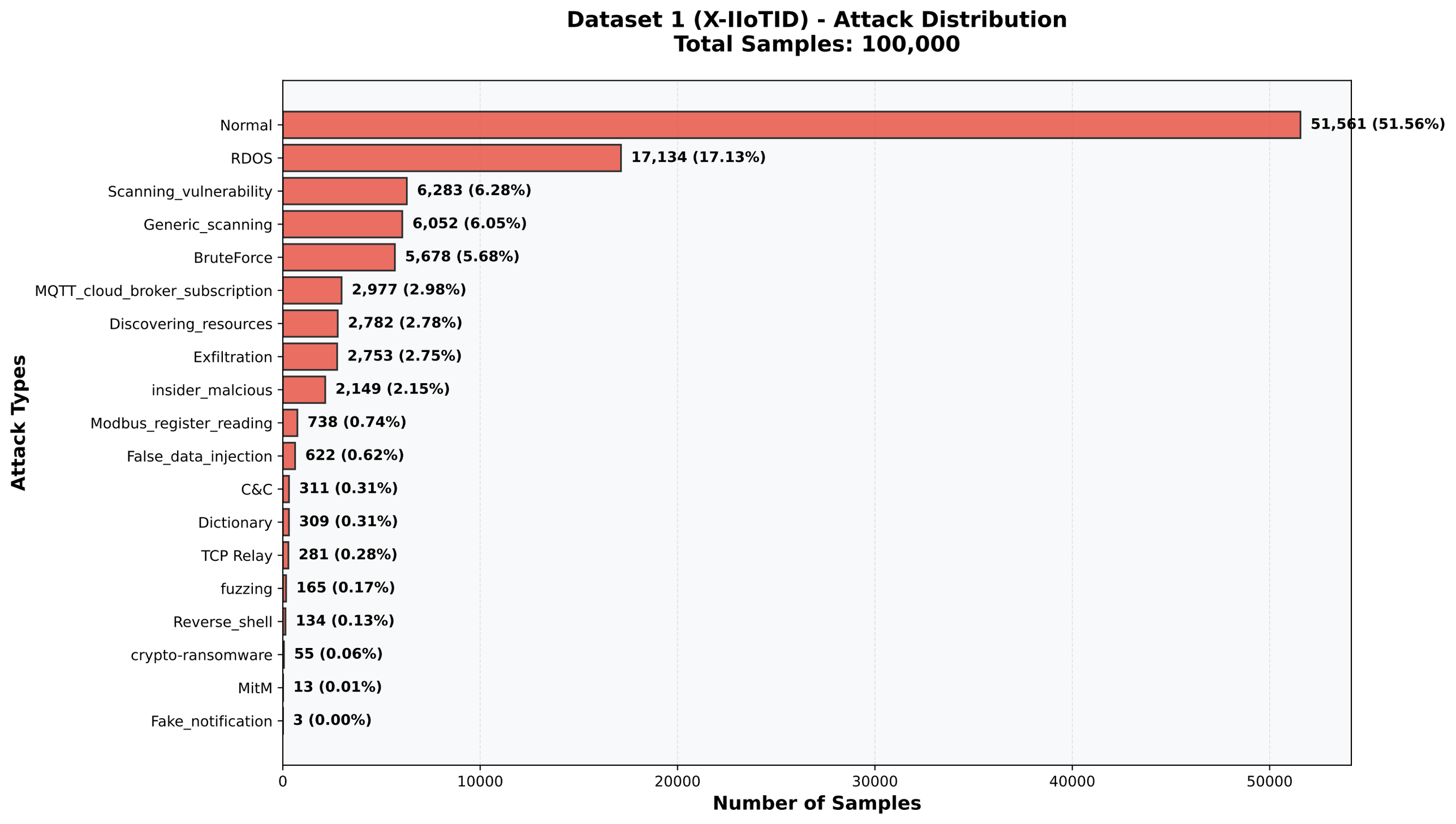
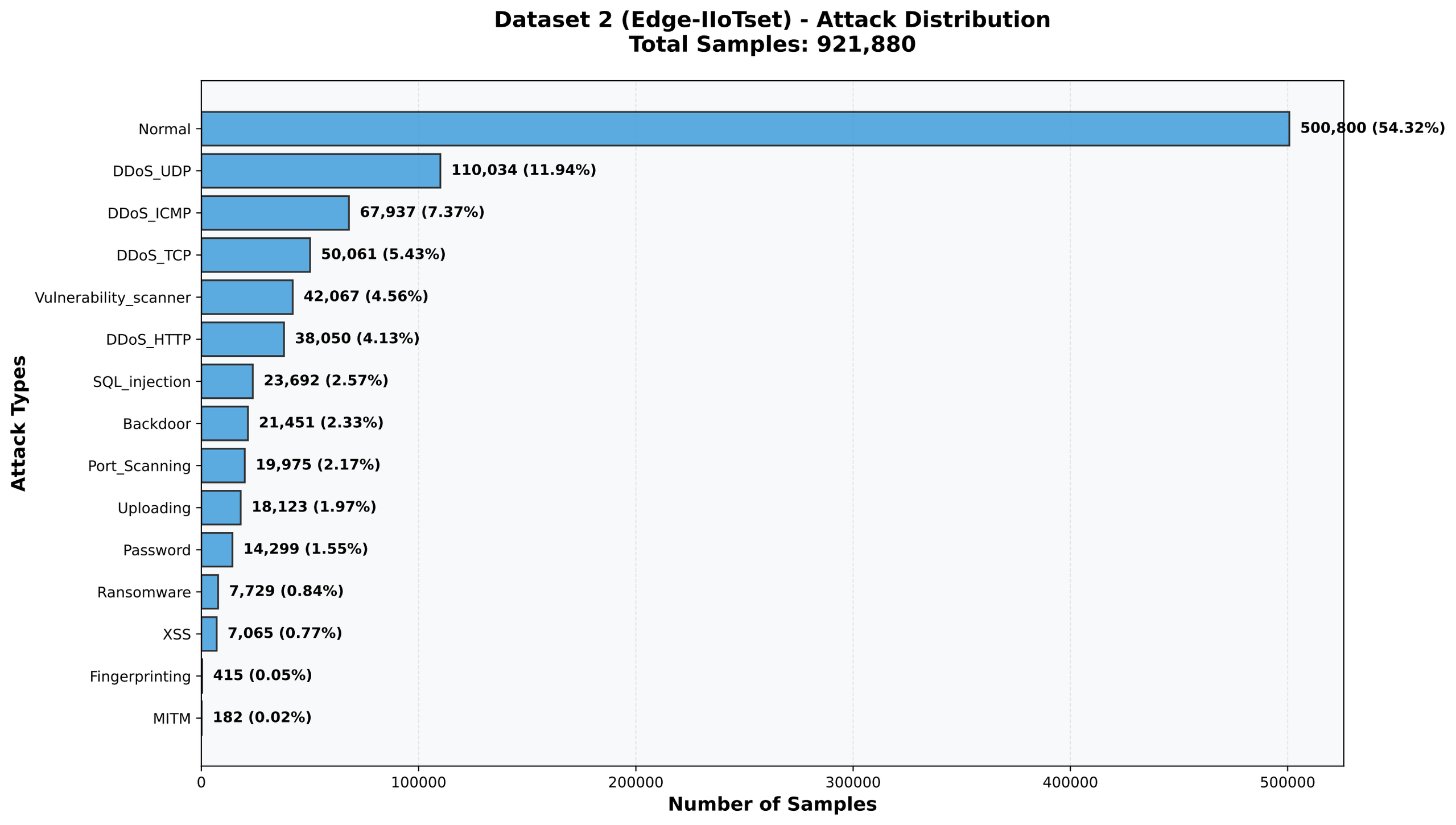
