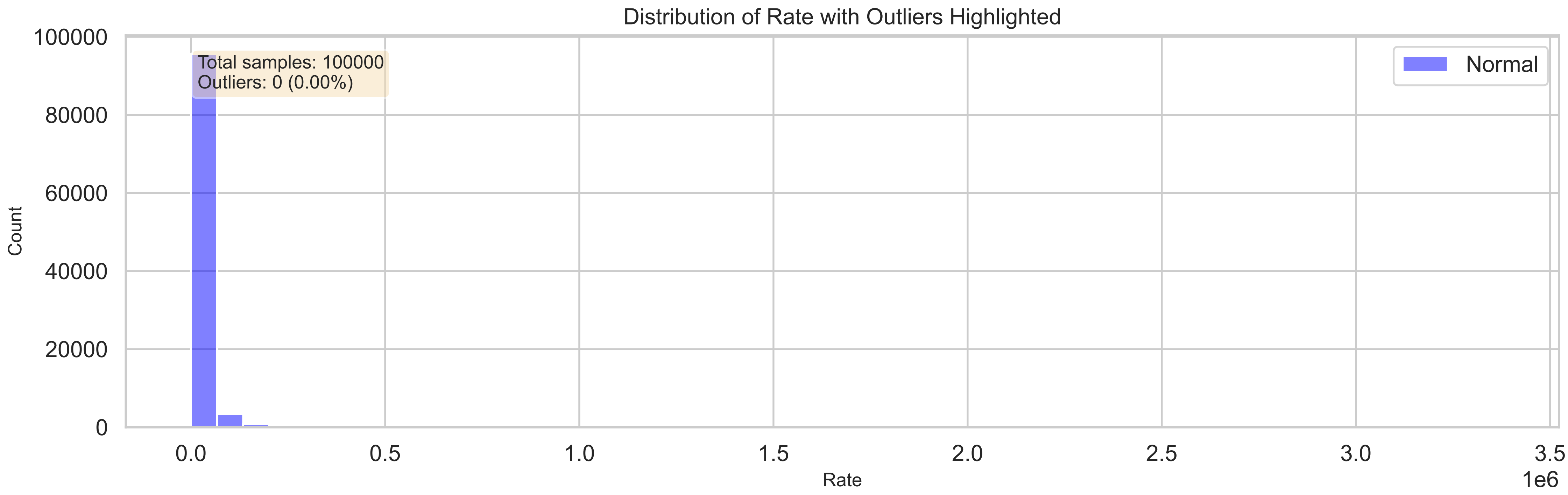
IoT Security Threat Detection for SMEs:

