2

Wi-Fi Tracking Threatens Users' Privacy in Fingerprinting Techniques

Lorenz Schauer

MOBILE AND DISTRIBUTED SYSTEMS GROUP, LMU MUNICH, MUNICH, GERMANY

1 Introduction

Due to the immense diffusion of smart mobile devices, the usage of Wi-Fi as state-of-the-art wireless communication standard has increased dramatically. Wi-Fi infrastructures are installed in many public spaces and buildings providing Internet access and local services. They also render Wi-Fi fingerprinting possible, which is one of the most auspicious technique for indoor positioning, first presented by Bahl and Padmanabhan (2000) in the RADAR system. Wi-Fi fingerprinting requires IEEE 802.11 active scans in the online phase, which are performed by users for position estimations. Hence, beside automatic Wi-Fi scans which usually take place every 2 min on average (Bonné et al., 2013), even more signals are sent out by mobile devices when fingerprinting is used. This leads to a serious risk for users' privacy, due to an increased Wi-Fi traffic which can be sniffed by any person in reach and without the users' consent or their awareness (Schauer et al., 2014).

However, only few works can be found in literature, where such privacy issues are investigated. For instance, Li et al. (2014) propose a privacy-preserving Wi-Fi fingerprinting localization scheme protecting both the data privacy of the localization service provider and the user's location. The authors realize an encrypted transmission of online fingerprints from mobile devices to the localization server. Thus, vectors of received signal strengths (RSS) are protected against sniffer attacks. However, the scanning process remains unencrypted and can be easily captured.

This chapter deals with privacy risks when using common Wi-Fi-based indoor positioning techniques, such as fingerprinting. In particular, we focus on the information, which can be extracted out of captured data from IEEE 802.11 active scans. Beside the technical background, we provide an overview of existing Wi-Fi tracking approaches. Furthermore, we describe investigated IEEE 802.11 protocol extensions and the mechanism of MAC address randomization, which was recently established in well-known mobile operating systems. However, these methods do not fulfill the requirements for an overall privacy-preserving positioning approach. Hence, we further investigate our

Geographical and Fingerprinting Data for Positioning and Navigation Systems. https://doi.org/10.1016/B978-0-12-813189-3.00002-2

previous work (Schauer et al., 2016a) proposing a fully passive scanning procedure for Wi-Fi fingerprinting.

In summary, the chapter is structured as follows: Section 2 reveals related work. The technical background for Wi-Fi tracking is described in Section 3 while Section 4 gives an overview of potentials and limitations. In Section 5 existing security mechanisms are presented, whereas Section 6 describes the privacy-preserving fingerprinting approach by Schauer et al. (2016a) and presents further investigations. Finally, Section 7 concludes the chapter giving hints on future work.

2 Related Work

A lot of researches have been done in recent years improving common Wi-Fi indoor positioning systems in terms of accuracy (Bell et al., 2010), precision (Martin et al., 2010), scalability (Ledlie et al., 2012), and efficiency (Sabek et al., 2015). Among these systems, Wi-Fi fingerprinting achieves adequate positioning results in literature (Farshad et al., 2013), but also suffers from high efforts for recording the radio map in the training phase. Therefore, many investigations have focused on reducing these efforts for creating the radio map, for example, Gunawan et al. (2012) or Koweerawong et al. (2013), rather than on solving privacy issues.

By contrast, a vast number of approaches for achieving location and communication privacy in location-based services (LBS) have been developed, such as Dorfmeister et al. (2015) or Yang et al. (2013), most of which are adapting well-established privacy concepts from other fields, such as k-anonymity (Sweeney, 2002), to the particular needs of LBS. However, with the majority of approaches dealing with outdoor LBS, privacy risks concerning a user's location caused by active Wi-Fi scans are not discussed any further. Especially, considering the terminal-based positioning approach of GPS, the mere acquisition of position updates in an outdoor environment usually does not pose any threats to a user's location privacy.

Privacy-preserving approaches are also well studied in the field of indoor positioning and wireless LANs. Jiang et al. (2007) analyze the problem of location privacy in wireless infrastructures and introduce a protocol to protect the user's location by obfuscating privacy compromising information leaking in Wi-Fi communications. They already consider silent attackers capturing Wi-Fi packets within communication range as the strongest attackers for users' privacy. Note that the approach by Schauer et al. (2016a) protects the positioning process against these silent sniffers.

Konstantinidis et al. (2015) propose a privacy-preserving indoor positioning approach for mobile devices protecting users against location tracking by the localization service. Furthermore, they discuss several positioning techniques, including Wi-Fi fingerprinting, and present a framework to protect the user's location against known privacy attacks. However, neither the scanning procedure nor any data sent from the mobile device, such as RSS vectors, are protected in particular.

Only few works can be found in literature, where privacy issues in case of Wi-Fi fingerprinting are investigated, such as Li et al. (2014), Schauer et al. (2016a), or Gschwandtner and Schindhelm (2011). The latter propose a privacy-preserving Wi-Fi fingerprinting approach using enhanced Wi-Fi beacons. The authors add all required positioning information into the vendor-specific elements of IEEE 802.11 beacon frames, such as AP positions, or essential parts of the recorded radio map. Thus, the client is able to calculate its' position locally on the device and no bidirectional communication with the localization service provider is required. The authors also describe the idea to listen on such enhanced beacon frames only, rather than sending probe requests for determining the current RSS vector. However, this is not further investigated, nor evaluated.

Besides our previous investigations (Schauer et al., 2016a), none of the mentioned works concentrate on the commonly used IEEE 802.11 active scanning process itself, which is proven to be a serious privacy risk (Lindqvist et al., 2009). Therefore, we deal with this aspect in the next section, where the technical background of Wi-Fi tracking is explained focusing on related privacy issues for Wi-Fi fingerprinting.

3 Technical Background

The wireless local area network technology, commonly known as Wi-Fi, is standardized as IEEE 802.11 (IEEE, 2007). The standard introduces three different frame types:

1. *Control frames*, in order to support the delivery process of data frames and to manage the medium access
2. *Data frames*, to transport user data for higher layers
3. *Management frames*, such as Beacon, Probe, Authentication, and Association frames, to exchange management information for connection establishment and maintenance

We focus on the latter, due to the fact that only management frames are involved in the 802.11 network discovery procedure, as shown in Fig. 1. From a mobile device's perspective, network discovery can be performed either passively (1) or actively (2). When performing passive scans, a client merely listens on beacon frames, which are periodically transmitted by access points over all channels they are currently operating on. The standard defines a periodic beacon interval of 100 ms. Hence, to receive a beacon frame transmitted on a certain channel, the client's radio must be set to the corresponding channel during transmission. For this purpose, the client iterates over all available channels and listens on each channel for a maximum duration defined in the IEEE 802.11 standard. Note, due to mismatching channels, clients may miss transmitted beacons using this procedure.

In order to bypass this problem and to enable a more efficient network discovery process, most clients—especially mobile devices—prefer active scanning, where clients actively send out probe request frames (2). This is done iteratively for each channel, with the client waiting for a certain period on each channel and listening for corresponding

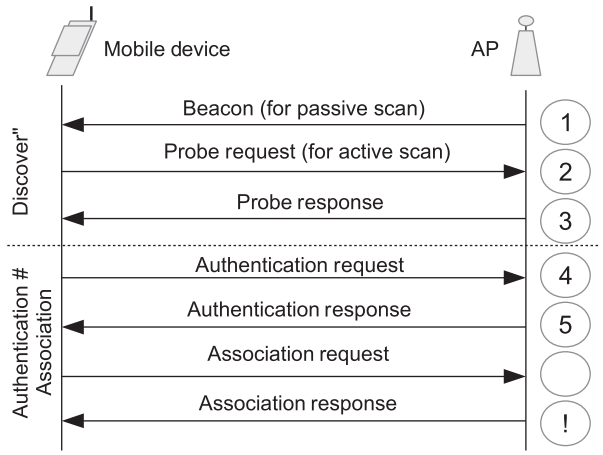


FIG. 1 Frames which are involved in the IEEE 802.11 discovery, authentication, and association process.

probe responses (3) According to Bonné et al. (2013), probe requests are sent out every 2 min on average, regardless of the connection status in order to discover available access points on the fly. Our own experiments confirm these results on average. However, the exact scanning interval depends on various parameters, such as the used chip set or the Wi-Fi driver.