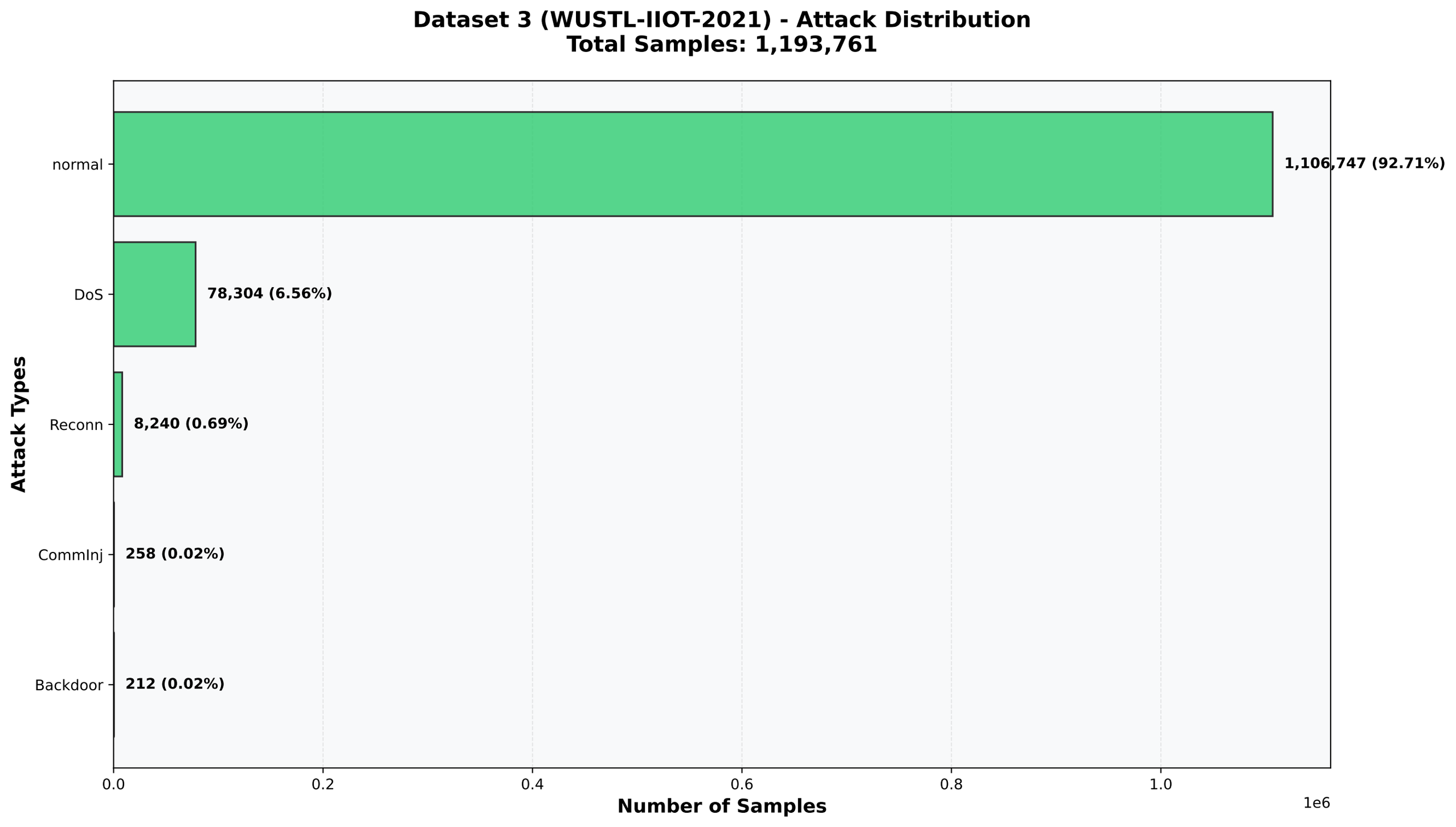
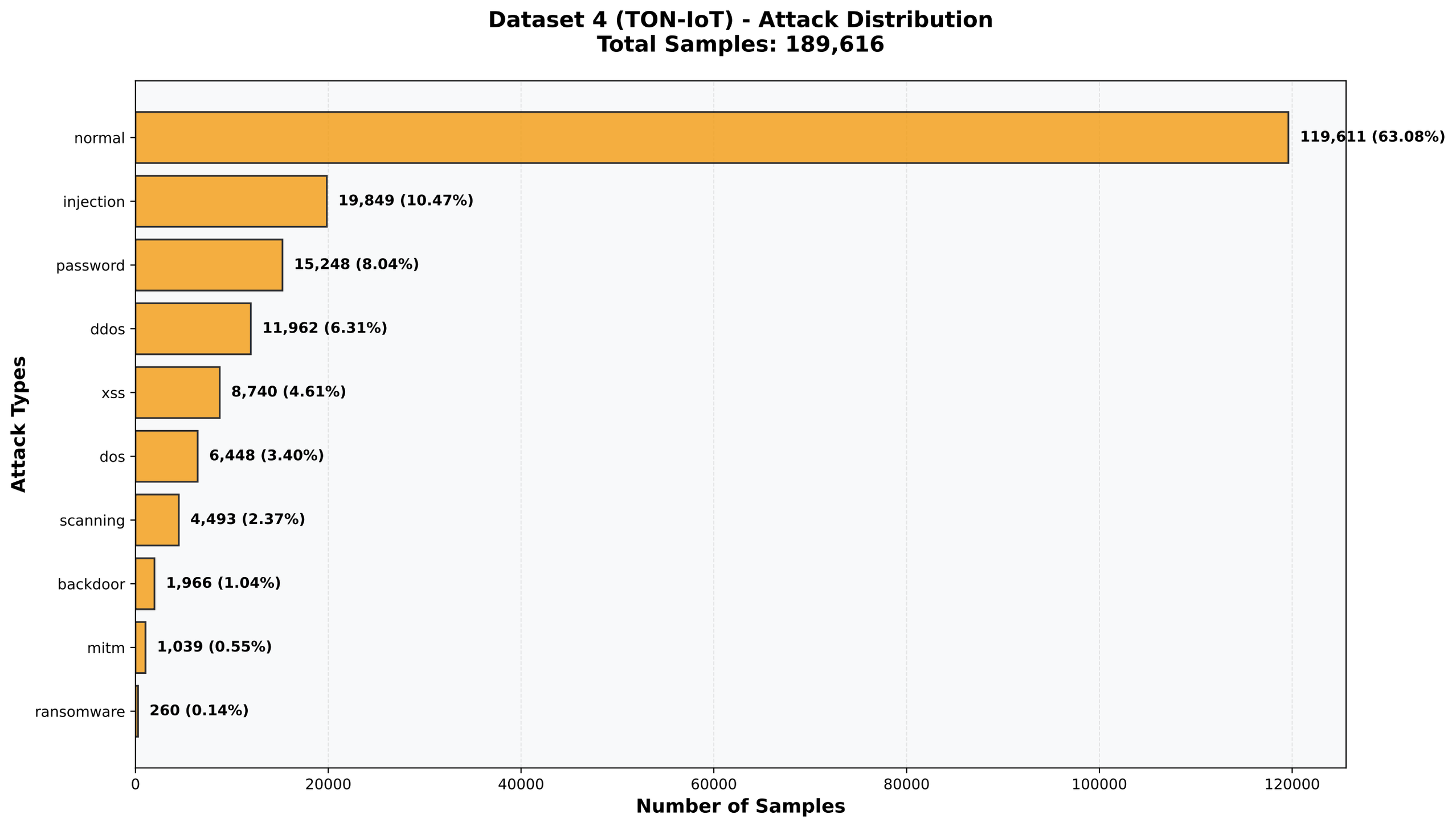

