SSD: A Robust RF Location Fingerprint Addressing Mobile Devices' Heterogeneity

A.K.M. Mahtab Hossain, Yunye Jin, Wee-Seng Soh, *Member*, *IEEE*, and Hien Nguyen Van

Abstract—Fingerprint-based methods are widely adopted for indoor localization purpose because of their cost-effectiveness compared to other infrastructure-based positioning systems. However, the popular location fingerprint, Received Signal Strength (RSS), is observed to differ significantly across different devices' hardware even under the same wireless conditions. We derive analytically a robust location fingerprint definition, the *Signal Strength Difference* (*SSD*), and verify its performance experimentally using a number of different mobile devices with heterogeneous hardware. Our experiments have also considered both Wi-Fi and Bluetooth devices, as well as both Access-Point(AP)-based localization and Mobile-Node (MN)-assisted localization. We present the results of two well-known localization algorithms (K Nearest Neighbor and Bayesian Inference) when our proposed fingerprint is used, and demonstrate its robustness when the testing device differs from the training device. We also compare these SSD-based localization algorithms' performance against that of two other approaches in the literature that are designed to mitigate the effects of mobile node hardware variations, and show that SSD-based algorithms have better accuracy.

Index Terms—Location fingerprint, signal strength difference (SSD), Wi-Fi, Bluetooth, indoor localization, positioning system, heterogeneous devices

1 Introduction

Accurate indoor location determination is an indispensable building block of various context-aware services and ubiquitous environments. Geometric approaches require antenna arrays with large number of array elements on transceivers to achieve good accuracy, which incur high hardware cost [2]. On the other hand, fingerprint-based approaches, utilizing signal parameters provided by off-the-shelf wireless devices, are widely adopted for indoor localization purpose for their cost-effectiveness.

In a typical fingerprint-based system, a set of "training locations" are chosen in the service area. During an offline "training phase," location-dependent signal parameters, most commonly Received Signal Strength (RSS) values, are measured and recorded at each training location as the fingerprint for that particular location. During the online localization phase, various methods utilizing the recorded data can be applied to estimate the target device's location when the online RSS values of the device are collected.

- *A.K.M. Mahtab Hossain is with the Internet Education and Research Laboratory (intERLab), Asian Institute of Technology (AIT), PO Box 4, Khlong Luang, Pathumthani 12120, Thailand. E-mail: mahtab@ait.ac.th.*
- *Y. Jin is with the Sense and Sense-Abilities Programme, Institute for Infocomm Research, A*STAR, Singapore, 1 Fusionopolis Way, #21-01 Connexis (South Tower), Singapore 138632. E-mail: yjin@i2r.a-star.edu.sg.*
- *W.-S. Soh is with the Department of Electrical and Computer Engineering, National University of Singapore, 4 Engineering Drive 3, #05-21, Singapore 117583. E-mail: elesohws@nus.edu.sg.*
- *H.N. Van is with the Department of Electrical and Computer Engineering, Center for Automation Research (CFAR), University of Maryland, College Park, A.V. William Building 115, Room #4413, MD 20740. E-mail: hien@umd.edu.*

Manuscript received 11 Mar. 2011; revised 7 Oct. 2011; accepted 14 Oct. 2011; published online 14 Nov. 2011.
For information on obtaining reprints of this article, please send e-mail to: tmc@computer.org, and reference IEEECS Log Number TMC-2011-03-0126. Digital Object Identifier no. 10.1109/TMC.2011.243.

Various commercially available hand-held devices and wireless Access Points (APs) are capable of reporting RSS. In general, the RSSs are mostly reported in dBm values. However, these devices usually come with many different hardware solutions, even for the same wireless technology. Regardless of whether a device's signal strengths as perceived by the APs are used to denote the device's location fingerprint, or the reverse approach in which the APs' signal strengths as perceived by the device (i.e., Mobile Node (MN)) are used, such fingerprints may differ significantly with the device's hardware even under the same wireless conditions [1], [3], [4], [5]. This is often observed in existing popular wireless technologies, such as Wi-Fi or Bluetooth. The presence of power control feature in some mobile devices further complicates the issue [3]. As a result, a positioning system that relies solely on RSS to define location fingerprints generally does not perform well across heterogeneous devices.

The need for a robust location fingerprint is obligatory for any fingerprint-based localization algorithm, no matter how sophisticated the algorithm is. In [1], we proposed a robust location fingerprint, namely, *Signal Strength Difference (SSD)*, which was shown to outperform the traditional RSS fingerprint in terms of robustness across heterogeneous mobile devices, both analytically and experimentally. In this paper, we analyze the robustness of SSD more elaborately, using several off-the-shelf Wi-Fi and Bluetooth devices.

In existing localization literature, we usually encounter two different approaches to collect the signal strength samples, namely, *AP-based*, where the RSS is measured at the AP, and *MN-assisted*, where the RSS is actually measured at the MN itself. In order to verify SSD's robustness, we need to consider both of these scenarios. However, we have only considered the AP-based scenario in [1], both for analysis and experiments. In this paper, we show that, regardless of whether the signal strength samples are

collected at the APs or at the MN, SSD is a more robust location fingerprint compared to the traditional RSS.

We consider two separate experimental testbeds for Wi-Fi and Bluetooth. The Bluetooth testbed follows the AP-based approach while the Wi-Fi testbed follows the MN-assisted approach. Both testbeds include concrete walls, cubicles, people, etc., representing the indoor environment more practically, compared to our initial lecture theater testbed in [1]. We also compare SSD with two other robust location fingerprints [5], [6] that are argued to mitigate the effects of MN's hardware variations.

The rest of the paper is organized as follows: We discuss our idea of defining a robust location fingerprint in Section 2. We provide a brief description of related works in Section 3. In Section 4, we present experimental findings supporting our claims. Finally, we depict in Section 5 the conclusions drawn and our future work.

2 Robust Location Fingerprint

Our research focuses on providing cost-effective location estimation in the indoor environment, utilizing existing infrastructure. Due to widespread availability of Wi-Fi and Bluetooth networks within buildings, we choose both these RF wireless technologies for our analysis and experiments.

RSS is the most common RF signal parameter used as location fingerprints for Wi-Fi since it was first proposed in [7]. For Bluetooth, both "Received Signal Strength Indicator (RSSI)" and "Link Quality (LQ)" have been previously used as location fingerprints; but generally, positioning systems that are solely based on Bluetooth have reported poor accuracy. A detailed analysis of the available Wi-Fi and Bluetooth signal parameters can be found in [8] and [9], respectively. Based on their analysis, it is apparent that all these signal parameters have specific usage according to their own respective technologies, which may render them inappropriate as location fingerprints. Among all the signal parameters available, RSS is argued to be the most viable option as location fingerprint for both Wi-Fi [8] and Bluetooth [9]. In this section, we deduce our location fingerprint, the SSD, and analytically prove its superiority over RSS, in terms of the system's overall robustness against heterogeneous mobile devices.

In existing localization literature based on location fingerprints, the signal strength samples are either collected at the APs, or at the MN that needs to be located. The AP-based approach has the advantage of not requiring any modification of the MNs' devices before the latter can be tracked. On the other hand, the MN-assisted approach could better ensure the security and privacy of the MN. In both approaches, the samples' signal strength values collected over a small time-window are generally averaged to obtain the traditional RSS location fingerprint.

The RSS location fingerprint is influenced by a particular transmitter-receiver pair's hardware-specific parameters, such as antenna gains. Consequently, having a different transmitter-receiver pair compared to the training phase would likely produce a different RSS signature at the same location [4].

In this section, we show that, rather than utilizing the absolute signal strength (i.e., RSS) as location fingerprint, the differences of signal strengths perceived at the APs or at the MN would actually provide a more stable location signature for any mobile device irrespective of its hardware used. We contend that, in this way, the transmitter-receiver pair's hardware effect is mitigated.

