Beijing Forest Studio

北京理工大学信息系统及安全对抗实验中心

Wireless Traffic Dataset for Krack and Kr00k Attacks in WPA2

PhD Student : Munna Md Minhazul Islam 蒙纳

导师： Professor LUO SENLIN

2025年09月27日

Outline

1. Overview of WPA2 and its significance
2. Attack Simulation & Methodology
3. WPA2-KKID Dataset
4. Model Training & Performance
5. Strength & Limitations

TIPO - PCDL

T	目标	Public dataset of KRACK/Kr00k traffic; fill WPA2 dataset gap; support ML-based IDS research
I	输入	Raw pcap traffic from controlled testbed; 5.5 M+ frames; both normal and attack sessions; engineered features
P	处理	Simulate attacks; capture traffic; extract features; heuristically label (EAPOL replay for KRACK, zero-key frames for Kr00k); balance & clean data; train various ML models.
O	输出	Public WPA2-KKID dataset (pcap & CSV); 34 features; labeled normal/attack samples; open access for researchers; foundation for IDS development

P	问题	Before this work, publicly accessible datasets on KRACK and Kr00k vulnerabilities were virtually nonexistent; existing collections like AWID3 include multiple attacks but lack public availability and a fully described framework.
C	条件	Lab-based environment with single AP and limited clients; ~10-minute capture sessions; WPA2-Personal on 2.4 GHz; controlled but not diverse
D	难点	Detecting subtle attack signatures; accurate labeling; handling class imbalance; limited device diversity and real-world variability.
L	水平	Computers & Security, (JCR 1, CAS 2)

Overview of WPA2 and its significance

Encryption Strength

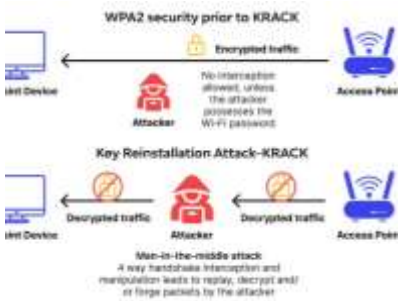
WPA2-Personal uses Advanced Encryption Standard (AES)

Prevalence& Popularity WPA2-Personal is the most common Wi-Fi security protocol worldwide

How It Works

WPA2-Personal relies on the 4-way handshake and Pre-Shared Key (PSK) is used for authentication

Public Wi-Fi hotspots, such as those in cafés and airports, commonly use WPA2-Personal due to its simplicity and ease of setup. However, this convenience comes with certain security risks, like susceptibility to man-in-the-middle attacks, especially if not combined with additional protections like a VPN.

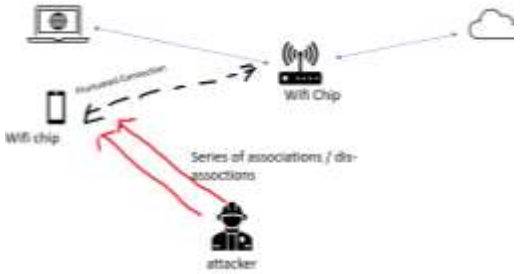


KRACK Attack: Real-World Exploitation

discovered in 2017

Kr00k Vulnerability

identified in 2019



To protect against WPA2 vulnerabilities, it's essential to use strong, complex passwords, and switch to WPA3 if possible, as it offers better security. Utilizing a VPN can help encrypt traffic on unsecured networks, while regularly updating router firmware ensures protection against known vulnerabilities.

Overview of WPA2 and its significance

	KRACK (Key Reinstallation Attack)	Kr00k (Key Zeroing Attack)
Attack level	Network-level attack	Device-level attack
How it works	KRACK exploits a flaw in the WPA2 4-way handshake by forcing a reinstallation of the encryption key, allowing attackers to intercept and decrypt network traffic.	When a device disconnects from Wi-Fi, the encryption key (used in WPA2) is reset to all zeros. An attacker can force this disconnect and capture encrypted packets, which now use a zero key, making them easy to decrypt.
Impact	Allows eavesdropping, packet injection, and in some cases, hijacking connections	Decryption of previously encrypted Wi-Fi traffic
Password exposure	No (PSK not revealed)	No (PSK not revealed)

Overview of WPA2 and its significance

SL NO.	Attack	Normal Traffic	Malicious Traffic
1	Deauth	1,587,527	38,942
2	Disas	1,938,585	75,131
3	(Re)Assoc	1,838,430	5,502
4	Rogue_AP	1,971,875	1,310
5	Krack	1,388,498	49,990
6	Kr00k	2,708,637	186,173
7	SSH	2,428,688	11,882
8	Botnet	3,169,167	56,891
9	Malware	2,181,148	131,611
10	SQL Injection	2,595,727	2,629
11	SSDP	2,641,517	5,456,395
12	Evil Twin	3,673,854	104,827
13	Website spoofing	2,263,446	405,121

AWID3 Dataset

- Publicly available datasets for these two attacks are scarce.
- The dataset contains 254 features, a majority of which have missing values (NaNs).
- The raw dataset necessitates extensive cleaning before it can be used for machine learning.
- The traffic generation framework lacks documentation for ensuring reproducibility.
- The documentation lacks any description of the feature selection or data labeling methodologies.

Overview of WPA2 and its significance

Devices and Tools for the Test bed

Node	Brand	OS	CPU/RAM	IP Address	MAC Address
Access Point (Wireless Router)	ASUS RT-AC68U	Linux 3.0.0.4.386_51733	Dual-core 800 MHz	192.168.3.2	0C:8D:98:23:SE:21
Windows STA (Desktop)	Custom	Windows 10 Pro (20H2)	Core i5, 8GB DDR4	192.168.3.45	D4:25:8B:E2:8F:92
Mobile STA (Smartphone)	iPhone 6s	iOS 14.2	Apple A9, 2GB	192.168.3.24	88:55:B5:55:B2:D6
Attacker (Laptop)	HP Omen 15	Kali Linux 2024.4	Core I7, 16GB	192.168.3.113	2C:DB:07:14:D3:BD
Mobile STA (Smartphone)	Samsung Note 4	Android 6.0.2	Snapdragon, 3GB	192.168.3.139	A4:08:DC:3C:9A:01