Dataset 1 (X-IIoTID) - Attack Distribution
Total Samples: 100,000
Attack Types
Normal
RDOS
Scanning_vulnerability
Generic_scanning
BruteForce
MQTT_cloud_broker_subscription
Discovering_resources
Exfiltration
insider_malcious
Modbus_register_reading
False_data_injection
C&C
Dictionary
TCP Relay
fuzzing
Reverse_shell
crypto-ransomware
MitM
Fake_notification
51,561 (51.56%)
17,134 (17.13%)
6,283 (6.28%)
6,052 (6.05%)
5,678 (5.68%)
2,977 (2.98%)
2,782 (2.78%)
2,753 (2.75%)
2,149 (2.15%)
738 (0.74%)
622 (0.62%)
311 (0.31%)
309 (0.31%)
281 (0.28%)
165 (0.17%)
134 (0.13%)
55 (0.06%)
13 (0.01%)
3 (0.00%)
0
10000
20000
30000
40000
50000
Number of Samples


Dataset 2 (Edge-IIoTset) - Attack Distribution
Total Samples: 921,880
Attack Types
Normal
DDoS_UDP
DDoS_ICMP
DDoS_TCP
Vulnerability_scanner
DDoS_HTTP
SQL_injection
Backdoor
Port_Scanning
Uploading
Password
Ransomware
XSS
Fingerprinting
MITM
500,800 (54.32%)
110,034 (11.94%)
67,937 (7.37%)
50,061 (5.43%)
42,067 (4.56%)
38,050 (4.13%)
23,692 (2.57%)
21,451 (2.33%)
19,975 (2.17%)
18,123 (1.97%)
14,299 (1.55%)
7,729 (0.84%)
7,065 (0.77%)
415 (0.05%)
182 (0.02%)
0
100000
200000
300000
400000
500000
Number of Samples


Dataset 3 (WUSTL-IIOT-2021) - Attack Distribution
Total Samples: 1,193,761
Attack Types
normal
DoS
Reconn
CommInj
Backdoor
1,106,747 (92.71%)
78,304 (6.56%)
8,240 (0.69%)
258 (0.02%)
212 (0.02%)
0.0
0.2
0.4
0.6
0.8
1.0
1e6
Number of Samples


Dataset 4 (TON-IoT) - Attack Distribution
Total Samples: 189,616
Attack Types
normal
injection
password
ddos
xss
dos
scanning
backdoor
mitm
ransomware
119,611 (63.08%)
19,849 (10.47%)
15,248 (8.04%)
11,962 (6.31%)
8,740 (4.61%)
6,448 (3.40%)
4,493 (2.37%)
1,966 (1.04%)
1,039 (0.55%)
260 (0.14%)
0
20000
40000
60000
80000
100000
120000
Number of Samples

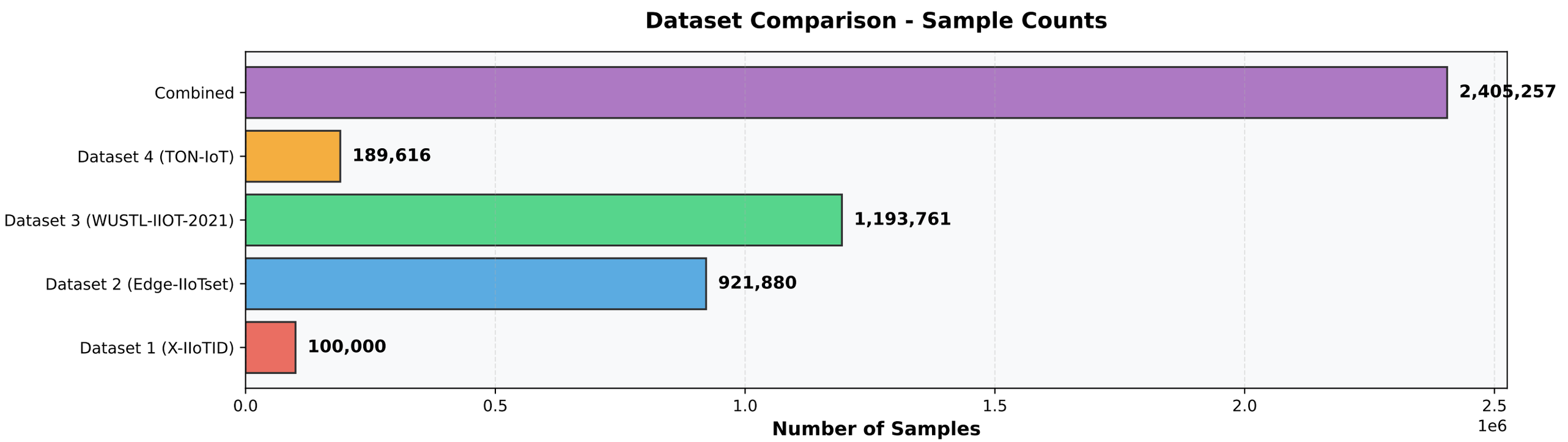
Combined Dataset - Attack Distribution by Source
Total Samples: 2,405,257

