Forensic Science International: Digital Investigation 32 (2020) 300907

Contents lists available at ScienceDirect

Forensic Science International: Digital Investigation

journal homepage: www.elsevier.com/locate/fsidi

DFRWS 2020 EU — Proceedings of the Seventh Annual DFRWS Europe

EMvidence: A Framework for Digital Evidence Acquisition from IoT Devices through Electromagnetic Side-Channel Analysis

Asanka Sayakkara*, Nhien-An Le-Khac, Mark Scanlon. *Forensics and Security Research Group, University College Dublin, Ireland*

Keywords
Digital forensics
Electromagnetic side-channels
software framework
Internet-of-things (IoT)
Machine learning

* Corresponding author.

E-mail addresses: asanka.sayakkara@ucdconnect.ie (A. Sayakkara), an.lekhac@ucd.ie (N.-A. Le-Khac), mark.scanlon@ucd.ie (M. Scanlon).

1. Introduction

Digital forensics involves data acquisition from digital devices in order to help progress corporate, civil and legal investigations. The emergence of the Internet of Things (IoT) has revolutionised the potential of digital forensics by opening up vast new sources of evidence. While IoT devices can provide invaluable data for digital investigations, acquisition of data from IoT devices is not a straightforward task. They are manufactured by various companies with custom hardware and software designs. As a result, IoT devices lack standard interfaces and forensic acquisition methods. This can often result in a device requiring a memory chip-off procedure in order to access its data (Watson and Dehghantanha, 2016).

EM side-channel analysis (EM-SCA) is a branch in information security where the unintentional electromagnetic (EM) emissions from computing devices (Kocher et al., 1999). This has been used for various purposes including software behaviour detection, software modification detection, malicious software identification, and data extraction (Sayakkara et al., 2019a). The possibility of applying EM-SCA in digital forensic investigation scenarios involving IoT devices has been proposed recently (Sayakkara et al., 2019b). When it is difficult or impossible to acquire forensic evidence from an IoT device, observing EM emissions of the device can provide valuable information to an investigator. This work addresses the challenge of making EM-SCA a practical reality to digital forensic investigators by introducing a software framework called EMvidence. The framework is designed to facilitate extensibility through an EM plug-in model.

2. EMvidence forensic framework

The EMvidence framework consists of a main core with multiple default modules and facilitates the addition of third-party plug-ins depending on future requirements. Its main component is its core GUI that provides the default interface to a user. It also manages the modules and plug-ins by establishing communication between them in a coordinated fashion. Together with the core GUI, the framework comes with three default software modules that are essential to the normal operations of the framework; data acquisition, data visualisation, and report generation. Furthermore, depending on the requirements, third-party users can quickly develop and integrate plug-ins to the core GUI. Such plug-ins may provide various data analysis capabilities such as software behaviour detection, cryptographic key recovery, etc. The source code of the framework and its associated plug-ins are available at a *Github* repository.[1]

EMvidence supports two types of data acquisition methods. Firstly, the observation of EM emission signals can be made for a predefined period of time without any interaction or communication with the device under test (DUT) from a few centimetres away. This is the approach used in a digital forensic investigation scenario. Secondly, EM signals can be acquired while actively interacting with the DUT in scenarios where it is safe to communicate with the device through an interface, e.g., universal serial bus (USB), universal asynchronous receiver/transmitter (UART), or Ethernet. Fig. 1 illustrates the acquisition of EM traces in a coordinated fashion from a target device. Once such data are used to build ML models, they can be incorporated into the EMvidence framework to inspect similar IoT devices in investigative scenarios.

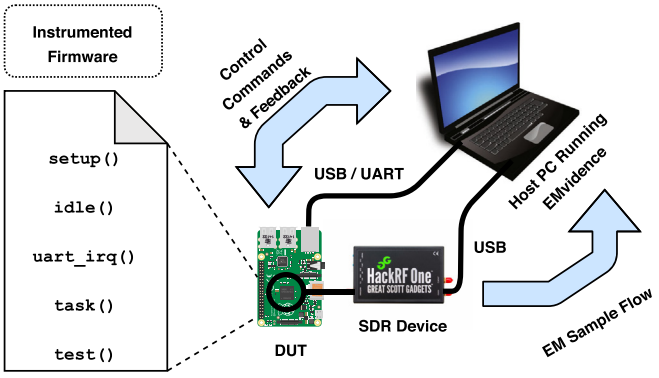

Fig. 1. Controlled/Instrumented EM signal Acquisition.

3. Evaluation

Consider a hypothetical scenario where an IoT device has been deployed in

[1] https://github.com/asanka-code/EMvidence

https://doi.org/10.1016/j.fsidi.2020.300907

a building as a part of an intruder detection system. The device consists of a sensor that detects movements within a specified space of the premises. The device consists of two actuators, an alarm and a door lock, that it can control independently. Furthermore, the device is connected to a GSM module in order to send and receive SMS. The device's firmware is programmed to continuously read the motion sensor to detect intrusions into the premises. Upon detection, it can perform one of three tasks – locking the door, firing an alarm, or sending a text message to the owner. At any time, the device can be disabled by pressing a physical button that puts the device into an idle state.