A Machine Learning Approach Using CIC-IoT Dataset

STAGE 3, STEP 3: ANOMALY IDENTIFICATION CRITERIA

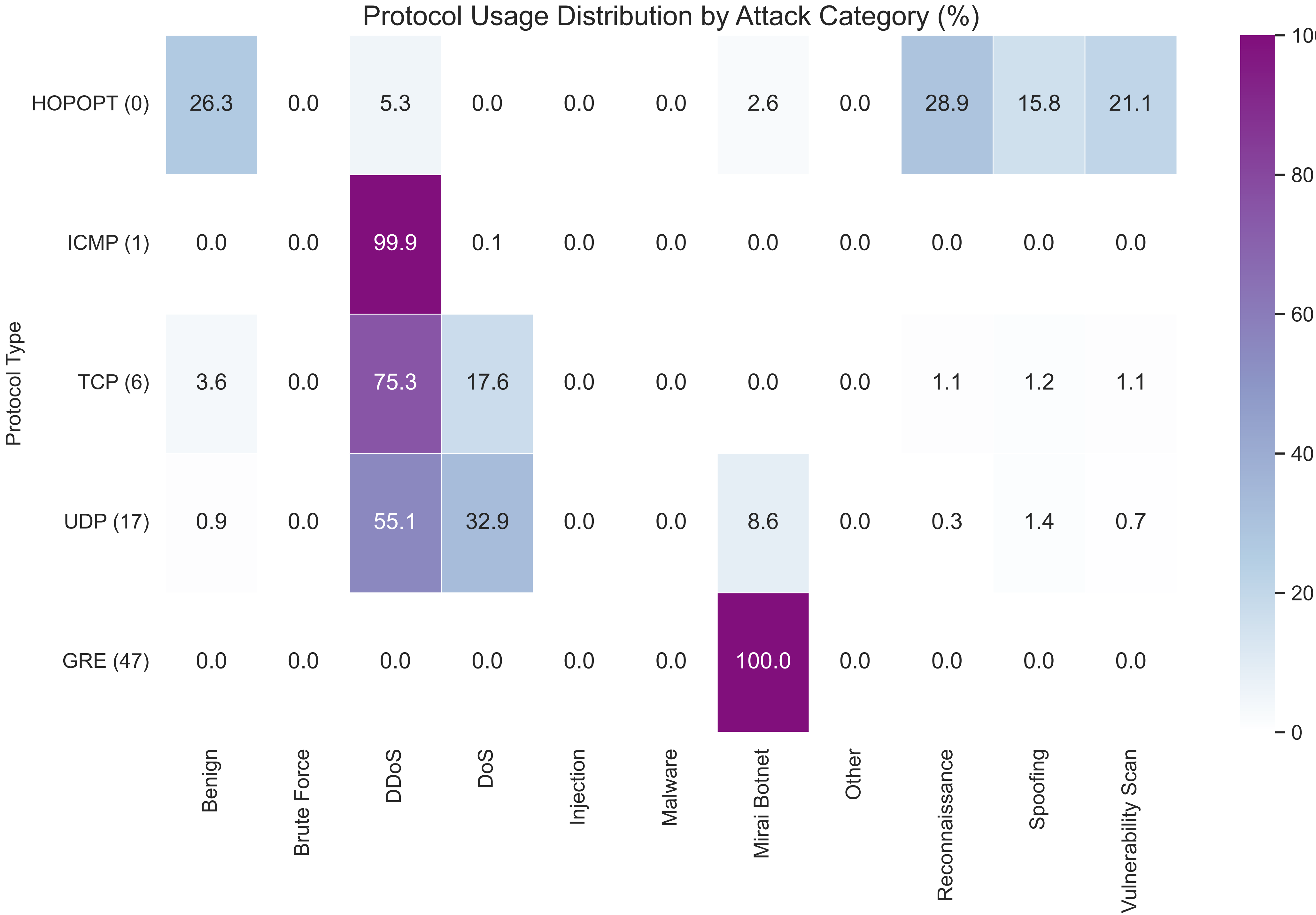
This report analyzes key anomaly identification criteria for IoT security threat detection, focusing on statistical outliers, unusual protocol usage, suspicious timing patterns, irregular packet size distributions, and unexpected flag combinations.



Distribution of Rate with Outliers Highlighted
Total samples: 100000
Outliers: 0 (0.00%)
Normal
100000
80000
60000
40000
20000
0
Count
0.0
0.5
1.0
1.5
2.0
2.5
3.0
3.5
1e6
Rate

These histograms illustrate the distribution of key traffic metrics with statistical outliers highlighted in red. Outliers are identified using Z-score methodology, with values beyond 3 standard deviations from the mean

