RF Fingerprinting for 802.15.4 Devices: Combining Convolutional Neural Networks and RF-DNA

Bernard Lebel

Thales Canada Inc. - TRT
Québec, Québec, Canada
Email: bernard.lebel[at]ca.thalesgroup.com

Louis N. Bélanger,
M. A. Haji Bagheri Fard,
and Jean-Yves Chouinard

Université Laval
Québec, Québec, Canada
Emails: louis.belanger[at]gel.ulaval.ca
mohammad-amin.haji-bagheri-fard.1@ulaval.ca
jean-yves.chouinard[at]gel.ulaval.ca

Abstract—**Wireless communications have traditionally relied on the content of the message for authenticating the sender. In protocols relying on the IEEE 802.15.4 standard, such as Zigbee, it is possible for an attacker with the right knowledge and tools to emit crafted packets that will be interpreted by the receiver as being properly identified and thus, inject arbitrary data. One way of protecting oneself from this type of attack is the use of radio frequency fingerprinting through a technique called Radio Frequency Distinct Native Attribute (RF-DNA). This approach has been demonstrated to be efficient for wireless devices of different models but still lacks accuracy when trying to identify a rogue device of the same model as the lawful emitter. This is even more of a challenge when attempting to conduct the fingerprinting using a low-cost yet flexible software defined radio. To address this challenge, the current work-in-progress attempts to train a convolutional neural network in order to be able to discriminate a legitimate device from a rogue device. Initial results show promising performance but a larger dataset of devices is required to be conclusive, which will be the focus of future work.**

Keywords–RF-DNA; Wireless Security; Physical Layer; Neural Networks; Machine Learning.

I. Introduction

A tide of electronic devices traditionally used in isolated small-scale hard-lined networks were augmented with full networking capabilities in the recent years. This mass of newly connected devices comprises industrial controllers, Internet Protocol (IP) cameras, sensors, actuators, and many others collectively forming what is called the Internet of Things (IoT). It is known for some of those devices to rely on wireless communications to operate. Protocols using the standards IEEE 802.15.4 [1] and IEEE 802.11 [2] are popular choices in the IoT [3].

Wireless communications can be used as an entry point to a private and/or restricted network where a malicious actor may interfere with the proper functioning of a system from a distant location. Moreover, attacks have been demonstrated (e.g., [4], [5]) with potential impacts including denial-of-service (DoS), impersonation attacks and Man-in-the-Middle (MitM) amongst others. Implementations of security measures (e.g., Wired Equivalent Privacy (WEP), Wireless Protected Access (WPA) and WPA2 [6]) usually rely on network layers at or above the data-link (MAC) layer of the open systems interconnection (OSI) model [7]. Those layers have been known to be susceptible to manipulations coming from an attacker, sometimes requiring only open-source tools (e.g., Aircrack-ng [8] or KillerBee [9] for IEEE 802.11 and IEEE 802.15.4 respectively) with commercial-off-the-shelf (COTS) wireless adapters that behave as rogue devices. A rogue device can be defined as an illegitimate device that behaves outside of what the communication protocol normally states in order to inject arbitrary traffic into a wireless network and forge data packets to contain misleading data intended to interfere with other devices or the communication itself.

A countermeasure to this problem is to implement security, and more pointedly, authentication at the physical (PHY) layer itself of the OSI model. It has been demonstrated that devices generating radio-signals involuntarily alter the desired theoretical signal due to the physical limitations of the device characteristics that are not part of the communication protocol but rather are due to the electronics of the device itself [10]. Those imperfections are usually within the normal threshold tolerated by a given protocol and do not interfere with the communication itself and constitute the RF fingerprints of a device. The RF fingerprints can be used to authenticate the emitter of a message as they differ across devices.