Probe request frames contain unencrypted device-specific information, such as the client's MAC address, supported rates, the destination's network name (SSID), and other management information. With an empty SSID field, the probe request is interpreted as a broadcast message and all access points in reach on the corresponding channel will reply with a probe response. In case a specific network name is contained in the SSID field, only the access point offering the specified network will respond. These so-called directed probe requests are required in case of hidden networks. However, various mobile devices send out directed probe requests for each SSID contained in their preferred network list (PNL) in every active scan. The PNL stores all SSIDs to which the mobile client has tried to connect in the past in order to reconnect to known networks automatically.

All management frames can simply be captured by any Wi-Fi card in reach set into monitor mode. This mode allows to read the content of these frames on application level. Hence, information about stored network names and the device-specific MAC address is accessible for silent attackers within an area of interest. In case of continuous active scans, for example, performed during common Wi-Fi fingerprinting inside buildings, the user's locations as well as complete trajectories can be tracked (Musa and Eriksson, 2012; Schauer et al., 2016b). In summary, it becomes obvious that continuous Wi-Fi active scans in conjunction with indoor positioning systems lead to privacy issues, as it is also stated in the following section.

4 Potentials and Limitations of Wi-Fi Tracking

Due to the increasing percentage of modern mobile devices, Wi-Fi tracking has gathered high interest in both scientific and commercial world. In this section, we discuss potentials and limitations of this technique and present both current research and commercial projects using Wi-Fi tracking for different purposes. A related overview is also given by Schauer (2018).

As stated earlier, IEEE 802.11 probe request frames contain management data, which can be used to gather general information about the crowd in an area of interest. Using the first three bytes of a captured MAC address, one can perform an organizationally unique identifier lookup to determine the manufacturer ID of the sender. Hence, the distribution of present mobile devices according to their Wi-Fi chip set is easily accessible, which is a valuable information about the crowd (Schauer et al., 2014). However, take into account that the results only reflect the composition of detected devices, rather than the real distribution. For instance, a remarkable dominance of Apple devices is often seen in real-world scenarios, for example, in Musa and Eriksson (2012), due to the fact, that these devices show a higher scan frequency.

Due to accessible information about preferred network names, Wi-Fi probes have also been used in literature to determine social links and relationships in the crowd by comparing sets of captured SSIDs from mobile devices (Barbera et al., 2013; Cunche et al., 2012). The authors of the latter come to the conclusion that social relationships can be easily detected and analyzed by just using simple Wi-Fi tracking techniques. Hence, this poses a huge privacy risk for mobile users.

Beside relationships, other types of context information have already been extracted out of Wi-Fi tracking data. In general, such information describes the user's current situation, whereas *location*, *activity*, *time*, and *identity* belong to the four primary types, according to Abowd et al. (1999). Among these, location information plays a key role. It forms the basis of LBS and is often used for activity recognition (Lara and Labrador, 2013). Therefore, many investigations focus on inferring location and trajectory information from captured probe requests (e.g., Musa and Eriksson, 2012 or Schauer et al., 2016b).

However, this is still a challenging task, due to the arbitrary nature of probe request bursts and existing fluctuations in received signal strength indicators (RSSI). Therefore, simplistic positioning approaches are not sufficient for achieving reliable localization results (Schauer et al., 2016b; Bonné et al., 2013). The authors have conducted various empirical experiments in real-world scenarios indicating that pure RSSI values are not suitable to track mobile devices in crowded environments. Hence, many researchers have started to use probabilistic approaches, such as particle filter (Bartoletti et al., 2014) or hidden Markov models (HMM) in combination with the Viterbi algorithm (Musa and Eriksson, 2012; Trogh et al., 2015), in order to track spatiotemporal trajectories using Wi-Fi data.

Activity information, as another primary type of users' context, has also been extracted out of captured probe requests. Among others, Schauer and Linnhoff-Popien (2017) infer the current mobility state of a person by analyzing Wi-Fi traces, which is an important activity information. The authors create a two-state HMM and perform the Viterbi algorithm to detect dwelling and motion periods of mobile users. Muthukrishnan et al. (2009) take deterministic distance functions into account and use the extracted motion information to improve state-of-the-art signal-based localization algorithms. Shen et al. (2016) also determine dwell times and use this information for a feature-based room-level localization approach, which is able to detect the correct room in case of indistinguishable fingerprints.

An extensive work for extracting identity knowledge from Wi-Fi tracking data is presented by Ruiz-Ruiz et al. (2014). The authors perform a real-world study in a hospital and compute several spatio and/or temporal features for classification tasks. They conclude that realistic information reflecting the users' behavior in such a complex environment can be extracted out of passively recorded Wi-Fi data from mobile phones. Beside these works dealing with the extraction of various types of context information, it has proven that Wi-Fi tracking can further be used to determine crowd density and flows (Schauer and Werner, 2015), the users' proximity (Maier et al., 2015), or even for measuring waiting periods in human queues (Wang et al., 2014). It is important to take into account that all of this information can be inferred without the user's consent or even awareness.

Due to low costs and its simplicity, Wi-Fi tracking has also been discovered as an innovative business technology in recent years. Various companies and start-ups (e.g., *42reports*, *sensalytics*, or *walkbase*) have started to use this technology offering different services, such as localization, density and flow analysis, motion inference, return of investment calculations, or marketing optimization. These services are mainly used by retailers or shop owners who try to find similar analytic tools for their business like those which are already well-adopted in online shops. Hence, shoppers in the physical world get tracked for marketing purposes just like in the Internet. The only precondition is that they carry a Wi-Fi enabled mobile device. No complex hardware or software is required, which makes it easy for companies and other persons to use this technique for different purposes.

In summary, it can be seen that Wi-Fi tracking provides great potentials to gather personal information from an unknown crowd and without requiring the users' permission. Hence, this shows high risks concerning the privacy of most smartphone users and their implicit right on personal data. Take into account that in case of common fingerprinting techniques, even more Wi-Fi signals can be captured, due to continuous IEEE 802.11 active scans. But what are the consequences for our modern society where Wi-Fi is the de facto standard for a stable and fast connection to the Internet, which can be seen as a basic need? Which kind of security mechanisms exist or have to be developed in the near future to protect user's privacy completely against Wi-Fi sniffing in real-world scenarios? These questions are discussed in the following section.

5 Security Mechanisms Against Wi-Fi Tracking

Any person who carries a Wi-Fi enabled mobile device can be tracked by passive sniffers without his or her awareness. Hence, the most effective way to bypass this issue is to disable the Wi-Fi interface anytime or at least when the user is on the way. Sure, this is not a suitable solution, due to the loss of Internet connectivity, especially within buildings where mobile data connections may degrade. Beside this, most people are not willing to turn their interfaces on and off all the time. By contrast, the majority of mobile users keep Wi-Fi enabled, as it is shown for an airport environment by Schauer et al. (2014).

Beside this, users could remove network names from the PNL. This would decrease the amount of SSIDs and hence, social profiling like in Barbera et al. (2013) becomes more difficult. However, this does not prevent that users can be generally tracked by Wi-Fi sniffers. In contrast to these simple tricks, more sophisticated security mechanisms are presented in literature. Within these works, two major categories can be identified: the first proposes changes or extensions to the existing IEEE 802.11 protocol in terms of network discovery. The second focuses on MAC address randomization to obfuscate the hardware identifier rendering device recognition challenging.

5.1 Protocol Extensions

One work of this category is presented by Lindqvist et al. (2009). The authors propose a modified 802.11 protocol for the network discovery process which bypass directed probe requests. Furthermore, SSID information is completely obfuscated by using encryption mechanisms. With this concept, a passive attacker should not be able to gather knowledge about saved network names, and thus, social links cannot be determined. However, the software of both client and access point has to be modified requiring small changes at every participating device. Note that a randomization of the client's MAC address is solely discussed, rather than implemented.

By contrast, Greenstein et al. (2008) present a complete and stand-alone protocol, which is called *SlyFi*. The goal is to obfuscate all of the explicit identifiers, such as the MAC address or SSIDs. For this reason, SlyFi applies two mechanisms. Like before, they use encryption techniques for data transmissions. Hence, information about explicit identifiers is no longer accessible by Wi-Fi sniffers and cannot be used for device recognition. Again, this approach requires software modifications for clients and involved access points. In contrast to the former, the complete IEEE 802.11 protocol has to be replaced by the proposed SlyFi protocol.