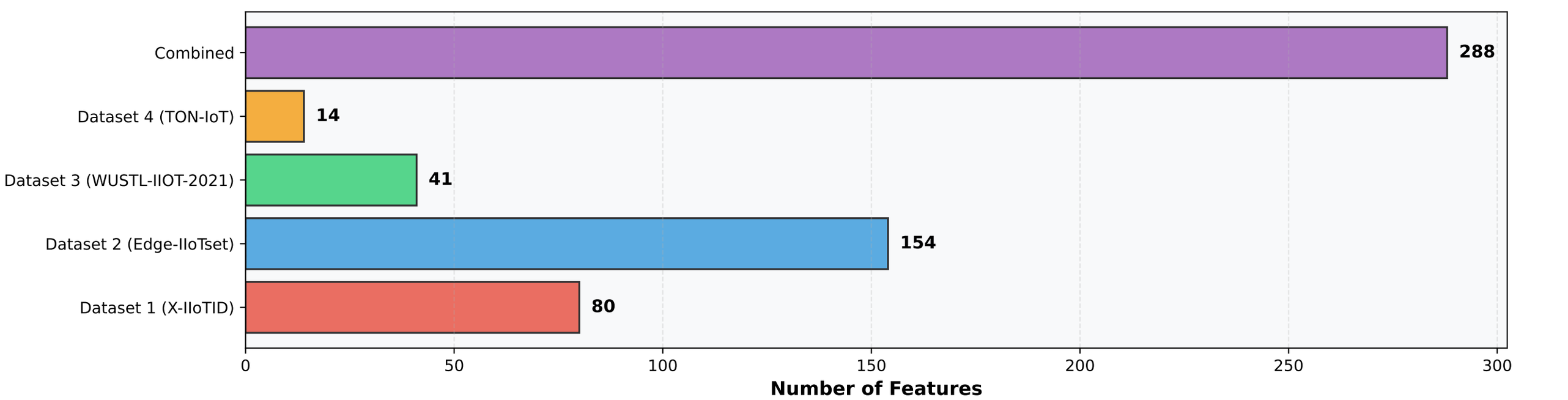
Attack Types

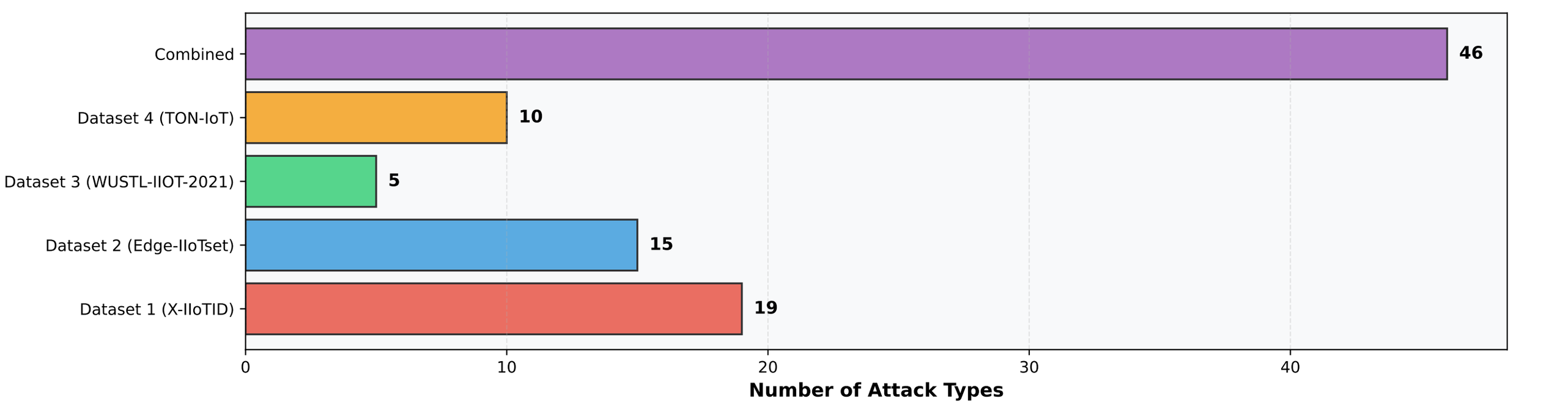
normal 1,106,747 119,611 1,226,358
Normal 51,561 500,800 552,361
DDoS_UDP 110,034 110,034
DoS 78,304 78,304
DDoS_ICMP 67,937 67,937
DDoS_TCP 50,061 50,061
Vulnerability_scanner 42,067 42,067
DDoS_HTTP 38,050 38,050
SQL_injection 23,692
Backdoor 21,663
Port_Scanning 19,975
injection 19,849
Uploading 18,123
RDOS 17,134
password 15,248
Password 14,299
ddos 11,962
xss 8,740
Recon 8,240
Ransomware 7,729
XSS 7,065
dos 6,448
Scanning_vulnerability 6,283
Generic_scanning 6,052
BruteForce 5,678
scanning 4,493
MQTT_cloud_broker_subscription 2,977
Discovering_resources 2,782
Exfiltration 2,753
insider_malicious 2,149
backdoor 1,966
mitm 1,039
Modbus_register_reading 738
False_data_injection 622
Fingerprinting 415
C&C 311
Dictionary 309
TCP Relay 281
ransomware 260
CommInj 258
MITM 182
fuzzing 165
Reverse_shell 134
crypto-ransomware 55
MitM 13
Fake_notification 3

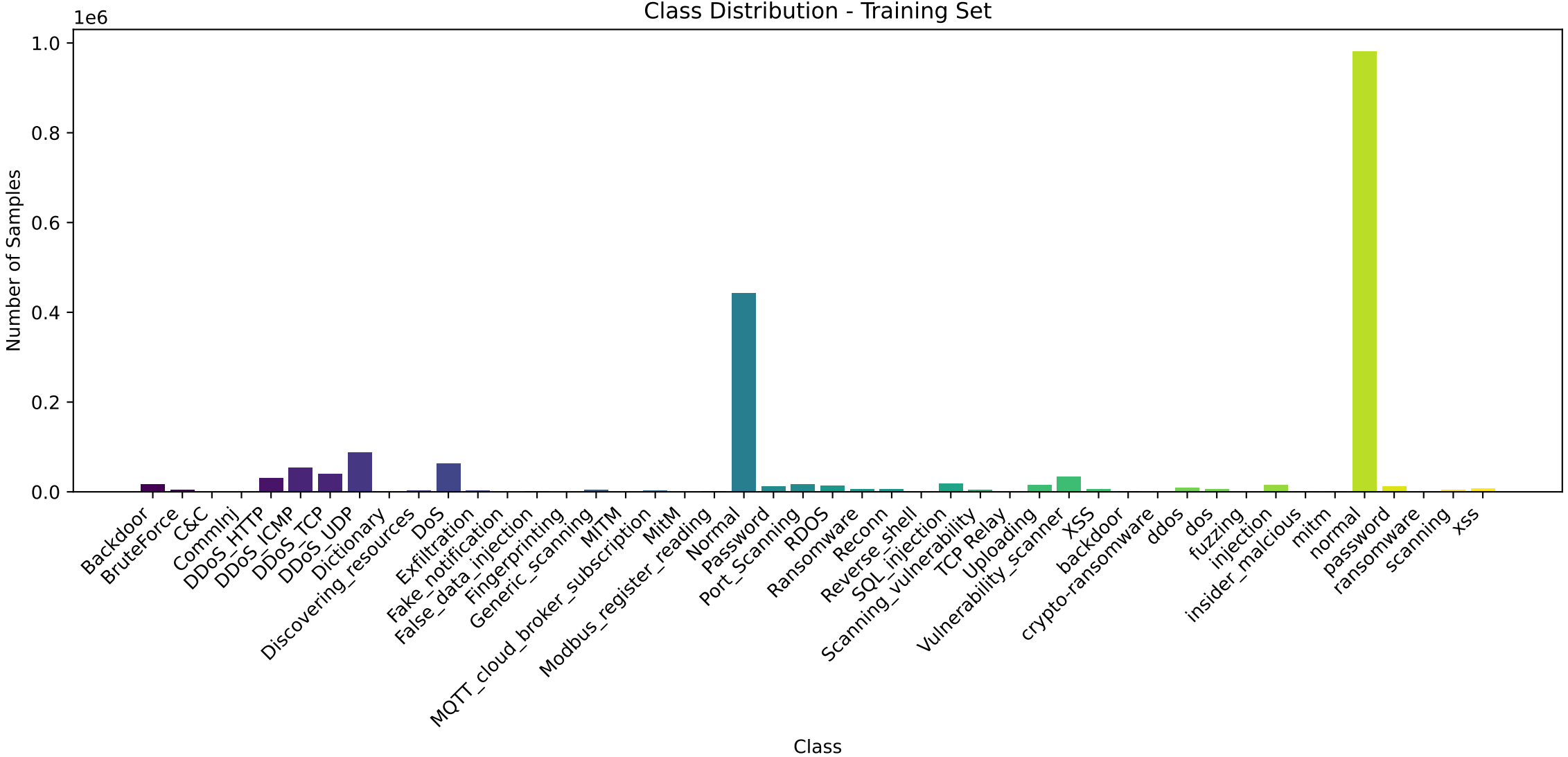
Dataset 1 (X-IIoTID)
Dataset 2 (Edge-IIoTset)
Dataset 3 (WUSTL-IIOT-2021)
Dataset 4 (TON-IoT)

