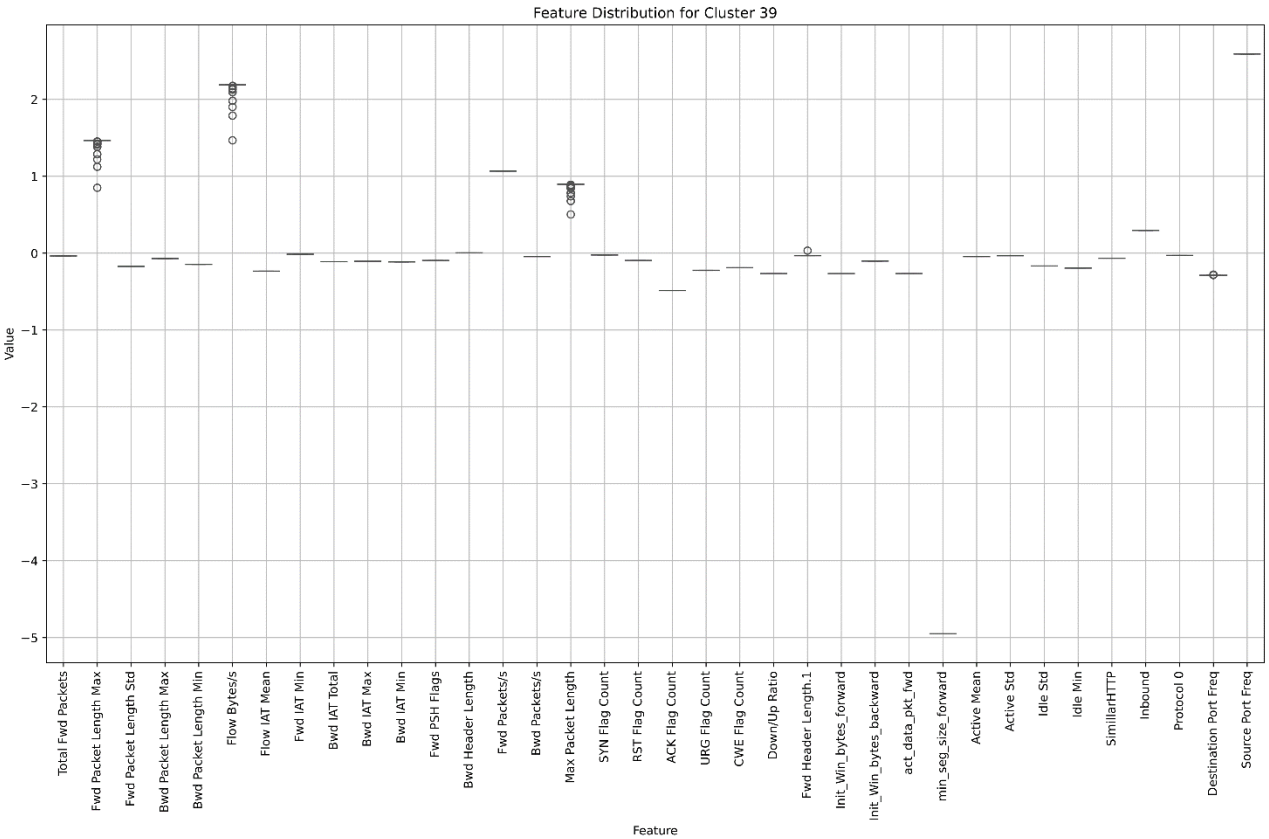

Feature Distribution for Cluster 34
Value
Feature
Total Fwd Packets
Fwd Packet Length Max
Fwd Packet Length Std
Bwd Packet Length Max
Bwd Packet Length Min
Flow Bytes/s
Flow IAT Mean
Fwd IAT Min
Bwd IAT Total
Bwd IAT Max
Bwd IAT Min
Fwd PSH Flags
Bwd Header Length
Fwd Packets/s
Bwd Packets/s
Max Packet Length
SYN Flag Count
RST Flag Count
ACK Flag Count
URG Flag Count
CWE Flag Count
Down/Up Ratio
Fwd Header Length.1
Init_Win_bytes_forward
Init_Win_bytes_backward
act_data_pkt_fwd
min_seg_size_forward
Active Mean
Active Std
Idle Std
Idle Min
SimillarHTTP
Inbound
Protocol 0
Destination Port Freq
Source Port Freq