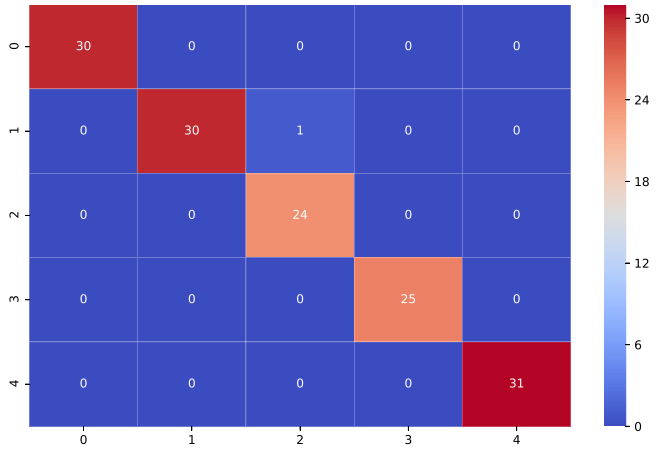

Fig. 2. Confusion Matrix of the IoT Device State Classifier.

The hypothetical IoT device was emulated by using an *Arduino Leonardo* board. A neural network classifier based on multi-layer perceptron (MLP) architecture was selected to distinguish each state of the IoT device on the EMvidence framework. Fig. 2 illustrates the confusion matrix of the classification results. The classifier was able to achieve an average F1-score of 99% in distinguishing the 5 IoT device states. This indicates that a pre-trained model can identify internal software states of IoT devices. For example, if it was identified that the device is in the *idle* state at the time investigators arrived at the scene, it is clear that someone deliberately turned the device into idle state in order to stop it from triggering the intruder alarm. In that case, fingerprints on the button of the IoT device could potentially help to identify the insider.

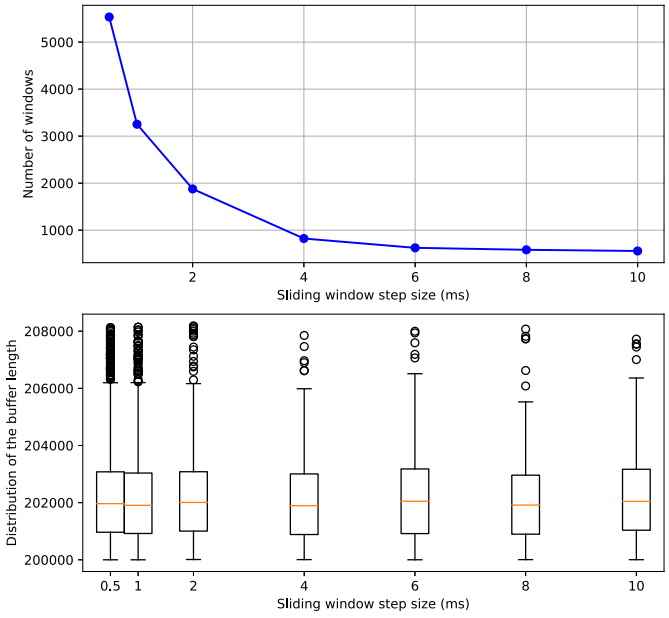

Fig. 3. Effect of the Sliding Window Step Size to the Data Collection Buffer over 10 ms.

The overhead of processing EM signals in real-time is evaluated as follows. While EM signal capture device is set to 20 MHz sampling rate, a sliding window with a fixed width of 10 ms was used to slide through the real-time I-Q data feed. Each data window was then preprocessed in real-time to generate the features and fed to a neural network-based binary classifier to detect the presence of ECC cryptography operations. The step size of the sliding window, i.e., the amount of overlap between consecutive windows, was varied between 0.5 ms and 10 ms for different independent trials. In all experiments, the total signal capturing duration was fixed to 10 s.

In Fig. 3, the graph on the top illustrates the number of windows produced against the sliding window step size while the graph at the bottom illustrates the statistics of the number of data samples waiting in the real-time buffer until the sliding window had processed them. It is evident that even though the number of sliding windows to process increases with the reduction of sliding window step size, it does not to incur any considerable overhead to the real-time processing buffer. The production and consumption of the EM samples were in an equilibrium.

4. Conclusion

With the ever-increasing applications of IoT systems in domestic and industrial environments, digital forensic investigations increasingly require the extraction of digital evidence from them. The most forensically useful information in IoT devices are currently extracted by intrusive inspections of hardware that makes them less forensically sound. This work presented the design of EMvidence, a framework for digital forensic investigators and researchers to leverage unintentional EM radiation from IoT devices as an information source. EMvidence is designed in a manner that it can be easily extended with new functionalities to keep up with the dynamism of IoT devices. Experimental demonstrations proved that ML classifiers can be used to gain useful insights in IoT investigative scenarios.

References

Kocher, P., Jaffe, J., Jun, B., 1999. Differential power analysis. In: Advances in Cryptology (CRYPTO '99). Springer, p. 789.

Sayakkara, A., Le-Khac, N.A., Scanlon, M., 2019a. A survey of electromagnetic side-channel attacks and discussion on their case-progressing potential for digital forensics. Digit. Invest. 29, 43–54. https://doi.org/10.1016/j.diin.2019.03.002.

Sayakkara, A., Le-Khac, N.A., Scanlon, M., 2019b. Leveraging electromagnetic side-channel analysis for the investigation of IoT devices. Digit. Invest. 29, S94–S103. https://doi.org/10.1016/j.diin.2019.04.012.

Watson, S., Dehghantanha, A., 2016. Digital forensics: the missing piece of the internet of things promise. Comput. Fraud Secur. 2016 (6), 5–8.