Also, forging the RF fingerprints of a victim is a challenge in itself for an attacker. Indeed, it requires identifying and mimicking those fingerprints. This in itself is not a trivial problem as the attacking device would also need to prevent its own RF fingerprints from leaking into the resulting signal. This adds a layer of protection that relies on an intrinsic property of the emitter itself (what it is) rather than a preshared secret key (what is known) or an authentication token (what is possessed). As those informations have been known to be stolen, cracked or guessed, they may come short for critical infrastructure protection. Thus, the approach is complementary to the other methods and can strengthen the confidence in the authenticity of the identity of an emitter.

Ramsey et al. [11] used an approach called the Radio Frequency Distinct Native Attribute (RF-DNA) to the Zigbee protocol which uses the IEEE 802.15.4 standard. This approach relies on calculating statistics (i.e., variance (σ^2), skewness (γ) and kurtosis (κ) on physical characteristics (instantaneous phase (ϕ_i), frequency (f_i) and amplitude (a_i))

ISBN: 978-1-61208-605-7

of subregions of an incoming signal.

Additionally, Ramsey et al. [11] demonstrated that it is possible to use a COTS software-defined radio (SDR) to obtain satisfactory results for discriminating an impersonator from a legitimate device. A SDR is a device capable of acquiring a radio-signal in a wide range of frequencies and which delegates the processing of this signal to a software implementation rather than using specialized hardware to do so. This allows a user to have access to a wide range of protocols and frequencies using a single device. It requires a software implementation of the protocol stack and that the communication occurs within the SDR frequency range and bandwidth. SDR vary greatly in terms of price range but some solutions, such as the USRP B200mini from Ettus Research [12] are fairly low-cost when compared to high-end lab equipment and have a smaller size factor. One concern of using a SDR for acquiring the signal to extract its RF-DNA is to ensure sufficient bandwidth can be achieved to capture the hardware-specific variations. The results obtained in [11] supported that a low-cost SDR, such as the B200mini, was enough to discriminate devices based on the comparison of their RF-DNA. The true verification rate (TVR) (i.e., how often a packet was accepted when it came from a legitimate device) neared 100% while the rogue acceptance rate (RAR) (i.e., how often the spoofing devices were accepted as legitimate ones) dropped to 0% for devices that were of different models. To maintain a TVR of >90% in the case where the devices were of the same model, the RAR ranged between 32% and 54%.

The current work seeks to lower the RAR while maintaining or increasing the TVR in the event of a rogue device using the same model as a legitimate one to communicate with another station using the IEEE 802.15.4 standard. The proposed approach seeks to improve the performance of the decision model from [11] that combined a multiple discriminant analysis/maximum likelihood (MDA/ML) process for dimensionality reduction and Bayesian decision criteria for classification. Instead, it is proposed to train a convolutional neural network (CNN) [13], [14] to recognize the devices without requiring that the dimensionality be reduced.

A CNN is a machine learning model that works by attempting to train a set of filters used for convolutions on the input signal to highlight the most discriminant features that can be spatially distributed in that signal. The response to the input signal of each filter at each location across the signal is the output of a convolutional layer (CL). This output is then passed to a subsampling layer which is responsible of reducing the number of outputs by pooling a given region together using a given function (e.g., maximum value or the average of values). One or more fully-connected layer make for the last layers of network and are responsible for the classification itself.

CNNs have been demonstrated to be robust to data translation and are able to take into account a level of spatial distribution of a dataset [15], [16], [17], [18]. This may constitute an advantage in the context of RF fingerprinting as signals are distributed in time and slight spatial translations may occur in the captured data. The ongoing work and preliminary results aim at validating the use of CNN in that context.

The next section presents the methodology used for acquiring RF signals for analysis and the extraction of features following the RF-DNA methodology and the structure of the CNN used for analysis. Section III describes preliminary results obtained in the ongoing work. Section IV summarizes the preliminary results, discuss implications and research concerns. Finally, section V presents future work and next steps.