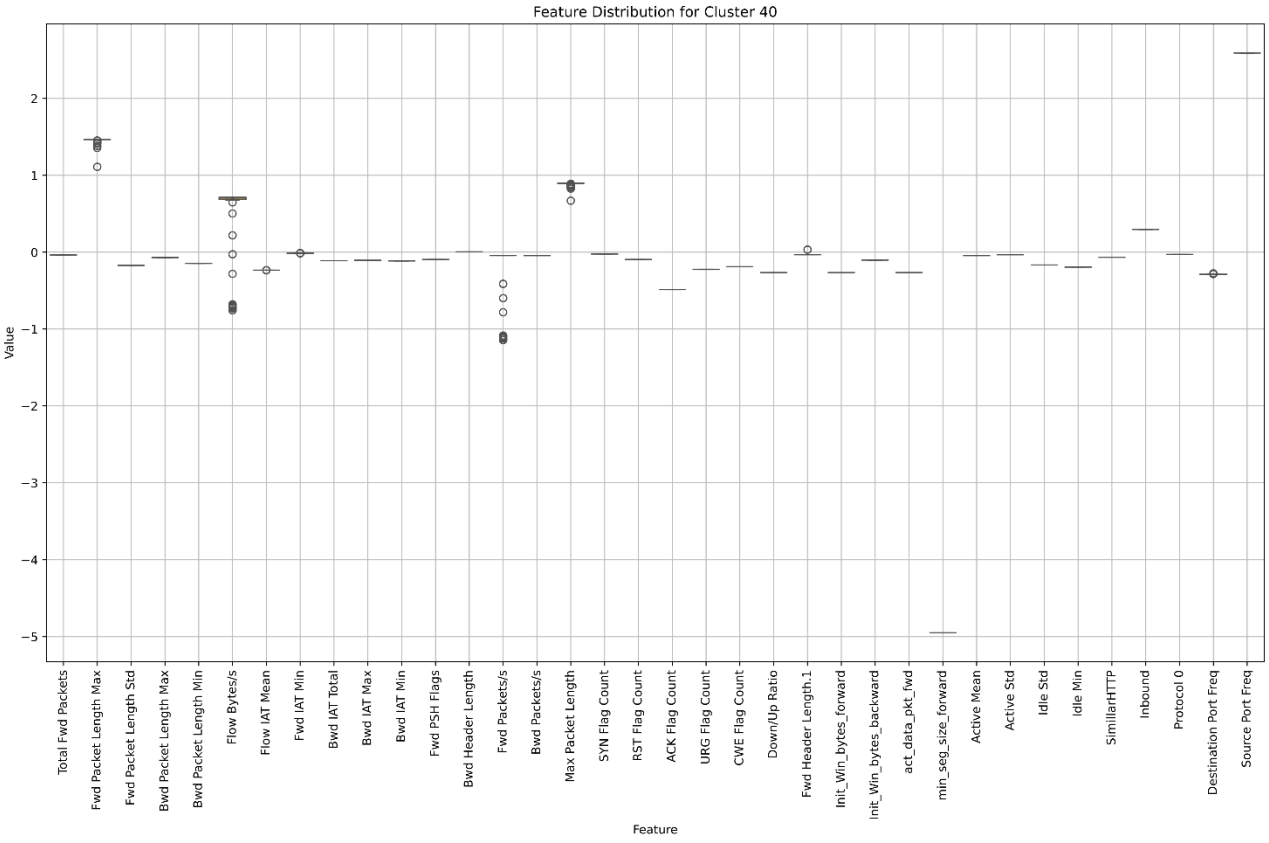

Feature Distribution for Cluster 35
Value
Feature
Total Fwd Packets
Fwd Packet Length Max
Fwd Packet Length Std
Bwd Packet Length Max
Bwd Packet Length Min
Flow Bytes/s
Flow IAT Mean
Fwd IAT Min
Bwd IAT Total
Bwd IAT Max
Bwd IAT Min
Fwd PSH Flags
Bwd Header Length
Fwd Packets/s
Bwd Packets/s
Max Packet Length
SYN Flag Count
RST Flag Count
ACK Flag Count
URG Flag Count
CWE Flag Count
Down/Up Ratio
Fwd Header Length.1
Init_Win_bytes_forward
Init_Win_bytes_backward
act_data_pkt_fwd
min_seg_size_forward
Active Mean
Active Std
Idle Std
Idle Min
SimillarHTTP
Inbound
Protocol 0
Destination Port Freq
Source Port Freq