Suppose $P(d)$ and $P(d_0)$ denote the received signal strengths at an arbitrary distance d and a close-in reference distance d_0 from the transmitter, respectively, for a particular transmitter-receiver pair. From the log-normal shadowing model [10], we get

$$\left[\frac{P(d)}{P(d_0)}\right]_{\mathrm{dB}} = -10\beta \log\left(\frac{d}{d_0}\right) + X_{\mathrm{dB}}. \tag{1}$$

The first term on the Right Hand Side (RHS) of (1) defines the path loss component (β is the path loss exponent), while the second term reflects the variation of the received power at a certain distance ($X_{\mathrm{dB}} \sim N(0, \sigma_{dB}^2)$). Equation (1) can be rewritten as

$$P(d)|_{\mathrm{dBm}} = P(d_0)|_{\mathrm{dBm}} - 10\beta \log\left(\frac{d}{d_0}\right) + \mathrm{X}_{\mathrm{dB}}. \tag{2}$$

Depending on the hardware used at both the AP and the MN, the perceived power at a reference distance (i.e., $P(d_0)$) varies, as a result of hardware-specific parameters, such as antenna gains. Therefore, the perceived RSS at a distance d is also hardware-dependent. This explains why RSS is not a robust location fingerprint, although it is commonly used in the existing literature.

To simplify our discussion, let us first focus on the AP-based approach, where the MN is the transmitter, while the AP is the receiver. Rather than using absolute RSS values as location fingerprints, the difference of the RSS values observed by two APs (i.e., *SSD*) can be used to define a more robust signature for a transmitting mobile device. In order to explain analytically, let $P(d_1)$ and $P(d_2)$ denote the RSSs of a mobile device's transmitted signal as perceived at two different APs (AP_1 and AP_2) which are at distances d_1 and d_2 from the mobile device, respectively. We assume that, all the APs have the same hardware properties, since it is quite common for an institution to choose the same brand and model for all their APs in the building. Consequently, using (2), we can write the following for AP_1 and AP_2, respectively

$$P(d_1)|_{\mathrm{dBm}} = P(d_0)|_{\mathrm{dBm}} - 10\beta_1 \log\left(\frac{d_1}{d_0}\right) + [\mathrm{X}_1]_{\mathrm{dB}} \tag{3}$$

and

$$P(d_2)|_{\mathrm{dBm}} = P(d_0)|_{\mathrm{dBm}} - 10\beta_2 \log\left(\frac{d_2}{d_0}\right) + [X_2]_{\mathrm{dB}}. \tag{4}$$

Subtracting (4) from (3), we obtain

$$\left[\frac{P(d_1)}{P(d_2)}\right]_{\mathrm{dB}} = -10\beta_1 \log\left(\frac{d_1}{d_0}\right) + 10\beta_2 \log\left(\frac{d_2}{d_0}\right) + [X_1 - X_2]_{\mathrm{dB}}. \tag{5}$$

Equation (5) denotes SSD's expression, which is free from $P(d_0)$. Based on the above analysis, we claim that SSD is more robust against device hardware variations, compared to traditional RSS in denoting the location fingerprint

when the signal strength samples are collected at the APs. In the following sections, we explain it in a more detailed way. We also inspect the case of MN-assisted localization where the signal strength samples are actually collected at the MN. Here, we have applied the shadowing model of RF propagation in our analysis which is a common practice in existing indoor localization literature for the sake of analytical tractability [11], [12]. The shadowing model has also been used to model indoor RF propagation in popular "Wireless Communications" textbooks [10], [13]. Nevertheless, we also provide an alternative analysis using multipath propagation channel model in Appendix A, which can be found on the Computer Society Digital Library at http://doi.ieeecomputersociety.org/10.1109/TMC.2011.243, where SSD could also be shown to be free from the effect of variations in MN's hardware-dependent transmitter/receiver power gain for both AP-based and MN-assisted localization.

Note that, although it is common for Wi-Fi communication infrastructure in most campus and industrial buildings to have APs with the same brand and model, it is not a mandatory condition for our proposed SSD fingerprint to work in practice. As we will show in the following analysis, as long as each AP remains constant for both the training phase and the localization phase, the proposed scheme is able to eliminate the hardware differences caused by device heterogeneity.

2.1 Signal Strength Samples Collected at APs (AP-Based Approach)

Consider the same scenario as above but with the assumption that the reference power, i.e., $P(d_0)$ of (2), can be evaluated using the free space propagation model as follows [13]:

$$P(d_0)|_{\mathrm{dBm}} = 10\log\left(\frac{P_{\mathrm{MN}}G_{\mathrm{MN}}G_{\mathrm{AP}_i}\lambda_{\mathrm{MN}}^2}{16\pi^2 d_0^2 L}\right), \tag{6}$$

where P_{MN} is the MN's transmitted power, G_{MN} is the MN's antenna gain, G_{AP_i} is the ith AP's antenna gain, L is the system loss factor, and λ_{MN} is the transmitted carrier's wavelength.

Using (6), both (3) and (4) can be rewritten, respectively, as

$$\begin{aligned} P(d_1)|_{\mathrm{dBm}} &= 10\log\left(\frac{P_{\mathrm{MN}}G_{\mathrm{MN}}G_{\mathrm{AP}_1}\lambda_{\mathrm{MN}}^2}{16\pi^2 d_0^2 L_1}\right) \\ &\quad - 10\beta_1\log\left(\frac{d_1}{d_0}\right) + [X_1]_{\mathrm{dB}} \end{aligned} \tag{7}$$

and

$$\begin{aligned} P(d_2)|_{\mathrm{dBm}} &= 10\log\left(\frac{P_{\mathrm{MN}}G_{\mathrm{MN}}G_{\mathrm{AP}_2}\lambda_{\mathrm{MN}}^2}{16\pi^2 d_0^2 L_2}\right) \\ &\quad - 10\beta_2\log\left(\frac{d_2}{d_0}\right) + [X_2]_{\mathrm{dB}}. \end{aligned} \tag{8}$$

In order to compute SSD, subtract (8) from (7), we have

$$\begin{aligned} \left[\frac{P(d_1)}{P(d_2)}\right]_{\mathrm{dB}} &= 10\log\left(\frac{G_{\mathrm{AP}_1}L_2}{G_{\mathrm{AP}_2}L_1}\right) - 10\beta_1\log\left(\frac{d_1}{d_0}\right) \\ &\quad + 10\beta_2\log\left(\frac{d_2}{d_0}\right) + [X_1 - X_2]_{\mathrm{dB}}. \end{aligned} \tag{9}$$

The expression of SSD for the AP-based approach in (9) does not contain any MN-dependent term. Therefore, the SSD would be entirely free from any influence caused by the MNs' hardware variations. Moreover, even if different APs have different antenna gains and system loss factors, as long as these settings for each individual AP remain consistent across both training and localization, SSD will achieve consistency between the offline and online fingerprints.

2.2 Signal Strength Samples Collected at MN (MN-Assisted Approach)

We consider the same scenario as above, except that the signal strength is now measured at the MN rather than at the APs. Subsequently, (7) and (8) take the following forms, respectively,

$$\begin{aligned} P(d_1)|_{\mathrm{dBm}} &= 10\log\left(\frac{P_{\mathrm{AP}_1}G_{\mathrm{AP}_1}G_{\mathrm{MN}}\lambda_{\mathrm{AP}_1}^2}{16\pi^2 d_0^2 L_1}\right) \\ &\quad - 10\beta_1\log\left(\frac{d_1}{d_0}\right) + [X_1]_{\mathrm{dB}} \end{aligned} \tag{10}$$

and

$$\begin{aligned} P(d_2)|_{\mathrm{dBm}} &= 10\log\left(\frac{P_{\mathrm{AP}_2}G_{\mathrm{AP}_2}G_{\mathrm{MN}}\lambda_{\mathrm{AP}_2}^2}{16\pi^2 d_0^2 L_2}\right) \\ &\quad - 10\beta_2\log\left(\frac{d_2}{d_0}\right) + [X_2]_{\mathrm{dB}}. \end{aligned} \tag{11}$$

In order to compute SSD in this scenario, subtract (11) from (10), we have

$$\begin{aligned} \left[\frac{P(d_1)}{P(d_2)}\right]_{\mathrm{dB}} &= 10\log\left(\frac{P_{\mathrm{AP}_1}G_{\mathrm{AP}_1}\lambda_{\mathrm{AP}_1}^2 L_2}{P_{\mathrm{AP}_2}G_{\mathrm{AP}_2}\lambda_{\mathrm{AP}_2}^2 L_1}\right) \\ &\quad - 10\beta_1\log\left(\frac{d_1}{d_0}\right) + 10\beta_2\log\left(\frac{d_2}{d_0}\right) \\ &\quad + [X_1 - X_2]_{\mathrm{dB}}. \end{aligned} \tag{12}$$

Again, in the MN-assisted approach, the SSD is entirely free from the influence caused by MNs' hardware variations. Although the SSD expression is affected by different APs' configurations such as power settings, antenna characteristics, and operated channels, as long as the configuration for each individual AP remains consistent across both training and localization phases, the SSD will achieve consistency between the offline and online fingerprints.

