On the Practicality of UHF RFID Fingerprinting: How Real is the RFID Tracking Problem?

Davide Zanetti, Pascal Sachs, and Srdjan Capkun

Department of Computer Science,
ETHZ, Zurich, Switzerland
{zanettid,sachsp,capkuns}@inf.ethz.ch

Abstract. In this work, we demonstrate the practicality of people tracking by means of physical-layer fingerprints of RFID tags that they carry. We build a portable low-cost USRP-based RFID fingerprinter and we show, over a set of 210 EPC C1G2 tags, that this fingerprinter enables reliable identification of individual tags from varying distances and across different tag placements (wallet, shopping bag, etc.). We further investigate the use of this setup for clandestine people tracking in an example Shopping Mall scenario and show that in this scenario the mobility traces of people can be reconstructed with a high accuracy.

Keywords: RFID, physical-layer identification, fingerprinting, tracking, privacy.

1 Introduction

Radio Frequency IDentification (RFID) technology has raised a number of privacy concerns in many different applications, especially when considering consumer privacy [17]. A person carrying several tags – attached to various objects like books, passports, medicines, medical devices, and clothes – can be subject to clandestine tracking by any reader in the read range of those tags; it has been shown that the read range of RFID tags can be extended up to 50 m [19]. Even if some objects are only temporarily with a person (e.g., a shopping bag), they will enable tracking of a person's behavior for shorter periods (e.g., during a morning or during a visit to a shopping mall). Other objects, such as wallets, personal bags, and medical devices will be frequently or permanently carried by people, thus allowing people being tracked over wider time periods.

Solutions that prevent a (clandestine) reader to communicate with tags were proposed on a logical level, and typically rely on the use of pseudonyms and access control mechanisms [1, 4, 8, 9, 20, 31]. Although effective on the logical level, these solutions do not prevent physical-layer identification of RFID tags. A number of features have been identified that allow physical-layer identification of RFID tags of different manufacturers, but also of individual RFID tags from the same manufacturer and model [6, 21–23, 27, 28, 34]. So far, physical-layer identification has been demonstrated in laboratory conditions, using high-sampling oscilloscopes and low-noise peripherals. This equipment can be costly and is rather impractical for real world tracking.

In this work, we present a low-cost, USRP-based RFID fingerprinter and show that physical-layer fingerprinting of RFID tags is feasible even with this portable setup. For

S. Fischer-Hübner and N. Hopper (Eds.): PETS 2011, LNCS 6794, pp. 97–116, 2011.

tag identification, we use timing features that rely on the extraction of tags backscatter frequencies [23, 34]. We tested our setup on a tag population composed of 210 EPC class-1 generation-2 (C1G2) RFID tags [11] of 12 different models and 3 manufacturers. EPC C1G2 tags are the *de facto* standard passive UHF tags and the most present in the current market. Our results show that this setup and features enable reliable identification of individual tags from varying distances and across different tag placements (wallet, jacket, shopping bag, backpack). The used feature allows the extraction of $\lfloor 2^{5.4} \rfloor$ RFID tag fingerprints independently of the population size (i.e., this feature results in approx. 5.4 bits of entropy). Since people will typically carry several tags, this will allow the creation of a large number of composite fingerprints, thus enabling, in a number of scenarios, highly precise people tracking (e.g., a set of 5 tags provides approx. 22 bits of entropy).

We investigate the use of our setup for clandestine people tracking in an example Shopping Mall scenario and show that in this scenario the mobility traces of people can be reconstructed with a high accuracy.

Although solutions that prevent a (clandestine) reader to communicate with tags at the physical layer exist (e.g., tag kill and sleep functions, Faraday cages, active jammers, and "clipped" tags [18]), the provided privacy comes at the price of tag functionality (e.g., the kill function permanently disables tags and therefore possible after-sales services or long-term deployments) or requires additional efforts (e.g., user interaction or extra hardware) that could make those solutions impractical and unattractive.

Therefore, the proposed setup and feature break people's privacy by enabling the tracking and mobility trace reconstruction of people carrying RFID tags. This privacy breach occurs disregarding of the RFID tag content (e.g., serial number) and with no need for interpreting the information transmitted by the RFID tags (which could be protected, e.g., encrypted, by logical-level mechanisms). People's privacy could be further compromised by means of side-channel information (e.g., a priori knowledge about target people) that builds the associations between tag fingerprints and objects to which they are attached, and between composite fingerprints and people's identities.

The rest of this paper is organized as follows. In Section 2, we define the people tracking scenario and our problem statement. In Section 3, we introduce the considered RFID tag population and physical-layer identification technique. In Section 4, we present our low-cost RFID fingerprinter, while in Section 5 we detail the performed experiments and summarize the collected data. We present the evaluation results in terms of tag distinguishability and fingerprint stability of our fingerprinter in Section 6, while we discuss their implications on tag holders' privacy in Section 7. We make an overview of background and related work in Section 8 and conclude the paper in Section 9.

2 Scenario and Problem Statement

In our study, we consider a scenario in which an attacker aims at tracking people carrying several passive UHF RFID tags over a limited period of time and within a bounded area (e.g., a mall). We assume that the attacker has the ability to position several physical-layer identification devices, i.e., *fingerprinters*, at strategic locations in the considered area. A fingerprinter profiles a person by (i) collecting RF signals from the set of tags assumed to be on a person, (ii) extracting the fingerprints for each tag in the set based on specific RF signal characteristics, or *features*, and finally, (iii) creating a profile, which is the collection of all tag fingerprints for the considered set of tags. The created profiles are then used for people tracking, which can reveal information about people's behavior (e.g., people are likely to visit shop A after they have visited shop B).

A number of works considered the threat of RFID-based tracking real [1, 8, 9, 17, 20]; however, some reservations still remain as to whether tracking is practical or confined only to laboratory environments. In this work we investigate how feasible and practical is RFID-based tracking in real-world scenarios. We consider that tracking will be practical if people's profiles (i.e., RFID fingerprints) can be reliably extracted in dynamic settings (i.e., when tags are on people, in wallets, bags, pockets, and when people are moving), if the fingerprinters can be built as compact, possibly low-cost devices, and if the profiles allow people's traces to be reconstructed with high accuracy. In the rest of the paper we will show that with the proposed fingerprinter setup and with the used features these three conditions are fulfilled.

3 RFID Tags, Signal Features and Tag Fingerprints

In our work, we evaluate the feasibility of people tracking by using our low-cost fingerprinter (Section 4) on a tag population composed of 210 EPC class-1 generation-2 (C1G2) RFID tags [11] of 12 different models and 3 manufacturers. EPC C1G2 tags are the *de facto* standard passive UHF tags and the most present in the current market. Those tags are mainly conceived for item- and pallet-level barcode replacement, which (especially for item-level tagging) makes them pervasive into everyday life.

3.1 EPC C1G2 Background

The communication between RFID readers and tags is half-duplex. A reader transmits commands and data to a tag by modulating an RF signal. The tag replies using a backscattered signal modulated by modifying the reflection coefficient of its antenna. Readers use pulse-interval encoding (PIE) and phase-reversal amplitude shift keying (PR-ASK) modulation to transmit data and commands to tags. Tags backscatter information by modulating an RF signal using ASK and/or PSK modulation and either FM0 baseband or Miller modulation as data encoding. The frequency range of RF signals is defined from 860 to 960 MHz. Readers transmit data at a maximum rate between 40 and 160 kbps. The tag backscatter link frequency (BLF, i.e., the tag data rate) is selected by the readers; the EPC C1G2 specification defines a BLF range between 40 and 640 kHz.

