Feature-Xniffer: On-the-fly Extraction and Compression of Network Traces for IoT Forensics

Fabio Palmese
DEIB, Politecnico di Milano
Milan, Italy
fabio.palmese@polimi.it

Alessandro E. C. Redondi
DEIB, Politecnico di Milano
Milan, Italy
alessandroenrico.redondi@polimi.it

Matteo Cesana
DEIB, Politecnico di Milano
Milan, Italy
matteo.cesana@polimi.it

***Abstract*—This demo presents Feature-Xniffer, a framework for real-time network traffic feature extraction and compression in IoT forensics scenarios. The tool operates on Wi-Fi Access Points, computing statistical features from network traffic as it flows, thus eliminating the need for storing raw PCAP files. Feature-Xniffer also implements lossy compression techniques (Scalar Quantization, Vector Quantization, and PCA) to significantly reduce storage requirements while maintaining forensic capabilities. We demonstrate its effectiveness in IoT forensics tasks such as device identification and human activity recognition.**

I. Introduction

The Internet of Things (IoT) has transformed our daily lives, with smart devices becoming ubiquitous in homes, offices, and industrial environments. While these devices offer unprecedented convenience and automation, they also introduce new challenges in the field of digital forensics. IoT devices continuously generate network traffic that may serve as valuable evidence in forensic investigations, potentially revealing information about device activities, user behaviors, and environmental conditions [1]. However, the standard approach to network traffic analysis in forensic investigations faces significant limitations that hinder its effectiveness in IoT scenarios. Traditional IoT forensics methodologies involve capturing raw network packets in PCAP files for subsequent analysis, facing two major challenges in IoT environments: first, the sheer volume of data generated by IoT devices quickly consumes substantial storage resources; second, the post-processing of large files introduces significant delays in forensic analysis, hampering timely investigation responses. To tackle the problem, we propose Feature-Xniffer, a comprehensive solution that combines real-time feature extraction with sophisticated compression techniques, enabling sustainable long-term forensic monitoring of IoT environments without sacrificing analytical capabilities. Our framework directly operates in Wi-Fi Access Points, capturing packets and aggregating them into time windows, producing statistical values on the fly. In addition, the tool incorporates specific lossy compression techniques to further optimize storage efficiency, including Scalar Quantization, Vector Quantization, and transform coding based on Principal Component Analysis. This demo presents Feature-Xniffer, with a focus on the impact of lossy compression on forensics capabilities.

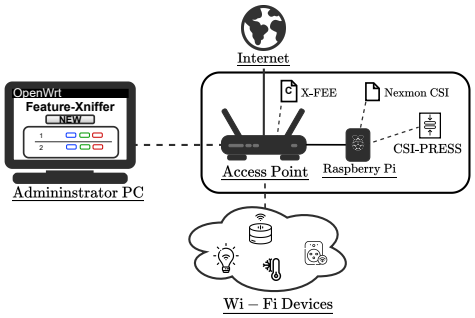

Fig. 1: Sketch of the Feature-Xniffer architecture

II. Framework Architecture

Feature-Xniffer builds upon our previous tool [2], maintaining its core components while introducing new compression modules. The system consists of:

- A web-based user interface integrated with the OpenWrt LuCI framework, providing intuitive configuration
- A Feature Extraction and Compression Engine *(X-FEE)* that processes network packets in real-time and produces compressed TCP/IP-based features.
- A Physical-layer Feature Extraction Engine *(P-FEE)* relying on the Nexmon CSI tool [3] for extraction of Wi-Fi Channel State Information, coupled with *CSI-Press* for real-time lossy compression. This module is executed in a Raspberry Pi as the Nexmon CSI tool is compatible with a limited number of network interface cards.

The key innovation in Feature-Xniffer is the implementation of different compression techniques that operate on the extracted features:

- **Scalar Quantization (SQ):** Reduces the precision of individual features by mapping them to a smaller set of quantized values.
- **Vector Quantization (VQ):** Groups similar feature vectors together and represents them with a single codeword from a codebook. We use k-means clustering for this goal.
- **Principal Component Analysis (PCA):** Reduces feature dimensionality by projecting data onto a lower-dimensional space that preserves maximum variance.