File Name	Attacks	Attack Tools	Total Frames	Attack Frames	Duration
Krack	Channel MitM Attack, Key Reinstallation	Hostapd	1,656,984	289,424	5/10
Kr00k	TK Reinstallation	Aircrack-ng	3,625,550	677,812	5/10

Attack Simulation & Methodology

Test Bed

Attack Scenario & Packet Analysis

Internet

AP(WPA2)

Linux STA

Mobile STAs

Windows STA

Approx. 10 minutes

Normal Traffic Capture
Attack Traffic Capture

Capture Attack- Wireshark

eapol || wlan.fc.type_subtype == 0x08

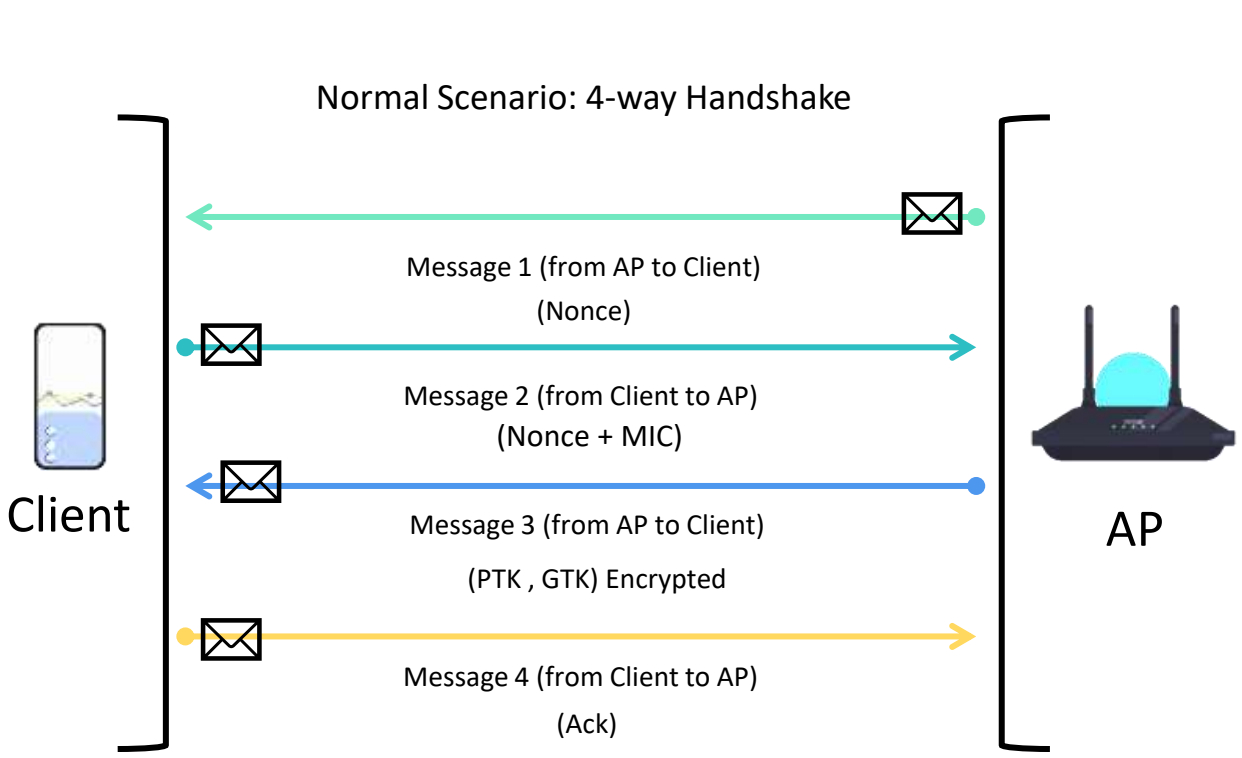
Tcpdump Analyzer

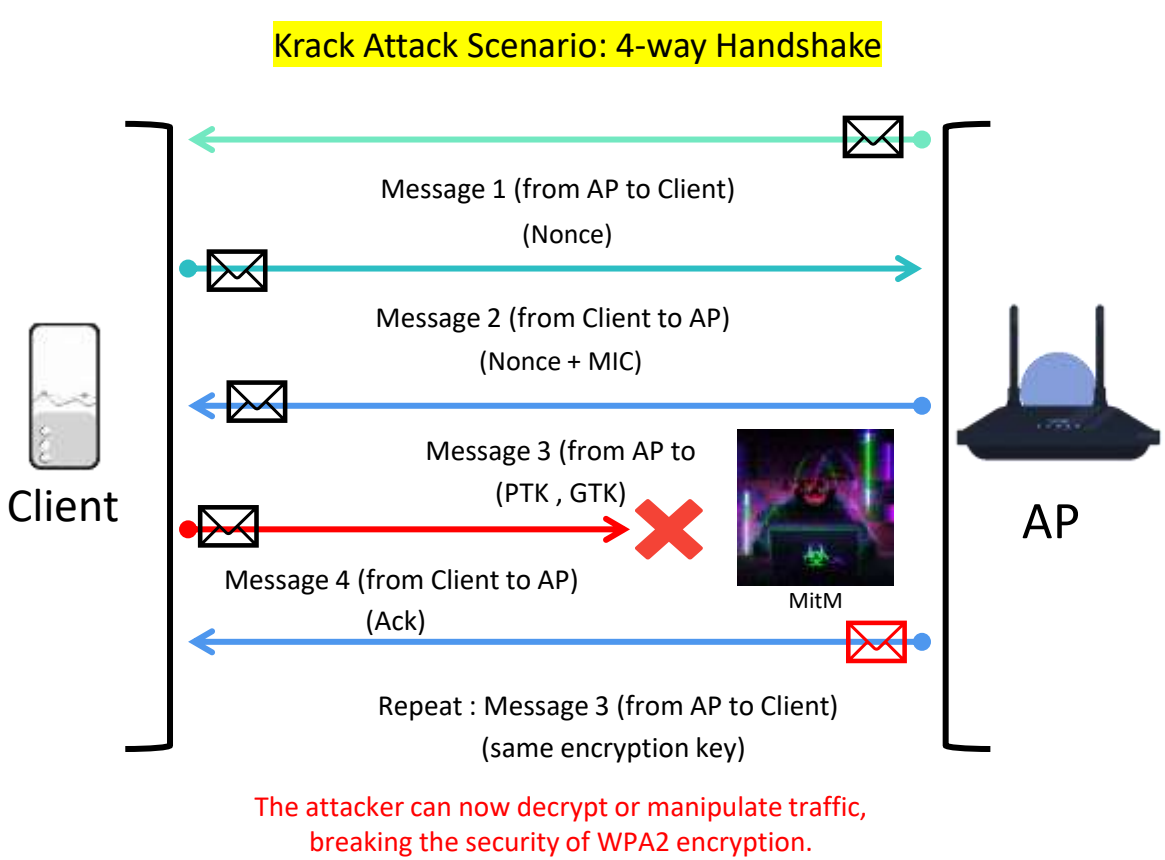
open source tcpdump to capture and analyze pcap files