The communication sequence between a reader and a tag during the tag inventorying process with no collisions is shown in Figure 1. The reader challenges the tag with a set of commands to select a particular tag population (*Select*), to initiate an inventory round (*Query*), and to request the transmission of the tag's identification (EPC) number (*Ack*). The tag replies first with an RN16 packet[1] (after the reader's Query) and then with an EPC packet (after the reader's Ack) containing the identification number.

[1] RN16 packets are sent as a part of the anti-collision protocol used during tag inventorying.

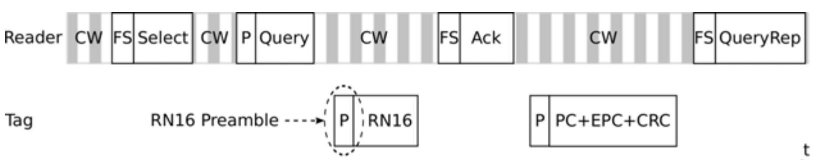


Fig. 1. EPC tag inventory sequence. P, FS, and CW stand for preamble, frame-sync, and continuous wave respectively.

3.2 Signal Features and Tag Fingerprints

Physical-layer device identification relies on random hardware impairments in the analog circuitry components introduced at the manufacturing process. Those impairments then manifest in the transmitted signals making them measurable.

To facilitate the adoption of RFID tags on a large-scale, tag manufacturers tend to optimize both the tag manufacturing process and the size of tag embedded integrated circuits in a effort to reduce the overall tag cost. Although the RFID tag market has been growing in the past years, high-speed processes and low-complexity integrated circuits may increase the possibility of finding tags' internal components affected by hardware impairments, as well as of finding impairments which create measurable and substantial differences between tags.

In our study, we consider random hardware impairments in the tags' local oscillator. According to the EPC C1G2 specification, the backscatter link frequency (BLF) at which tags communicate is defined within a range between 40 and 640 kHz with a frequency tolerance between $\pm4\%$ and $\pm22\%$ depending on the selected BLF. As shown by Periaswamy et al. [23] and Zanetti et al. [34], the relatively large BLF tolerances allowed by the EPC specification can represent a distinguishing factor between different tags of the same model and manufacturer. Additionally, it has been shown [34] that the BLF is not affected by the tag-reader distance and mutual position; this can allow tag distinguishability disregarding tags' location and position. Therefore, the signal feature we consider for tag identification is the backscatter link frequency at which each tag transmits data. We extract this signal feature from the fixed preamble of the RN16 packets sent by tags during tag inventorying. This is done not to introduce any data-dependent bias in our evaluation, since the RN16 preamble is fixed for all tags. Tag fingerprints are built from N acquired RN16 preambles, i.e., a tag fingerprint is a one-dimensional value corresponding to the average BLF over N RN16 preambles collected for a certain tag.

4 Low-Cost RFID Fingerprinter

In our study, we build and deploy a compact and low-cost fingerprinter that challenges tags to initiate an EPC C1G2 inventory round, collects tags' responses, i.e., RN16 packets, and builds tag fingerprints based on the backscatter link frequency (BLF) that it extracts from the RN16 preambles. Our fingerprinter is composed of a Universal Software Radio Peripheral 2 (USRP2) platform and an RFX900 daughterboard by Ettus

Research [2], as well as of a host PC providing signal processing through the GNU Radio toolkit [3]. The block diagram of our low-cost fingerprinter is shown in Figure 6 (Appendix A).

Our fingerprinter consists of a transmitter, a receiver and a feature extraction module. It uses a bistatic antenna configuration to minimize the leakage from the transmitter to the receiver. The chosen antennas are circularly polarized, which allows our fingerprinter to power up (and then communicate with) a tag thus minimizing the impact of the tag orientation. The transmitter outputs commands and data at the baseband frequency according to the pulse-interval encoding (PIE) and phase-reversal amplitude shift keying (PR-ASK) modulation (as defined in the EPC C1G2 specification [11]). The carrier frequency that is used for upmixing the baseband signal is 866.7 MHz[2] and, after the final amplification stage, the nominal transmission power is 29.5 dBm (including the antenna gain). The receiver is based on a direct-conversion I/Q demodulator[3]. After quadrature downmixing, the tag backscatter baseband signal is first converted into the digital domain with a nominal sampling rate of 10 MS/s (for each of the I and Q channels) and 14-bit resolution, and then low-pass filtered. For each channel, the feature extraction module processes the baseband tag signal to extract the BLF from the RN16 preambles. The extraction is a streaming-like process: the module continuously monitors the incoming signal for RN16 packets. When one is detected, the length of the preamble is measured and the BLF is computed and recorded.

5 Performed Experiments and Collected Data

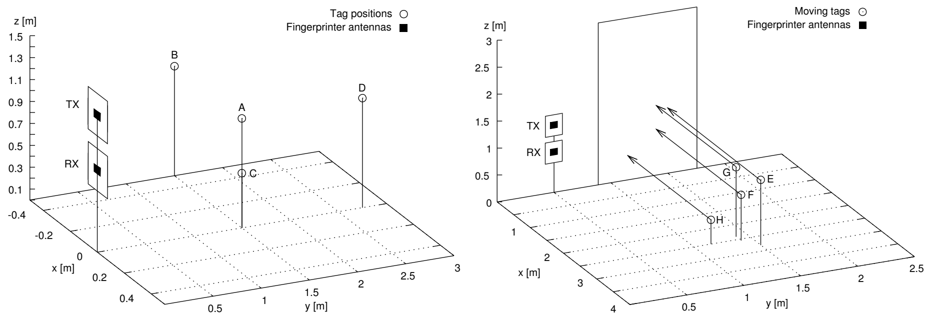
We base our experiments on the interaction between a reader and a tag population that is used for inventorying purposes as defined in the EPC C1G2 specification [11]. We use our fingerprinter to challenge RFID tags (i.e., to initiate an inventory round), collect tags' replies (i.e., RN16 packets), and extract the specified signal feature (i.e., the backscatter link frequency, BLF) to obtain tag fingerprints.

Our tag population is composed of 210 EPC C1G2 RFID tags of 12 different models and 3 manufacturers: Alien Technology ALN9540, ALN9562, ALN9640 and ALN9654, Avery Dennison AD821, AD833, AD224 and AD824, and UPM Raflatac Dogbone (3 different integrated circuit models) and ShortDipole. The selected tag models present different characteristics in terms of antenna size and material, embedded integrated circuit, and application. Table 5 (Appendix B) summarizes the considered models and their main characteristics.

In order to increase the possibility of finding the largest distinguishing characteristic, for all experiments we select the BLF which, according to the EPC C1G2 specification, presents the largest allowed frequency tolerance. The selected nominal BLF is thus equal to 426 kHz and presents a maximal allowed frequency tolerance equal to $\pm 22\%$.