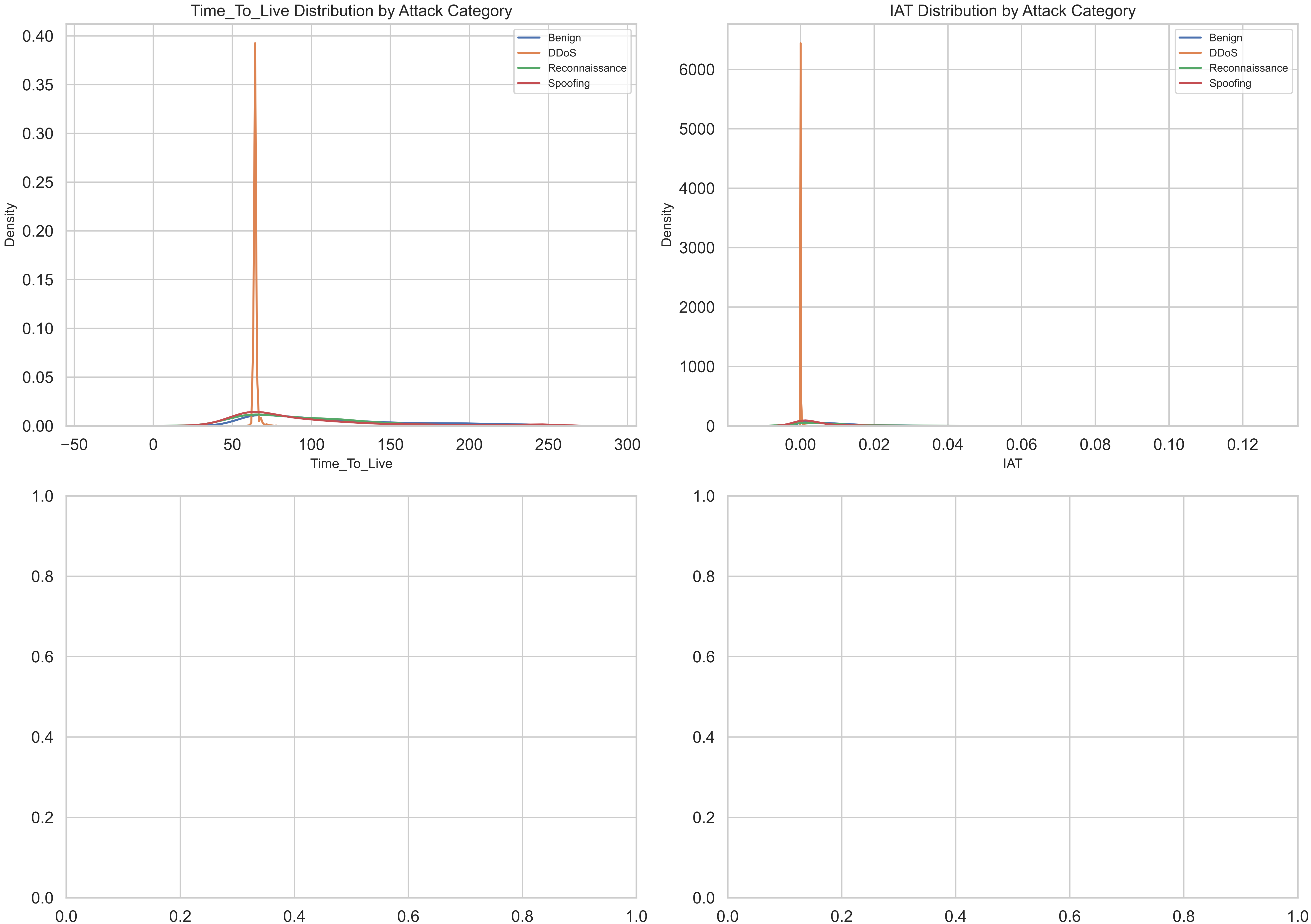
considered anomalous. The distinct separation between normal traffic patterns (blue) and outliers (red) demonstrates how statistical analysis can effectively detect potentially malicious traffic. For SMEs, this approach provides a computationally efficient method for identifying suspicious network activity without requiring extensive historical datasets or complex modeling techniques. By establishing baseline distributions for normal traffic, organizations can implement simple statistical filters that flag significant deviations for further investigation. These statistical outlier detection methods are particularly effective for identifying volumetric attacks such as DDoS, which typically generate traffic patterns that fall well outside normal operating parameters.