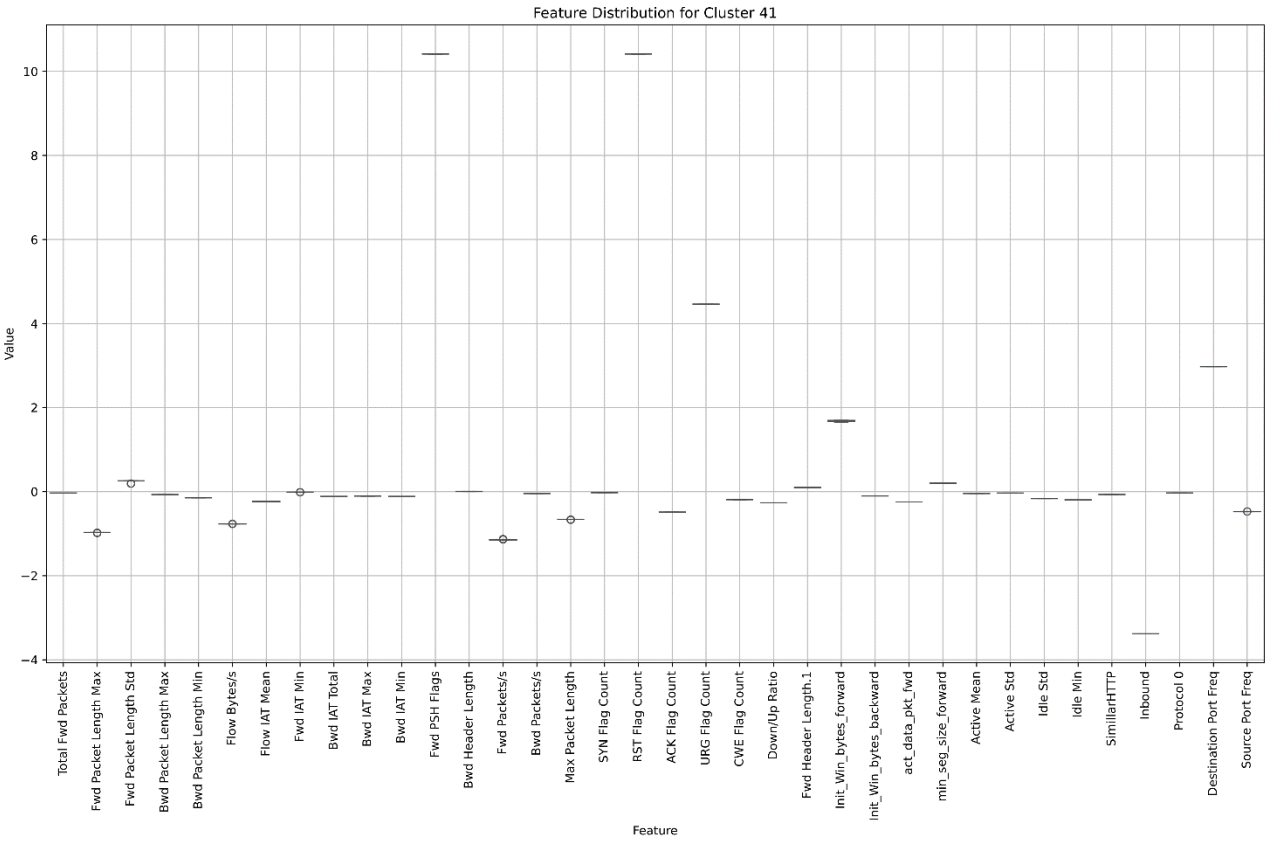

Feature Distribution for Cluster 36
Value
Feature
Total Fwd Packets
Fwd Packet Length Max
Fwd Packet Length Std
Bwd Packet Length Max
Bwd Packet Length Min
Flow Bytes/s
Flow IAT Mean
Fwd IAT Min
Bwd IAT Total
Bwd IAT Max
Bwd IAT Min
Fwd PSH Flags
Bwd Header Length
Fwd Packets/s
Bwd Packets/s
Max Packet Length
SYN Flag Count
RST Flag Count
ACK Flag Count
URG Flag Count
CWE Flag Count
Down/Up Ratio
Fwd Header Length.1
Init_Win_bytes_forward
Init_Win_bytes_backward
act_data_pkt_fwd
min_seg_size_forward
Active Mean
Active Std
Idle Std
Idle Min
SimillarHTTP
Inbound
Protocol 0
Destination Port Freq
Source Port Freq