[2] The chosen carrier frequency corresponds to channel 6, band 2, of the ETSI EN 302 208 regulations [12], which define 10 channels of 200 KHz @ 2W ERP between 865.6 and 867.6 MHz.

[3] The phase of the tag backscatter signal is not predictable or controllable, as it varies with the distance to the tag; the I/Q demodulator allows the reception of a backscatter signal regardless of the distance to the tag.



(a) (b)

Fig. 2. Considered positions of the fingerprinter antennas and of the tags. In our experiments, fingerprinter antennas (TX and RX) are fixed, while tag responses are acquired (a) from different fixed locations (A-D, Table 1) and (b) when tags are moving (E-H, Table 2).

5.1 Performed Experiments

For all the tags in our population, we use our fingerprinter to initiate an inventory round and extract the BLF while tags are at a fixed location (on a stand). Figure 2(a) shows the considered positions of the fingerprinter transmitting (TX) and receiving (RX) antennas and of the tags (position A). Table 1 – configuration 3 summarizes the fingerprinter and tag settings for this experiment.

For a subset of tags in our population placed on a stand, we use our fingerprinter to extract the BLF under 16 different configurations of tag and antenna positions, acquisition sampling rate, tag temperature, transmission power, and fingerprinter hardware. The different configurations are summarized in Table 1 (configurations 1 to 16). The considered positions of the fingerprinter TX and RX antennas and of the tags are shown in Figure 2(a). In terms of tag position, we explore different tag distances to the fingerprinter antennas (up to 2.75 m), as well as different tag vertical and lateral positions. We also explore 3 different transmission powers (from 17.5 to 23 dBm), 3 different acquisition sampling rates (from 5 to 20 MS/s), and 5 different temperatures (from 10 to 50°C). Additionally, we consider 3 different fingerprinter hardware configurations (changing USRP2 platform, USRP daughterboard, antennas, and host PC) and swap the position of the TX and RX antennas. Finally, we explore time effects by acquiring RN16 preambles and extracting BLF one month after the beginning of this experiment.

For a subset of tags in our population, we use our fingerprinter to extract the BLF while tags are carried by a person. For this experiment, we investigate 6 different configurations of tag location (backpack, wallet, jacket, shopping bag), tag holder's activity (standing, walking), and number of carried tags (from 1 to 5). The fingerprinter is configured as detailed in Table 1 – configurations 17-22, while the different tag configurations are summarized in Table 2. The considered positions of the fingerprinter TX and RX antennas and of the tags are shown in Figure 2(b).

Table 1. Varied parameters for the different configurations - tags placed on a stand

	Tag position		Antennas	TX	Temp.[2]	Sampling	Fingerprinter
Config.	Fig. 2(a)	(x,y,z)-axis	position	power[1]		rate	hardware
		[m]	(TX,RX) [m]	[dBm]	[°C]	[MS/s]	set[3]
1	A	(0, 1.5, 1.0)	(1.25, 0.75)	21	22	5	1
2	‖	‖	‖	‖	‖	20	‖
3	‖	‖	‖	‖	‖	10	‖
4[4]	‖	‖	‖	‖	‖	‖	‖
5	‖	‖	(0.75, 1.25)	‖	‖	‖	‖
6	B	(-0.5, 1.5, 1.0)	(1.25, 0.75)	‖	‖	‖	‖
7	C	(0, 1.5, 0.5)	‖	‖	‖	‖	‖
8	D	(0, 2.75, 1.0)	‖	23	‖	‖	‖
9	A	(0, 1.5, 1.0)	‖	17.5	‖	‖	‖
10	‖	‖	‖	23	‖	‖	‖
11	‖	‖	‖	21	‖	‖	2
12	‖	‖	‖	‖	‖	‖	3
13	‖	‖	‖	‖	10	‖	1
14	‖	‖	‖	‖	30	‖	‖
15	‖	‖	‖	‖	40	‖	‖
16	‖	‖	‖	‖	50	‖	‖
17-22	Tag on a person, see Table 2		‖	23	22	‖	‖

[1] Power before the TX antenna. For fingerprinter sets 1 and 3, the TX antenna has a gain of 8.5 dBi, while for set 2 this is equal to 6 dBi.

[2] Temperature variations of ±2°C.

[3] Set 2: same host PC as set 1, but different USRP, USRP daughterboard and antennas. Set 3: same USRP, USRP daughterboard and antennas as set 1, but different host PC.

[4] Same as configuration 3, but fingerprints obtained from RN16 preambles collected 1 month after the RN16 preambles collected for configuration 3.

Table 2. Varied parameters for the different configurations - tags on a person

Configuration	Tag location		Tag holder's	# of tags during
	Fig. 2(b)		activity	acquisition
17	E	Backpack	walking away from TX/RX antennas	1
18	F	Wallet	‖	‖
19	G	Jacket	walking towards TX/RX antennas	‖
20	H	Shopping bag	‖	‖
21	‖	‖	standing in front of TX/RX antennas	5
22	‖	‖	walking towards TX/RX antennas	‖

5.2 Collected Data

Using our fingerprinter, we performed the experiments described in Section 5.1. Table 3 summarizes the data that we collected, represented in a form of datasets.

Data collection was performed over one month, one tag at the time (unless otherwise indicated, i.e., for data collection under configurations 21 and 22 – Table 2), 200 extracted BLFs in a row, in an indoor, RF noisy environment with active Wi-Fi and GSM networks. The nominal environment temperature was approx. 22°C. We increased the tag temperature by means of a heat gun, while we lowered it by decreasing the overall environment temperature. Temperatures were measured with an infrared thermometer[4]. We note a ±2°C variations for the given temperatures. We sped up the acquisition process by adjusting the aforementioned EPC inventory sequence (Figure 1) in a way to collect several RN16 packets in the same inventory round and by not requesting the tag's identification (EPC) number[5]. Giving the considered acquisition sequence, the theoretical upper bound for BLF acquisition is approx. 1250 extracted BLFs per second (we discuss the fingerprinter acquisition speed in Section 7.3).

6 Evaluation of Tag Distinguishability and Fingerprint Stability

In this section, we first review the metrics that we used to evaluate the tag distinguishability and the fingerprint stability. Then, we present the results for those evaluations obtained by the proposed signal feature over the considered tag population.

6.1 Evaluation Metrics

To evaluate the tag distinguishability and the fingerprint stability, we compute the entropy of the probability distribution of the tag fingerprints given the selected signal feature. For each tag and configuration, fingerprints are built from N extracted BLFs. Table 4 summarizes the computed entropies for the different analysis we performed.

We compute the entropy of the fingerprint probability distribution in order to show how many bits of information are contained within that distribution. To compute the entropy, we consider bins of width equal to the double of the average standard deviation of the signal feature in the dataset and count the number of fingerprints that fall into the different bins. We then apply the standard entropy formula [29].

Additionally, for each performed analysis, we define an entropy upper bound[6] by computing its theoretical maximum given the EPC C1G2 specification [11], i.e., the maximum number of information bits that could be learned from the BLF feature considering the maximal allowed frequency tolerance as defined in the EPC specification (±22% around the nominal BLF) and giving the bin width of the considered analysis.

[4] Temperature was measured on the tag front surface. Tags were heated up from the back surface and, for each considered temperature, for at least 5 minutes before data acquisition.