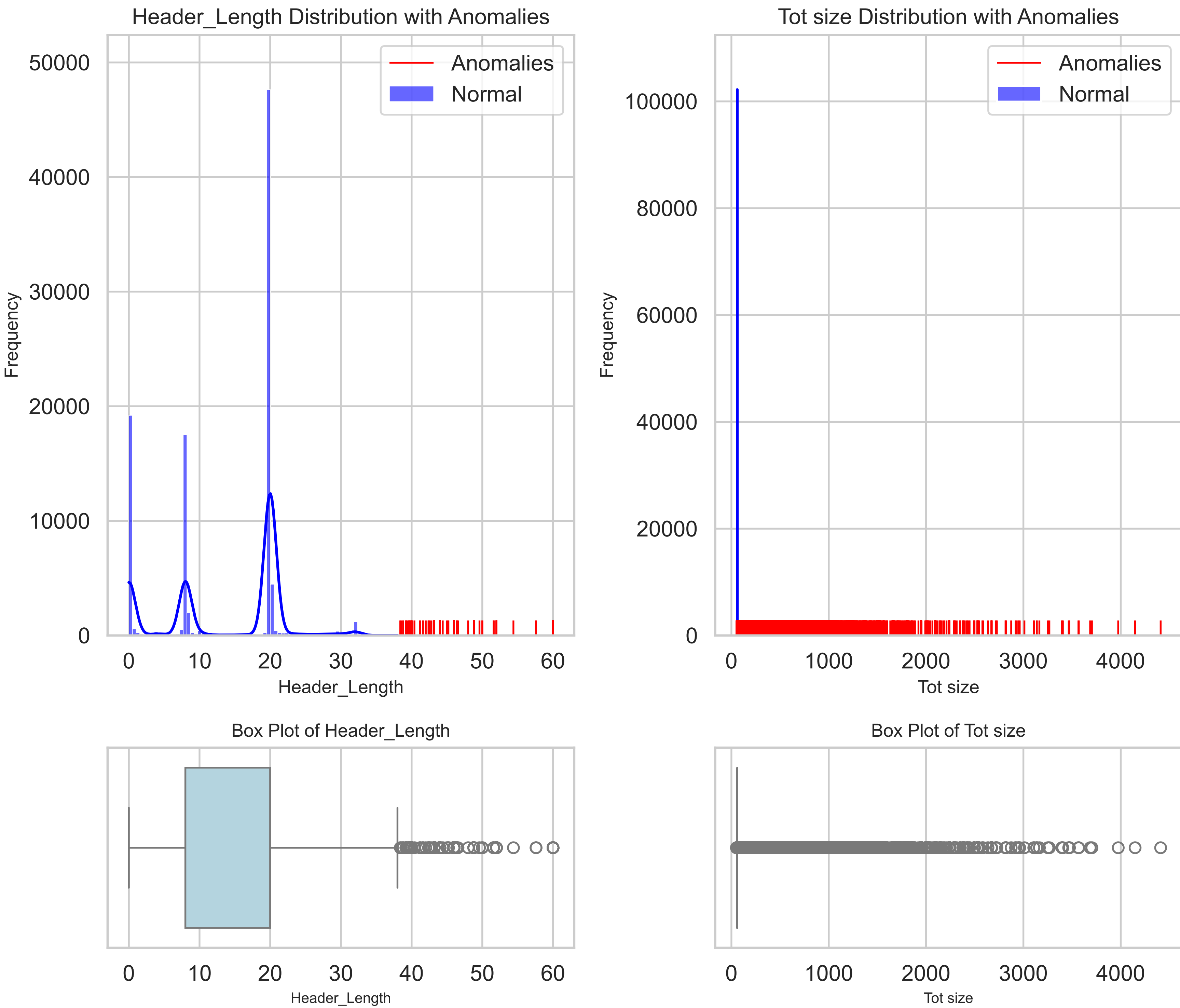
Protocol Usage Distribution by Attack Category (%)
Protocol Type
HOPOPT (0)
ICMP (1)
TCP (6)
UDP (17)
GRE (47)
Attack Category
Benign
Brute Force
DDoS
DoS
Injection
Malware
Mirai Botnet
Other
Reconnaissance
Spoofing
Vulnerability Scan
26.3 0.0 5.3 0.0 0.0 0.0 2.6 0.0 28.9 15.8 21.1
0.0 0.0 99.9 0.1 0.0 0.0 0.0 0.0 0.0 0.0 0.0
3.6 0.0 75.3 17.6 0.0 0.0 0.0 0.0 1.1 1.2 1.1
0.9 0.0 55.1 32.9 0.0 0.0 8.6 0.0 0.3 1.4 0.7
0.0 0.0 0.0 0.0 0.0 0.0 100.0 0.0 0.0 0.0 0.0
0
20
40
60
80
100