II. Methodology

The emitting devices were 4 Atmel RZUSBStick, or RZ for short. This device is capable of sending Zigbee packets containing arbitrary data and is also able to communicate through the Zigbee protocol [9]. Arbitrary data was sent periodically at a frequency of 40 packets/sec on channel 26 (i.e., 2.480 GHz). The acquisition was conducted through a USRP B200mini from Ettus Research at a center frequency of 2.480 GHz with a bandwidth of 20 MHz. The data collection was conducted in a RF shielded box to prevent outside interferences with the measurements. Each device was placed at the exact same location for each data collection.

The raw signal from the IEEE 802.15.4 preambles of each communication was extracted. Following the work done by Ramsey et al. [11], the preambles were split into 32 equal regions, each comprising 80 samples per region plus the full preamble itself of 2560 samples, for a total of 33 subregions per preamble captured. For each sample, the instantaneous phase (ϕ_i), instantaneous frequency (f_i) and instantaneous amplitude (a_i) were evaluated. Their variance (σ^2), skewness (γ) and kurtosis (κ) were calculated for each subregion. This amounted to a total of 297 features per preamble composed of 33 subregions × 3 RF characteristics × 3 statistics. The number of preambles collected is presented in Table I.

TABLE I. Samples per device.

	RZUSBStick1	*RZUSBStick2*	*RZUSBStick3*	*RZUSBStick4*
Preambles	7148	7094	6832	7086

Extracted features were standardized following (1). Standardization is required to constrain values of features within a comparable range.

$$z_i = \frac{x_i - \overline{x}}{s^2} \tag{1}$$

z_i is the standardized score, x_i the input value, $\overline{x}$ the mean and s^2 the variance. To apply the standardization, the features are structured along a 3 × 33 × 3 matrix as presented in (2) for each collected preamble.

$$\begin{bmatrix} \begin{bmatrix}\sigma^2_{1_\phi} & \sigma^2_{1_f} & \sigma^2_{1_a}\end{bmatrix} & \dots & \begin{bmatrix}\sigma^2_{33_\phi} & \sigma^2_{33_f} & \sigma^2_{33_a}\end{bmatrix} \\ \begin{bmatrix}\gamma_{1_\phi} & \gamma_{1_f} & \gamma_{1_a}\end{bmatrix} & \dots & \begin{bmatrix}\gamma_{33_\phi} & \gamma_{33_f} & \gamma_{33_a}\end{bmatrix} \\ \begin{bmatrix}\kappa_{1_\phi} & \kappa_{1_f} & \kappa_{1_a}\end{bmatrix} & \dots & \begin{bmatrix}\kappa_{33_\phi} & \kappa_{33_f} & \kappa_{33_a}\end{bmatrix} \end{bmatrix} \tag{2}$$

The values s^2 and $\overline{x}$ are calculated along all collected preambles for the 33 subregions. The result is two matrices containing the s^2 and $\overline{x}$ values for the 33 subregions across all collected preambles. The resulting matrix is shown in (3) for $\overline{x}$. s^2 follows the same structure.

$$\begin{bmatrix} \overline{\sigma^2_\phi} & \overline{\sigma^2_f} & \overline{\sigma^2_a} \\ \overline{\gamma_\phi} & \overline{\gamma_f} & \overline{\gamma_a} \\ \overline{\kappa_\phi} & \overline{\kappa_f} & \overline{\kappa_a} \end{bmatrix} \tag{3}$$

The set of features for each of the 33 subregions per preamble was standardized according to its RF characteristics and statistics.

A. *Convolutional Neural Network*

The 297 standardized features were passed on to a CNN constituted of 2 CL with 32 3×1 filters and 64 3×1 filters. Each CL output was connected to a subsampling layer (3×1 average pooling function with 2-step strides). The last subsampling layer was followed by a fully connected layer of 1024 neurons trained with a 0.75 dropout chance before connecting to the output layer. Optimization was conducted using the adaptive moment estimation (ADAM) optimizer with a learning rate of 0.01. Batch size was set at 128 preambles per mini-batch. The implemented model is presented in Figure 1.