[5] This procedure is also valid for multiple-tag acquisitions. For each tag, several RN16 packets are collected before moving to the next tag. This also provides the association between extracted BLFs and tags.

[6] The entropy upper bound is computed by assuming the fingerprint distribution as uniform [13].

Table 3. Collected data

Dataset	Model	# tags	# extracted BLFs per tag	Conf. (Table 1)	Total # extracted BLFs per tag
1	ALN9640	100	200	3	200
2[1]	ALN{9540, 9562, 9640[2], 9654}	40	200	3	200
	AD{224, 821, 824, 833}	‖	‖	‖	‖
	ShortDipole, Dogbone[3]	‖	‖	‖	‖
3	ALN9640[2]	10	200	3-22	4000
4	ALN9640	100	200	1,2	400

[1] For each model, 10 tags are considered.

[2] Tags randomly selected among the 100 used in datasets 1 and 4.

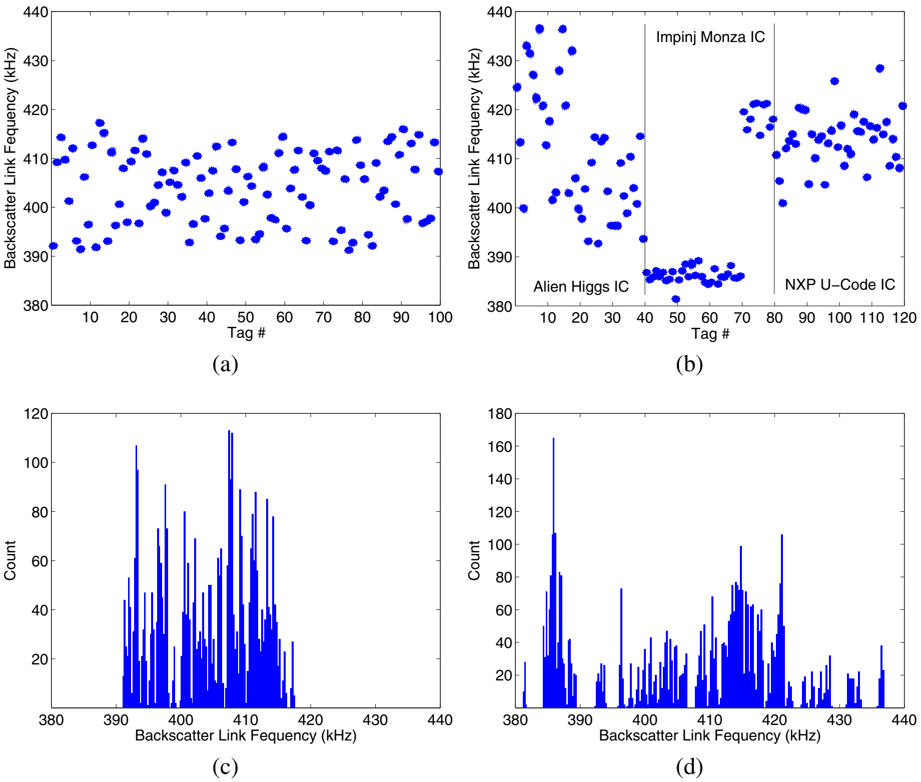
[3] For Dogbone tags, 3 different integrated circuit models are considered.

6.2 Tag Distinguishability

In this section, we analyze the tag distinguishability of the proposed feature based on the fingerprint probability distribution of two datasets: dataset 1, which contains 20,000 extracted BLFs for 100 same-model (and same-manufacturer) tags, and dataset 2, which contains 24,000 extracted BLFs for 120 tags of 12 different models.

Figure 3(a) and 3(b) show the computed fingerprints for the 100 same-model and the 120 different-model tags respectively. Each fingerprint is obtained by averaging 5 extracted BLFs ($N = 5$), resulting in 40 fingerprints per tag. Tag distinguishability depends only on the variations of the BLF within each tag and between different tags. For both sets of tags, we can observe a certain degree of distinguishability. First, the fingerprint variations within each tag are relatively small (average standard deviation of approx. 120 and 196 Hz for the 100 same-model and the 120 different-model tags respectively). Second, fingerprints of different tags are located in different frequency areas. However, we note that (i) fingerprints of different tags also overlap (i.e., different tags present a similar BLF), which reduces the possibility, or even prevent to distinguish those tags, and (ii) that the overall frequency range is less than the maximal frequency range allowed by the EPC C1G2 specification (between 332 and 520 kHz given the $\pm 22\%$ tolerance around the nominal BLF), which indicates that the actual fingerprint entropy will not correspond to its potential upper bound. Additionally, we note that different tag models could also be distinguished, in particular when considering tags embedding Impinj Monza IC.

Figure 3(c) and 3(d) show the empirical fingerprint distributions for the 100 same-model and the 120 different-model tags respectively. The entropy result based on the empirical distribution of 120 different-model tags suggests that we could learn 6.78 bits of information about a single UHF RFID tag. For the 100 same-model tags, this value is equal to 6.32 bits. The difference between these two results simply lies in the larger frequency range exploited by several models with respect to one single model. The entropy upper bound considering the maximal allowed BLF tolerance is, for same-model tags, equal to 9.45 bits and, for different-model tags, to 9.38 bits.



(a)

(b)

(c)

(d)

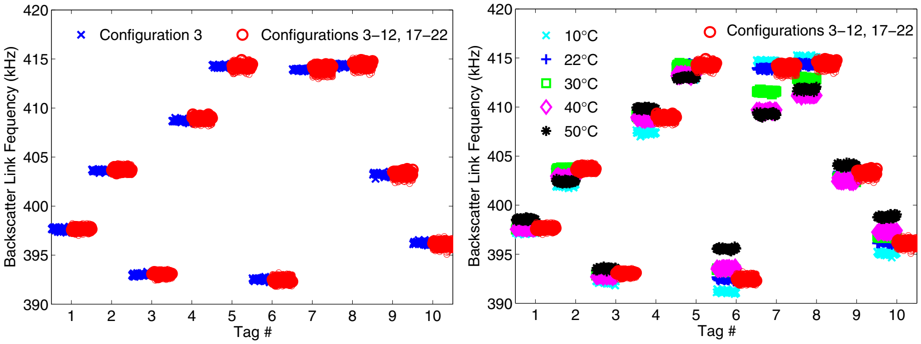
Fig. 3. Fingerprints for (a) 100 same-model tags and (b) 120 tags of 12 different models. Fingerprint distribution for (c) 100 same-model tags and (d) 120 tags of 12 different models. For each tag, 40 fingerprints are considered ($N = 5$).

We evaluate the impact of the number N of extracted BLFs over which we average to obtain the tag fingerprints by computing the entropy based on the empirical distribution of the 100 same-model tags obtained for different values of N. The results of the analysis for $N = 1, 2, 5, 10, 20$ are 5.39, 5.81, 6.32, 6.67, and 6.97 bits respectively.

6.3 Fingerprint Stability

In the previous section, we have analyzed the tag distinguishability under a fixed configuration of fingerprinter and tag settings. In this section, we evaluate the stability of the proposed signal feature under different settings, i.e., we analyze the impact of different settings on the tag distinguishability. More specifically, we evaluate:

1. The entropy of the proposed feature under 16 different configurations of tag position (with respect to the fingerprinter antennas) and location (on a stand, on a person), antenna position, transmission power, fingerprinter hardware, and, when tags are carried by a person, tag holder's activity (walking, standing) and the number of carried tags (Table 1 – configurations 3-12 and 17-22, and Table 2).