5.2 MAC Address Randomization

Continuous tracking approaches mainly use the client's MAC address as explicit identifier in order to recognize and trace a particular device over time. Therefore, prior privacy-preserving methods, such as Hu and Wang (2005), or Jiang et al. (2007) proposed randomization of the client's MAC address to obfuscate this identifier rendering unambiguous

tracking challenging. However, these solutions have only been investigated theoretically, rather than applying them in real scenarios.

This was changed in 2014, when Apple—as one of the leading companies for mobile platforms—has started to integrate a MAC address randomization mechanism into the mobile operating system *iOS 8* in order to protect users against Wi-Fi sniffing. Hence, for the first time, a simple but also comprehensive privacy-preserving solution was applied to common devices. The mechanism keeps the real hardware address and broadcasts a randomized version during the standard network discovery process. Due to this approach, it is getting difficult to trace an iOS device over time. However, when Apple introduced the first version, MAC address randomization was only activated if the device has changed completely into sleep mode. This occurs only when both mobile data and location services are disabled, which is rarely given in real-world scenarios. Hence, most users were not protected by the proposed randomization mechanism (Beasley, 2014). Therefore, some improvements have been made and released with *iOS 9* where the mechanism was extended to location and auto join scans enabling MAC address randomization also for devices being in active mode (Skinner and Novak, 2015). However, the mechanism is disabled if a device is associated (connected) with an access point, and thus, the real MAC address of the mobile device is sent out by probe requests, like before.

Beside Apple, other well-known providers (e.g., Windows or Google) realized that Wi-Fi sniffing is an existing privacy issue, and hence, they started to develop their own MAC address randomization mechanisms. Since *Windows 10* (Wang, 2015) or *Android 6.0* (Android Developers, 2015) such approaches are integrated in the corresponding operating systems. However, due to the diversity of available mobile devices, the operability depends on the hardware driver. Hence, many users may still not have randomization enabled.

Furthermore, it is clearly stated in literature that randomization of the client's MAC address does not fulfill an overall privacy protection mechanism and it does not prevent users to be tracked by Wi-Fi sniffers. Recent experimental studies (e.g., by Vanhoef et al., 2016 or Martin et al., 2017) show that between 50% and 100% of all devices have been successfully recognized and traced, despite the usage of randomized MAC addresses. Also Pang et al. (2007) have proven that implicit identifiers and certain characteristics of 802.11 data transmissions are sufficient to recognize 64% of mobile devices with obfuscated hardware addresses.

Hence, it can be seen that the mentioned approaches are just a first and necessary step in order to complicate Wi-Fi tracking and to protect users against sniffers. However, and up to now, these attacks cannot be completely avoided by using the existing mechanisms, such as MAC address randomization. An overall protection is only given, if users deactivate their Wi-Fi adapters, which is not an adequate solution, due to the loss of network connectivity. Furthermore, many services and indoor positioning approaches (e.g., Wi-Fi fingerprinting) are not working with disabled Wi-Fi interfaces. Therefore, we present an alternative and privacy-preserving solution for Wi-Fi fingerprinting.

6 Privacy-Preserving Wi-Fi Fingerprinting

In order to preserve the users' privacy, a fully passive positioning process for Wi-Fi fingerprinting is described in this section. The basic idea is that a mobile device just listens for beacon frames in the online phase and determines a valid RSSI fingerprint, rather than performing any IEEE 802.11 active scans. Note that the content of this section is based on our previous work (Schauer et al., 2016a) and presents further investigations.

6.1 Basic Concept

In the online phase of common Wi-Fi fingerprinting systems, IEEE 802.11 active scans are usually performed in order to determine the RSSI vector of all access points in reach. However, this procedure involves probe request frames, and hence, continuous Wi-Fi tracking becomes easily possible. Therefore, a fully passive fingerprint creation is proposed as follows: For a specific time interval Δt, the mobile device listens for incoming beacon frames while iterating over the possible radio channels, switching channels after another interval Δt_{h}. This channel hopping is necessary to capture signals from all access points in reach operating on different channels.

The fingerprint vector v is filled with the mean RSSI values $\overline{v}_i$ of each access point i. Note that different values of Δt impact both the duration of determining v and the amount of information contained in a single fingerprint, which may influence the position accuracy as it is able to planish the impact of an observed RSSI outlier. Hence, Δt can be adapted to different mobility states depending on constraints concerning duration and accuracy of a position fix. For instance, when a persons is moving, Δt should be much smaller than for users who are dwelling.

For the online phase in Wi-Fi fingerprinting approaches, there are two ways of retrieving a location estimation based on a recent online RSSI vector: a deterministic way of comparing the input vector with distinct entries of the radio map, and a probabilistic approach considering the probability of a measurement given the prior knowledge. For evaluation, we investigate both ways which are described in the following section.

6.1.1 Deterministic Approach

According to SMARTPOS by Kessel and Werner (2011), we consider both weighted and nonweighted k-nearest neighbors (kNN) classifiers in signal space for deterministic location estimations during the online phase. Based on the Euclidean distance $d_i = \mathrm{dist}(v, r_i)$ between a passively measured RSSI vector v and a specific record r_i of the fingerprint database, we determine the k nearest candidates of possible user positions. Using the nonweighted kNN classifier, the centroid of these k positions is calculated and returned to the user as his/her current location. In addition, the weighted kNN method multiplies an individual weight w_i to each of the k position candidates, with w_i being calculated as:

$$w_i = \left(d_i \sum_{j=1}^{k} \frac{1}{d_j} \right)^{-1} \tag{1}$$

The user's current location estimate is then calculated as the sum of weighted k position candidates. Note that the authors of SMARTPOS obtained better results using the weighted procedure. For a thorough comparison, we also investigate both types of kNN classifiers for our passive Wi-Fi fingerprinting and compare the results.

Whenever an online RSSI vector is compared to those stored in the radio map, it is possible that any two vectors differ in their lengths, due to RSSI values of a certain access points only being contained in one of them and vice versa. Thus, one has to find a consistent way of dealing with missing values in order to compute the k-nearest neighbors. These values can either simply be ignored or be set to a predefined minimum value (e.g., −100 dBm; Kessel and Werner, 2011). Both ways have their qualification. When ignoring matchless entries, important information for accurate location estimations may be lost, while negative effects caused by changes occurring in the setup of access points may be kept low. On the other hand, a fixed minimum value such as −100 dBm punishes comparisons between strong RSSI and missing values, and favors comparisons between weak RSSI and missing values. However, this is to be expected in real-world scenarios, due to the fact, that strong signals should be measured again at the corresponding position, while weak signals may be missed cause of strong fluctuations of radio signals within buildings. We investigate the impact of treating missing values in our evaluation, described in Section 6.2.2. Take into account that SMARTPOS shows better results when ignoring missing values.

6.1.2 Probabilistic Approach

For probabilistic location estimations, we use a naive Bayes classifier in order to classify RSSI measurements into certain rooms of our building. Unlike before, this approach returns a specific room number to the user rather than a certain position fix as a coordinate-pair. To this end, room information has to be saved together with corresponding RSSI vectors in the radio map during offline phase.

More specifically, our naive Bayes classifier is based on the Bayes theorem and assigns the most probable class to a problem instance represented by a feature vector. In our case, we treat rooms as classes, and vectors of RSSI measurements as problem instances and feed them to Bayes theorem:

$$P(R|\nu) = \frac{P(\nu|R) \cdot P(R)}{P(\nu)} \tag{2}$$

calculating the posteriori probability $P(R|\nu)$ of being in a certain room R under the condition that fingerprint ν is observed. The probability $P(R)$ is the prior probability, which is based on our knowledge of frequencies in the training set, and hence, it can be easily estimated by counting the occurrence of each room. $P(\nu|R)$ is the likelihood function determining the probability of observing ν in case of being in room R, and $P(\nu)$ is called the evidence which can be calculated assuming a normal distribution with mean μ and the standard deviation σ for each one-dimensional parameter.