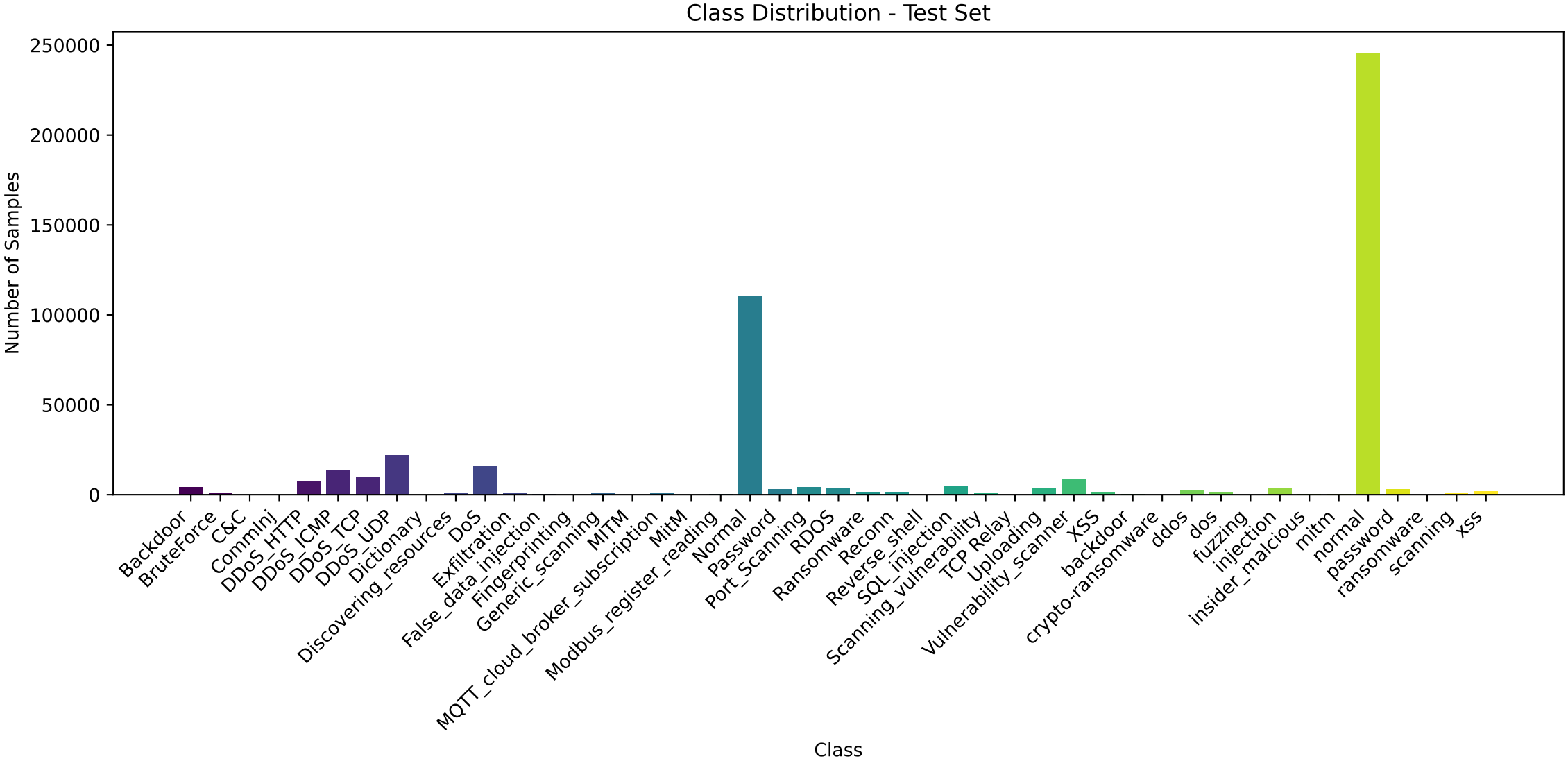
0.2 0.4 0.6 0.8 1.0 1.2
1e6
Number of Samples


Dataset Comparison - Sample Counts
Combined
Dataset 4 (TON-IoT)
Dataset 3 (WUSTL-IIOT-2021)
Dataset 2 (Edge-IIoTset)
Dataset 1 (X-IIoTID)
2,405,257
189,616
1,193,761
921,880
100,000
0.0
0.5
1.0
1.5
2.0
2.5
1e6
Number of Samples


Dataset Comparison - Feature Counts
Combined
Dataset 4 (TON-IoT)
Dataset 3 (WUSTL-IIOT-2021)
Dataset 2 (Edge-IIoTset)
Dataset 1 (X-IIoTID)
288
14
41
154
80
0
50
100
150
200
250
300
Number of Features