Furthermore, even if the APs were to switch to different channels (e.g., changing from channel 1 to channel 11 for 802.11 g) from the training phase, the changes in the λ's of (12) will not be significant [14]. It should also be noted that, the samples gathered at the MN can be derived from the beacon frames that come from the APs [4]. Since these frames are generally sent using some default power setting, we can approximate that $P_{\mathrm{AP}_1} \approx P_{\mathrm{AP}_2}$.

Although the SSD is robust against device heterogeneity, an important tradeoff needs to be made when it is used to replace the RSS as a location fingerprint—the SSD fingerprint vector is always one dimension lower than the RSS fingerprint vector for the same number of APs. In order to understand this, suppose there are N APs within the range of a mobile device. Since each AP yields one RSS reading, the resulting RSS fingerprint vector has N elements. On the

other hand, although there are $\binom{N}{2}$ different SSD values resulting from the N RSS readings, only $(N-1)$ of these are independent. Hence, an SSD fingerprint vector contains only $(N-1)$ elements. The smaller dimensionality potentially puts it at a disadvantage compared to an RSS fingerprint vector (if all else remains the same). This implies that, if the same device were to be used for both training and online localization phases, then the use of RSS fingerprint vectors could yield better localization accuracy than SSD fingerprint vectors. Nevertheless, it was found in [11] that when N is large ($N > 5$), an increase in RSS fingerprint vector's dimensionality no longer results in any significant improvement of the localization accuracy. Therefore, the effect arising from the slightly smaller dimensionality of the SSD fingerprint vector should also become insignificant when N is large.

In many practical scenarios, a localization system is intended to track heterogeneous devices, and hence, we would expect the user devices to be frequently different from the training device. As reported in [1], [3], [4], and [5], different devices tend to report quite different RSS values at the same location. Under such circumstances, the use of RSS as a location fingerprint usually results in significant deterioration of the localization accuracy. The SSD, in contrast, is able to maintain its good localization accuracy across heterogeneous devices. As we will show in our experimental results in Section 4.4, with four APs, it is observed that the localization accuracy obtained from using SSD fingerprints is only slightly lower than using RSS fingerprints when the same device is used for both training and online localization phases. However, in the more practical case in which different devices are used for training and online localization respectively, the SSD outperforms RSS significantly even though the SSD fingerprint vector has a smaller dimensionality.

3 Related Work

The current research efforts for indoor positioning systems can largely be divided into two main categories: 1) those that require specialized hardware (e.g., RF tags, ultrasound receivers) and extensive deployment of dedicated infrastructure solely for localization purpose [15], [16], [17], and 2) those that utilize the location-dependency of easily measurable signal parameters (e.g., received signal strength). The latter approach aims to build a positioning system by leveraging on an existing infrastructure (e.g., Wi-Fi networks) [7], [18], [19], [20], [21], [22], [23], [24], [25] in a cost-effective way.

A few works in the second category above exploit RSS directly for distance estimation. One example is [19], in which distance and path loss exponent are jointly estimated in a least square approach. Another example is [26], which converts Signal Strength Differences into distance measurements using signal propagation model. Then a hyperbolic positioning algorithm [26] based on the distance differences is suggested to locate a Base Transceiver Station (BTS) in cellular communications environment. However, the robustness of these schemes in the presence of hardware heterogeneity is not addressed. In contrast, our work focuses on fingerprint-based positioning systems. Due to space limitation, we only provide an overview of some existing approaches under this category. More in-depth discussions can be found in [27] and [28]. Location fingerprinting techniques became popular with RADAR [7], mainly because of the unavailability of appropriate radio signal propagation models for indoor environments. It also opened the door for many different approaches to be applied for the indoor localization problem. For example, Nibble [18] is one of the first systems to use a probabilistic approach for location estimation. Local linearization technique and factor graph have been employed in [29] to model the mapping between RSS and location, based on the training data. To date, Ekahau's Positioning Engine [30] claims to be the most accurate location system with 1 metre average accuracy and a short training period. Note that, Ekahau is a commercial positioning system, and there exists no scientific evaluation to verify their claim.

The effects of different devices' hardware variations on RF location fingerprint have gained little attention in the localization literature so far. As discussed before, existing works generally use the same mobile device during both training and testing phases, thereby, invoking similar setups (i.e., transmitter-receiver pair) in both cases. However, [1], [3], [4], [5], [31], [32] have observed that the location fingerprints (i.e., RSSs) produced by using different mobile devices vary quite significantly from one another even under the same wireless conditions. Haebarlen et al. [4] try to accommodate various devices by having a benchmark training database taken with only one device. For other devices, they require a set of linear RSS conversion formulae, which translate the RSSs of those devices into the benchmark device's RSSs. These linear conversion formulae are obtained by laboriously experimenting with each supported device to discover its RSS relationship with that of the benchmark device. Kjærgaard [31] follows a similar approach, and also discusses a method to tune the parameters of the linear conversion formula automatically instead of manual calibration. In their subsequent work, Kjærgaard [32] identifies that the hearability problem (i.e., a mismatch of the set of APs that can be heard by different NICs) might affect the performance of their scheme, and suggests that the NIC used for collecting fingerprints (i.e., their benchmark device in the formula) should be the one that can hear the most APs. Tao et al. [3] utilize signal strength difference as a location fingerprint like our approach. Their motivation was to find the locations of rogue machines with different hardware configurations and varying transmitting powers. They have only provided experimental results based on the idea, without any intuition or analysis about why the differences in signal strengths could work successfully in their scenarios. On the contrary, our work gives both the detailed analysis and the experimental results as to why the SSD could be regarded as a *robust* location fingerprint, for both AP-based and MN-assisted localization approaches.

Around the same time when we proposed the use of SSD as a robust location fingerprint in [1], another work [5] also attempted to use a fingerprint that is related to signal strength differences. Specifically, it explored the use of normalized logarithmic signal-strength ratios, termed as

TABLE 1
The List of Wi-Fi and Bluetooth Devices Used as MN and AP in Our Experimental Testbeds

	Technology	MN Devices	AP Devices
Testbed 1	Wi-Fi	Intel PRO/Wireless 2200BG Atheros AR242x 802.11abg	Linksys WRT54G Cisco Aironet 1200
Testbed 2	Bluetooth	Ranger's BT-2100 (Class 1) Billionton's USBBT02-B (Class 2) Acer n300 PDA (Class 2) Motorola V3xx Phone (Class 2)	Ranger's BT-2100

Hyperbolic Location Fingerprint (HLF) [5] that is experimentally motivated. However, it does not provide any theoretical analysis as to why the HLF mitigates the hardware variation effects. Although the logarithm function is monotonic, the relationship between HLF and SSD is not as subtle as it seems. To appreciate the difference, one needs to take an analytical approach. The HLF approach is equivalent to taking the logarithm of (3) and (4), and then combining them. It can be easily seen that the resulting expression is not totally free from $P(d_0)$, unlike our SSD's expression in (5). Since heterogeneous devices are likely to have different $P(d_0)$, the HLF is unable to fully mitigate hardware variations from a theoretical standpoint. As will be demonstrated in our experimental results in Section 4.4.3, the SSD indeed outperforms the HLF when heterogeneous training and testing devices are used.

Another method, termed Ecolocation [6], uses ordered sequence of RSS measurements rather than the absolute RSSs to constitute a unique location fingerprint. If $P(d_i)$ and $P(d_j)$ denote the RSSs at AP_i and AP_j, which are at distances d_i and d_j from the MN, respectively, then a *constraint* of the sequence is defined as

$$P(d_i) > P(d_j) \Rightarrow d_i < d_j. \tag{13}$$

First, the constraint set for each grid point is calculated using the RHS of (13). Only the locations of *reference nodes* (i.e., APs) are required in this phase—no signal strength collection surveys are necessary. During location determination phase, the ordered sequence of RSSs collected at the APs is translated into the ordered sequence of distances using (13), and subsequently matched against the constraint set of each grid point calculated beforehand. The centroid of the grid points where the maximum number of constraints are matched is returned as the location estimate. We believe that, owing to MNs' hardware variations and varying transmission powers, both $P(d_i)$ and $P(d_j)$ should be affected in a similar way. Therefore, the constraint (13) is expected to remain intact over different MNs. Consequently, Ecolocation could be robust against MNs' hardware variations as well. Hence, we also compare SSD's performance with that of Ecolocation in our experiments.