The naive Bayes classifier is simple to use, given the fact that it always assumes conditional independent features. Hence, we are allowed to express the probability of being in a certain room R in case of observing fingerprint v consisting of n access points' RSSI values v_i as follows:

$$P(R|v_1,\ldots,v_n) = \frac{1}{Z}P(R)\prod_{i=1}^{n}P(v_i|R) \tag{3}$$

with evidence $Z = P(v)$ treated as a constant in this case, because the values of RSSI measurements are known. Therefore, and due to the fact that we are only interested in the most probable room R_j with $j \in 1,\ldots,|R|$ using the maximum-a-posteriori decision rule, the naive Bayesian classifier can be directly derived from Eq. (3) and is expressed as follows:

$$R = \underset{R_j}{\operatorname{argmax}} \left\{ P(R_j)\prod_{i=1}^{n}P(v_i|R_j) \right\} \tag{4}$$

Hence, applying an online measured RSSI vector of length n to Eq. (4), the most probable room R_j out of all labeled rooms R is returned by the proposed naive Bayes classifier. Note that a similar probabilistic estimator is also used in SMARTPOS, but is not described by the authors how they treat missing values in this case.

Being linked by means of multiplication, however, observed RSSI values that are lacking their counterpart in the radio map for a certain room R_j would rigorously lead to zero-probability of R_j. In our case, this would lead to false classifications, due to the fact that a room will consequently show the probability of zero even if only one RSSI value is missing. In order to solve this problem, a small sample correction is added to all probability estimations guaranteeing that no probability is ever set to zero. These corrections are called pseudocounts or additive smoothing, which is commonly used with naive Bayes classifiers in order to treat missing values. In our case, we use Laplace smoothing for all of our measurements v_i taken in a certain room R_j and smooth $P(v_i|R_j)$ with a pseudocount $\gamma = 1$. This is done according to the following formula:

$$\hat{P}(v_i|R_j) = \frac{\text{count}(v_i)_{R_j} + \gamma}{|v|_{R_j} + \gamma \cdot (v)_{R_j}} \tag{5}$$

where $\text{count}(v_i)_{R_j}$ is the number of occurrences of the measurement v_i in room R_j, $|v|_{R_j}$ is the amount of all measurements made in room R_j, and $(v)_{R_j}$ is the domain of all measurements observed in room R_j. By using this Laplace smoothing technique, we are able to consider all of our rooms for classification even when the measurement data differ from the corresponding entries in the radio map. Thus, Eq. (4) is still correct for the whole set of possible rooms and returns the most probable room as the user's location estimate derived from a certain measured RSSI vector.

6.2 Evaluation

The proposed passive Wi-Fi fingerprinting approach was implemented for a mobile device using the existing wireless infrastructure in our office environment. For evaluation, it is compared to common active Wi-Fi fingerprinting in terms of well-known performance metrics, which are also used by SMARTPOS. Both deterministic and probabilistic location estimations are performed and confronted with the results of active scanning. Furthermore, we enhance our previous investigations and consider both static and dynamic users. We investigate how passive Wi-Fi fingerprinting differs in terms of accuracy and precision when users are moving. In order to compare the results of active and passive Wi-Fi fingerprinting, we perform both types of online scans using identical parameter setups.

6.2.1 Implementation and Setup

We will now give a closer look to our implementation and experimental setup. As first step of Wi-Fi fingerprinting, a database of RSSI measures on certain reference points has to be built up during an offline training phase. We use common active scans for recording the radio map, as this step is not sensitive to users' privacy. The active scans are performed directly by an application on a Samsung Galaxy S2 (I9100), which is used as our test device. At each reference point, we perform 20 active Wi-Fi scans for each of the four main orientations 0, 90, 180, and 270 degrees. A series of scans is always annotated with the position of the corresponding reference point on the map and the user's orientation when the fingerprint was taken. For radio map generation in the deterministic approach, each entry in the fingerprint database represents the vector of the means of 20 consecutively measured RSSI values per reachable access point. For the probabilistic approach, we determine the fingerprint as normal distribution over the measurements of each access point and add the corresponding room label information. In total, we recorded 332 fingerprints on 83 reference points located within one aisle of our office building and its main corridor, as shown in Fig. 2A. The distance between two consecutive reference points is always under 1.5 m.

For a fully privacy preserving fingerprint approach, any active bidirectional communication with a central location server has to be avoided. Instead, the radio map can either be locally stored on the device or might be transmitted in the beacon frames, as successfully shown by Gschwandtner and Schindhelm (2011). For our evaluation the radio map is directly stored on our mobile test device, which is used for location estimations. Due to the fact that beacon frames are only used for management purposes, they are not forwarded to the application layer, and thus usually cannot be read or further processed by user-level programs. In order to make them usable for passive fingerprinting, the device's Wi-Fi card has to be set into monitor mode, which is not possible with all common phones. Hence, we first rooted the phone and installed a patch for the Wi-Fi card using the Android application package of Bcmon[1], in order to be able to use 802.11 monitor mode for recording beacons. It is important to note that when the Wi-Fi card is set into

[1] https://code.google.com/p/bcmon/

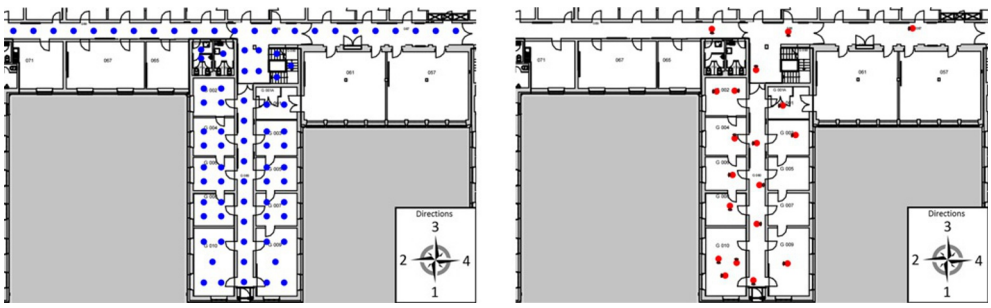
monitor mode, the device is only listening and does not send out any packages that could be captured by a malicious party or infrastructure provider. Hence, this can be seen as the highest level of protecting users from Wi-Fi tracking, while still being able to offer indoor positioning and navigation.

In order to create passive fingerprints during the online phase, the mobile device listens on incoming beacon frames for a specific time interval Δt, switching between the most commonly used Wi-Fi channels 1, 6, and 11 after Δt_{h}. If not explicitly stated otherwise, we set Δt to 3 s, and Δt_{h} to 1 s in our experiments. All necessary information, such as hardware addresses of access points and corresponding RSSIs, is then extracted from the resulting dump file and a fingerprint is constructed containing the mean values $\overline{v}_i$ of the observed RSSI values for each seen access point i.

For the creation of common active fingerprints the same application as for generating the radio map was used. Take into account that one active scan required about 4.5 s on average. Thus, our passive approach with $\Delta t = 3$ needs less time to collect the data necessary for formulating a position fix query. To allow for direct comparisons of the active and passive fingerprinting approaches, we successively apply both methods during the online phase on the same device at 19 randomly chosen locations across the test environment, as indicated in Fig. 2B. For determining the user's orientation at the moment a fingerprint was taken, we use a digital compass derived from the smartphone's accelerometer and magnetometer sensor readings. Both the orientation information and the observed fingerprint are compared fed to the locally stored radio map using either deterministic or probabilistic location estimation. The following sections indicate the achieved results, separately.

6.2.2 Deterministic Location Estimation

We now investigate the position accuracy and precision using various deterministic methods. More precisely, we calculate the mean, minimum and maximum positioning



(A) (B)

FIG. 2 Schematic overview of our test setup indicating reference points and the locations of position fixes with users' orientation. (A) Reference points for the radio map marked as *dots*. (B) Locations of position fixes marked as *dots* with user's orientation shown as *black pinnacles*.