(a) (b)

Fig. 4. Fingerprint visualization for 10 randomly selected ALN9640 tags and different settings ($N = 5$). For each tag in (a), the set of fingerprints on the left is composed of 40 fingerprints of 1 fixed configuration, while the set on the right of 640 fingerprints of 16 different configurations. For each tag in (b), the set of fingerprints on the left is composed of 200 fingerprints of 5 different temperatures, while the set on the right of 640 fingerprints of 16 different configurations.

2. The entropy of the proposed feature given different acquisition sampling rates (Table 1 – configurations 1-3).
3. The effect of temperature on tag fingerprints (Table 1 – configurations 3, 13-16).

Figure 4(a) shows the fingerprints of the selected 10 tags under 16 different configurations of fingerprinter and tag settings ($N = 5$, 40 fingerprints for each tag and configuration). For each tag, two sets of fingerprints are shown: 40 fingerprints (the set on the left) obtained under one single configuration (Table 1 – configuration 3) and 640 fingerprints (the set on the right) obtained under 16 different configurations of fingerprinter and tag settings (Table 1 – configurations 3-12 and 17-22). We observe an increase on the BLF variation within each tag when comparing those two sets: the average standard deviation within each tag increases from approx. 120 to 150 Hz. Although this increase (less than 30 Hz) seems relatively small when compared to the considered frequency range (approx. 30 kHz for the 100 same-model tags), the entropy for the 100 same-model tag decreases from 6.32 (Section 6.2) to 5.39 bits[7]. Similarly, the entropy upper bound decreases from 9.45 to 8.41 bits.

In order to evaluate the impact of the acquisition sampling rate, we compute the entropy based on the empirical distribution of the 100 same-model tags obtained for RN16 preambles acquired at different rates. The results of the analysis for 5, 10, and 20 MS/s are 6.19, 6.32, and 6.49 bits respectively.

Figure 4(b) shows the fingerprints of the selected 10 tags under 20 different configurations ($N = 5$, 40 fingerprints for each tag and configuration). For each tag, two sets of fingerprints are shown: 200 fingerprints (the set on the left) obtained under 5 different temperatures (Table 1 – configurations 3, 13-16) and 640 fingerprints (the set

[7] We compute this entropy over dataset 1 (100 tags, 1 configuration), but considering the standard deviation under the stability analysis of dataset 3 (10 tags, 16 configurations), i.e., 150 Hz. This allows us to compare entropies and evaluate the effect of different configurations.

Table 4. Computed entropies (with 95% confidence interval) for the performed analysis

Dataset	Sampling rate [MS/s]	N	Config. (Table 1)	Standard deviation [Hz]	Entropy (empirical dist.) [bits]	Entropy (upper bound) [bits]
1	10	1	3	273.32 (270.14;275.99)	5.39 (5.38;5.42)	8.27 (8.25;8.29)
〃	〃	2	〃	192.63 (189.89;195.19)	5.81 (5.78;5.83)	8.77 (8.75;8.79)
〃	〃	5	〃	120.21 (117.05;123.31)	6.32 (6.29;6.35)	9.45 (9.42;9.49)
〃	〃	10	〃	83.45 (81.14;86.02)	6.67 (6.62;6.71)	9.97 (9.94;10.02)
〃	〃	20	〃	56.58 (54.06;58.99)	6.97 (6.91;7.02)	10.54 (10.48;10.60)
2	10	5	3	196.05 (180.38;211.80)	6.78 (6.75;6.80)	9.38 (9.35;9.41)
3	10	5	3-12,17-22	149.57 (140.42;159.72)	5.39[1] (5.37;5.42)	8.41[1] (8.41;8.41)
4	5	5	1	134.12 (129.78;138.40)	6.19 (6.14;6.24)	9.29 (9.24;9.34)
〃	20	〃	2	109.35 (106.44;112.65)	6.49 (6.45;6.52)	9.59 (9.55;9.63)

[1] Computed for dataset 1 (100 tags) given the standard deviation of dataset 3 (10 tags).

on the right) obtained under 16 different configurations of fingerprinter and tag settings (Table 1 – configurations 3-12 and 17-22). Differently from the previous results, temperature seems to have a relatively large impact on the BLF variation within each tag, especially when considering the limit temperatures in our analysis (10 and 50°C). We note that tags are not equally affected by temperature and that we could not observe any common trend (i.e., a relation between temperature and BLF variation) that would facilitate the mitigation of the temperature effect on tag fingerprints.

7 Implications on Tag Holders' Privacy

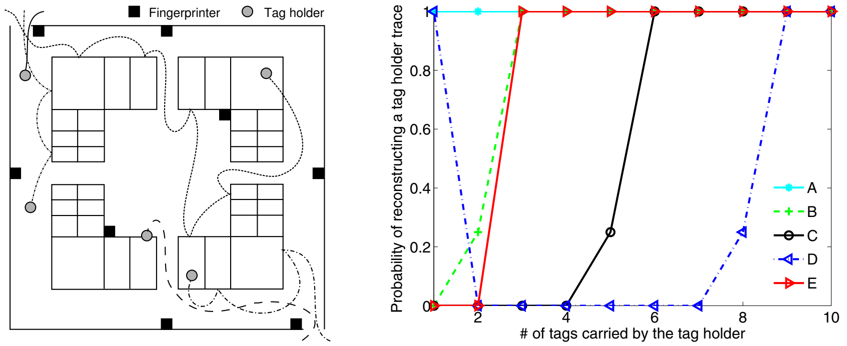
In this section, we first discuss the implications on people's privacy given the obtained results, in particular with respect to people tracking. Then, we discuss possible countermeasures against clandestine tracking and fingerprinter requirements for practical tracking.

7.1 People Tracking: Breaking Tag Holders' Privacy

The results of our work show that we can learn 5.39 bits of information about a single RFID tag by only observing the data rate at which it transmits[8]. This information can be extracted independently of the tag position and location, fingerprinter hardware and antennas position, transmission power, tag holder's activity, and number of carried tags.

The relatively low distinguishability (per tag) can be improved when considering sets of tags. Our fingerprinter extracts $b = 5.39$ bits of information for each tag, i.e, when individually considered, a maximum of $n = \lfloor 2^b \rfloor$ tags can be uniquely distinguished.

[8] The amount of information could be further increased by considering sets of tags composed of different tag models and manufacturers, an higher acquisition sampling rate, and a larger number of acquired signals over which the tag fingerprints are obtained.



(a) (b)

Fig. 5. (a) A possible shopping mall scenario and (b) the upper bound probability of reconstructing a tag holder's trace as a function of the number of tags carried by that tag holder. Curve A represents a population size of $P = 3000$, where $p_T = \mathcal{N}(5, 1)$, the tag entropy $b = 5.39$, and each tag holder has been profiled once, i.e., $E_H = 1$. Curves B, C, D, and E are similar to A, but they consider $p_T = \mathcal{N}(2, 1)$, $P = 5,000,000$, $b = 1$, and $E_H = 10$ respectively.