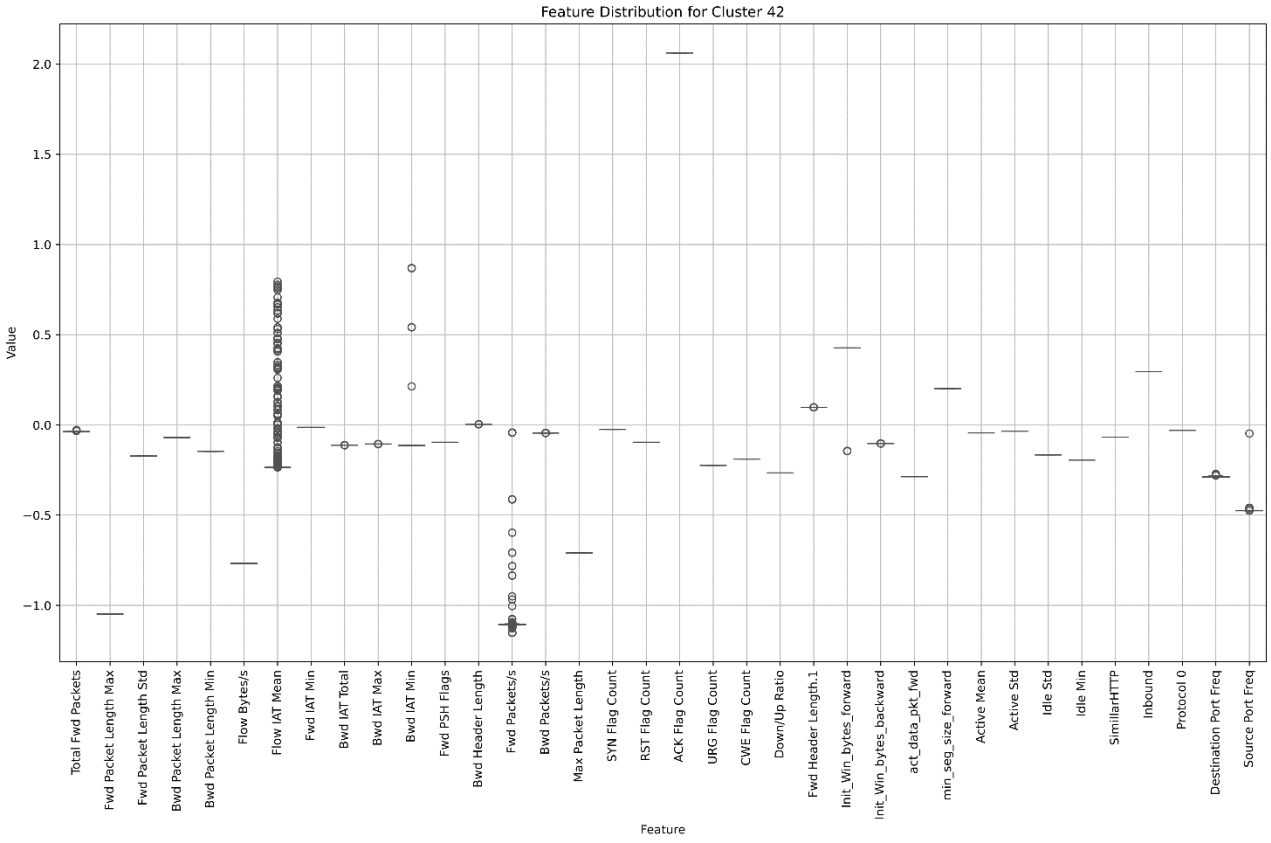

Feature Distribution for Cluster 37
Value
Feature
Total Fwd Packets
Fwd Packet Length Max
Fwd Packet Length Std
Bwd Packet Length Max
Bwd Packet Length Min
Flow Bytes/s
Flow IAT Mean
Fwd IAT Min
Bwd IAT Total
Bwd IAT Max
Bwd IAT Min
Fwd PSH Flags
Bwd Header Length
Fwd Packets/s
Bwd Packets/s
Max Packet Length
SYN Flag Count
RST Flag Count
ACK Flag Count
URG Flag Count
CWE Flag Count
Down/Up Ratio
Fwd Header Length.1
Init_Win_bytes_forward
Init_Win_bytes_backward
act_data_pkt_fwd
min_seg_size_forward
Active Mean
Active Std
Idle Std
Idle Min
SimillarHTTP
Inbound
Protocol 0
Destination Port Freq
Source Port Freq

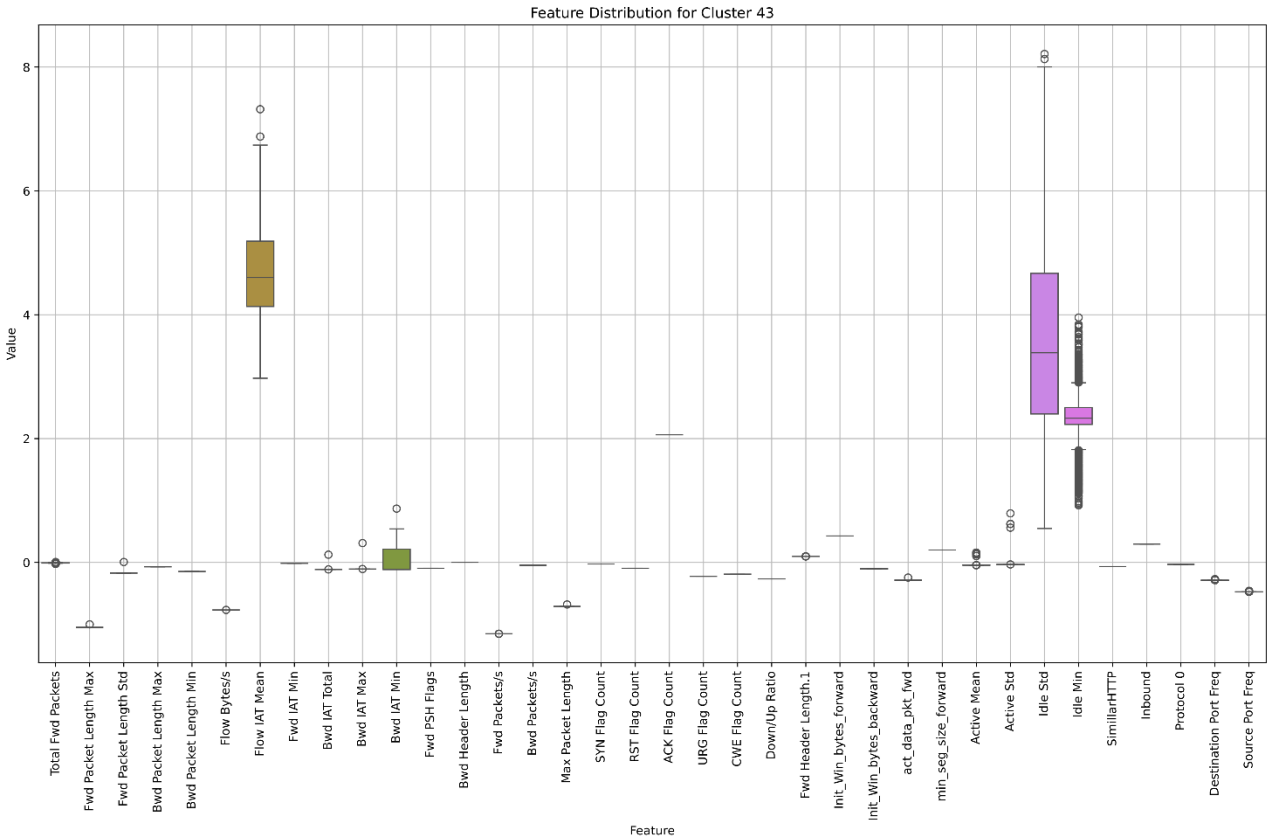

Feature Distribution for Cluster 38
Value
Feature
Total Fwd Packets
Fwd Packet Length Max
Fwd Packet Length Std
Bwd Packet Length Max
Bwd Packet Length Min
Flow Bytes/s
Flow IAT Mean
Fwd IAT Min
Bwd IAT Total
Bwd IAT Max
Bwd IAT Min
Fwd PSH Flags
Bwd Header Length
Fwd Packets/s
Bwd Packets/s
Max Packet Length
SYN Flag Count
RST Flag Count
ACK Flag Count
URG Flag Count
CWE Flag Count
Down/Up Ratio
Fwd Header Length.1
Init_Win_bytes_forward
Init_Win_bytes_backward
act_data_pkt_fwd
min_seg_size_forward
Active Mean
Active Std
Idle Std
Idle Min
SimillarHTTP
Inbound
Protocol 0
Destination Port Freq
Source Port Freq