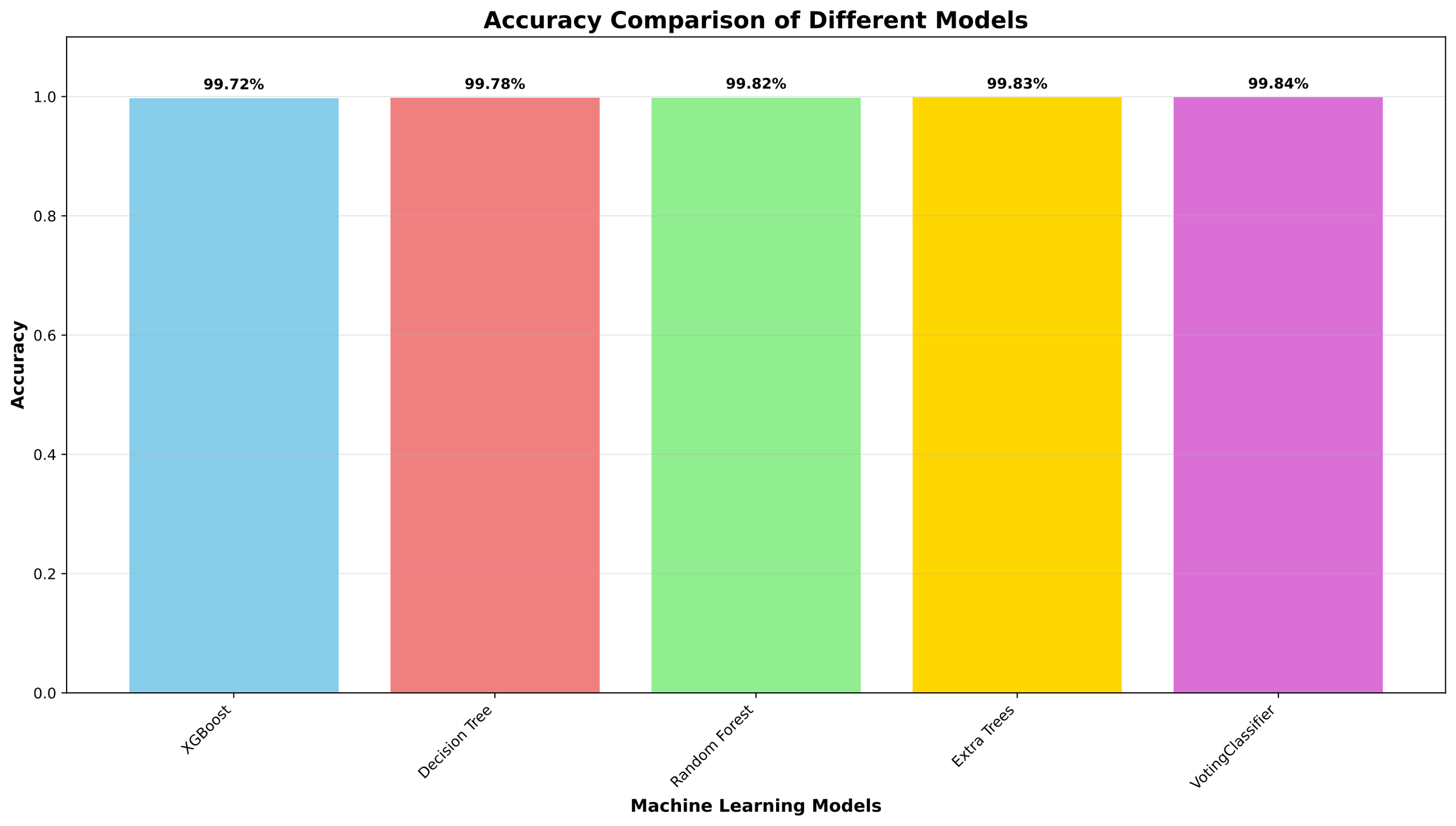

Dataset Comparison - Attack Type Diversity
Combined
Dataset 4 (TON-IoT)
Dataset 3 (WUSTL-IIOT-2021)
Dataset 2 (Edge-IIoTset)
Dataset 1 (X-IIoTID)
46
10
5
15
19
0
10
20
30
40
Number of Attack Types

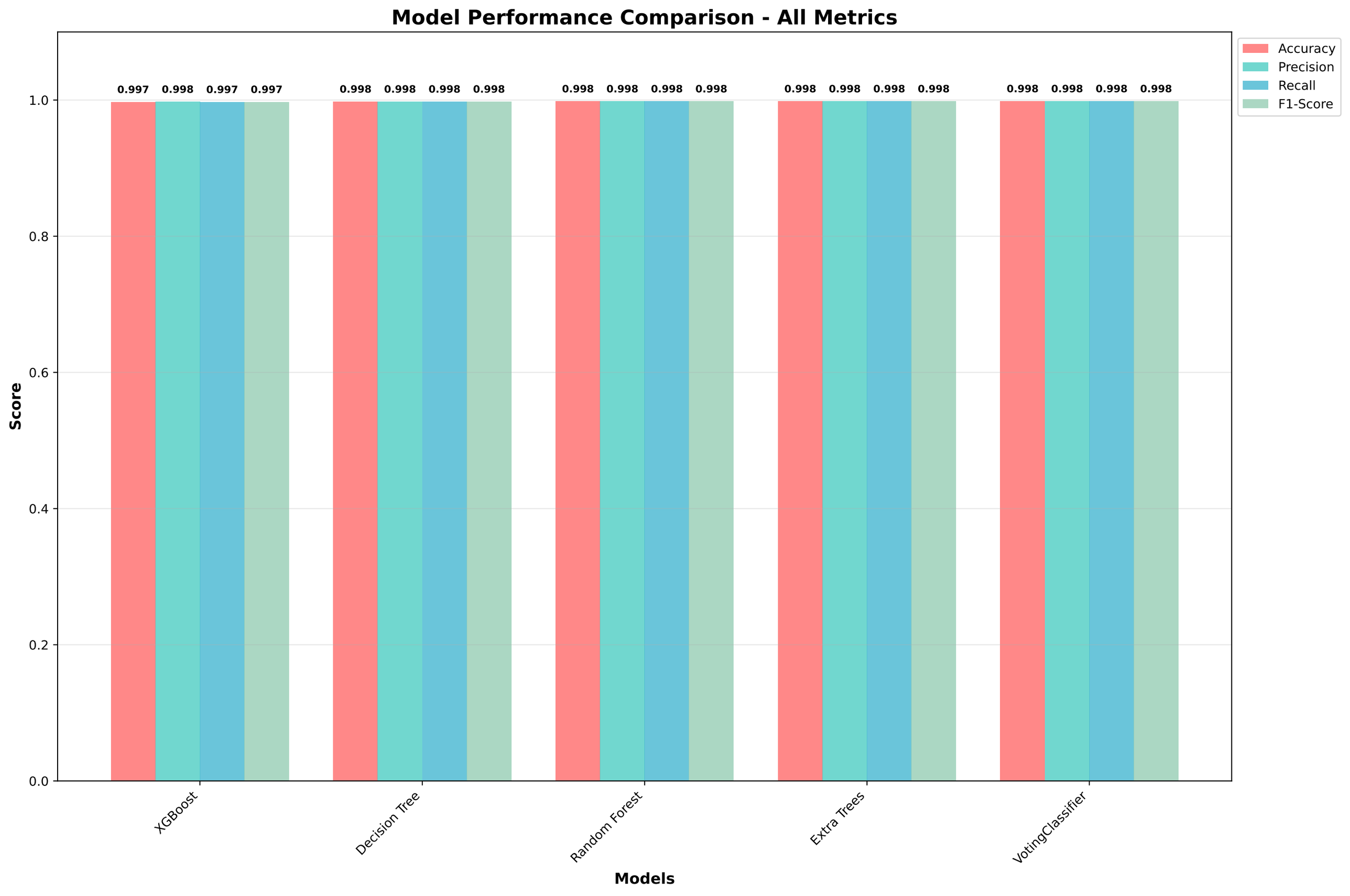

Class Distribution - Training Set
1e6
1.0
0.8
0.6
0.4
0.2
0.0
Number of Samples
Backdoor
BruteForce
C&C
CommInj
DDoS_HTTP
DDoS_ICMP
DDoS_TCP
DDoS_UDP
Dictionary
Discovering_resources
DoS
Exfiltration
Fake_notification
False_data_injection
Fingerprinting
Generic_scanning
MITM
MQTT_cloud_broker_subscription
MitM
Modbus_register_reading
Normal
Password
Port_Scanning
RDOS
Ransomware
Reconn
Reverse_shell
SQL_injection
Scanning_vulnerability
TCP Relay
Uploading
Vulnerability_scanner
XSS
backdoor
crypto-ransomware
ddos
dos
fuzzing
injection
insider_malcious
mitm
normal
password
ransomware
scanning
xss
Class


Class Distribution - Test Set
Number of Samples
250000
200000
150000
100000
50000
0
Backdoor
BruteForce
C&C
CommInj
DDoS_HTTP
DDoS_ICMP
DDoS_TCP
DDoS_UDP
Dictionary
Discovering_resources
DoS
Exfiltration
False_data_injection
Fingerprinting
Generic_scanning
MITM
MQTT_cloud_broker_subscription
MitM
Modbus_register_reading
Normal
Password
Port_Scanning
RDOS
Ransomware
Reconn
Reverse_shell
SQL_injection
Scanning_vulnerability
TCP Relay
Uploading
Vulnerability_scanner
XSS
backdoor
crypto-ransomware
ddos
dos
fuzzing
injection
insider_malcious
mitm
normal
password
ransomware
scanning
xss
Class