This heatmap reveals atypical protocol usage patterns across different attack categories, with color intensity
and annotations representing the percentage of traffic within each protocol type. By analyzing the protocol distribution matrix, security analysts can identify unusual protocol utilization that may indicate malicious activity. For example, a sudden increase in ICMP traffic (protocol 1) could signal reconnaissance or DDoS attacks,
while unusual GRE tunneling traffic might indicate data exfiltration attempts. The visualization shows clear protocol preferences for different attack types - DDoS attacks heavily leverage ICMP and UDP, while reconnaissance
activities predominantly use TCP. For SMEs, monitoring protocol distribution changes can provide early warning
of potential attacks with minimal computational overhead. Security systems can establish baseline protocol

distributions for normal operations, then trigger alerts when significant deviations occur.


Time_To_Live Distribution by Attack Category
IAT Distribution by Attack Category
Benign
DDoS
Reconnaissance
Spoofing
Density
Time_To_Live
IAT

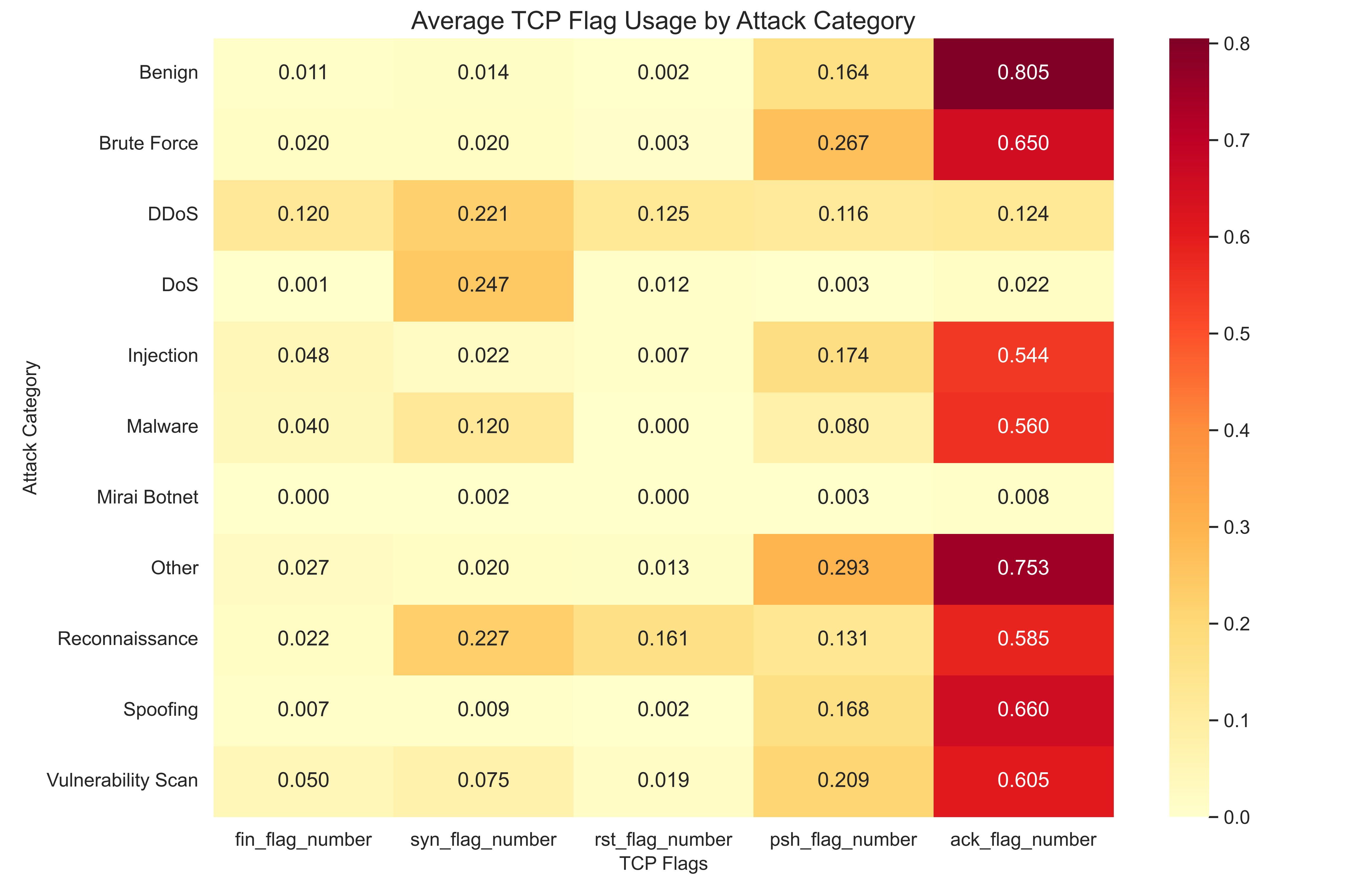
These density plots reveal how timing patterns differ across various attack categories, demonstrating a powerful anomaly detection criterion. The distinct shapes of these distributions highlight characteristic temporal signatures - DDoS attacks typically show sharp peaks with very short inter-arrival times, while reconnaissance activities display more dispersed patterns as they methodically probe networks. Benign traffic generally exhibits more natural, wider