4 Experimental Study

We first describe our experimental testbeds and data collection procedure in Sections 4.1 and 4.2, respectively. Then, we list in Section 4.3 some assumptions that we have made for our experiments. Finally, in Section 4.4, we present our results and findings.

4.1 Testbed Setup

We have two experimental testbeds. Table 1 lists the devices used in our testbeds.

Testbed 1 is a Wi-Fi testbed located inside a laboratory of our campus (see Fig. 1) that spans over an area of $382\ \mathrm{m}^2$. It has three separate rooms (divided by walls), where one is a discussion room, and the other two include many small cubicles.

We have mainly used Linksys WRT54G routers as our APs. The locations of these APs are marked as stars in Fig. 1. Additionally, we have also utilized measurements from the Cisco Aironet 1200 series routers for some of our experiments that provide wireless connectivity in the building. Note that, this testbed follows an *MN-assisted* approach where the MN itself retrieves the signal strength information.

Testbed 2 is a Bluetooth testbed located within another laboratory of our campus (see Fig. 2) that spans over an area of $214\ \mathrm{m}^2$, and includes many small cubicles. In this testbed, we have used four Aopen MP945 Mini PCs as our APs which are placed near the ceilings. The locations of these APs are marked as stars in Fig. 2. Each MP945 is equipped with BT-2100 Class 1 Bluetooth adapter which scans for Bluetooth packets by issuing inquiries periodically. This testbed emulates the *AP-based* positioning system where the signal strengths are actually measured at the AP side.

4.2 Data Collection Procedure

In our two testbeds, there are 466 and 337 training points or grids, respectively. The training process involves placing the mobile device at each training point, and collecting data. Our front-end of the signal strength collection program has a Java Graphical User Interface (GUI), which allows the user to load the map and click on the location to be trained conveniently. We have collected our measurements during afternoons over 10 working days. As mentioned before, the settings and surroundings of both testbeds include concrete walls, cubicles, movement of people, etc., and represent the indoor environment more practically compared to our initial testbed in [1].

In Testbed 1, we have utilized tcpdump [33] to capture the signal strength at the MN. We first put the MN's NIC into "monitor mode," and continuously cycle through the non-overlapping Wi-Fi channels 1, 6 and 11, where it stays on each channel for 10 ms. Concurrently, we run tcpdump

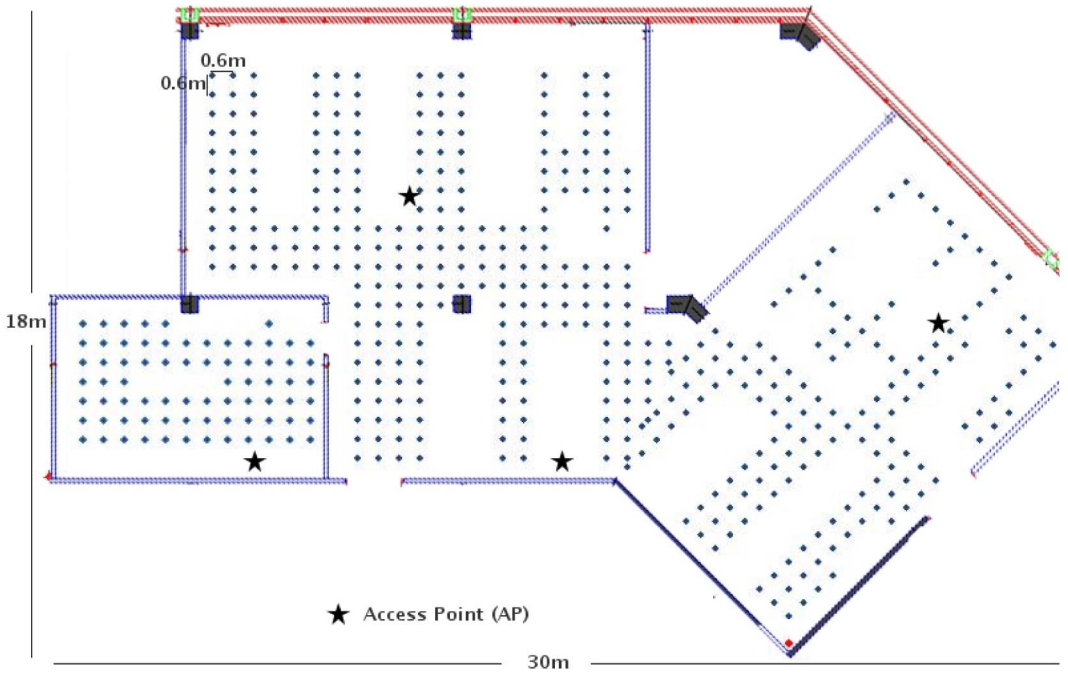


Fig. 1. Our Wi-Fi experimental testbed (MN-assisted localization)—all the training locations are marked as shaded points.

to snoop all the 802.11 packets from the air. Later on, we ran some scripting programs on the tcpdump's actual output to retrieve the required RSS information from our desired APs. In the case of Bluetooth, we log onto the mini PCs using Secure Shell (SSH) and make the APs issue Bluetooth inquiries which the mobile device responds to. The Bluetooth signal strength information retrieval program is written utilizing the HCI API of BlueZ [34] protocol stack. In each case (Bluetooth or Wi-Fi), the packet information is transferred to our central server's database from the APs (i.e., mini PCs) or the MN. The central server is also responsible for calculating the location during the testing phase. Our signal strength collection programs are invoked externally from the Java program when we click on the locations to be trained on the map. Note that, our Bluetooth adapters provide the absolute RSS values of the inquiry response packets, rather than the RSSI values as stipulated by the Bluetooth Core specification. At each location, stationing ourselves with an MN for 1 minute would give enough samples (200 to 300) in case of Wi-Fi for every AP, whereas for Bluetooth, we would have to stand for 2 or 3 minutes to gather the same number of samples. We aim to

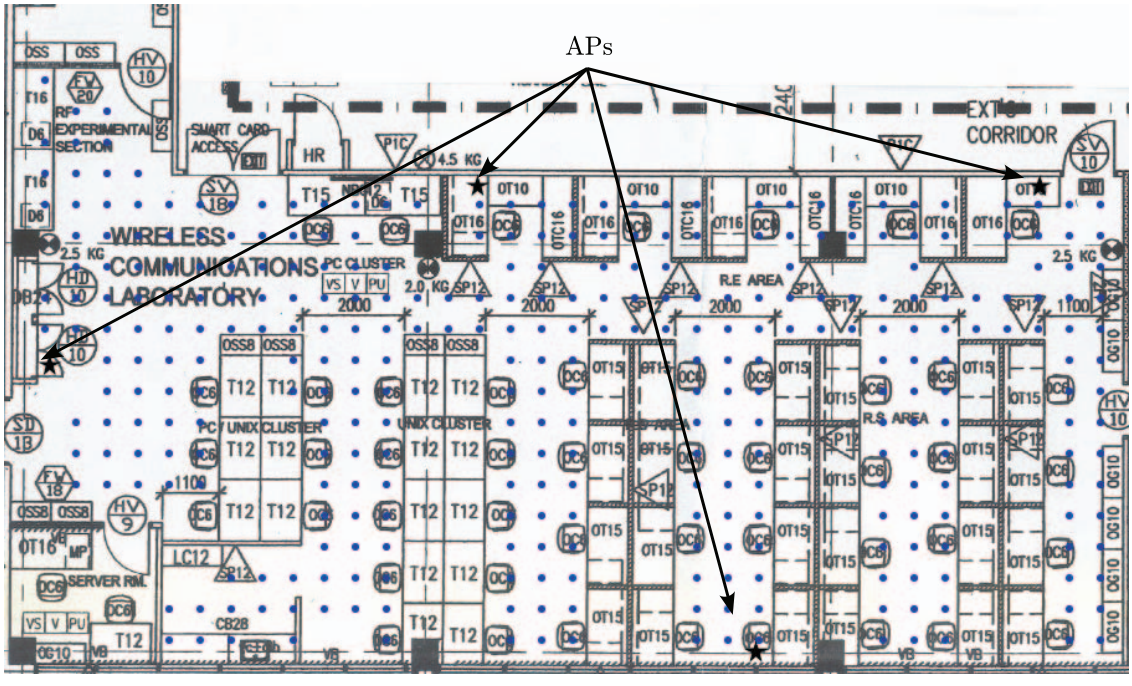


Fig. 2. Our Bluetooth experimental testbed (AP-based localization)—all the training locations are marked as shaded points.

take many samples at a particular location because we want to prove statistically that SSD is better in the experimental results. However, collecting 20 samples per location was observed to provide comparable results to even collecting 200 to 300 samples per location in case of SSD.