As a consequence, a set S composed of T tags can be uniquely distinguished among other $S_T = \binom{n+T-1}{T} = \frac{(n+T-1)!}{T!(n-1)!}$ sets. For example, a set composed of 5 tags can be uniquely distinguished among other 1.2 million sets of 5 tags. Larger sets provide more information (for $T = 5$, approx. 22 bits) and lead to a larger distinguishability of people carrying several tags, even with relatively low distinguishability per tag.

To show the impact of our technique on tag holders' privacy, we evaluate the probability that the attacker can correctly reconstruct a customer's path in a shopping mall. Reconstructed paths, or traces, can be used to derive customers' behavior and trend and, ultimately, to optimize the location of shops and facilities in the mall.

We consider a scenario in which several fingerprinters are disseminated in a shopping mall (Figure 5(a)). Tag holders, i.e., customers carrying tags, are subject to profiling when passing near the fingerprinters. Each profile is composed of the profiling time and location, and of the set of fingerprints obtained from the carried tags. A tag holder's trace is composed of all the profiles built by the disseminated fingerprinters that relate to that tag holder over a period of interest. We note that the number of tags carried by a customer may increase over time, i.e., the more he/she buys, the more tags he/she carries. Considering this scenario, we evaluate the probability of entirely reconstructing a tag holder's trace given all profiles built over the period of interest. We define as P the size of the customer population which has been profiled over the considered period.

The anonymity set $k_{S,T}$ represents how many tag holders within a population of size P carry the same set S of T tags (fingerprints). $k_{S,T}$ depends on the population size P, the distribution p_T of the number of carried tags per customer within P, the number of carried tag T, the distribution p_S of the possible tag sets, and the tag entropy b. An anonymity set $k_{S,T} = 2$ means that each profile referring to a specific set S of T tags could be potentially related to 2 different tag holders. It is possible to derive the minimal population size in order to find at least 2 customers carrying the same set S of T tags. For example, giving $p_T = \mathcal{N}(5, 1)$, $p_S = \mathcal{U}(1, S_T)$, and $b = 5.39$ bits, the minimal population size necessary to find at least 2 customers carrying the same set S of T tags is 149,480, 3.2 million, and 66 billion for $T = 2, 5, 8$ tags respectively.

For a tag holder carrying a set S of T tags and having an anonymity set of $k_{S,T}$, the probability p_R of reconstructing that tag holder's trace is computed as $(k_{S,T})^{-E}$, where E is the total number of profiles referring to the considered set of tags S (i.e., all the profiles built for all the customers carrying that set S). Figure 5(b) shows the upper bound probability[9] $\overline{p_R}$ of reconstructing a tag holder's trace as a function of the number of tags T carried by that tag holder (curve A) and for a different distribution of the number of carried tags p_T (curve B), population size P (curve C), tag entropy b (curve D), and number of profiles built for each tag holder in the considered population E_H (curve E). Since p_R is derived from the anonymity set, this is affected by the tag entropy, the population size, the distribution of the number of carried tags within that population, and the number of carried tags by the consider tag holder. In general, for the same b, P, p_T, and E_H, increasing the number of carried tags T increases p_R: the more shopping, the less anonymity[10]. Differently, increasing the population size, decreasing the tag entropy, or having a population with a smaller number of carried tags per customer increases the anonymity set and therefore reduces p_R. Additionally, p_R is also affected by the total number of profiles built for all the customers carrying the same set of tags: the more profiles, the larger the number of possible profile combinations that a certain tag holder's trace could match, and therefore, the less p_R. Finally, we note that p_R could be increased by considering information like spatial and temporal correlation of profiles.

Therefore, our fingerprinter and selected signal feature allow, in fact, people profiling and clandestine tracking. Temperature effects on tag fingerprints can be neglected when tags maintain a similar temperature over the different profilings, for example, like in a shopping mall where temperature control is used.

7.2 Countermeasures: How to Preserve Tag Holders' Privacy

Countermeasures against physical-layer identification can be categorized into solutions that prevent tag-reader communication or that prevent physical-layer identification.

Tag kill and sleep functions, Faraday cages, and active jammers [17] are solutions that prevent any reader[11] to communicate with a tag, thus eliminating any possible physical-layer identification. Permanently killing tags will guarantee privacy, but at the price of tag functionality. Sleep functions and active jammers will preserve long-term tag functionality, but the required additional measures in order to guarantee privacy (e.g., user interaction, tag access control, or extra hardware) could make those solutions unattractive (especially given the deployment model of RFID tags, in particular when considering item-level tagging). Faraday cages are the most simple and effective

[9] The upper bound probability is computed by assuming p_S as uniform.

[10] Exceptions can occur depending on the size of the group of all customers carrying T tags and the entropy b. As shown in Figure 5(b) - curve D, $\overline{p_R}$ decreases when increasing T from 1 to 2, since the small size of the group of all customers carrying 1 tag allows to reconstruct all traces, while the bigger size of the group of all customers carrying 2 tags provides some anonymity.

[11] Preventing only clandestine readers will not provide any benefit, since the communication between a tag and a legitimate reader can be easily eavesdropped.

solutions to guarantee privacy by temporarily preventing tag-reader communication, but, although shielded wallets and shopping bags could be easily deployed, other RFID-enabled devices (e.g., medical devices) may require additional efforts that could make those solutions impractical.

Solutions that prevent physical-layer identification aim at removing or reducing the effect of the random hardware impairments in the analog circuitry components introduced at the manufacturing process that make physical-layer identification possible. Although very effective, those solutions require first the (possibly hard) task to identify the components that make devices identifiable, and then to adjust the manufacturing process accordingly, which may introduce additional costs that could make those solutions unattractive. In addition, such solutions do not guarantee that a new discriminant feature will never be exploited in future.

Achieving effective and practical countermeasures against unauthorized physical-layer identifications remains an open issue that needs to be addressed.

7.3 RFID Fingerprinter Requirements

Besides tag distinguishability, requirements for a practical use of an RFID fingerprinter for people tracking include acquisition speed, system cost, read range, and size.

Giving the acquisition sequence as detailed in Section 5.2 and the selected EPC C1G2 settings (nominal BLF equal to 426 kHz and 4-subcarrier Miller encoding [11]), the theoretical upper bound for the BLF acquisition speed is approx. 1250 BLFs per second. Besides the well-known factors affecting the tag read rate like tag position, orientation, surrounding material, etc., the communication and computation capabilities of our fingerprinter also influence the actual acquisition speed. If for a sampling rate of 5 MS/s the acquisition speed is close to the theoretical upper bound (approx. 1220 BLF/s), for higher sampling rates the larger amount of data to transmit and process reduces the actual acquisition speed. For 10 and 20 MS/s, the acquisition speed is reduced to approx. 390 and 75 BLF/s respectively[12]. We note that, since tags share the same medium, the EPC C1G2 specification provides a medium access control mechanism to limit tag collisions, which, in fact, reduces the overall acquisition speed. Although for 10 MS/s and 5 tags we find a relatively low acquisition speed equal to approx. 85 BLF/s, this was enough to acquire the necessary tag signals in all our experiments.