Attack Simulation & Methodology

KRACK Attack

Exploiting the 4-Way Handshake in WPA2



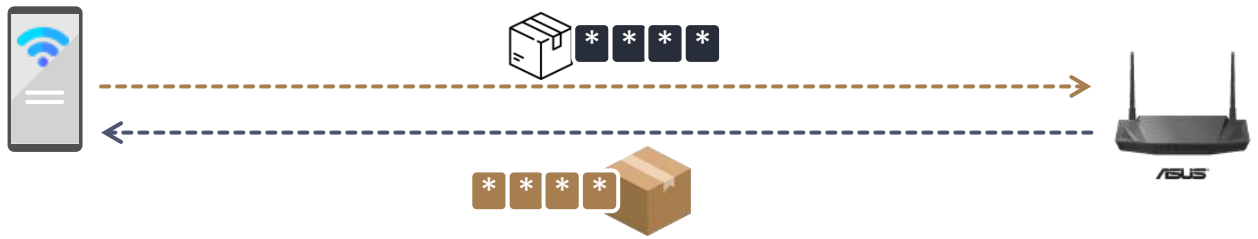


The attacker can now decrypt or manipulate traffic, breaking the security of WPA2 encryption.

Attack Simulation & Methodology

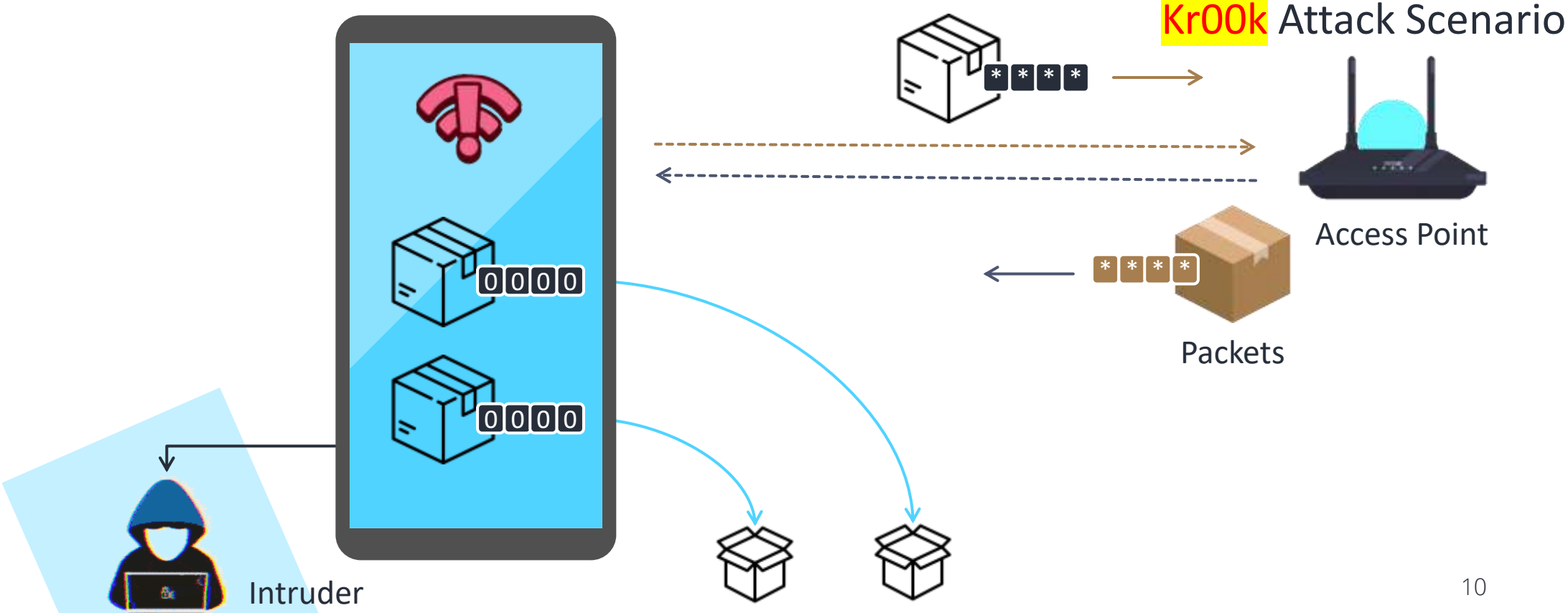
Kr00k Attack