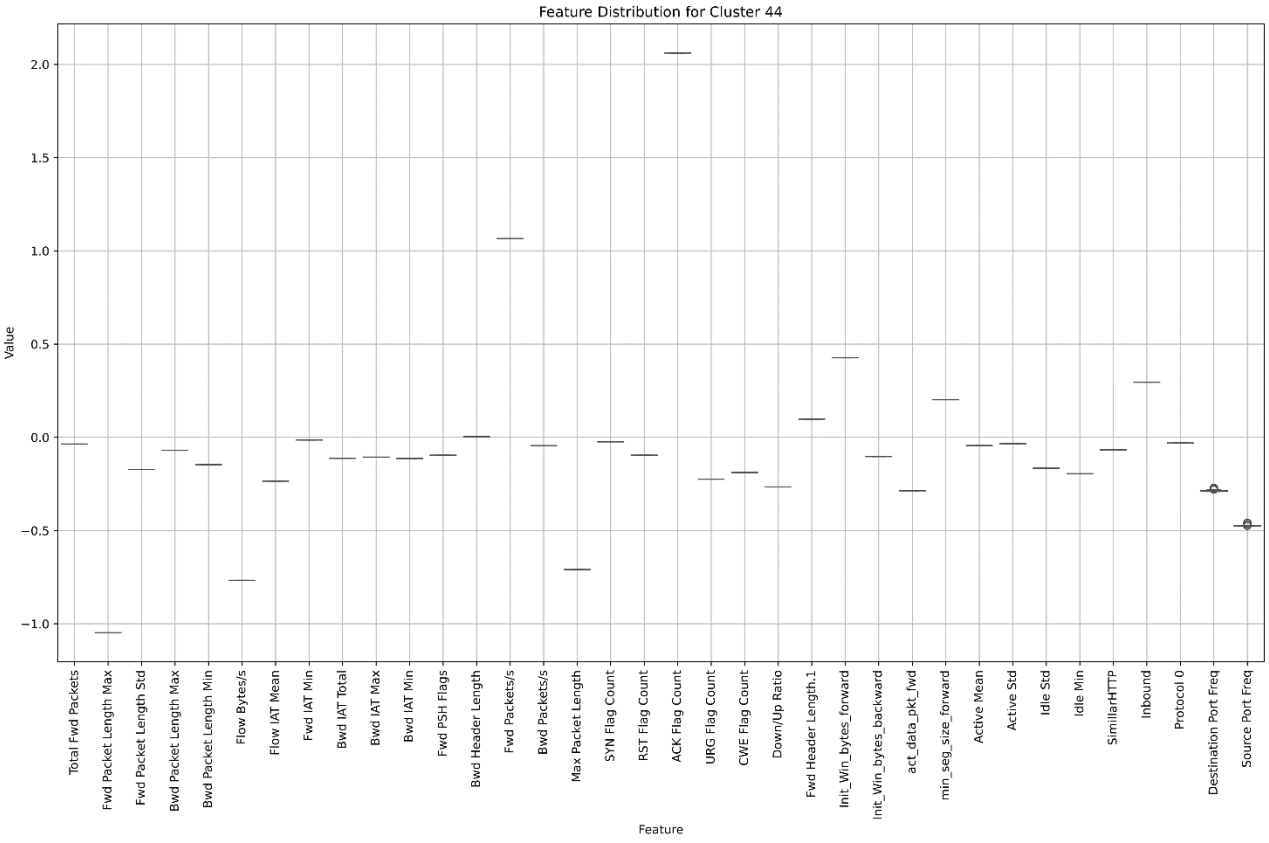

Feature Distribution for Cluster 39
Value
Feature
Total Fwd Packets
Fwd Packet Length Max
Fwd Packet Length Std
Bwd Packet Length Max
Bwd Packet Length Min
Flow Bytes/s
Flow IAT Mean
Fwd IAT Min
Bwd IAT Total
Bwd IAT Max
Bwd IAT Min
Fwd PSH Flags
Bwd Header Length
Fwd Packets/s
Bwd Packets/s
Max Packet Length
SYN Flag Count
RST Flag Count
ACK Flag Count
URG Flag Count
CWE Flag Count
Down/Up Ratio
Fwd Header Length.1
Init_Win_bytes_forward
Init_Win_bytes_backward
act_data_pkt_fwd
min_seg_size_forward
Active Mean
Active Std
Idle Std
Idle Min
SimillarHTTP
Inbound
Protocol 0
Destination Port Freq
Source Port Freq