III. PRELIMINARY RESULTS

Collected results were analyzed according to two scenarios. Scenario 1 explored training a CNN to discriminate between the 4 known devices. This scenario is meant to demonstrate the general performance of a CNN in the context of RF-DNA. Scenario 2 seeks to replicate the case where an algorithm is trained to be specialized in recognizing if a given preamble belongs to a specific device.

A. *Scenario 1: Differentiation*

The collected preambles were randomized. The full dataset was divided with 80% used for training, 10% for validation and 10% for testing. The output layer has 4 classes, one for each known device. The resulting confusion matrix is reported in Table II. The calculated accuracy is 95.86%.

TABLE II. CONFUSION MATRIX FOR INTERDEVICE CLASSIFICATION.

		Input Labels			
		RZ1	RZ2	RZ3	RZ4
Predicted	RZ1	0.947	0.006	0.002	0.037
	RZ2	0.013	0.951	0.018	0.007
	RZ3	0.002	0.031	0.975	0.004
	RZ4	0.038	0.012	0.006	0.953

The high accuracy obtained for this task demonstrates that CNNs are especially well adapted for ingesting RF-DNA inputs for device classification. Work is still in progress to establish a baseline based on current literature to achieve a comparison between the proposed approach and the one described in [11]. However, the represented case in Scenario 1 is valid only if an algorithm can be trained on all known devices and is expected to find the correct match in a pool of devices that was used during training. In practice, this method is ineffective in the context of rogue device identification as the attacking device is usually not known before the attack occurs. This nullifies the chances that the model can be trained with all expected devices in a certain area. This problem is addressed in the next scenario.

B. *Scenario 2: One vs All*

This scenario aims at filling the gap from the previous one where a model was trained to identify if a preamble originates from one unique device or not. 80% of the dataset was used for training, 10% for validation and 10% for testing. In the first phase, all preambles from one device were considered as being ”Good” (approx. 25% of the total dataset) and preambles from the remaining devices were considered as ”Bad” (approx. 75% of the total dataset), generating an output layer of 2 classes. During training, labels were balanced according to the proportion of the dataset they represented to compensate for the unbalanced dataset. Results are reported in Table III.

TABLE III. CONFUSION MATRIX FOR ONE-VS-OTHERS CLASSIFICATION.

	Others	RZ1
Others	0.986	0.063
RZ1	0.013	0.937

$Acc = 0.973$

	Others	RZ2
Others	0.991	0.080
RZ2	0.009	0.920

$Acc = 0.972$

	Others	RZ3
Others	0.994	0.046
RZ3	0.006	0.954

$Acc = 0.984$

	Others	RZ4
Others	0.977	0.046
RZ4	0.023	0.954

$Acc = 0.987$

As the training set contained samples from each device, it has been postulated that the predictor would be confused if a new device was introduced and requested predictive measures, showing proof of overfitting. To test this hypothesis, a test was conducted by training the expert systems on the dataset but withholding all data from RZ4. The test dataset was evaluated using inputs only from RZ4. If the system did not overfit, attribution would show nearly only *others* attribution. During a subsequent phase of the ongoing work, the results will be compiled for test cases where RZ1, RZ2 or RZ3 is excluded instead of RZ4. Results are presented in Table IV.

TABLE IV. CONFUSION MATRIX FOR ONE-VS-OTHERS WITH A NEW DEVICE (RZ4) EXCLUDED FROM TRAINING SET.

	RZ4
RZ4	0.168
RZ1	0.832

$Acc = 0.168$

	RZ4
RZ4	0.878
RZ2	0.122

$Acc = 0.878$

	RZ4
RZ4	0.955
RZ3	0.045

$Acc = 0.955$

As the results show, the trained model is achieving an accuracy of 95.5% for RZ3 but lower than 87.8% for RZ2 and is very poor (16.8%) for RZ1. The results from introducing a new device during the test phase demonstrate that the model has trouble differentiating devices that may have a more similar