Exploiting Disassociation to Decrypt Wireless Traffic



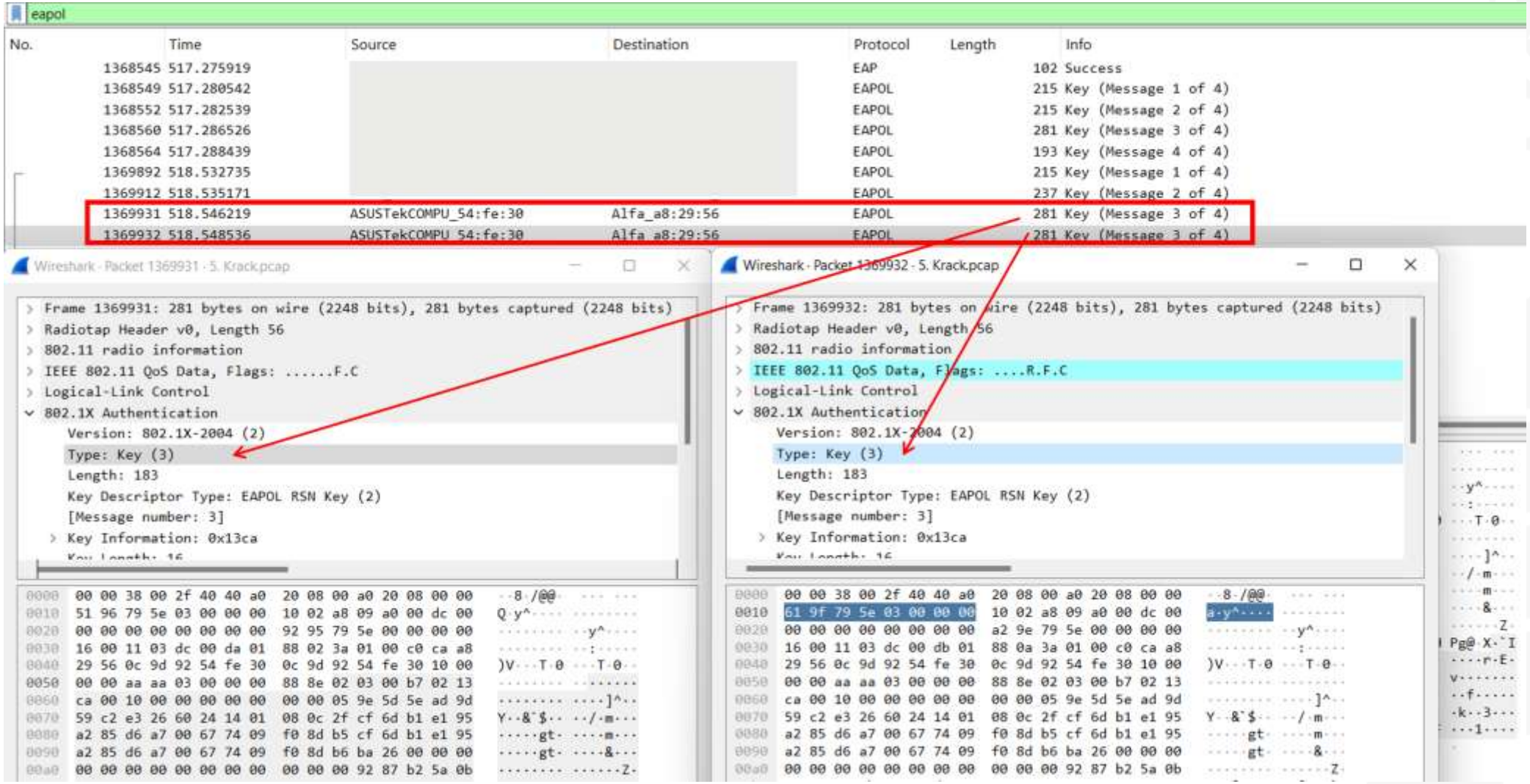
Active Eavesdropping

Kr00k Attack Scenario



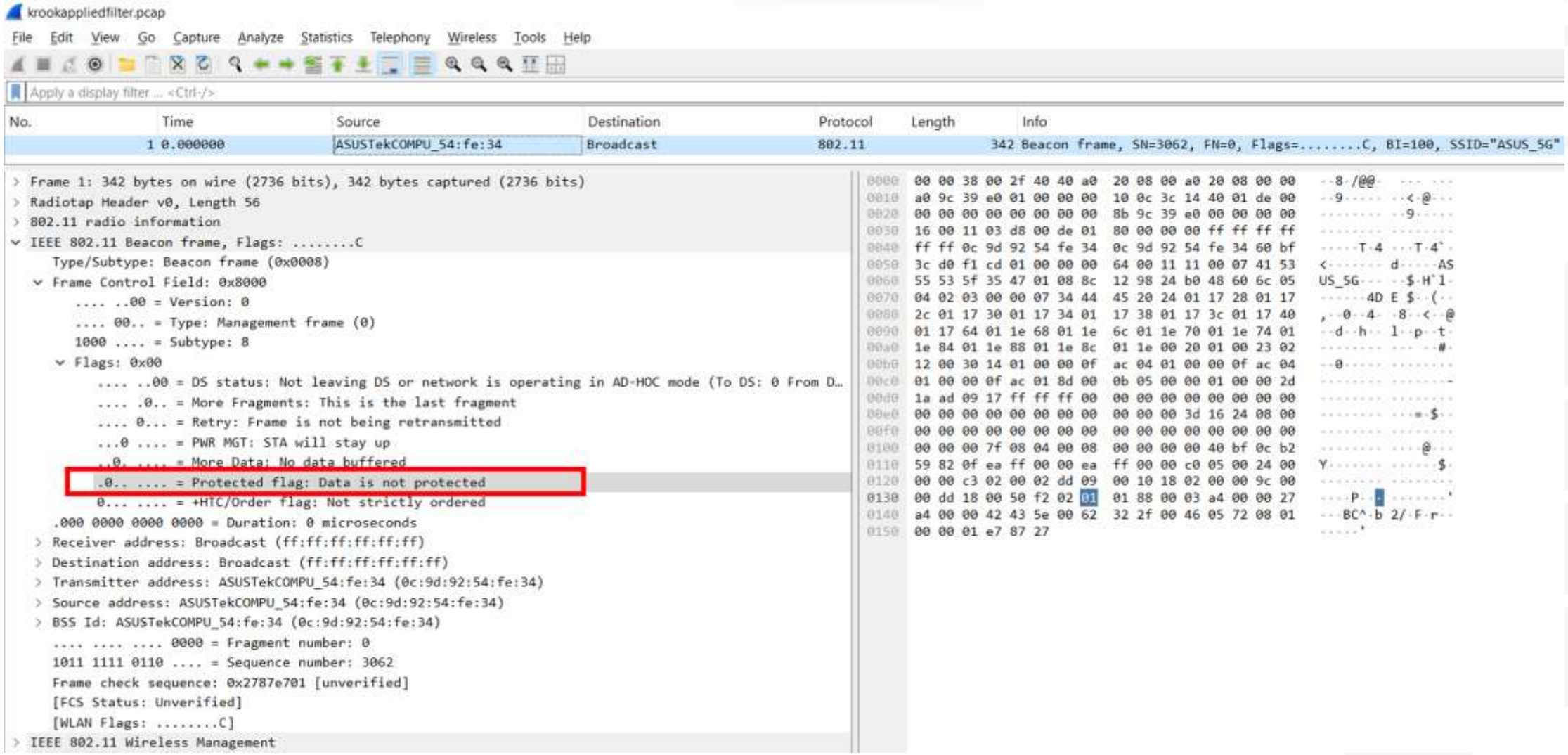
Attack Simulation & Methodology

Detecting KRACK Attacks in PCAP Files



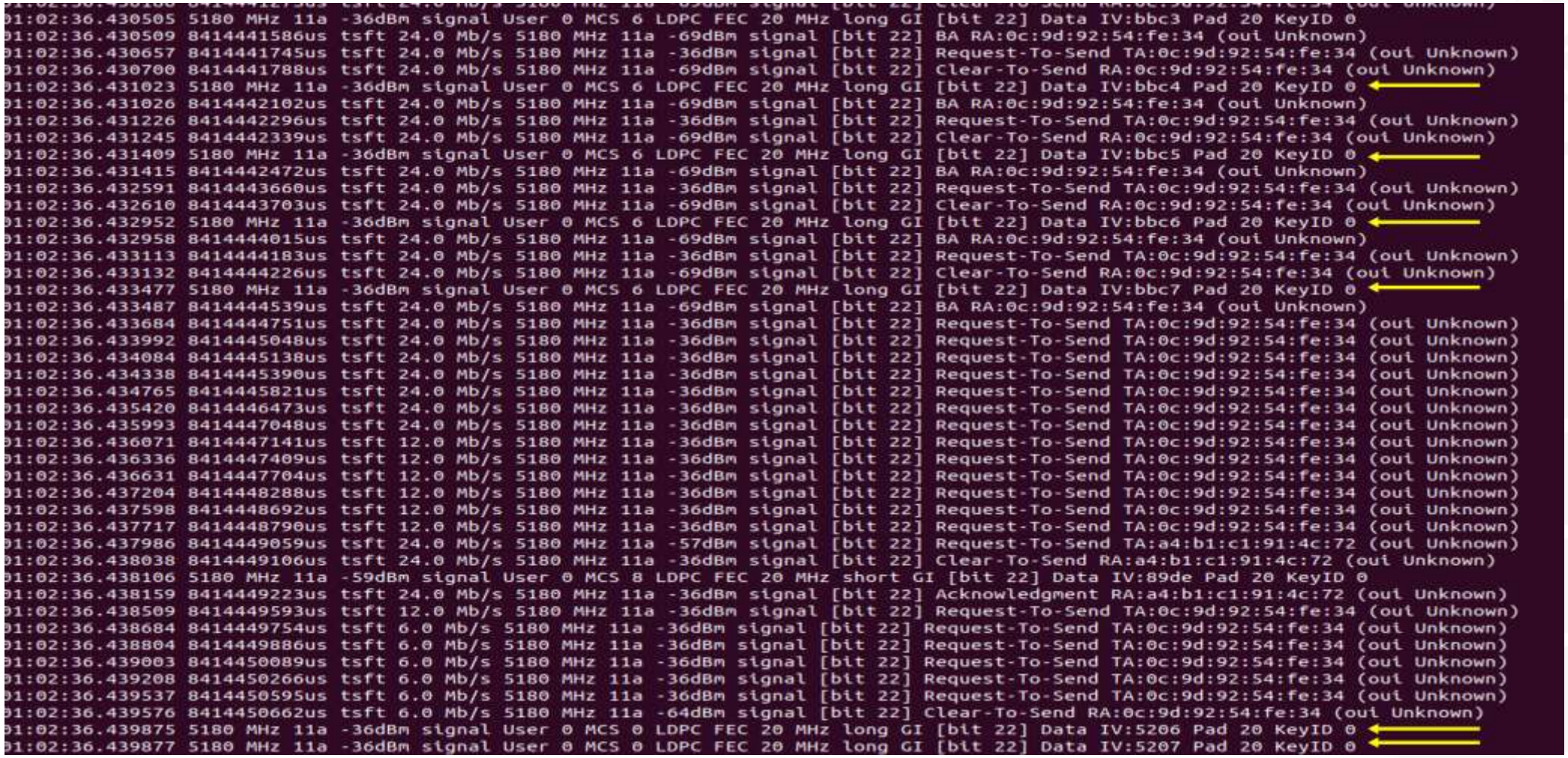
Attack Simulation & Methodology