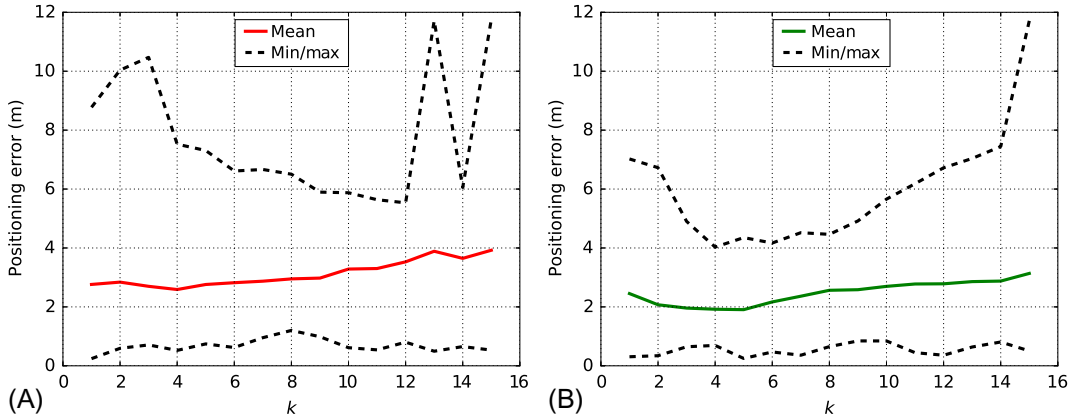


FIG. 3 Comparison of active versus passive Wi-Fi fingerprinting considering missing values, weighted *k*NN, and user orientation. (A) Active scanning. (B) Passive scanning.

error, as well as the standard deviation considering all of the 19 active versus the 19 passive position fixes while iterating over different values of k. Fig. 3 indicates the results of this real-world experiment when both the user's orientation and missing RSSI values are considered and a weighted kNN approach is used for deterministic location estimation. The best results are obtained with $k = 4$, showing a mean positioning error of 1.92 m and a standard deviation of 0.92 m for passive scans, and 2.59 m with a standard deviation of 1.68 m in case of common active scanning. Hence, these results indicate that the passive approach performs more accurately within our test set. Furthermore, and as expected, it can be observed that for both scan types, the mean positioning error tends to increase for higher values of k.

In order to investigate the impact of the used parameters, we now successively compare the results obtained by ignoring missing values, using nonweighted kNN, and completely neglecting the user's orientation. Eventually, the optimal parameter setting that results in the lowest average positioning error will be determined and discussed. Fig. 4 indicates the results for active and passive fingerprinting, when missing RSSI values are ignored, but relative weighting and orientation are still considered for location estimation. It is clearly shown that missing values should be treated by applying a minimal value. Otherwise, as shown in Fig. 4, the obtained values show unfeasible positioning results both for the active and the passive approach. In our case, we observe a mean position error of over 10 m for both scan types and $k < 10$. The standard deviation is greater than 8 m for active and greater than 6 m for passive scanning, which is not suitable for most indoor position scenarios. These observations are contrary to SMARTPOS, where the authors decided to ignore missing values in order to achieve slightly better positioning results.

As next step, we use a nonweighted kNN while considering the user's orientation and treating missing RSSI values. Again, the obtained results are shown in Fig. 5 for both scan types. It can be seen that the mean positioning error is a bit higher for each value of k and

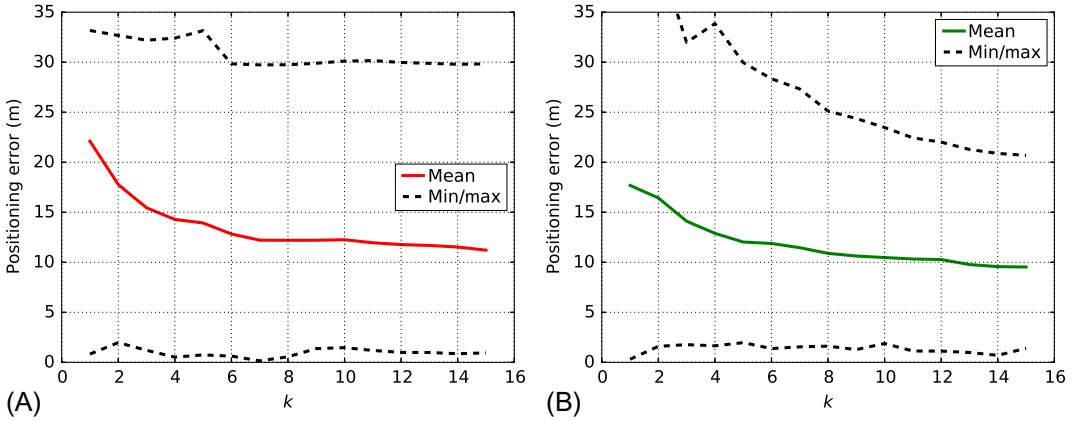


FIG. 4 Comparison of active versus passive Wi-Fi fingerprinting using weighted *k*NN, and users' orientation, but ignoring missing values. (A) Active, ignoring missing values. (B) Passive, ignoring missing values.

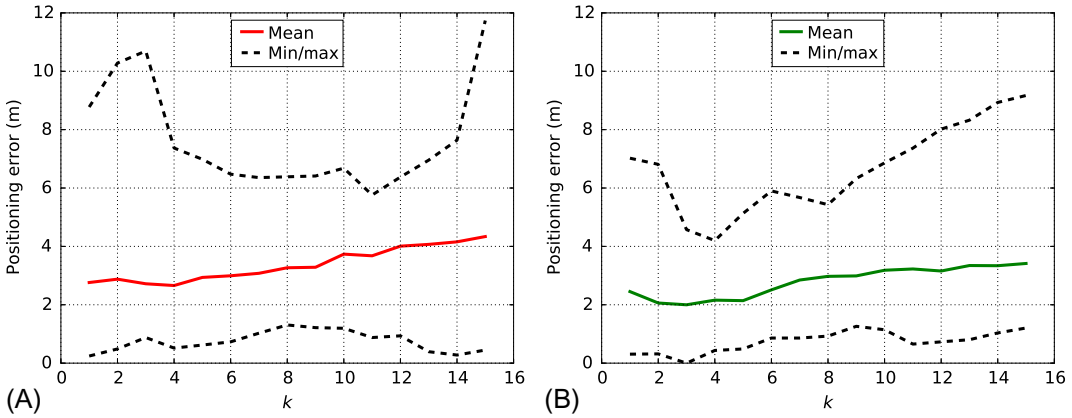


FIG. 5 Comparison of active versus passive Wi-Fi fingerprinting using weighted and nonweighted *k*NN. (A) Active, with nonweighted *k*NN. (B) Passive, with nonweighted *k*NN.

for both approaches when using nonweighted *k*NN instead of weighted *k*NN. With $k = 4$, the mean accuracy lies at 2.66 m for active and at 2.16 m for passive fingerprinting with a standard deviation of 1.2 and 1.7 m, respectively. Hence, our passive approach returns more accurate position fixes even when a nonweighted *k*NN location estimator is used. Overall, weighted *k*NN is to be preferred for both scan types, like in SMARTPOS.

As a last parameter, we investigate the impact of considering or ignoring the user's orientation. Thus, we apply weighted *k*NN, consider missing values, but now ignore the orientation information for our location estimation. Fig. 6 shows the corresponding results for active and passive fingerprinting. It can be observed that the overall positioning error on average is slightly lower for both scan types, especially for higher values of *k*, and again

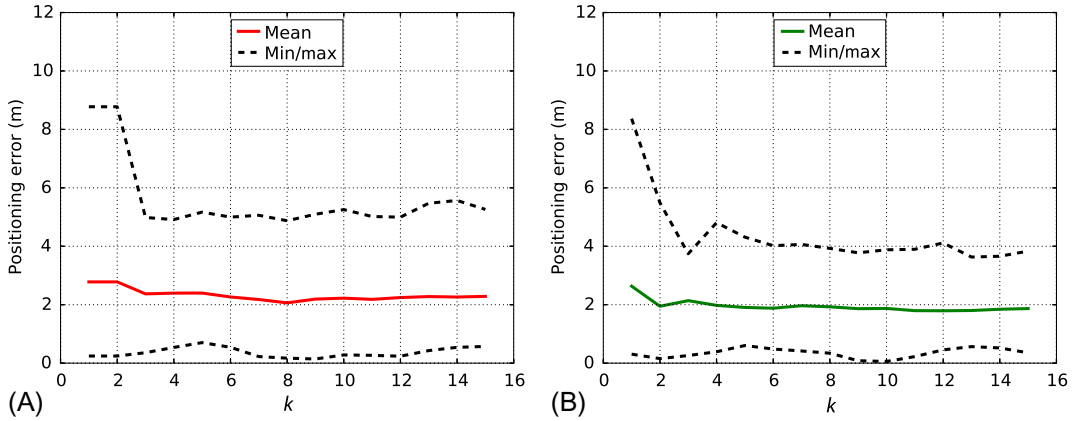


FIG. 6 Comparison of active versus passive Wi-Fi fingerprinting when considering and ignoring the users' orientation. (A) Active, without orientation. (B) Passive, without orientation.