distributions. For SMEs implementing security monitoring, these timing-based anomaly indicators are particularly valuable because they can detect sophisticated attacks that might evade signature-based detection. By establishing baseline timing distributions for normal network traffic, even organizations with limited security resources can implement effective monitoring for temporal anomalies. Timing-based detection is also effective against zero-day attacks, as the abnormal timing patterns often remain consistent even when packet contents are novel.



Header_Length Distribution with Anomalies
Anomalies
Normal
Frequency
50000
40000
30000
20000
10000
0
0
10
20
30
40
50
60
Header_Length
Tot size Distribution with Anomalies
Anomalies
Normal
Frequency
100000
80000
60000
40000
20000
0
0
1000
2000
3000
4000
Tot size
Box Plot of Header_Length
0
10
20
30
40
50
60
Header_Length
Box Plot of Tot size
0
1000
2000
3000
4000
Tot size

These visualizations demonstrate how irregular packet size distributions can serve as effective anomaly indicators in IoT network traffic. The histograms (blue) display the distribution of normal packet sizes, while the red rug plots at the bottom mark anomalous values that fall outside the expected range (calculated using the interquartile range method). The box plots below each histogram provide additional context for identifying outliers. Abnormal packet sizes often indicate malicious activity - unusually small packets might signify scanning or probing attacks, while exceptionally large packets could indicate buffer overflow attempts or covert channel communications.