Detecting Kr00k Attacks in PCAP Files



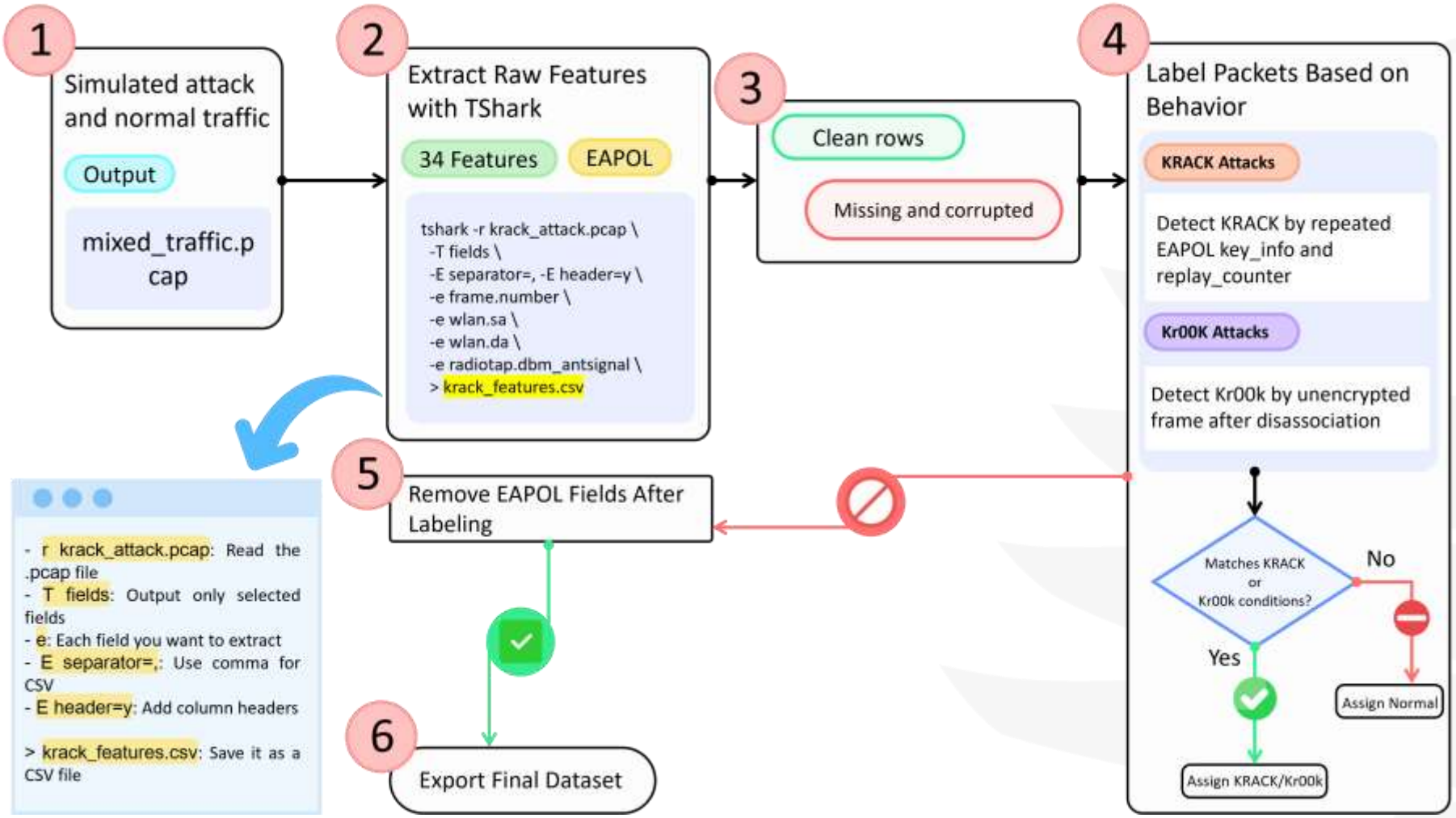
Attack Simulation & Methodology

Detecting Kr00k Attacks in PCAP Files



WPA2-KKID Dataset

TShark Feature Extraction Workflow Diagram



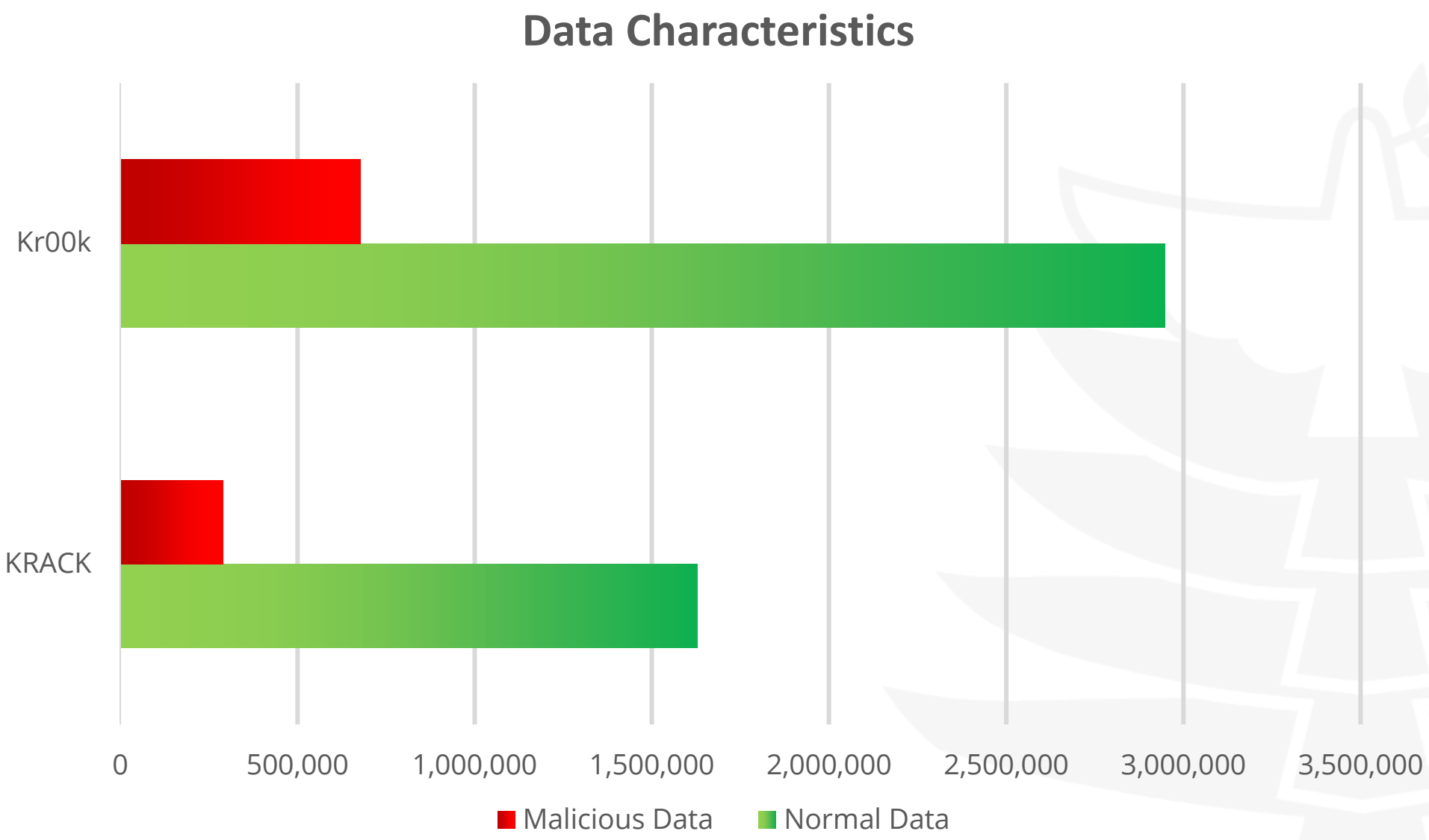
Feature Description

SL No.	Node
1	frame.encap_type
2	frame.len
3	frame.number
4	frame.time
5	frame.time_delta
6	frame.time_delta_displayed
7	frame.time_epoch
8	frame.time_relative
9	radiotap.channel.flags.cck
10	radiotap.channel.flags.ofdm
11	radiotap.channel.freq
12	radiotap.dbm_antsignal
13	radiotap.length
14	radiotap.present.tsft
15	radiotap.rxflags
16	radiotap.timestamp.ts