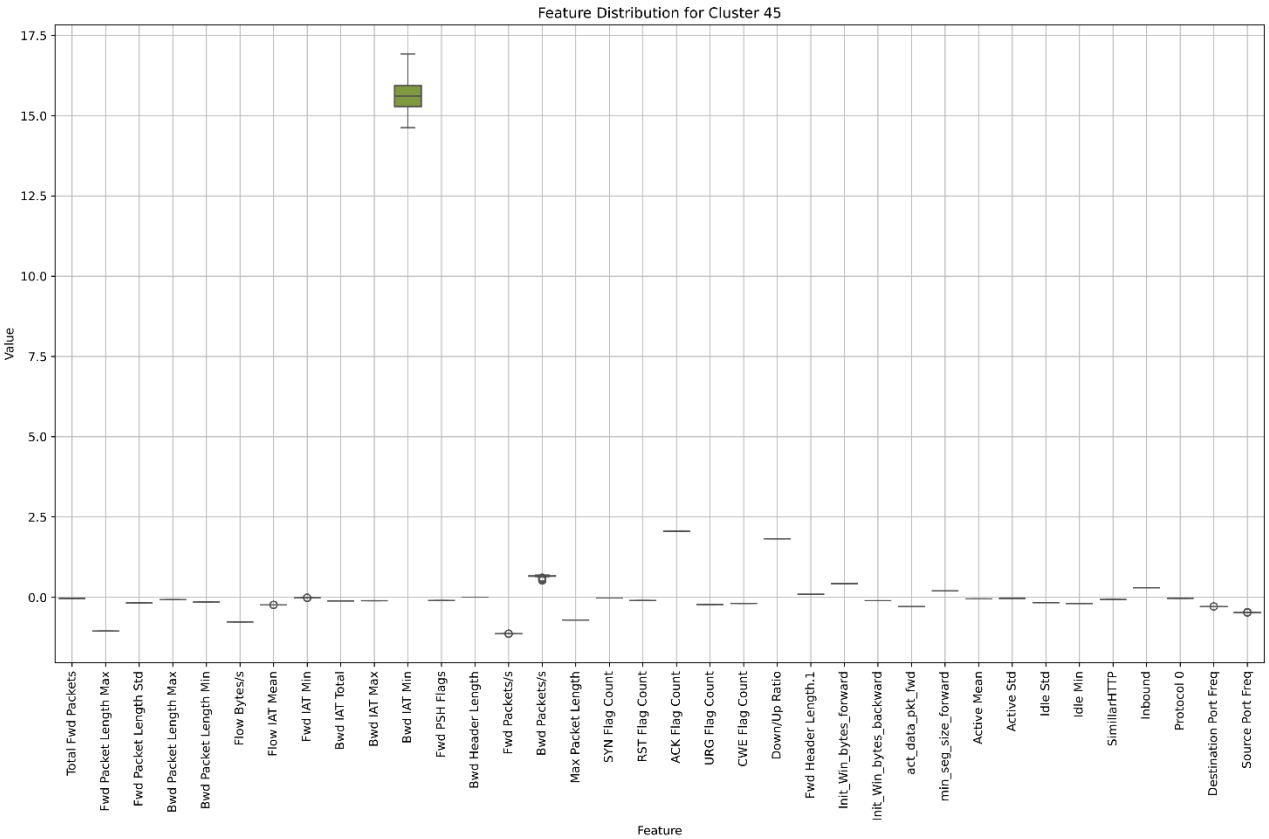

Feature Distribution for Cluster 40
Value
Feature
Total Fwd Packets
Fwd Packet Length Max
Fwd Packet Length Std
Bwd Packet Length Max
Bwd Packet Length Min
Flow Bytes/s
Flow IAT Mean
Fwd IAT Min
Bwd IAT Total
Bwd IAT Max
Bwd IAT Min
Fwd PSH Flags
Bwd Header Length
Fwd Packets/s
Bwd Packets/s
Max Packet Length
SYN Flag Count
RST Flag Count
ACK Flag Count
URG Flag Count
CWE Flag Count
Down/Up Ratio
Fwd Header Length.1
Init_Win_bytes_forward
Init_Win_bytes_backward
act_data_pkt_fwd
min_seg_size_forward
Active Mean
Active Std
Idle Std
Idle Min
SimillarHTTP
Inbound
Protocol 0
Destination Port Freq
Source Port Freq