For SMEs, monitoring packet size distributions requires minimal computational resources while providing valuable security insights. By establishing baseline packet size profiles for different device types and communication patterns, even basic security monitoring can detect potentially malicious traffic through simple statistical methods.


Average TCP Flag Usage by Attack Category
Attack Category
Benign
Brute Force
DDoS
DoS
Injection
Malware
Mirai Botnet
Other
Reconnaissance
Spoofing
Vulnerability Scan
fin_flag_number
syn_flag_number
rst_flag_number
psh_flag_number
ack_flag_number
TCP Flags
0.011 0.014 0.002 0.164 0.805
0.020 0.020 0.003 0.267 0.650
0.120 0.221 0.125 0.116 0.124
0.001 0.247 0.012 0.003 0.022
0.048 0.022 0.007 0.174 0.544
0.040 0.120 0.000 0.080 0.560
0.000 0.002 0.000 0.003 0.008
0.027 0.020 0.013 0.293 0.753
0.022 0.227 0.161 0.131 0.585
0.007 0.009 0.002 0.168 0.660
0.050 0.075 0.019 0.209 0.605
0.0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8

This heatmap visualizes how different attack categories utilize various TCP flag combinations, revealing distinctive signatures that can serve as anomaly indicators. The annotations show the average frequency of
each flag type across different attack categories, with color intensity proportional to usage frequency. Unusual flag combinations often indicate malicious activity - for example, a high frequency of SYN flags without corresponding ACK flags typically indicates SYN flood attacks, while unusual combinations of RST and FIN flags might signal port scanning or connection hijacking attempts. For SMEs, monitoring TCP flag patterns provides an efficient method for detecting reconnaissance and exploitation attempts. By establishing
baselines for normal flag usage in legitimate traffic, organizations can implement rule-based detection for suspicious patterns. Unlike deep packet inspection, flag analysis requires minimal processing resources and
can be performed at line speed even in bandwidth-constrained environments.

Anomaly Identification Criteria Framework for IoT Security in SMEs

Criterion	Description	Key Indicators	SME Relevance	Implementa… Approac…
… in Traffic Patterns	Traffic metrics that deviate significantly from normal behavior	Z-score > 3 for traffic rates Isolation Forest anomaly score < -0.5 Traffic v[illegible] baseline	High	Simple statistical filters on ne…
…tocol Usage	Protocols being used in atypical ways or unexpected protocols	Protocols not in whitelist Abnormal protocol distribution Proto[illegible] hours	Medium-High	Protocol whitelisting and fr…
…ming Patterns	Inter-arrival times and request patterns that indicate automated activity	Regular, clockwork timing Extremely short IAT values Reque[illegible]bution	High	Time-series analysis o…
…Size Distributions	Packet sizes that fall outside the normal range for the protocol/service	Sizes outside IQR boundaries Uniform packet sizes (automation) Maximum sized packets (buffer overflow attempts)	Medium	Packet size profiling by…
…g Combinations	TCP flag combinations that violate protocol norms or indicate scanning	SYN flood patterns FIN without corresponding SYN/ACK ACK storms RST abuse patterns	Medium-High	Flag pattern rules in ne…

This comprehensive framework outlines five key anomaly identification criteria for IoT security monitoring in SME environments. Each criterion is evaluated based on its description, specific indicators, relevance to SMEs, and practical implementation approaches. The color coding indicates relevance level, with darker shades representing higher priority criteria for resource-constrained organizations. Statistical outliers and suspicious timing patterns are highlighted as particularly relevant for SMEs due to their high detection efficacy and relatively straightforward implementation requirements. This framework provides SMEs with a prioritized approach to anomaly detection, allowing them to focus limited security resources on the most effective criteria. By implementing these criteria in stages, starting with the highest-relevance approaches, organizations can establish a robust anomaly detection capability that balances security effectiveness with operational constraints.