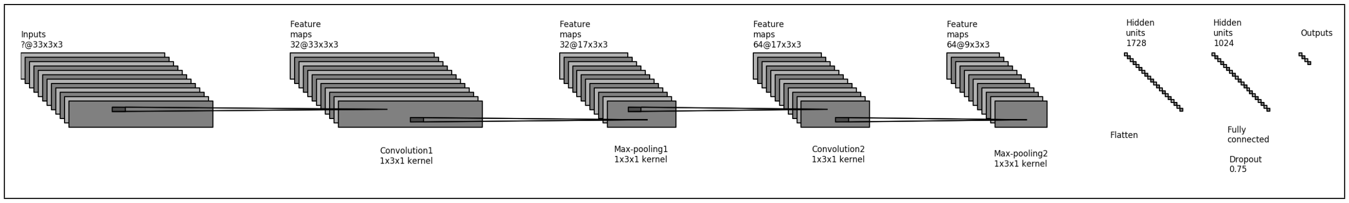

Figure 1. CNN structure.

RF-DNA such as RZ1 and RZ4. At this stage, more data from more devices is required before a conclusion can be achieved.

IV. DISCUSSION

Firstly, when it comes to differentiating between known devices onto which data exists and can be used for training, results show that CNN with standardization from features presented in [11] are effective, achieving a 95% accuracy. Moreover, results have shown that an approach of training a system to recognize itself from other systems performs well in the case where all other systems are known.

However, when exposed to devices which were never part of the learning process, results become unreliable. It is likely that to perform better, the model would need to train on a dataset with more devices. Also, the problem defined in this research specifically targets devices of the same model and manufacturer. It is possible that it is sufficient for categorizing devices from different manufacturers and future work will investigate this.

Also, Table IV shows that some devices may be more alike than others. For instance, it is possible that RZ1 may be more alike to RZ4 and thus, is harder to discriminate when the latter is excluded from the learning process but used only for testing. This supports the hypothesis that more devices are needed for a better predictive model.

Also, when attempting to conduct the learning process on different combinations of RF characteristics, it was noted that statistics on the amplitude lead to better predictive results. This differs from Ramsey et al. [11] whom instead pointed to phase and frequency as being the most useful for categorization. More data collection is required to determine if this could be due to *environmental conditions* that might have altered the transmissions in-between acquisition campaigns.

V. CONCLUSION

This work-in-progress has demonstrated the potential of using CNN in the context of RF fingerprinting while using affordable and flexible SDRs. Also, RF-DNA provides promising results when used to provide features to a CNN. More work will be carried on to tweak the hyperparameters of the model to achieve better results and collect more preambles from new devices. Also, baseline measures based on the state-of-the-art are being generated and will be used for assessing the success of the current approach. New features based on the scientific literature need to be identified and extracted from the RF signal. This would allow to have more elements on which devices from the same manufacturer and of the same model could be discriminated. Finally, ongoing work focuses on trying to compute the RF-DNA in real time on an embedded system in order to optimize the signal processing component. This would ensure a real-time computation and minimize the impact of the implementation of this method on a wireless communication itself.

ACKNOWLEDGMENTS

This research was supported by a Mitacs internship awarded to Mohammad Amin Haji Bagheri Fard and jointly funded by NSERC and Thales Canada Inc./TRT. The authors would also like to thank Mathieu Lvesque and Guillaume Godbout for their participation in data collection and Frederic Audet for his support.

REFERENCES

[1] IEEE Std 802.15.4, IEEE Standard for Low-Rate Wireless Networks, IEEE Computer Society Std., 2015, accessed on 2017-09-27. [Online]. Available: http://ieeexplore.ieee.org/iel7/6677511/6677512/06677513.pdf

[2] IEEE Std 802.11, Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications, IEEE Computer Society Std., 2016, accessed on 2017-09-27. [Online]. Available: http://ieeexplore.ieee.org/iel7/6837412/6837413/06837414.pdf

[3] A. Al-Fuqaha, M. Guizani, M. Mohammadi, M. Aledhari, and M. Ayyash, "Internet of Things: A Survey on Enabling Technologies, Protocols, and Applications," IEEE Communications Surveys & Tutorials, vol. 17, no. 4, 2015, pp. 2347–2376.