The system cost relates to the quality of the obtained fingerprints and the acquisition speed. With our fingerprinter, we were able to obtain reliable fingerprints for people tracking at a relatively low-cost: the overall cost of our fingerprinter (USRP2, USRP daughterboard, host PC, and antennas) is less than USD3200.

During our experiments, we tested tag-reader distances of up to 2.75 m. Although we did not evaluate larger distances (for this, an external amplifier increasing the fingerprinter transmission power would have been necessary), given the exploited signal feature and the obtained results, we can extend the tag distinguishability range to the actual tag read range (which can reach up to 50 m [19]).

[12] Those values could be increased by tuning some of the EPC C1G2 settings (e.g., by increasing the nominal BLF or using FM0 as data encoding scheme) and by optimizing the fingerprinter blocks having the highest demand of computational power (e.g., the signal filtering processes).

In terms of size, our fingerprinter fits in a briefcase: the USRP2 platform has sizes 21x17x5 cm, while a laptop can be used as host PC. We deployed planar antennas of sizes 37x37x4 cm (smaller could be used), which can be easily hidden in wall panels.

8 Related Work

Physical-layer fingerprinting (identification) of UHF RFID tags has been investigated in several works [21–23, 34]. Periaswamy et al. [22] studied physical-layer identification of UHF RFID tags as a mechanism to detect counterfeit tags. The authors used the tag minimum power response measured at multiple frequencies as discriminant feature. The authors considered a set of 100 tags from 2 manufacturers and collected tag signals with a middle/high-range acquisition setup in a clean environment (anechoic chamber). The results showed that same-model tags can be distinguished, but fingerprint stability was not considered. The same authors also proposed a method to enable ownership transfer of UHF RFID tags based on the same discriminant feature [21]. Timing characteristics (packet length) of the tag-to-reader communication are used by Periaswamy, Thompson and Romero [23] to identify (classify) UHF RFID tag. The authors considered a set of 30 tags from 3 manufacturers and collected tag signals with a high-range acquisition setup in a noisy environment (lab room). Results showed that tags can be correctly classified, depending on the considered model, with an accuracy between approx. 32 and 98%. Fingerprint stability was not considered. Zanetti et al. [34] studied physical-layer identification of UHF RFID tags using timing and spectral characteristics of tag signals. The authors considered a set of 70 tags from 3 manufacturers and collected tag signals with a high-range acquisition setup in a noisy environment (lab room). The results showed the existence of stable physical-layer fingerprints for distinguishing UHF RFID tags. The authors also evaluated the implications of the proposed fingerprinting techniques on users' privacy and as cloning detection mechanism.

In comparison to the above works, our work is the first to evaluate the practicality of UHF RFID fingerprinting for people tracking. More specifically, we deployed low-cost fingerprinters to challenge tags, collect tags' responses, and build fingerprints in a tracking-like scenario, i.e., in which tags are carried by people moving into a bounded area. In our study, we considered a larger tag population of 210 tags of 12 models and 3 manufacturers and a more complete fingerprint stability evaluation.

Besides the mentioned works on UHF RFID tags, physical-layer fingerprinting has been explored on different platforms such as VHF [10, 30, 32], Bluetooth [15], IEEE 802.11 [5, 14, 16, 33], IEEE 802.15.4 (ZigBee) [7, 24], and GSM [25, 26]. Physical-layer identification has also been considered for inductive coupled HF RFID devices [6, 27, 28], especially for detecting cloned or counterfeit HF RFID smart cards and electronic passports. The results showed that the proposed techniques enable identification of same model and manufacturer HF RFID devices, but at a very close proximity.

9 Conclusion

In this work, we investigated the practicality of people tracking by means of physical-layer fingerprints of RFID tags that they carry. We have constructed a compact

USRP-based RFID fingerprinter and have shown that using this fingerprinter people's RFID profiles (i.e., RFID fingerprints) can be reliably extracted in dynamic settings (i.e., when tags are on people, in wallets, bags, pockets, and when people are moving). We have further shown, in a representative mall scenario, that these profiles allow people's traces to be reconstructed with high accuracy. Effective and practical countermeasures against unauthorized physical-layer fingerprinting remain an open problem.

References

1. http://www.avoine.net/rfid/index.html
2. http://www.ettus.com/
3. http://www.gnu.org/software/gnuradio/
4. Berbain, C., Billet, O., Etrog, J., Gilbert, H.: An efficient forward private RFID protocol. In: Proc. ACM Conference on Computer and Communications Security, pp. 43–53 (2009)
5. Brik, V., Banerjee, S., Gruteser, M., Oh, S.: Wireless device identification with radiometric signatures. In: Proc. ACM International Conference on Mobile Computing and Networking (2008)
6. Danev, B., Heydt-Benjamin, T.S., Čapkun, S.: Physical-layer identification of RFID devices. In: Proc. USENIX Security Symposium (2009)
7. Danev, B., Čapkun, S.: Transient-based identification of wireless sensor nodes. In: Proc. ACM/IEEE Conference on Information Processing in Sensor Networks (2009)
8. Dimitriou, T.: A lightweight RFID protocol to protect against traceability and cloning attacks. In: Proc. International ICST Conference on Security and Privacy in Communication Networks (2005)
9. Duc, D.N., Park, J., Lee, H., Kim, K.: Enhancing security of EPCglobal Gen-2 RFID tag against traceability and cloning. In: Proc. Symposium on Cryptography and Information Security (2006)
10. Ellis, K., Serinken, N.: Characteristics of radio transmitter fingerprints. Radio Science 36, 585–597 (2001)
11. EPCglobal: UHF Class 1 Gen 2 Standard v. 1.2.0. Standard (2008)
12. ETSI: ETSI EN 302 208-1 (2006)
13. Guiasu, S., Shenitzer, A.: The principle of maximum entropy. The Mathematical Intelligencer 7, 42–48 (1985)
14. Hall, J., Barbeau, M., Kranakis, E.: Enhancing intrusion detection in wireless networks using radio frequency fingerprinting. In: Proc. Communications, Internet, and Information Technology (2004)
15. Hall, J., Barbeau, M., Kranakis, E.: Detecting rogue devices in Bluetooth networks using radio frequency fingerprinting. In: Proc. IASTED International Conference on Communications and Computer Networks (2006)
16. Jana, S., Kasera, S.K.: On fast and accurate detection of unauthorized wireless access points using clock skews. In: Proc. ACM International Conference on Mobile Computing and Networking (2008)
17. Juels, A.: RFID security and privacy: A research survey. IEEE Journal on Selected Areas in Communications 24(2) (2006)
18. Karjoth, G., Moskowitz, P.A.: Disabling RFID tags with visible confirmation: clipped tags are silenced. In: Proc. ACM Workshop on Privacy in the Electronic Society (2005)
19. Koscher, K., Juels, A., Kohno, T., Brajkovic, V.: EPC RFID tag security weaknesses and defenses: Passport cards, enhanced drivers licenses, and beyond. In: Proc. ACM Conference on Computer and Communications Security (2009)