4.3 Assumptions

Here, we list the assumptions that we have made for our experiments.

1. Whenever we have used RSS as location fingerprint for certain experiments, we have assumed it to be normally distributed at any particular location in our paper. Though some works defy this phenomenon, others lend support to it [35]. We assume each RSS value in the location fingerprint to be a normal random variable characterized by only its mean and standard deviation. Similar to [4], our experimental results also suggest that it is a reasonable approximation, as significant improvement cannot be achieved even if we were to utilize histogram representations of RSS. However, we have used the histogram representation for HLF and the histogram's bin size is selected to be 0.02 as suggested by Kjærgaard and Munk [5].
2. We have chosen two well-known algorithms in the localization literature, namely, K Nearest Neighbor (KNN) [7] and Bayesian Inference [4], in order to test our ideas. Our key intention is to show that our ideas are quite generic and can be helpful irrespective of the choice of algorithms. For the KNN algorithm, we chose the value of K empirically, similar to prior works [7]. While applying Bayes formula, the priori probabilities are assumed to be uniformly distributed.
3. In order to apply probabilistic models, one assumption that has widely been used is the independence of RSS values of different APs [18], [21]. This assumption is justifiable for a well-designed network where each AP runs on a non-overlapping channel. Kaemarungsi and Krishnamurthy have performed experiments in [35] to evaluate the correlation factor among the APs' RSS values in the presence of interference and they have strengthened this claim as well. Thus, we have also adopted their vindication.

4.4 Experimental Results and Findings

4.4.1 Justification of SSD as a Robust Fingerprint

For this experiment, we have chosen various mobile devices which are listed in Table 1 to inspect their effects on both RSS and SSD location fingerprints. In Testbed 1, we conducted the signal strength survey by plugging two different Wi-Fi NICs (Intel PRO/Wireless 2200BG and Atheros AR242x 802.11abg) into our laptop. Since our Testbed 1 emulates the MN-assisted localization scenario, we actually collected the signal strength samples at the MN rather than at the APs. In Testbed 2, we have selected four different Bluetooth devices and measured their signal strengths at the APs (i.e., mini PCs). The Acer n300 PDA and the Motorola V3xx phone have integrated Class 2 Bluetooth chips, whereas the USBBT02-B Class 2 adapter and Ranger's BT-2100 Class 1 adapter were plugged into a laptop during the experiments. In both testbeds, we have picked 20 random training points and stationed the devices at those locations, while ensuring that we have collected enough samples at the MN (Testbed 1) or APs (Testbed 2) for all the devices.

Fig. 3a shows the RSSs as perceived by two different Wi-Fi NICs (i.e., MNs) from packets transmitted by a Linksys WRT54G router (i.e., AP) while Fig. 3c depicts the SSDs between two such routers' signals as perceived by the two NICs. Only Linksys routers' data have been presented here for brevity, as Cisco routers' measurements have yielded similar trends. Fig. 3b shows the RSSs as perceived by a Bluetooth adapter (i.e., AP) from packets transmitted by several Bluetooth devices (i.e., MNs) while Fig. 3d depicts the SSDs of these MN devices' transmitted packets as perceived by two Bluetooth APs.

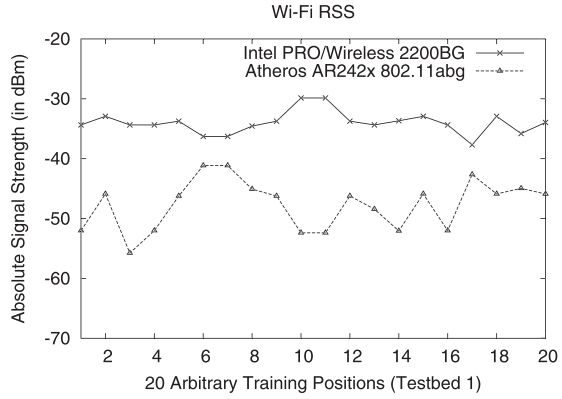
From our experimental results, we see that, the RSS perceived by a certain MN (MN-assisted) or AP (AP-based) varies significantly across different mobile devices at each training location. This has repercussion in their use as fingerprints because the RSS fingerprint vectors collected during the training phase will be strongly dependent on the mobile device used. Most existing works perform both their training and testing phases using the same device, thereby, ignoring this practical issue. On the contrary, the SSD remains quite consistent across different mobile devices in our experiments. This readily complies with our analysis in Section 2.

4.4.2 Comparison between SSD and RSS as Location Fingerprint

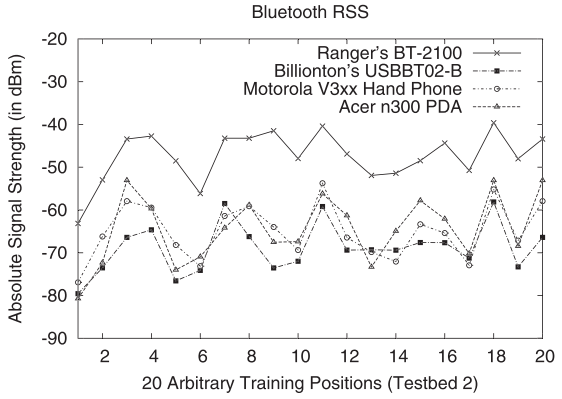
As pointed out in the previous section, the use of the same MN for both training and testing phases may have biased the reported results of the existing fingerprinting techniques. To investigate further, we conducted experiments inside both our MN-assisted Wi-Fi (Testbed 1) and AP-based Bluetooth (Testbed 2) testbeds to visualize the effects of MN's hardware variations.

In order to inspect the "same device" effect, we utilized Intel's NIC for both training and testing phases in Testbed 1. Among the 466 training grids as shown in Fig. 1, 200 of them are selected randomly as training points while the remaining 266 are kept for testing purpose. We then run our algorithms (i.e., KNN and Bayesian) to obtain the localization errors. We repeat this procedure for 101 times in order to obtain all the errors for different combinations of training and testing samples, and finally obtain the cumulative probability graph of Fig. 4a. In Testbed 2, we utilized Ranger's BT-2100 Class 1 adapter for both training and testing phases. In this particular testbed, 200 of the 337 training grids as shown in Fig. 2 are selected randomly as training points, while the remaining 137 are kept for testing purpose. We follow a similar approach as the one described for Testbed 1 in order to obtain the cumulative probability graph of errors in Fig. 4b. Note that, only Bayesian algorithm's results are presented here for brevity. The results obtained using the KNN algorithm have demonstrated similar trends.

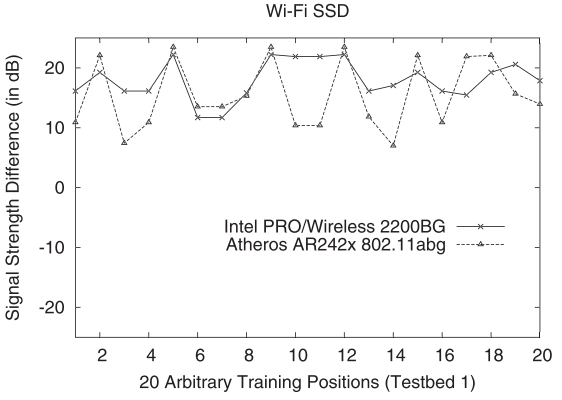
In order to inspect the "different device" effect, we utilized two different Wi-Fi NICs as listed in Table 1 for Testbed 1. The Intel NIC's collected data at 466 grids as



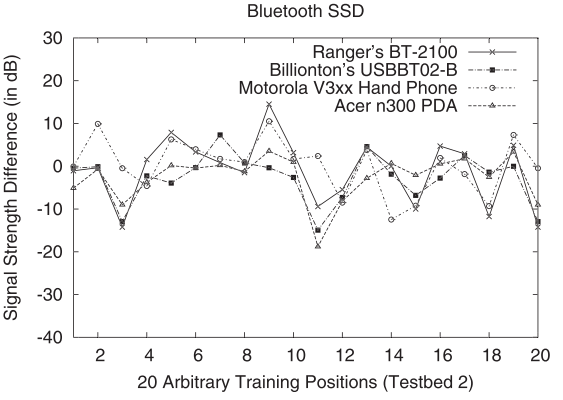
(a) RSS of an Wi-Fi AP's signals as perceived by each MN (MN-assisted).