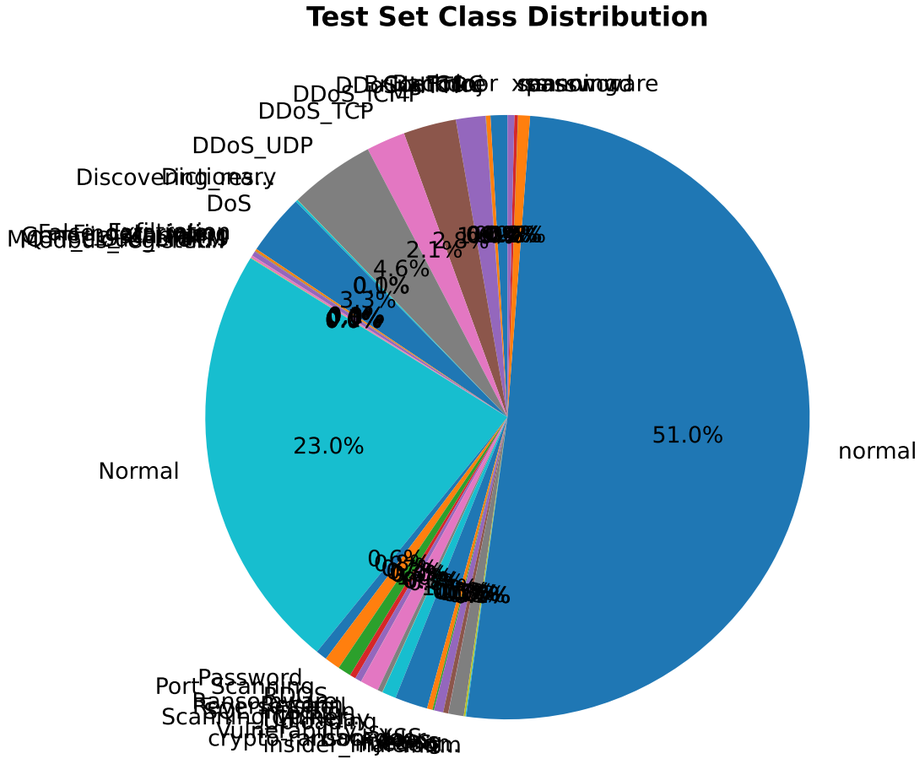
Model Performance Summary

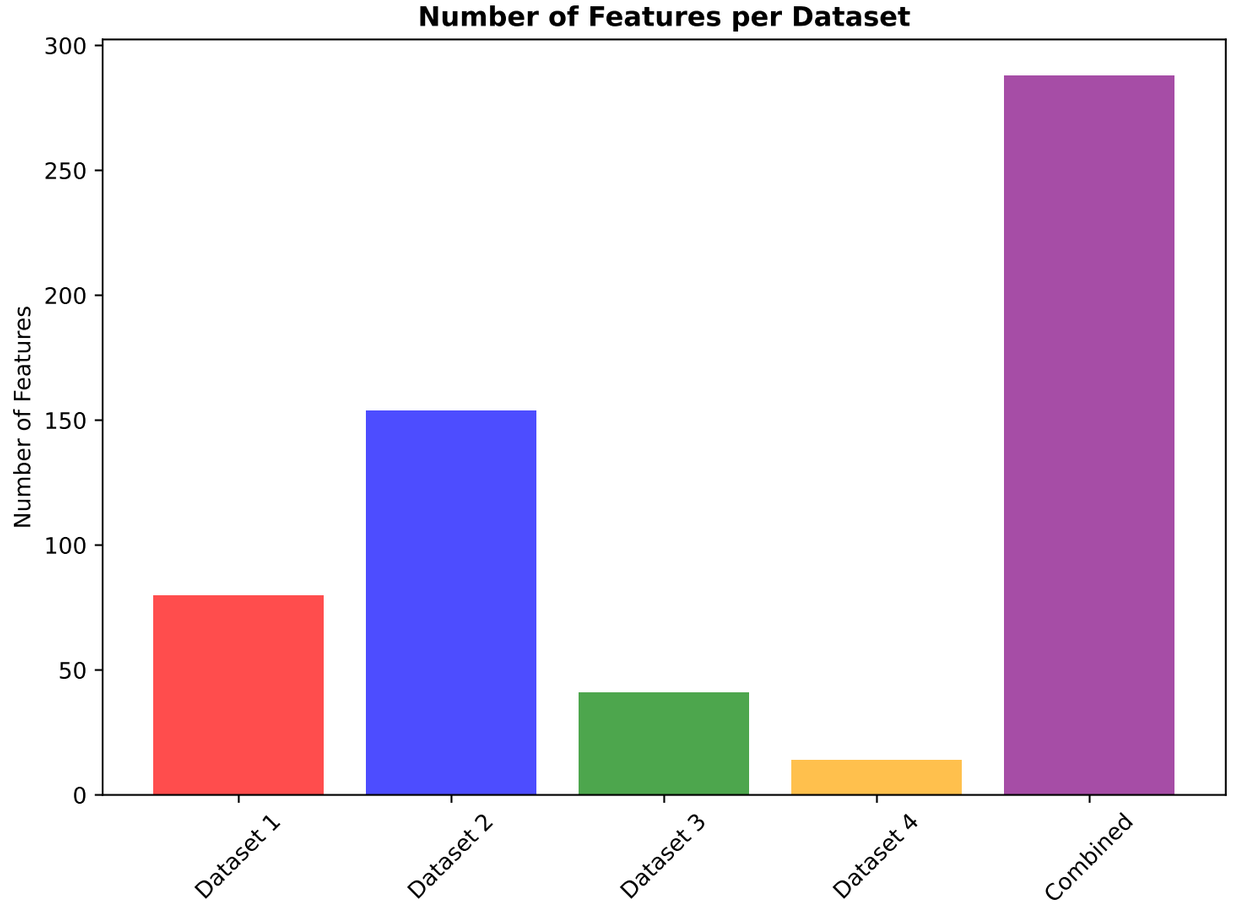
Model	Accuracy	Precision	Recall	F1-Score
XGBoost	0.9972	0.9975	0.9972	0.9973
Decision Tree	0.9978	0.9978	0.9978	0.9978
Random Forest	0.9982	0.9982	0.9982	0.9981
Extra Trees	0.9983	0.9982	0.9983	0.9982
VotingClassifier	0.9984	0.9984	0.9984	0.9984