SL No.	Node
17	wlan.duration
18	wlan.fc.ds
19	wlan.fc.frag
20	wlan.fc.order
21	wlan.fc.moredata
22	wlan.fc.protected
23	wlan.fc.pwrmgt
24	wlan.fc.type
25	wlan.fc.retry
26	wlan.fc.subtype
27	wlan.ra
28	wlan_radio.duration
29	wlan_radio.channel
30	wlan_radio.data_rate
31	wlan_radio.frequency
32	wlan_radio.signal_dbm

SL No.	Node
33	wlan_radio.phy
34	Label

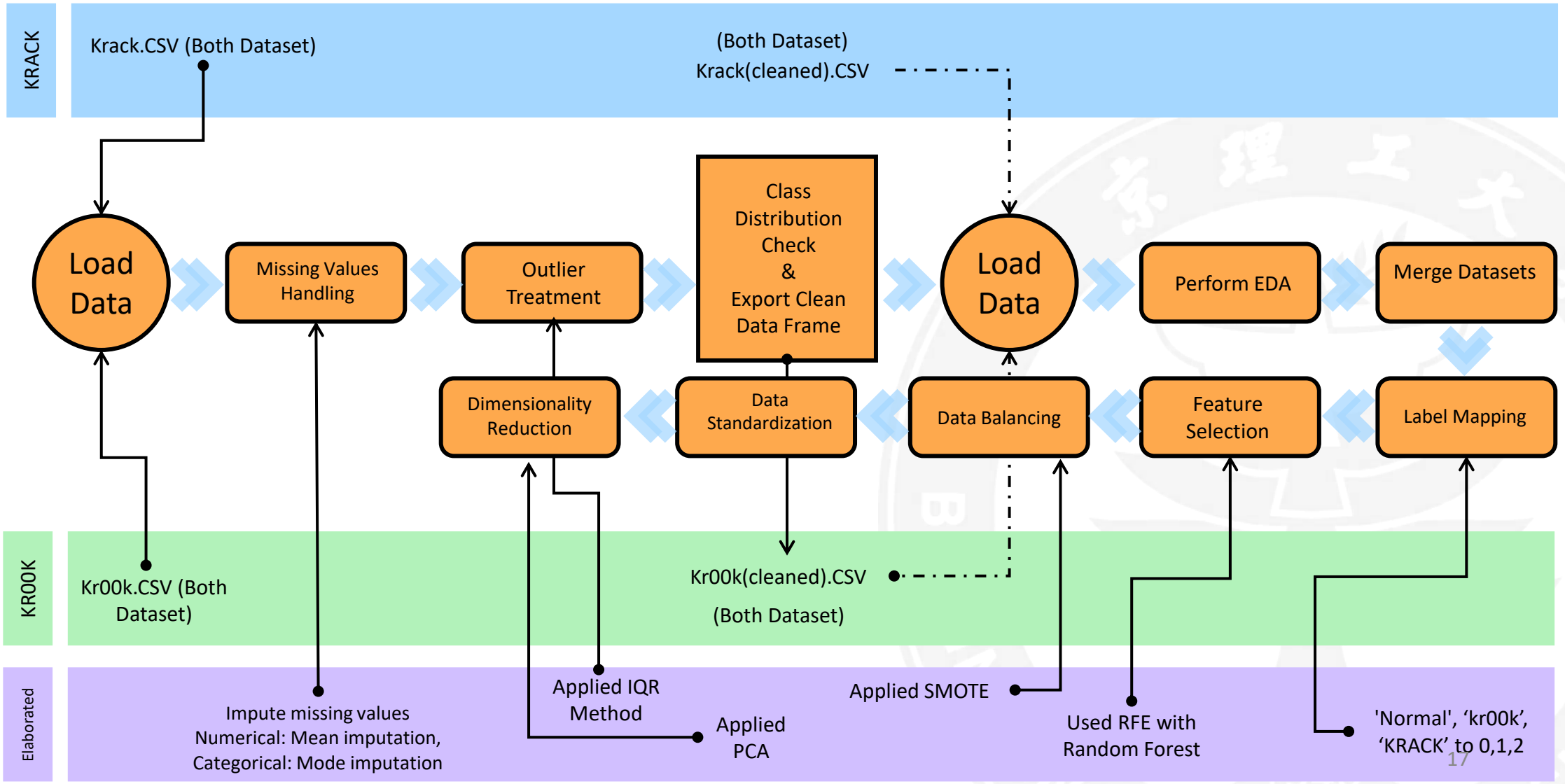
WPA2-KKID Dataset



Model Training & Performance

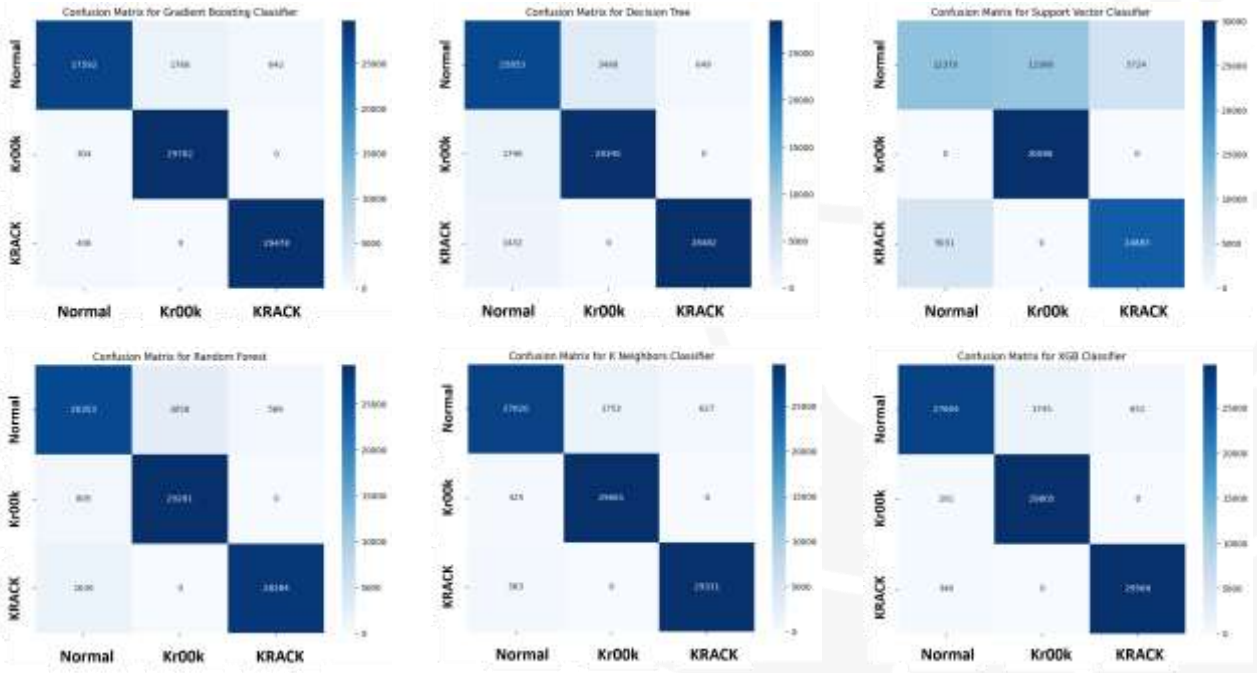
Data Preprocessing

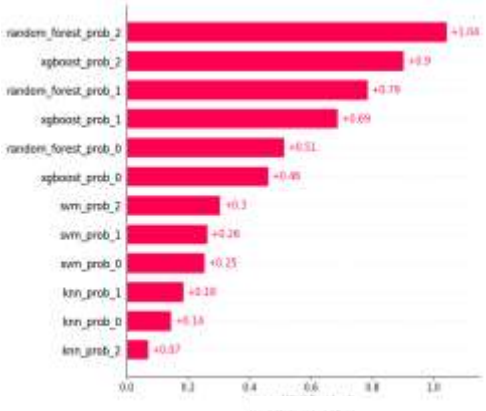
The Foundation of Machine Learning

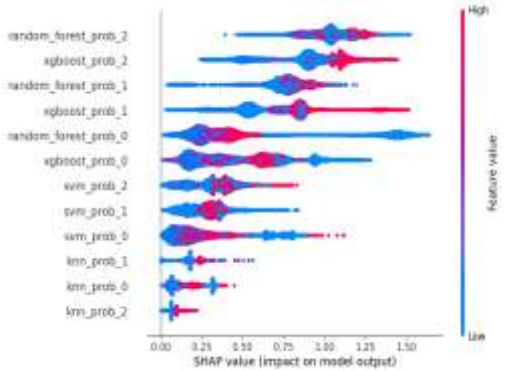


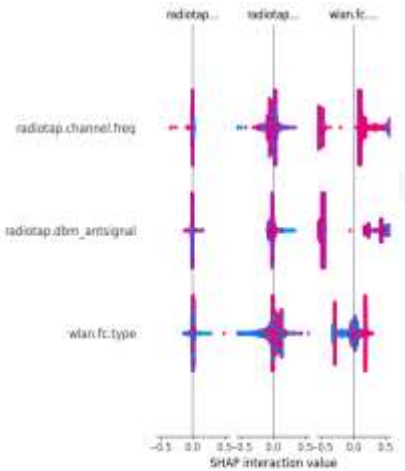
Model Training & Performance

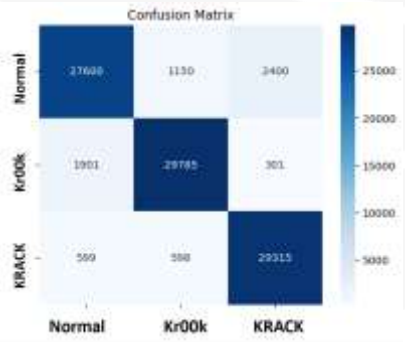
Method	Cross-Validation Accuracy	Accuracy	Precision	Recall	F1-Score
Decision Tree	88.76% ± 0.13	91.86%	91.94%	91.86%	91.86%
Random Forest	92.70% ± 0.19	93.24%	93.34%	93.24%	93.23%
Support Vector Classifier	76.49% ± 0.14	74.82%	74.67%	72.64%	72.64%
XGB Classifier	96.48% ± 0.06	96.64%	96.69%	96.62%	96.62%
Gradient Boosting	96.31% ± 0.06	96.50%	96.55%	96.48%	96.48%
K Neighbors Classifier	96.17% ± 0.07	96.24%	96.27%	96.22%	96.22%
Ensemble Method	**96.84% ± 0.26**	**97.00%**	**97.02%**	**97.05%**	**97.08%**











Strength & Limitations

Strengths & Limitations

A Data Scarcity.

Addresses scarcity of KRACK & Kr00k data and Comprehensive: >5.5M frames & 34 features

b Realistic Testbed Setup.

Reproducible testbed with diverse clients

C Reproducibility.

Dataset and ML-ready features are openly available on Figshare. Code for training and evaluation is shared on GitHub, promoting transparency and reproducibility.