(b) RSS of each Bluetooth MN's signals as perceived by an AP (AP-based).



(c) SSD between two Wi-Fi APs' signals as perceived by each MN (MN-assisted).



(d) SSD of each Bluetooth MN's signals as perceived by two APs (AP-based).

Fig. 3. Comparison between RSS and SSD for both Wi-Fi and Bluetooth considering various mobile devices.

shown in Fig. 1 are kept as training data while the Atheros NIC's collected data at 244 of the 466 grids are utilized for testing purpose. In Testbed 2, we have utilized four different Bluetooth devices as listed in Table 1 for collecting measurements at the 337 locations as shown in Fig. 2. We set aside Ranger's BT-2100 Class 1 adapter's data set as our training samples, while the remaining $(3 \times 337) = 1{,}011$ samples from the other three Class 2 devices are used for testing. The resulting cumulative probability graphs of localization errors are shown in Figs. 4c and 4d for Wi-Fi and Bluetooth, respectively.

The error performance when using the same device for both training and testing can be visualized in Figs. 4a and 4b for Wi-Fi and Bluetooth, respectively. In this case, the RSS-based algorithms perform slightly better than its SSD counterparts. As explained earlier in Section 2, the SSD fingerprint vector has a smaller dimensionality compared to the RSS fingerprint vector (3 versus 4 for the case of 4 APs). This puts SSD at a slight disadvantage when the same device is used for both training and online localization. Moreover, one may also argue that SSD has higher variance than RSS. Using (2) and (5), and assuming that X_1 and X_2 are independent and identically distributed Gaussian with variance σ^2_{dB}, RSS and SSD are distributed as

$$N\left(P(d_0)|_{\mathrm{dBm}} - 10\beta \log\left(\frac{d}{d_0}\right), \sigma^2_{\mathrm{dB}}\right)$$

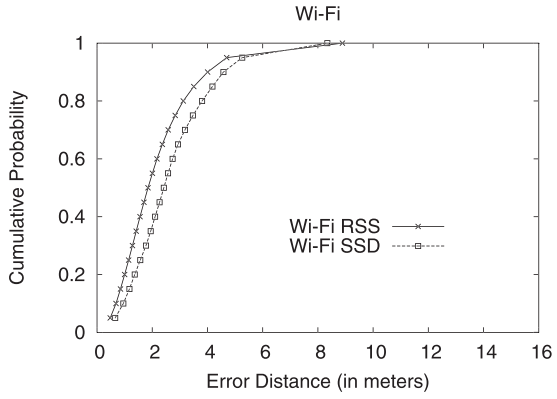
and

$$N\left(-10\beta_1 \log\left(\frac{d_1}{d_0}\right) + 10\beta_2 \log\left(\frac{d_2}{d_0}\right), 2\sigma^2_{\mathrm{dB}}\right),$$

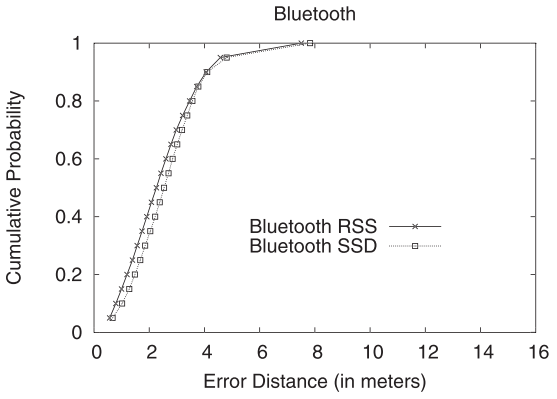
respectively. For the same device, we notice that the means of both RSS and SSD do not change, and the variance of RSS is actually lower than that of SSD.

However, in practical scenarios, a localization system is usually intended to track heterogeneous devices, and hence, the better performance of RSS only occurs occasionally when the user device happens to be the same as the device used for training. In practice, it is more often for the users to carry different devices from the training device. It can be easily seen from the Gaussian approximations of RSS and SSD that the mean of RSS varies depending on different MNs' hardware since it includes $P(d_0)$, while SSD's mean still remains the same. As we will see, the practical hardware dependency issue overshadows the disadvantage of the larger variance and smaller dimensionality of the SSD fingerprint, based on our experimental results shown below, using commonly found commercial devices.

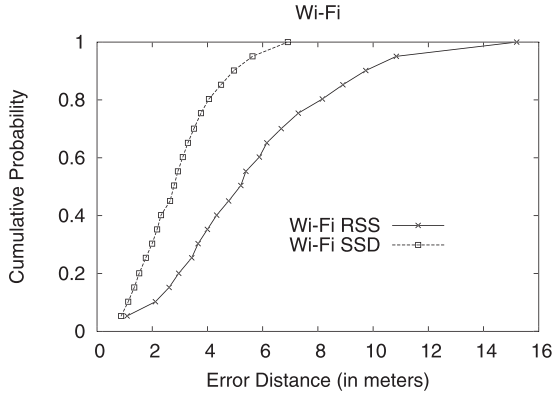
Let us investigate the more common scenario, where the user devices are different from the training device. From Figs. 4c and 4d, it is apparent that the hardware variations of the MN have adverse effects on the RSS-based localization's performance for both Bluetooth and Wi-Fi. We further notice that, this issue is prevalent regardless of



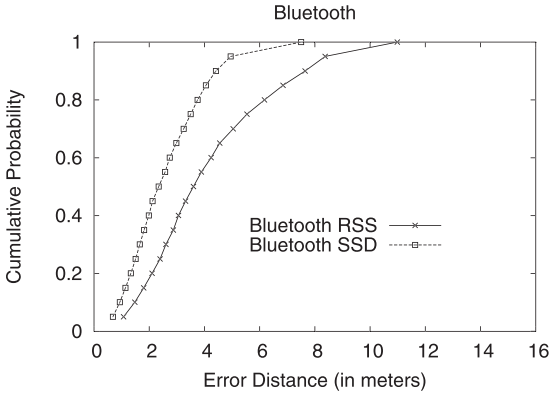
(a) Bayesian algorithm's performance (MN-assisted, same device for both training and testing phases).



(b) Bayesian algorithm's performance (AP-based, same device for both training and testing phases).



(c) Bayesian algorithm's performance (MN-assisted, different devices for training and testing phases).



(d) Bayesian algorithm's performance (AP-based, different devices for training and testing phases).

Fig. 4. Comparison of error performance using RSS versus SSD as location fingerprint for Wi-Fi and Bluetooth when the testing phase is conducted with either the *same* training device ((a) and (b)) or a *different* training device ((c) and (d)).

whether the RSS is measured at the APs for AP-based localization, or at the MN for MN-assisted localization. On the contrary, SSD based localization has much better accuracy than RSS-based localization in the presence of hardware variations in both our Wi-Fi and Bluetooth experiments (see Figs. 4c and 4d). It is also noteworthy to compare Fig. 4a against Fig. 4c, as well as Fig. 4b against Fig. 4d. As can be seen, the accuracy of SSD based localization remains almost the same in the respective comparisons. This implies that SSD-based localization is invariant to the mobile device being used, regardless of whether it is the same as the training device or not. This agrees with our analysis in Section 2 that SSD is free from hardware-dependent effects.