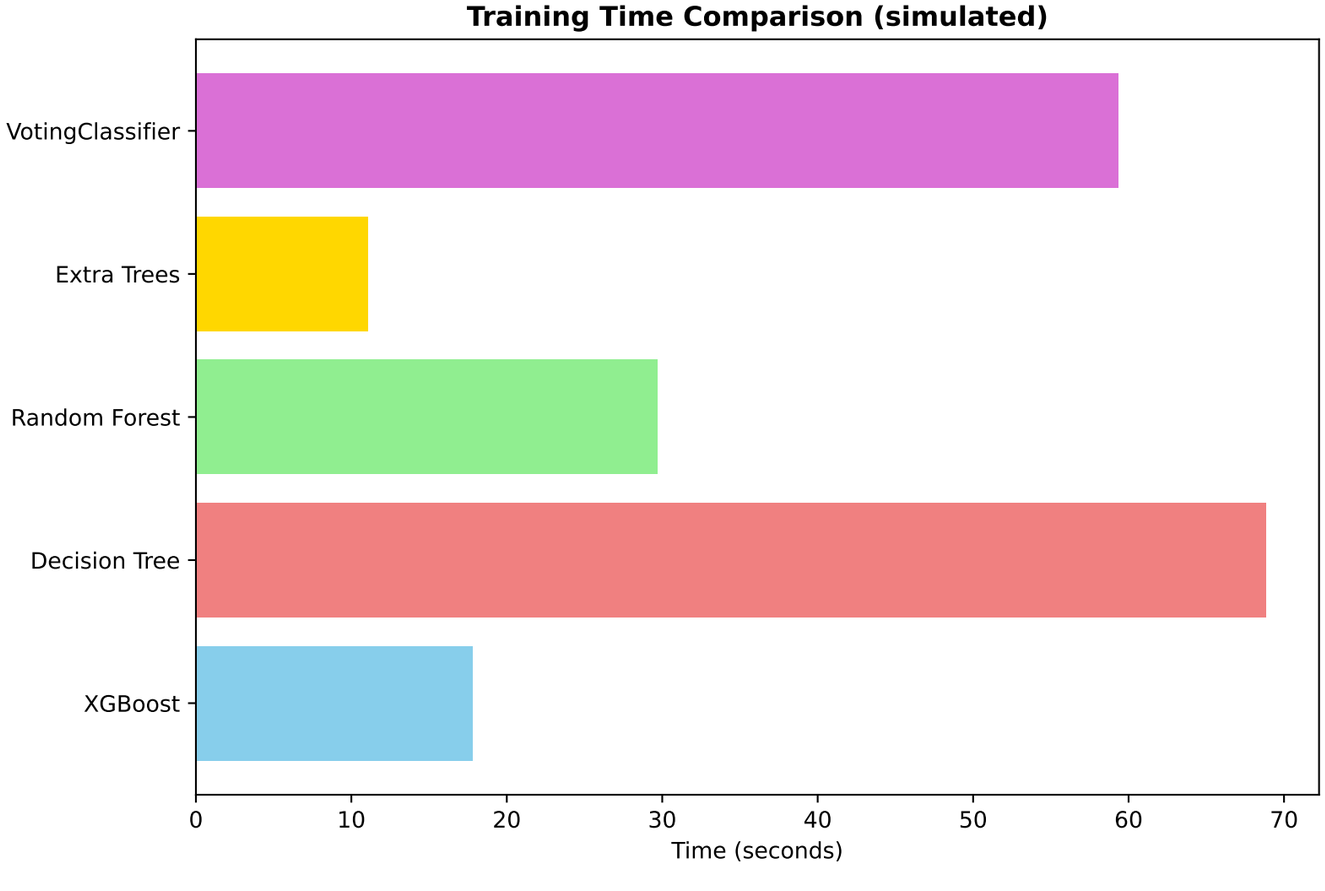

Accuracy Comparison of Different Models
99.72%
99.78%
99.82%
99.83%
99.84%
1.0
0.8
0.6
0.4
0.2
0.0
Accuracy
XGBoost
Decision Tree
Random Forest
Extra Trees
VotingClassifier
Machine Learning Models

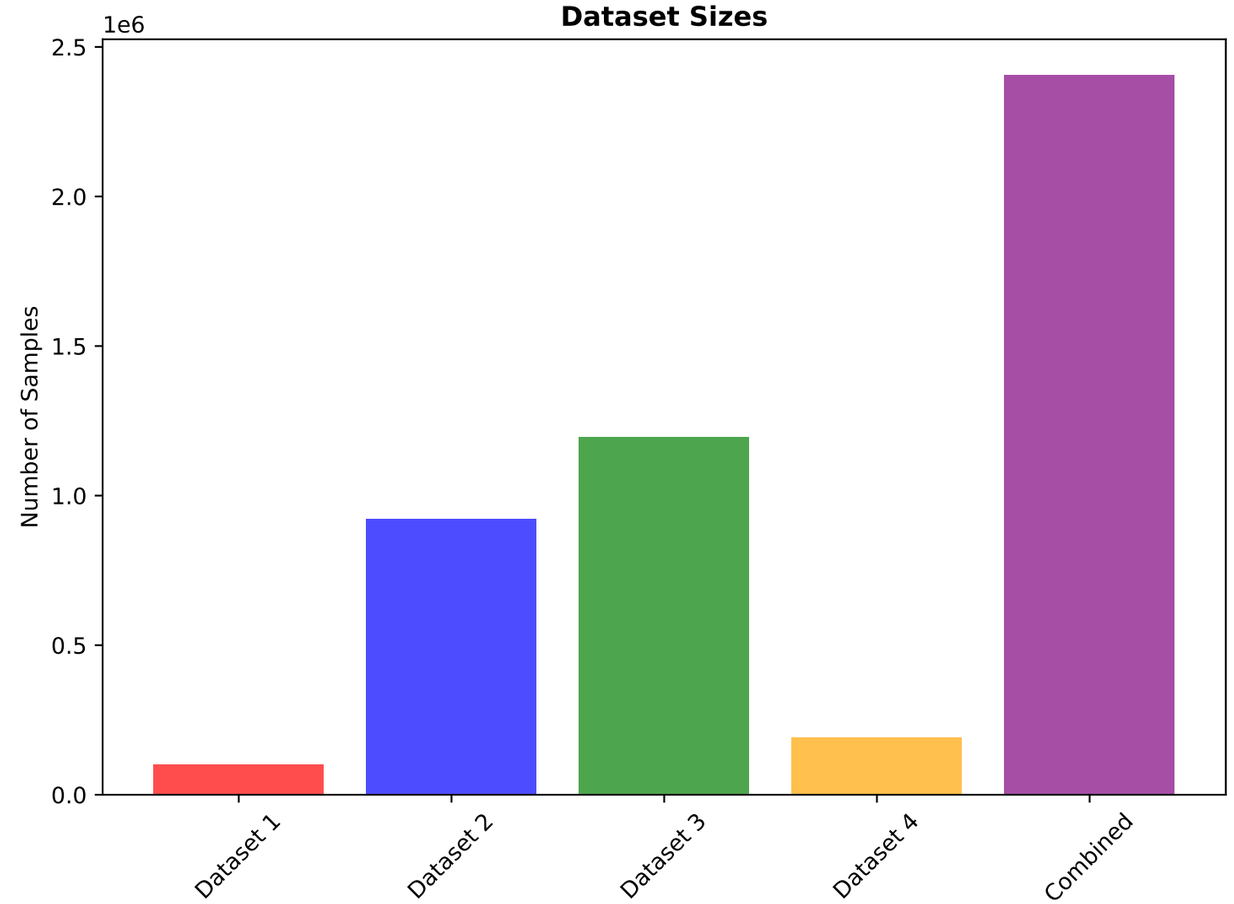

Model Performance Comparison - All Metrics
Accuracy
Precision
Recall
F1-Score
0.997 0.998 0.997 0.997
0.998 0.998 0.998 0.998
0.998 0.998 0.998 0.998
0.998 0.998 0.998 0.998
0.998 0.998 0.998 0.998
Score
1.0
0.8
0.6
0.4
0.2
0.0
XGBoost
Decision Tree
Random Forest
Extra Trees
VotingClassifier
Models

Detailed Analysis Statistics









Analysis Metadata and Summary

```
MACHINE LEARNING ANALYSIS REPORT
================================

Analysis Date: 2025-10-12 20:02:52

DATASETS USED:
• Dataset 1 (X-IIoTID): 100,000 samples, 80 features
• Dataset 2 (Edge-IIoTset): 921,880 samples, 154 features
• Dataset 3 (WUSTL-IIOT-2021): 1,193,761 samples, 41 features
• Dataset 4 (TON-IoT): 189,616 samples, 14 features

COMBINED DATASET:
• Total Samples: 2,405,257
• Total Features: 288
• Training Samples: 1,924,205
• Test Samples: 481,052
• Number of Classes: 45

MODELS EVALUATED:
• XGBoost Classifier
• Decision Tree Classifier
• Random Forest Classifier
• Extra Trees Classifier
• Voting Classifier (Ensemble)

BEST PERFORMING MODEL:
VotingClassifier
(Accuracy: 0.9984)

PREPROCESSING STEPS:
• Feature harmonization using UNION strategy
• Missing value imputation with median strategy
• Label encoding and standardization
• Outlier removal and feature engineering
• Train/test split with stratification
```