passive scanning results in a lower positioning error than the common approach. For active scanning with $k = 4$, the mean accuracy is at 2.40 m and a standard deviation of 1.52 m is obtained. This is a tiny improvement of 0.19 m in terms of accuracy and 0.16 m in terms of precision for active fingerprinting. In case of our passive approach with $k = 4$, we observe a little degradation of 0.06 m for both accuracy and precision when ignoring the user's orientation instead of considering it. However, for $k = 6$ a mean positioning error of 1.88 m and a standard deviation of 0.99 m is obtained, indicating a slightly improvement of 0.04 m. An interesting observation is that for higher values of k, the mean accuracy remains constant when ignoring orientation while it is increasing when considering the orientation information. The explanation is that in case of ignoring orientation, the RSSI vector of online measurements is compared to the complete database, rather than comparing only entries with corresponding orientation. Thus, more similar fingerprints are available for k-nearest neighbors, and hence, an increasing k is less likely to negatively influence the positioning result. However, when neglecting orientation information, deterministic location estimation requires four times more database comparisons, and thus, a position request takes more time to be served.

In summary, we obtain the best results for both scan types within our experiment when using weighted kNN, considering missing values instead of ignoring them, but ignoring the user's orientation. These findings are contrary to SMARTPOS, where the information of users' orientation helped to increase the positioning accuracy. In our case for passive fingerprinting, the mean positioning error remains constantly below 2 m, the standard deviation below 1.1 m for $k > 3$. In comparison, the best results for active scanning were achieved with $k = 8$, showing an accuracy on average of 2.06 m and a standard deviation of 1.8 m. Hence, with respect to these results based on deterministic location estimation, we conclude that our passive approach performs slightly better than common active Wi-Fi

fingerprinting within our test set. In the next section, we describe our evaluation using probabilistic location estimations.

6.2.3 Probabilistic Location Estimation

We now apply the naive Bayes classifier as described in Section 6.1.2 to the same 19 position fixes of our online phase. In order to use the classifier, we first partition our test environment into 19 different rooms and corridor segments as shown in Fig. 7. Each segment contains four to six reference points marked with the corresponding room label. A room segment is classically divided by its walls, except the segments mapped onto the corridors of the building, which are quite long and are hence further divided into several parts to allow for a fine-grained positioning.

Overall, we investigate the correctness of the classification result for each position fix in three categories:

1. *Correct:* The online RSSI vector is classified to the room where the user is actually located.
2. *Nearby:* The online RSSI vector is classified to a direct neighbor of the actual room.
3. *False:* The classification returns a room far away from the actual room.

As before, we evaluate the impact of user's orientation by considering (+o) and ignoring (−o) the orientation information in the database. This is done for both active and passive scans. The classification results for the complete test set are depicted in Table 1, where

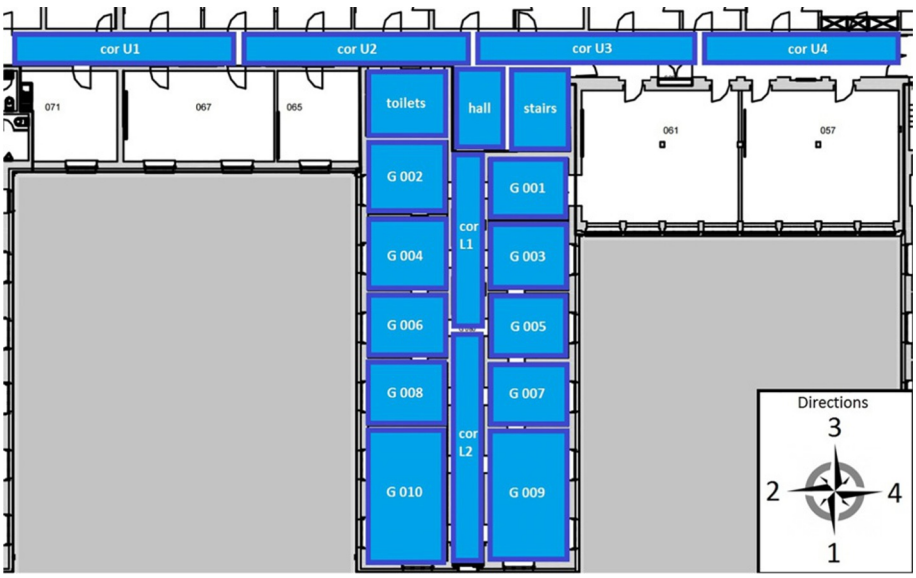


FIG. 7 Schematic overview of test environment divided into room segments.

Table 1 Classification Results for Active and Passive Scanning

	Active Scanning		Passive Scanning	
Real Room	**−o**	**+o**	**−o**	**+o**
hall	**hall**	*corU3*	*g001*	**hall**
corL2	**corL2**	**corL2**	**corL2**	**corL2**
g002	*g004*	*g004*	corU1	*toilets*
corL1	*g006*	*g006*	**corL1**	**corL1**
corU2	stairs	g004	g001	g003
g002	**g002**	*toilets*	g001	g001
g009	**g009**	**g009**	**g009**	**g009**
g008	**g008**	toilets	**g008**	g001
g006	**g006**	*corL2*	**g006**	**g006**
g003	*corL1*	*corL1*	**g003**	**g003**
g001	**g001**	**g001**	*g003*	**g001**
corL2	*g010*	*g010*	*g009*	**corL2**
corL1	*g007*	*g008*	*g007*	*g007*
corU2	*toilets*	*toilets*	**corU2**	g007
g010	g006	g006	*corL2*	*corL2*
g010	g006	g006	**g010**	*corL2*
g010	**g010**	**g010**	**g010**	**g010**
g004	toilets	g001	**g004**	*corL1*
corU3	g001	g001	**corU3**	g001
		Summary		
Correct	8	4	11	9
Nearby	5	8	5	5
False	6	7	3	5

emphases denote the correctness, such as bold values = correct, italic values = nearby, and underlined values = false.

It can easily be seen that the best results are obtained for both scan types when the information about users' orientation is ignored. This observation confirms to SMARTPOS, where the authors conclude that orientation information should not be used as a filter in a naive Bayesian estimator. When orientation is ignored, 42% of all rooms are classified correctly and 31% are false results using common active fingerprint. In comparison, when using our passive approach, 58% of all rooms are classified correctly with only 16% being misclassified. Based on these results, we conclude that passive Wi-Fi fingerprinting performs more accurately for both deterministic and probabilistic location estimations, and furthermore, it is capable to completely preserve mobile users' privacy during the whole positioning process, as the device's transceiver is only passively used for receiving beacons.

An explanation for the improvement in terms of accuracy is that even quick passive scans are able to aggregate more information about received signal strengths for

calculating an online fingerprint than common active scans which cycle through all Wi-Fi channels but typically only report one observed RSSI value per access point. We will review this assumption in the following section by performing longer scanning periods.

6.2.4 Considering User Movement

Up to this point, we have only been considering a user moving through our building and hence requiring a short passive scanning process. For this purpose, we set $\Delta t = 3$ and $\Delta t_{\mathrm{h}} = 1$. This means that for the creation of a passive Wi-Fi fingerprint, the mobile device listens for beacon frames for a period of 3 s, while staying on the most commonly used Wi-Fi channels for 1 s each. As a next step, we investigate whether the positioning results can be improved by allowing a longer period of time for both Δt and Δt_{h}. We thereby increase the information for calculating an RSSI vector, which in return can be expected to further reduce the possible impact of observed RSSI outliers as the number of overheard beacons grows. This directly applies to the scenario of a static user, for example, a person remaining in a certain location for a longer period of time, which of course is very common within buildings. Note that this type of a user's context (i.e., activity) can easily be inferred by modern mobile devices using its integrated sensors (Maier and Dorfmeister, 2013) or even by passive Wi-Fi captures (Schauer and Linnhoff-Popien, 2017). Hence, while the user remains static, for example, sitting in an office or cafeteria, we propose a longer-time period, experimentally set to $\Delta t = 20$ and $\Delta t_{\mathrm{h}} = 2$.