Fig. 2: CSI-PRESS User Interface with a configuration tab, the compression settings, the feature output, and an overview of the instantaneous storage impact.

The demo allows visitors to observe the effects of different compression settings on both storage requirements and forensic accuracy, providing real-time visualization of the extracted features and the impact on the following forensic task.

III. Demo Description

The demo is built on a Linksys WRT3200ACM access point running OpenWrt firmware with Feature-Xniffer installed. Several consumer smart home devices are connected to the network to generate realistic traffic patterns for analysis.
Visitors can interact with the demo through a dedicated user interface: they can (i) create custom configurations through the web interface, (ii) visualize network traffic features in real-time before and after compression, demonstrating the tool's ability to maintain forensic information while reducing data volume, and (iii) they can monitor the storage requirements with a comparative visualization of raw PCAP size versus the tool feature size (before/after compression). Figure 2 shows the demo user interface. Throughout the demonstration, we emphasize the practical benefits of Feature-Xniffer for real-world IoT forensics, including reduced storage costs, faster analysis, and the ability to maintain longer historical records for investigation. The demonstration also highlights how different compression techniques can be optimized for specific forensic applications, allowing users to balance storage efficiency against analytical precision based on their specific requirements.

IV. Results

We used the framework's compressed output for IoT forensic tasks such as IoT device identification and human activity recognition.

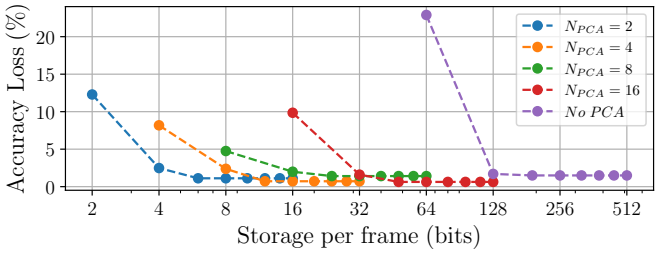

Fig. 3: Performance Loss when using PCA combined with Scalar Quantization. The original frame size is 1792 bits.

We report here the results of one of the considered scenarios: CSI-based human presence detection. After collecting and compressing the CSI values, the framework aggregates packets and extracts a single-valued feature A^* for each time window. The feature is computed over the values of the CSI amplitude and is used to feed a binary threshold classifier. To improve the classification, we opt for a differential feature relying on the difference between two consecutive windows, and apply a high-pass filter before classification. The classifier is trained on the original data, and is evaluated with compression: before the feature extraction, we apply PCA combined with SQ and VQ. Figure 3 reports the performance loss for the PCA + SQ case when compared with results obtained with the uncompressed data. Using 8 bits per frame (against 1792 bits uncompressed) is enough to obtain almost perfect results (less than 5% loss).

V. Conclusions

This demo has presented Feature-Xniffer, a complete framework allowing on-the-fly network feature extraction with advanced compression techniques for efficient storage of network traffic features in IoT forensic applications. The tool aggregates network packets and extracts statistical features on-the-fly, relying on both network/transport layer and physical layer characteristics. Moreover, implementing lossy compression methods (Scalar Quantization, Vector Quantization, and PCA), Feature-Xniffer addresses one of the most significant challenges in IoT forensics: the sustainable storage of forensic data in increasingly complex IoT environments. The live demo shows the tool with feature extraction in real-time and the impact of lossy compression on the classification performance.

Acknowledgements

This study was carried out within the PRIN project COMPACT, CUP: D53D23001340006

References

[1] M. Stoyanova et al., "A Survey on the Internet of Things (IoT) Forensics: Challenges, Approaches, and Open Issues," *IEEE Communications Surveys & Tutorials*, vol. 22, no. 2, pp. 1191–1221, 2020.
[2] F. Palmese, A. E. C. Redondi, and M. Cesana, "Designing a Forensic-Ready Wi-Fi Access Point for the Internet of Things," *IEEE Internet of Things Journal*, vol. 10, no. 23, pp. 20 686–20 702, 2023.
[3] F. Gringoli et al., "Free your CSI: A channel state information extraction platform for modern Wi-Fi chipsets," in *Proceedings of the 13th International Workshop on Wireless Network Testbeds, Experimental Evaluation & Characterization*, 2019.