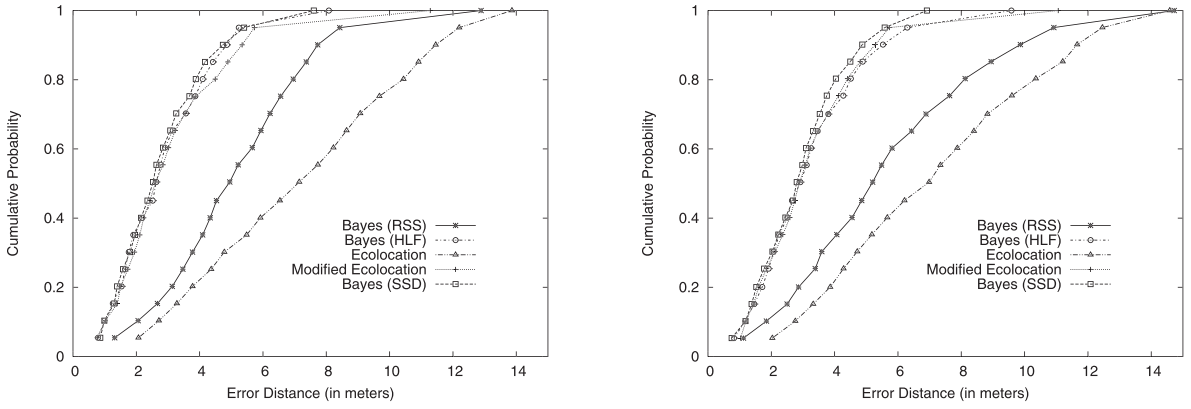
4.4.3 Comparison of SSD with Other Robust Location Fingerprints

In order to compare SSD with other robust location fingerprints, we consider two different combinations of training and testing data for both Wi-Fi and Bluetooth.

Among the 466 locations of our Wi-Fi testbed, we first keep the Atheros chipset's data collected at 244 locations as training samples. The Intel chipset's data collected at the remaining 222 locations are used for testing purpose. For the second combination, we just swap our training and testing data. In other words, the 222 locations' Intel data are used as training samples whereas the 244 locations' Atheros data are kept for testing. The results of the experiments can be visualized in Figs. 5a and 5b, respectively. Some numerical values (e.g., percentiles and average) of these two figures are listed in Table 2. The training and testing locations of both combinations are chosen in a way that they are spread uniformly over the whole testbed.

For our Bluetooth testbed, we first keep the Class 1 device's data collected at 171 among the 337 locations as training samples, whereas the three Class 2 devices' data collected at the remaining 166 locations are used for testing purpose. For the second combination, the 166 locations' Class 2 devices' data are used as training samples whereas the 171 locations' Class 1 device's data are kept for testing. The results of the experiments can be visualized in Figs. 6a and 6b, respectively. Some numerical values of these two figures are also listed in Table 3. Similar to our Wi-Fi testbed, the training and testing locations of both combinations are chosen in a way that they are spread uniformly over the whole testbed.

For the case of Wi-Fi, it is evident from Fig. 5 and Table 2 that, SSD based techniques are better than the other two schemes (HLF and Ecolocation) described in Section 3 that could also mitigate the MNs' hardware variation effects to some extent. Similar conclusions could also be drawn from our Bluetooth experimental results, as can be seen from



(a) Training dataset: Atheros, Testing dataset: Intel.

(b) Training dataset: Intel, Testing dataset: Atheros.

Fig. 5. Bayesian (RSS, HLF, and SSD are used as location fingerprints) and Ecolocation algorithm's performance for MN-assisted localization (Wi-Fi).

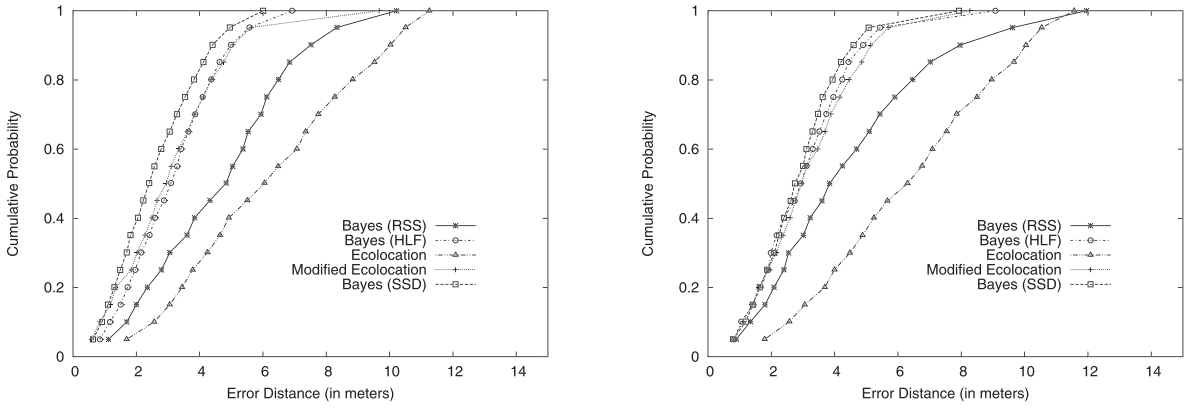
Fig. 6 and Table 3. Although we have utilized both KNN and Bayesian algorithms for performance comparison, we only show the figures for the Bayesian algorithm's results here for brevity. However, the numerical results from both Bayesian and KNN algorithms are listed in Tables 2 and 3 for Wi-Fi and Bluetooth, respectively.

As explained earlier in Section 3, the HLF relies on normalized logarithmic signal-strength ratios for location fingerprinting, which is shown analytically to be still vulnerable to MN's hardware heterogeneity. This explains why our SSD-based algorithms perform better than the HLF-based algorithms. Nevertheless, the HLF-based algorithms are still comparably more robust than the RSS-based algorithms.

Ecolocation performs even worse than the RSS-based algorithms for both our Wi-Fi and Bluetooth experiments. This can be attributed to the following reasons: 1) Ecolocation is mainly targeted at localizing inexpensive sensors and is shown to perform better than other localization algorithms found in wireless sensor networks [6]. Its main advantage lies in the fact that it requires no time-consuming signal strength collection surveys in the location space, whereas all the other algorithms considered in our experiments require the use of offline training data. 2) RSS measurements cannot be translated into distances accurately in the real world. Therefore, uncertainties could arise while using (13) as discussed in [6]. Moreover, since we only have four APs in each testbed, the number of

TABLE 2
Percentile Values and Averages of Errors (in Meter) when Various Fingerprints Are Considered for Wi-Fi (MN-Assisted)

Training Dataset: Atheros, Testing Dataset: Intel				
Algorithm (Fingerprint)	25^{th} Percentile	Median	90^{th} Percentile	Average
KNN (RSS)	3.42	5.04	7.90	5.06
KNN (HLF)	1.61	2.61	5.34	2.92
KNN (SSD)	1.60	2.58	5.15	2.86
Bayes (RSS)	3.47	4.95	7.72	5.04
Bayes (HLF)	1.64	2.63	4.86	2.80
Bayes (SSD)	1.58	2.53	4.74	2.73
Ecolocation	4.36	7.15	11.46	7.08
Modified Ecolocation	1.74	2.61	5.34	3.01
Training Dataset: Intel, Testing Dataset: Atheros				
KNN (RSS)	3.28	4.83	8.33	4.93
KNN (HLF)	1.81	2.72	5.18	2.99
KNN (SSD)	1.78	2.71	5.00	2.91
Bayes (RSS)	3.39	5.20	9.87	5.54
Bayes (HLF)	1.92	2.92	5.53	3.13
Bayes (SSD)	1.78	2.80	4.87	2.88
Ecolocation	4.28	6.98	11.66	6.99
Modified Ecolocation	1.85	2.90	5.29	3.10



(a) Training dataset: Class 1, Testing dataset: Class 2.

(b) Training dataset: Class 2, Testing dataset: Class 1.

Fig. 6. Bayesian (RSS, HLF, and SSD are used as location fingerprints) and Ecolocation algorithm's performance for AP-based localization (Bluetooth).

constraints (i.e., $\binom{4}{2}$) at each grid point is also quite limited. For fairer comparisons with the other schemes, we modify Ecolocation by making use of the offline training data, and call the resulting scheme "Modified Ecolocation." The constraint set for each grid point of the modified algorithm consists of the ordered sequence of RSS values collected during the training phase instead of the distance constraints as discussed in Section 3. The ordered sequence of RSSs collected during the online localization phase is now directly compared with each grid point's constraint set without the need for translation into distance constraints using (13). The experimental results show that the performance of our modified Ecolocation is significantly better than the original Ecolocation scheme, and it also outperforms the RSS-based algorithms. However, its performance is still inferior to our SSD-based algorithms.