A Restricted Hardware & Software Scope.

Testbed limited to a few devices, may not fully capture the variability of attacks on other chipsets, vendors, or operating systems

b Raw Traffic Release.

Due to privacy/security concerns, raw PCAPs are not shared; only feature-engineered datasets are provided.

C No WPA3 or Wider Attack Coverage.

Focused only on KRACK and Kr00k, without inclusion of other WPA2/WPA3 vulnerabilities (e.g., Dragonblood, Evil Twin).

Future Work

Expanding Dataset Diversity

Expand the dataset to include Evil Twin and Disassociation attacks, which exploit WPA2 handshake vulnerabilities, and capture data from diverse devices, chipsets, and network configurations to cover a broader range of attack scenarios in varied environments.

01

Beyond WPA2

Extend research to WPA3 and other protocols; explore side-channel and application-layer attacks

02

Improving Model Generalization

Explore advanced techniques for handling class imbalance, such as generative models and data augmentation, while testing the models on real-world network traffic to improve generalization across diverse environments and devices.

03

Integration with Network Monitoring Tools

Integrate the expanded dataset into network monitoring tools like Wireshark or Snort for seamless detection of KRACK, Kr00k, Evil Twin, and Disassociation attacks, while exploring cloud-based and distributed systems for scalable intrusion detection across large networks.

04

道德经

知人者智，自知者明。 胜人者有力，自胜者强。 知足者富。 强行者有志。 不失其所者久。 死而不亡者，寿。

谢谢！