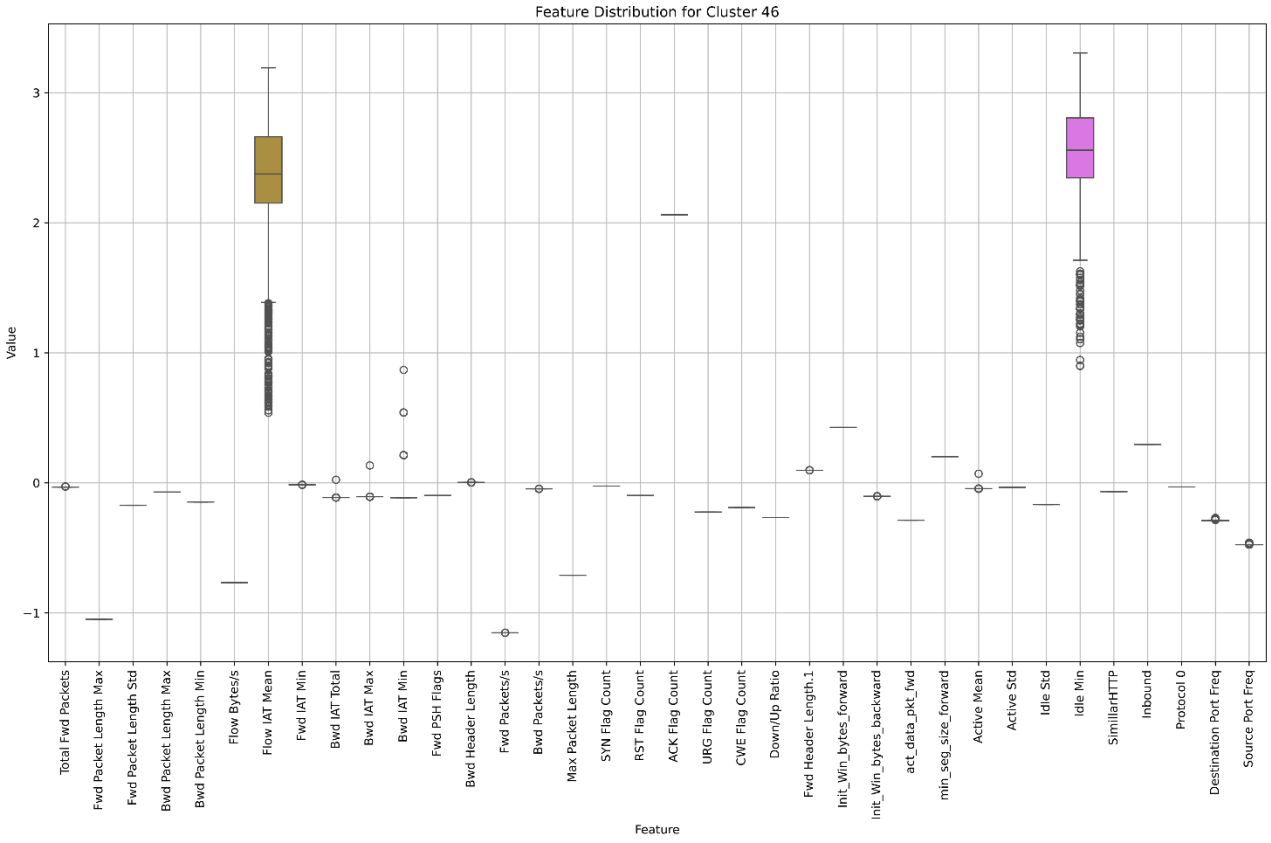

Feature Distribution for Cluster 41
Value
Feature
Total Fwd Packets
Fwd Packet Length Max
Fwd Packet Length Std
Bwd Packet Length Max
Bwd Packet Length Min
Flow Bytes/s
Flow IAT Mean
Fwd IAT Min
Bwd IAT Total
Bwd IAT Max
Bwd IAT Min
Fwd PSH Flags
Bwd Header Length
Fwd Packets/s
Bwd Packets/s
Max Packet Length
SYN Flag Count
RST Flag Count
ACK Flag Count
URG Flag Count
CWE Flag Count
Down/Up Ratio
Fwd Header Length.1
Init_Win_bytes_forward
Init_Win_bytes_backward
act_data_pkt_fwd
min_seg_size_forward
Active Mean
Active Std
Idle Std
Idle Min
SimillarHTTP
Inbound
Protocol 0
Destination Port Freq
Source Port Freq


Feature Distribution for Cluster 42
Value
Feature
Total Fwd Packets
Fwd Packet Length Max
Fwd Packet Length Std
Bwd Packet Length Max
Bwd Packet Length Min
Flow Bytes/s
Flow IAT Mean
Fwd IAT Min
Bwd IAT Total
Bwd IAT Max
Bwd IAT Min
Fwd PSH Flags
Bwd Header Length
Fwd Packets/s
Bwd Packets/s
Max Packet Length
SYN Flag Count
RST Flag Count
ACK Flag Count
URG Flag Count
CWE Flag Count
Down/Up Ratio
Fwd Header Length.1
Init_Win_bytes_forward
Init_Win_bytes_backward
act_data_pkt_fwd
min_seg_size_forward
Active Mean
Active Std
Idle Std
Idle Min
SimillarHTTP
Inbound
Protocol 0
Destination Port Freq
Source Port Freq


Feature Distribution for Cluster 43
Value
Feature
Total Fwd Packets
Fwd Packet Length Max
Fwd Packet Length Std
Bwd Packet Length Max
Bwd Packet Length Min
Flow Bytes/s
Flow IAT Mean
Fwd IAT Min
Bwd IAT Total
Bwd IAT Max
Bwd IAT Min
Fwd PSH Flags
Bwd Header Length
Fwd Packets/s
Bwd Packets/s
Max Packet Length
SYN Flag Count
RST Flag Count
ACK Flag Count
URG Flag Count
CWE Flag Count
Down/Up Ratio
Fwd Header Length.1
Init_Win_bytes_forward
Init_Win_bytes_backward
act_data_pkt_fwd
min_seg_size_forward
Active Mean
Active Std
Idle Std
Idle Min
SimillarHTTP
Inbound
Protocol 0
Destination Port Freq
Source Port Freq


Feature Distribution for Cluster 44
Value
Feature
Total Fwd Packets
Fwd Packet Length Max
Fwd Packet Length Std
Bwd Packet Length Max
Bwd Packet Length Min
Flow Bytes/s
Flow IAT Mean
Fwd IAT Min
Bwd IAT Total
Bwd IAT Max
Bwd IAT Min
Fwd PSH Flags
Bwd Header Length
Fwd Packets/s
Bwd Packets/s
Max Packet Length
SYN Flag Count
RST Flag Count
ACK Flag Count
URG Flag Count
CWE Flag Count
Down/Up Ratio
Fwd Header Length.1
Init_Win_bytes_forward
Init_Win_bytes_backward
act_data_pkt_fwd
min_seg_size_forward
Active Mean
Active Std
Idle Std
Idle Min
SimillarHTTP
Inbound
Protocol 0
Destination Port Freq
Source Port Freq