In order to compare the obtained positioning errors of dynamic and static users, we conduct another experiment. Our test device captures Wi-Fi data at a fixed position for a long duration of 30 min, which simulates a person sitting in an office. We distinguish between two types of users computing an online fingerprint as follows:

1. *Dynamic user:* Every 20 s, an RSSI vector is calculated based on a 3-s capture with $\Delta t_{\mathrm{h}} = 1$.
2. *Static user:* Every 20 s, an RSSI vector is calculated based on all information captured during the last 20 s with $\Delta t_{\mathrm{h}} = 2$.

Both types of calculated RSSI vectors are used for deterministic location estimation with weighted kNN, considering the user's orientation, and missing RSSI values. Fig. 8 depicts the results of the obtained positioning errors for this experiment.

It can clearly be seen that in case of a static user's device performing a longer scanning period, the achieved positioning error fluctuates considerably less and shows a higher accuracy on average of 2.2 m. By contrast, position fixes of the dynamic user suffer from high fluctuations in terms of their accuracy during the experiment. Hence, we conclude that the precision of our positioning approach is lower when using a shorter scanning period, which was expected. Thus, in order to obtain position fixes with higher precision when the user does not move, it is useful to determine the user's behavior in a first step in order to adjust the scanning period.

Finally, we investigate the accuracy of our naive Bayes classifier for a static user and compare the performed classifications with the results obtained in Table 1. For this

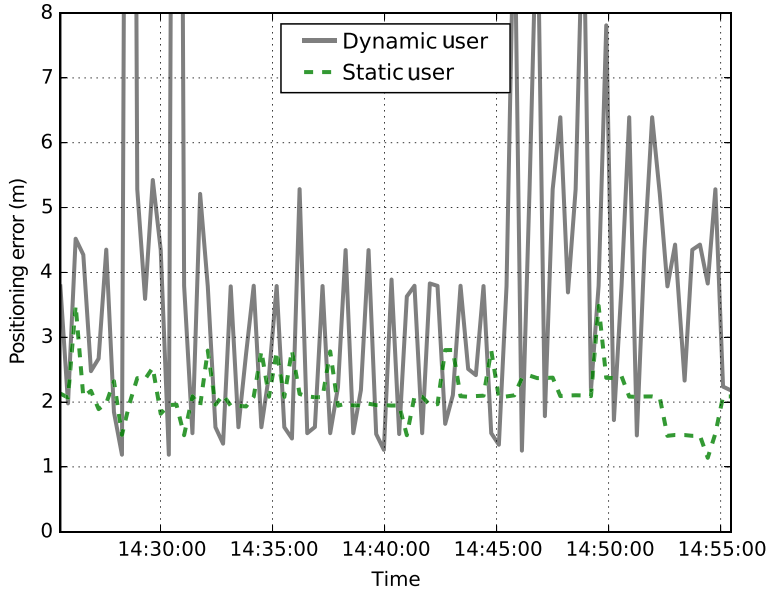


FIG. 8 Comparing positioning errors for dynamic and static users.

purpose, we repeat the 19 online position fixes with $\Delta t = 20$ and $\Delta t_{\mathrm{h}} = 2$. The obtained results indicate that one more room is classified correctly. Overall, 12 (63%) rooms were correct, 4 (21%) rooms nearby, and, still, 3 (16%) rooms falsely classified. This indicates a marginal improvement for static scenarios when probabilistic location estimations are performed. Hence, we conclude that for both deterministic and probabilistic fingerprinting, a longer scanning period can help to improve the accuracy and precision of our passive approach. This confirms the previous assumption that passive scans are able to aggregate more information for localization tasks than common IEEE 802.11 active scans, while fully preserving users' privacy.

In summary, the proposed approach presents one possible solution to overcome the privacy problem in IEEE 802.11 fingerprinting techniques. However, it requires root privileges and small driver manipulations to set the phone's Wi-Fi card into monitor mode. Furthermore, network connectivity is disabled in this mode, and hence, the proposed method is not applicable in real life for most smartphone users.

7 Conclusion and Future Work

In this chapter, we have focused on existing privacy problems for mobile users due to simple and low-cost Wi-Fi tracking techniques. Beside the technical background, we have demonstrated the wide range of possibilities to analyze an unknown crowd without the users' consent or awareness by just capturing probe requests. Overall, it has been shown

that every person who carries a Wi-Fi-enabled mobile device risks to be tracked and analyzed involuntarily. This threat even increases when performing Wi-Fi fingerprinting using additional active scans. Therefore, several exiting approaches have been presented for the protection of users' privacy against this attack. Beside protocol extensions and MAC address randomization, a fully passive Wi-Fi fingerprinting approach was discussed in this chapter.

However, it is clear that none of the described solutions is both practical and suitable for real-world scenarios. While protocol extensions need software manipulations, the randomization of the hardware address is not sufficient for completely obfuscating the identity of the device. On the one hand, our passive Wi-Fi fingerprinting approach ensures the highest level of privacy preservation, due to the fact that no signals are sent out by the mobile device. On the other hand, it requires software modifications and disables network connectivity.

Hence, we conclude that an overall and practical privacy-preserving approach for fingerprinting techniques is still missing. Further research has to be performed to prevent Wi-Fi sniffing completely while providing the full functionality. Therefore, future work should enhance MAC address randomization and extend it to other implicit and explicit identifiers in order to fully obfuscate the user's identity.

References

Abowd, G.D., Dey, A.K., Brown, P.J., Davies, N., Smith, M., Steggles, P., 1999. Towards a better understanding of context and context-awareness. In: International Symposium on Handheld and Ubiquitous Computing, pp. 304–307.

Android Developers, 2015. Android 6.0 Changes. Available from: https://developer.android.com/about/versions/marshmallow/android-6.0-changes.html#behavior-notifications (Accessed 12 July 2017).

Bahl, P., Padmanabhan, V.N., 2000. RADAR: an in-building RF-based user location and tracking system. In: INFOCOM 2000. Nineteenth Annual Joint Conference of the IEEE Computer and Communications Societies. Proceedings. IEEE, vol. 2, pp. 775–784.

Barbera, M.V., Epasto, A., Mei, A., Perta, V.C., Stefa, J., 2013. Signals from the crowd: uncovering social relationships through smartphone probes. In: Proceedings of the 2013 Conference on Internet Measurement Conference, pp. 265–276.

Bartoletti, S., Conti, A., Giorgetti, A., Win, M.Z., 2014. Sensor radar networks for indoor tracking. IEEE Wireless Commun. Lett. 3 (2), 157–160.

Beasley, M., 2014. More details on how iOS 8's MAC address randomization feature works (and when it doesn't). Available from: https://9to5mac.com/2014/09/26/more-details-on-how-ios-8s-mac-address-randomization-feature-works-and-when-it-doesnt/ (Accessed 12 July 2017).

Bell, S., Jung, W.R., Krishnakumar, V., 2010. WiFi-based enhanced positioning systems: accuracy through mapping, calibration, and classification. In: Proceedings of the 2nd ACM SIGSPATIAL International Workshop on Indoor Spatial Awareness, pp. 3–9.

Bonné, B., Barzan, A., Quax, P., Lamotte, W., 2013. WiFiPi: involuntary tracking of visitors at mass events. In: 2013 IEEE 14th International Symposium and Workshops on a World of Wireless, Mobile and Multimedia Networks (WoWMoM), pp. 1–6.

Cunche, M., Kaafar, M.A., Boreli, R., 2012. I know who you will meet this evening! Linking wireless devices using Wi-Fi probe requests. In: 2012 IEEE International Symposium on a World of Wireless, Mobile and Multimedia Networks (WoWMoM), pp. 1–9.

Dorfmeister, F., Wiesner, K., Schuster, M., Maier, M., 2015. Preventing restricted space inference in online route planning services. In: MOBIQUITOUS'15 Proceedings of the 12th EAI International Conference on Mobile and Ubiquitous Systems: Computing, Networking and Services. ACM.

Farshad, A., Li, J., Marina, M.K., Garcia, F.J., 2013. A microscopic look at WiFi fingerprinting for indoor mobile phone localization in diverse environments. In: 2013 International Conference on Indoor Positioning and Indoor Navigation (IPIN), pp. 1–10.