[4] T. Zillner and S. Strobl, "ZigBee exploited: The good the bad and the ugly," version, Tech. Rep., 2015, accessed on 2017-09-29. [Online]. Available: http://www.sicherheitsforschung-magdeburg.de/uploads/journal/MJS_045_Zillner_ZigBee.pdf

[5] R. Sankar, "mdk3," Sep. 2015, accessed on 2017-09-29. [Online]. Available: http://kalilinuxtutorials.com/mdk3/

[6] A. H. Lashkari, M. M. S. Danesh, and B. Samadi, "A survey on wireless security protocols (WEP, WPA and WPA2/802.11i)," in 2009 2nd IEEE International Conference on Computer Science and Information Technology, Aug. 2009, pp. 48–52.

[7] J. D. Day and H. Zimmermann, "The OSI reference model," Proceedings of the IEEE, vol. 71, no. 12, Dec. 1983, pp. 1334–1340.

[8] Mister X, "aircrack-ng: WiFi security auditing tools suite," Jul. 2017, accessed on 2017-09-27. [Online]. Available: https://github.com/aircrack-ng/aircrack-ng

[9] River Loop Security, "killerbee: IEEE 802.15.4/ZigBee Security Research Toolkit," Jul. 2017, accessed on 2017-09-27. [Online]. Available: https://github.com/riverloopsec/killerbee

[10] W. C. Suski II, M. A. Temple, M. J. Mendenhall, and R. F. Mills, "Using spectral fingerprints to improve wireless network security," in Global Telecommunications Conference, 2008. IEEE, 2008, pp. 1–5.

[11] B. W. Ramsey, T. D. Stubbs, B. E. Mullins, M. A. Temple, and M. A. Buckner, "Wireless infrastructure protection using low-cost radio frequency fingerprinting receivers," International Journal of Critical Infrastructure Protection, vol. 8, Jan. 2015, pp. 27–39.

[12] Ettus Research, "USRP B200mini-i," accessed on 2017-09-27. [Online]. Available: https://www.ettus.com/product/details/USRP-B200mini-i

[13] Y. LeCun, B. Boser, J. S. Denker, D. Henderson, R. E. Howard, W. Hubbard, and L. D. Jackel, "Backpropagation Applied to Handwritten Zip Code Recognition," Neural Computation, vol. 1, no. 4, Dec. 1989, pp. 541–551.

[14] Y. LeCun and Y. Bengio, “The handbook of brain theory and neural networks,” M. A. Arbib, Ed. Cambridge, MA, USA: MIT Press, 1998, ch. Convolutional Networks for Images, Speech, and Time Series, pp. 255–258.

[15] P. Y. Simard, D. Steinkraus, and J. C. Platt, “Best practices for convolutional neural networks applied to visual document analysis,” in 7th International Conference on Document Analysis and Recognition. Proceedings, vol. 2. Washington, DC, USA: IEEE Computer Society, 2003, pp. 958–963.

[16] D. C. Ciresan, U. Meier, L. M. Gambardella, and J. Schmidhuber, “Convolutional Neural Network Committees for Handwritten Character Classification,” in 2011 International Conference on Document Analysis and Recognition, Sep. 2011, pp. 1135–1139.

[17] S. Lawrence, C. L. Giles, A. C. Tsoi, and A. D. Back, “Face recognition: a convolutional neural-network approach,” IEEE Transactions on Neural Networks, vol. 8, no. 1, Jan. 1997, pp. 98–113.

[18] A. Krizhevsky, I. Sutskever, and G. E. Hinton, “ImageNet Classification with Deep Convolutional Neural Networks,” in Advances in Neural Information Processing Systems 25, F. Pereira, C. J. C. Burges, L. Bottou, and K. Q. Weinberger, Eds. Curran Associates, Inc., 2012, pp. 1097–1105.