Feature Distribution for Cluster 45
Value
Feature
Total Fwd Packets
Fwd Packet Length Max
Fwd Packet Length Std
Bwd Packet Length Max
Bwd Packet Length Min
Flow Bytes/s
Flow IAT Mean
Fwd IAT Min
Bwd IAT Total
Bwd IAT Max
Bwd IAT Min
Fwd PSH Flags
Bwd Header Length
Fwd Packets/s
Bwd Packets/s
Max Packet Length
SYN Flag Count
RST Flag Count
ACK Flag Count
URG Flag Count
CWE Flag Count
Down/Up Ratio
Fwd Header Length.1
Init_Win_bytes_forward
Init_Win_bytes_backward
act_data_pkt_fwd
min_seg_size_forward
Active Mean
Active Std
Idle Std
Idle Min
SimillarHTTP
Inbound
Protocol 0
Destination Port Freq
Source Port Freq




2. Multi-class classification: also here, as done in k-means, the combination of shap with xgboost was adopted to identify the contribution of each individual feature to the classification decision:

As with K-Means, the two methods also return consistent results here. Indeed, a quick glance at a few clusters reveals that the most significant features chosen by both models are identical. We can now integrate all the previously gathered data with the understanding of feature weights within the individual clusters to highlight the similarities and address the misclassifications in our DBSCAN clustering algorithm effectively.

*Detection of Similarities*

The next step a similarity analysis that will be conducted among the attacks, identifying common features, in order to better understand the nature of the attacks and detect recurring patterns in malicious traffic.

*Pure (or nearly pure) clusters (previously identified):*

| *Cluster* | 0 | 1 | 2 | 3 | 4 | 5 | 6 |
|---|---|---|---|---|---|---|---|
| *Label* | ddos_dns | ddos_dns | ddos_dns | ddos_dns | ddos_dns | ddos_dns | ddos_dns |

| *Cluster* | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 |
|---|---|---|---|---|---|---|---|---|---|
| *Label* | ddos_dns | benign | benign | benign | ddos_ntp | ddos_dns | ddos_ntp | ddos_ntp | ddos_ntp |

| *Cluster* | 16 | 17 | 18 | 19 |
|---|---|---|---|---|
| *Label* | benign | benign | benign | ddos_ntp |

| *Cluster* | 20 | 21 | 22 | 23 | 24 | 27 | 28 | 32 |
|---|---|---|---|---|---|---|---|---|
| *Label* | ddos_ntp | benign | benign | benign | ddos_ntp | ddos_ntp | benign | benign |

| *Cluster* | 33 | 34 | 35 | 36 | 41 | 43 | 46 |
|---|---|---|---|---|---|---|---|
| *Label* | benign | ddos_dns | ddos_dns | benign | benign | ddos_syn | ddos_tftp |

*To be classified:*

| *Cluster* | 25 | 26 | 29 | 30 | 31 | 37 | 38 | 39 | 40 | 42 | 44 | 45 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|

- *Clusters 25, 42, 44, 45:* they are all characterized by an almost identical distribution of labels between ddos_syn and ddos_tftp. As previously analyzed, the nearly identical distributions of ddos_syn and ddos_tftp labels in these clusters suggest that these attack types share significant similarities in the dataset's feature space. By closely examining the cluster boxplots (intra-cluster) and the SHAP plots, we observe that the most important shared features are ACK flag count, Flow IAT Mean, and FWD Packets/s. Cluster 45 is also included in this discussion, albeit with many fewer samples involved. Furthermore, the t-SNE graph reveals that the ddos_syn and ddos_tftp clusters overlap in a sparse manner, indicating that the DBSCAN algorithm struggled to distinguish between these attack types effectively.
- *Clusters 26, 29:* For these two clusters, the similarity in the contingency table values suggests a commonality in the feature distributions within these clusters.

The provided boxplots and SHAP graphs for Clusters 26 and 29 visually reinforce this observation, indicating that these clusters share similar feature distributions. Specifically, features like Total Fwd Packets, Max Packet Length, Flow Bytes/s, FWD Packets/s, and Source Port Frequency exhibit similar distributions in both clusters, further demonstrating their significance.

- *Clusters 30,31:* These clusters have an almost identical distribution of the percentage between the labels (both containing ddos_syn and ddos_tftp at 44.01% and 55.99% for one cluster, and 42.02% and 57.98% for the other). Examining the features, both the boxplots and SHAP graphs confirm this strong similarity, highlighting the same important features: Fwd Packet Length Max, Flow Bytes/s, Max Packet Length, Fwd Header Length.1, and Source Port Frequency. The similarity in feature importance and distribution indicates that the underlying characteristics of the attacks within these clusters may overlap significantly. For instance, the flow and packet metrics exhibit similar patterns for ddos_syn and ddos_tftp attacks, making it challenging for the clustering algorithm to differentiate between them based solely on these features. As evidence of this, looking at the t-SNE graph it is noticeable how the two labels overlap a lot.
- *Clusters 37,38,39,40:* The boxplots and SHAP graphs for Clusters 37, 38, 39, and 40 visually reinforce the observation that these clusters share similar feature distributions. The boxplots show that features like Source Port Frequency and Flow Bytes/s have similar distributions across these clusters, while the SHAP graphs highlight the impact of these features on the model's output, further demonstrating their significance.