Greenstein, B., McCoy, D., Pang, J., Kohno, T., Seshan, S., Wetherall, D., 2008. Improving wireless privacy with an identifier-free link layer protocol. In: Proceedings of the 6th International Conference on Mobile Systems, Applications, and Services, pp. 40–53.

Gschwandtner, F., Schindhelm, C.K., 2011. Spontaneous privacy-friendly indoor positioning using enhanced WLAN beacons. In: 2011 International Conference on Indoor Positioning and Indoor Navigation (IPIN), pp. 1–8.

Gunawan, M., Li, B., Gallagher, T., Dempster, A.G., Retscher, G., 2012. A new method to generate and maintain a WiFi fingerprinting database automatically by using RFID. In: 2012 International Conference on Indoor Positioning and Indoor Navigation (IPIN), pp. 1–6.

Hu, Y.C., Wang, H.J., 2005. A framework for location privacy in wireless networks. In: ACM SIGCOMM Asia Workshop.

IEEE, 2007. IEEE Std 802.11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications. IEEE Computer Society, New York, NY.

Jiang, T., Wang, H.J., Hu, Y.-C., 2007. Preserving location privacy in wireless LANs. In: Proceedings of the 5th International Conference on Mobile Systems, Applications and Services, pp. 246–257.

Kessel, M., Werner, M., 2011. SMARTPOS: accurate and precise indoor positioning on mobile phones. In: Proceedings of the 1st International Conference on Mobile Services, Resources, and Users, Barcelona, pp. 158–163.

Konstantinidis, A., Chatzimilioudis, G., Zeinalipour-Yazti, D., Mpeis, P., Pelekis, N., Theodoridis, Y., 2015. Privacy-Preserving Indoor Localization on Smartphones. IEEE Transactions on Knowledge and Data Engineering 27 (11), 3042–3055. IEEE.

Koweerawong, C., Wipusitwarakun, K., Kaemarungsi, K., 2013. Indoor localization improvement via adaptive RSS fingerprinting database. In: 2013 International Conference on Information Networking (ICOIN), pp. 412–416.

Lara, O.D., Labrador, M.A., 2013. A survey on human activity recognition using wearable sensors. IEEE Commun. Surv. Tutorials 15 (3), 1192–1209.

Ledlie, J., Park, J.G., Curtis, D., Cavalcante, A., Camara, L., Costa, A., Vieira, R., 2012. Molé: a scalable, user-generated WiFi positioning engine. J. Locat. Based Serv. 6 (2), 55–80.

Li, H., Sun, L., Zhu, H., Lu, X., Cheng, X., 2014. Achieving privacy preservation in WiFi fingerprint-based localization. In: INFOCOM, 2014 Proceedings IEEE, pp. 2337–2345.

Lindqvist, J., Aura, T., Danezis, G., Koponen, T., Myllyniemi, A., Mäki, J., Roe, M., 2009. Privacy-preserving 802.11 access-point discovery. In: Proceedings of the Second ACM Conference on Wireless Network Security, pp. 123–130.

Maier, M., Dorfmeister, F., 2013. Fine-grained activity recognition of pedestrians travelling by subway. In: 5th International Conference on Mobile Computing, Applications and Services (MobiCASE 2013), vol. 130. Springer, Paris, France, pp. 122–139.

Maier, M., Schauer, L., Dorfmeister, F., 2015. ProbeTags: privacy-preserving proximity detection using Wi-Fi management frames. In: 2015 IEEE 11th International Conference on Wireless and Mobile Computing, Networking and Communications (WiMob), pp. 756–763.

Martin, E., Vinyals, O., Friedland, G., Bajcsy, R., 2010. Precise indoor localization using smart phones. In: Proceedings of the International Conference on Multimedia, pp. 787–790.

Martin, J., Mayberry, T., Donahue, C., Foppe, L., Brown, L., Riggins, C., Rye, E.C., Brown, D., 2017. A Study of MAC Address Randomization in Mobile Devices and When It Fails. Proceedings on Privacy Enhancing Technologies 4, 365–383. De Gruyter Open. ArXiv preprint arXiv:1703.02874.

Musa, A.B.M., Eriksson, J., 2012. Tracking unmodified smartphones using Wi-Fi monitors. In: Proceedings of the 10th ACM Conference on Embedded Network Sensor Systems, pp. 281–294.

Muthukrishnan, K., van der Zwaag, B., Havinga, P., 2009. Inferring motion and location using WLAN RSSI. In: Mobile Entity Localization and Tracking in GPS-Less Environments, pp. 163–182.

Pang, J., Greenstein, B., Gummadi, R., Seshan, S., Wetherall, D., 2007. 802.11 user fingerprinting. In: Proceedings of the 13th Annual ACM International Conference on Mobile Computing and Networking, pp. 99–110.

Ruiz-Ruiz, A.J., Blunck, H., Prentow, T.S., Stisen, A., Kjærgaard, M.B., 2014. Analysis methods for extracting knowledge from large-scale WiFi monitoring to inform building facility planning. In: 2014 IEEE International Conference on Pervasive Computing and Communications (PerCom), pp. 130–138.

Sabek, I., Youssef, M., Vasilakos, A.V., 2015. ACE: an accurate and efficient multi-entity device-free WLAN localization system. IEEE Trans. Mobile Comput. 14 (2), 261–273.

Schauer, L., 2018. Analyzing the digital society by tracking mobile customer devices. In: Digital Marketplaces Unleashed. Springer, pp. 467–478.

Schauer, L., Linnhoff-Popien, C., 2017. Extracting context information from Wi-Fi captures. In: Proceedings of the 10th International Conference on PErvasive Technologies Related to Assistive Environments, pp. 123–130.

Schauer, L., Werner, M., 2015. Analyzing pedestrian flows based on Wi-Fi and Bluetooth captures. EAI Endorsed Trans. Ubiquit. Environ. 1, e4.

Schauer, L., Werner, M., Marcus, P., 2014. Estimating crowd densities and pedestrian flows using Wi-Fi and Bluetooth. In: Proceedings of the 11th International Conference on Mobile and Ubiquitous Systems: Computing, Networking and Services, pp. 171–177.

Schauer, L., Dorfmeister, F., Wirth, F., 2016a. Analyzing passive Wi-Fi fingerprinting for privacy-preserving indoor-positioning. In: 2016 International Conference on Localization and GNSS (ICL-GNSS), pp. 1–6.

Schauer, L., Marcus, P., Linnhoff-Popien, C., 2016b. Towards feasible Wi-Fi based indoor tracking systems using probabilistic methods. In: 2016 International Conference on Indoor Positioning and Indoor Navigation (IPIN), pp. 1–8.

Shen, J., Cao, J., Liu, X., Wen, J., Chen, Y., 2016. Feature-based room-level localization of unmodified smartphones. In: Smart City 360°, pp. 125–136.

Skinner, K., Novak, J., 2015. Privacy and your app. In: Apple Worldwide Dev. Conf. (WWDC).

Sweeney, L., 2002. K-anonymity: a model for protecting privacy. Int. J. Uncertain. Fuzziness Knowl. Based Syst. 10 (5), 557–570. ISSN 0218-4885.

Trogh, J., Plets, D., Martens, L., Joseph, W., 2015. Advanced real-time indoor tracking based on the Viterbi algorithm and semantic data. Int. J. Distrib. Sens. Netw. 11 (10), 271818.

Vanhoef, M., Matte, C., Cunche, M., Cardoso, L.S., Piessens, F., 2016. Why MAC address randomization is not enough: an analysis of Wi-Fi network discovery mechanisms. In: Proceedings of the 11th ACM on Asia Conference on Computer and Communications Security, pp. 413–424.

Wang, W., 2015. Wireless networking in Windows 10. In: Windows Hardware Engineering Community conference (WinHEC).

Wang, Y., Yang, J., Chen, Y., Liu, H., Gruteser, M., Martin, R.P., 2014. Tracking human queues using single-point signal monitoring. In: Proceedings of the 12th Annual International Conference on Mobile Systems, Applications, and Services, pp. 42–54.

Yang, D., Fang, X., Xue, G., 2013. Truthful incentive mechanisms for k-anonymity location privacy. In: INFOCOM 2013 Proceedings IEEE, pp. 2994–3002.