20. Lee, Y.K., Batina, L., Singelée, D., Verbauwhede, I.: Low-cost untraceable authentication protocols for RFID. In: Proc. ACM Conference on Wireless Network Security (2010)
21. Periaswamy, S.C.G., Thompson, D.R., Di, J.: Ownership transfer of RFID tags based on electronic fingerprint. In: Proc. International Conference on Security and Management (2008)
22. Periaswamy, S.C.G., Thompson, D.R., Di, J.: Fingerprinting RFID tags. IEEE Transactions on Dependable and Secure Computing PrePrints (99) (2010)
23. Periaswamy, S.C.G., Thompson, D.R., Romero, H.P., Di, J.: Fingerprinting radio frequency identification tags using timing characteristics. In: Proc. Workshop on RFID Security - RFIDsec Asia (2010)
24. Rasmussen, K., Čapkun, S.: Implications of radio fingerprinting on the security of sensor networks. In: Proc. International ICST Conference on Security and Privacy in Communication Networks (2007)
25. Reising, D.R., Temple, M.A., Mendenhall, M.J.: Improved wireless security for GMSK-based devices using RF fingerprinting. International Journal of Electronic Security and Digital Forensics 3, 41–59 (2010)
26. Reising, D.R., Temple, M.A., Mendenhall, M.J.: Improving intra-cellular security using air monitoring with RF fingerprints. In: Proc. IEEE Wireless Communications and Networking Conference (2010)
27. Romero, H.P., Remley, K.A., Williams, D.F., Wang, C.M.: Electromagnetic measurements for counterfeit detection of radio frequency identification cards. IEEE Transactions on Microwave Theory and Techniques 57(5), 1383–1387 (2009)
28. Romero, H.P., Remley, K.A., Williams, D.F., Wang, C.M., Brown, T.X.: Identifying RF identification cards from measurements of resonance and carrier harmonics. IEEE Transactions on Microwave Theory and Techniques 58(7), 1758–1765 (2010)
29. Shannon, C.: A mathematical theory of communication. The Bell System Technical Journal 27, 379–423 (1948)
30. Shaw, D., Kinsner, W.: Multifractal modeling of radio transmitter transients for classification. In: Proc. IEEE Conference on Communications, Power and Computing (1997)
31. Spiekermann, S., Evdokimov, S.: Privacy enhancing technologies for RFID - A critical investigation of state of the art research. In: Proc. IEEE Privacy and Security (2009)
32. Ureten, O., Serinken, N.: Detection of radio transmitter turn-on transients. Electronic Letters 35, 1996–1997 (2007)
33. Ureten, O., Serinken, N.: Wireless security through RF fingerprinting. Canadian Journal of Electrical and Computer Engineering 32(1) (Winder 2007)
34. Zanetti, D., Danev, B., Čapkun, S.: Physical-layer identification of UHF RFID tags. In: Proc. ACM Conference on Mobile Computing and Networking (2010)

Appendix A: Low-Cost Fingerprinter Block Diagram

The block diagram of our low-cost fingerprinter is shown in Figure 6.

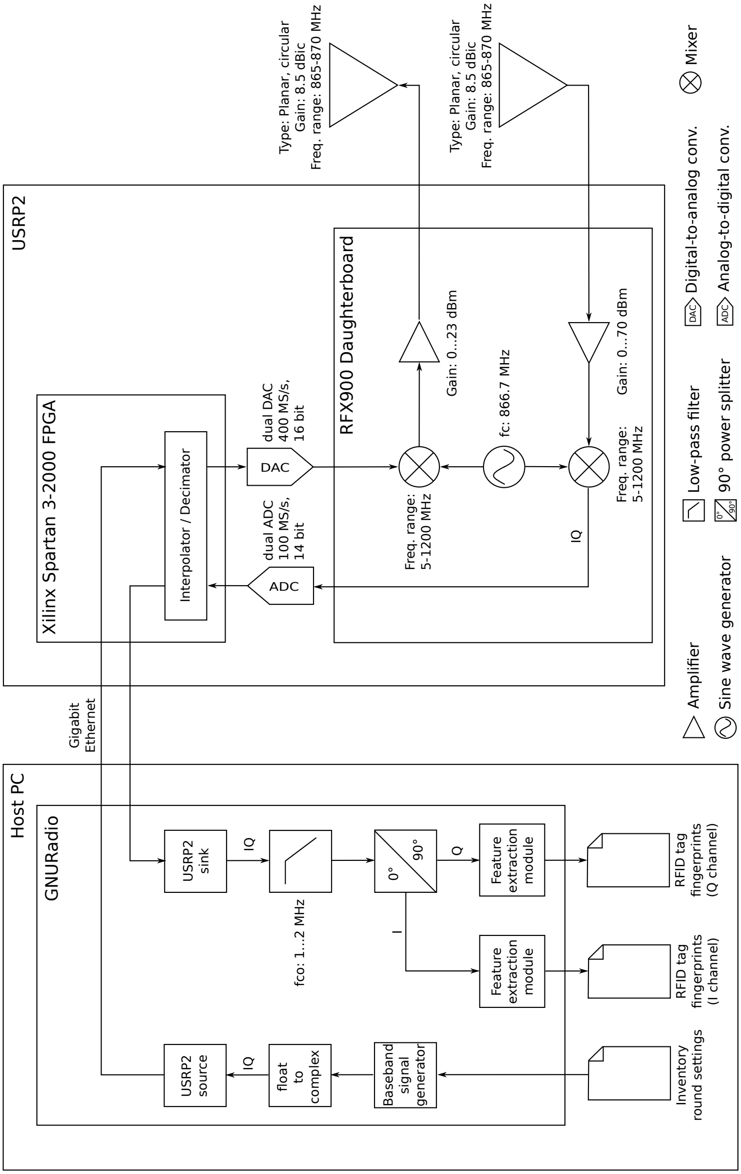


Fig. 6. Block diagram of our low-cost fingerprinter

Appendix B: Considered Tag Models

In our study, we consider a tag population composed of 210 EPC C1G2 RFID tags of 12 different models and 3 manufacturers. Table 5 summarizes the considered models and their main characteristics.

Table 5. Considered tag models and their main characteristics

Model	Manufacturer	IC	IC characteristics	Antenna size [mm]	Antenna material	Application (tagging)
ALN9540	Alien Technology	Alien Higgs-2	96-bit EPC num.	94.8 x 8.1	Cu	Cartoon, pallet
ALN9562	‖	‖	‖	70 x 19	‖	‖
ALN9640	‖	Alien Higgs-3	96/480-bit EPC num. 512-bit user memory	94.8 x 8.1	‖	‖
ALN9654	‖	‖	‖	93 x 19	‖	‖
AD821	Avery Dennison	Impinj Monza1	96-bit EPC num.	72 x 30	Al	Item, carton, pallet
AD833	‖	Impinj Monza3	‖	38 x 93.5	‖	‖
AD224	‖	NXP U-Code Gen2 XM	96/240-bit EPC num 512-bit user memory	95 x 7.4	‖	‖
AD824	‖	‖	‖	30 x 50	‖	Item
Dogbone	UPM Raflatac	Impinj Monza2	96-bit EPC num.	93 x 23	Al	Item, carton, pallet
Dogbone	‖	Impinj Monza4	128/480-bit EPC num. 512-bit user memory	86 x 24	‖	‖
Dogbone	‖	NXP U-Code Gen2 XM	96/240-bit EPC num 512-bit user memory	93 x 23	‖	‖
ShortDipole	‖	‖	‖	92 x 11	‖	‖