5 Conclusions and Future Work

In this paper, we define a robust location fingerprint, the SSD, which provides a more robust location signature compared to the traditional RSS in the presence of mobile node hardware heterogeneity. Both our theoretical analysis and experimental studies have shown that, regardless of whether the signal strength samples are collected at the APs (AP-based localization) or at the MN (MN-assisted localization), SSD-based localization algorithms outperform those based on the traditional RSS fingerprints, as well as

TABLE 3
Percentile Values and Averages of Errors (in Meter) when Various Fingerprints Are Considered for Bluetooth (AP-Based)

Training Dataset: Class 1, Testing Dataset: Class 2				
Algorithm (Fingerprint)	25^{th} Percentile	Median	90^{th} Percentile	Average
KNN (RSS)	1.92	3.15	6.19	3.43
KNN (HLF)	1.87	2.97	5.40	3.09
KNN (SSD)	1.44	2.33	4.83	2.62
Bayes (RSS)	2.79	4.84	7.52	4.61
Bayes (HLF)	1.97	3.10	4.99	3.09
Bayes (SSD)	1.49	2.41	4.41	2.57
Ecolocation	3.78	6.05	10.03	6.07
Modified Ecolocation	1.84	2.95	5.05	2.99
Training Dataset: Class 2, Testing Dataset: Class 1				
KNN (RSS)	2.04	3.16	5.94	3.43
KNN (HLF)	1.72	2.85	5.26	3.00
KNN (SSD)	1.61	2.55	4.63	2.70
Bayes (RSS)	2.40	3.84	7.96	4.37
Bayes (HLF)	1.91	2.93	4.89	2.97
Bayes (SSD)	1.87	2.75	4.60	2.84
Ecolocation	4.00	6.30	10.04	6.23
Modified Ecolocation	1.90	2.96	5.14	3.10

several other techniques that are designed to mitigate the effects of MNs' hardware variations. This conclusion could not be drawn in our early work in [1] where only AP-based analysis was carried out. In this paper, we also considered two different testbeds for Wi-Fi and Bluetooth which emulate MN-assisted and AP-based localization, respectively. The settings and surroundings of both testbeds represent an indoor environment more practically compared to our initial lecture theater testbed of [1] which only considered an AP-based localization approach.

We point out two future directions. First, although previous works on Bluetooth-based localization have largely provided discouraging results [36], or required the aid of additional wireless technologies [20], our experience with Bluetooth shows that it is a promising technology as well that requires more investigation. Second, more experiments could be conducted in testbeds with different setup and size to explore SSD's viability across different settings. Moreover, investigating the impact of testbed's grid size, and the sample collection procedure's effects (e.g., fewer samples at each grid) on our SSD-based algorithms could certainly provide interesting future work directions.

Acknowledgments

This work was published in part at IEEE MASS 2007 [1]. It was supported by National Research Foundation project grant NRF2007IDM-IDM002-069 on "Life Spaces."

References

[1] M. Hossain, H. Nguyen Van, Y. Jin, and W.-S. Soh, "Indoor Localization Using Multiple Wireless Technologies," *Proc. IEEE Int'l Conf. Mobile Adhoc and Sensor Systems (MASS),* http://www.ece.nus.edu.sg/stfpage/elesohws/mass07.pdf, Oct. 2007.

[2] H. Krim and M. Viberg, "Two Decades of Array Signal Processing Research: The Parametric Approach," *IEEE Signal Processing Magazine,* vol. 13, pp. 67-94, 1996.

[3] P. Tao, A. Rudys, A.M. Ladd, and D.S. Wallach, "Wireless LAN Location-Sensing for Security Applications," *Proc. Second ACM Workshop Wireless security (WiSe '03),* pp. 11-20, Sept. 2003.

[4] A. Haeberlen, E. Flannery, A.M. Ladd, A. Rudys, D.S. Wallach, and L.E. Kavraki, "Practical Robust Localization over Large-Scale 802.11 Wireless Networks," *Proc. ACM MobiCom,* pp. 70-84, 2004.

[5] M.B. Kjærgaard and C.V. Munk, "Hyperbolic Location Fingerprinting: A Calibration-Free Solution for Handling Differences in Signal Strength," *Proc. IEEE Sixth Ann. Int'l Conf. Pervasive Computing and Comm. (PerCom '08),* Mar. 2008.

[6] K. Yedavalli, B. Krishnamachari, S. Ravula, and B. Srinivasan, "Ecolocation: A Sequence Based Technique for RF Localization in Wireless Sensor Networks," *Proc. Fourth Int'l Symp. Information Processing in Sensor Networks (ISPN '05),* Apr. 2005.

[7] P. Bahl and V.N. Padmanabhan, "RADAR: An In-Building RF-Based User Location and Tracking System," *Proc. IEEE INFOCOM,* pp. 775-784, Mar. 2000.

[8] J. Bardwell, "A Discussion Clarifying Often-Misused 802.11 WLAN Terminologies," http://www.connect802.com/download/techpubs/2004/you_believe_D100201.pdf, 2011.

[9] M. Hossain and W.S. Soh, "A Comprehensive Study of Bluetooth Signal Parameters for Localization," *Proc. IEEE 18th Int'l Symp. Personal, Indoor and Mobile Radio Comm. (PIMRC),* http://www.ece.nus.edu.sg/stfpage/elesohws/pimrc07.pdf, Sept. 2007.

[10] T.S. Rappaport, *Wireless Communication - Principles and Practice.* Prentice Hall, 1996.

[11] K. Kaemarungsi and P. Krishnamurthy, "Modeling of Indoor Positioning Systems Based on Location Fingerprinting," *Proc. IEEE INFOCOM,* pp. 1012-1022, Mar. 2004.

[12] C. Chang and A. Sahai, "Estimation Bounds for Localization," *Proc. IEEE First Ann. Comm. Soc. Conf. Sensor and Ad Hoc Comm. and Networks (SECON '04),* pp. 415-424, Oct. 2004.

[13] S.R. Saunders, *Antennas and Propagation for Wireless Communication Systems.* John Wiley & Sons, 1999.

[14] M.S. Gast, *802.11 Wireless Networks: The Definitive Guide.* O'Reilly & Assoc., 2002.

[15] R. Want, A. Hopper, V. Falco, and J. Gibbons, "The Active Badge Location System," *ACM Trans. Information Systems,* vol. 10, no. 1, pp. 91-102, Jan. 1992.

[16] A. Ward, A. Jones, and A. Hopper, "A New Location Technique for the Active Office," *IEEE Personal Comm.,* vol. 4, no. 5, pp. 42-47, Oct. 1997.

[17] N. Priyantha, A. Chakraborty, and H. Balakrishnan, "The Cricket Location-Support System," *Proc. ACM MobiCom,* pp. 32-43, Aug. 2000.

[18] P. Castro, P. Chiu, T. Kremenek, and R.R. Muntz, "A Probabilistic Room Location Service for wIreless Networked Environments," *Proc. Third Int'l Conf. Ubiquitous Computing (UbiComp '01),* pp. 18-34, Sept. 2001.

[19] X. Li, "RSS-Based Location Estimation with Unknown Pathloss Model," *IEEE Trans. Wireless Comm.,* vol. 5, no. 12, pp. 3626-3633, 2006.

[20] Y. Gwon, R. Jain, and T. Kawahara, "Robust Indoor Location Estimation of Stationary and Mobile Users," *Proc. IEEE INFOCOM,* pp. 1032-1043, Mar. 2004.

[21] M.A. Youssef, A. Agrawala, and A.U. Shankar, "WLAN Location Determination via Clustering and Probability Distributions," *Proc. IEEE First Int'l Conf. Pervasive Computing and Comm. (PERCOM '03),* Mar. 2003.

[22] E. Elnahrawy, X. Li, and R.P. Martin, "The Limits of Localization Using Signal Strength: A Comparative Study," *Proc. IEEE First Ann. Comm. Soc. Conf. Sensor and Ad Hoc Comm. and Networks (SECON '04),* 2004.

[23] A. Ladd, K. Bekris, G. Marceau, A. Rudys, L. Kavraki, and D. Wallach, "Robotics-Based Location Sensing Using Wireless Ethernet," Technical Report TR02-393, Dept. of Computer Science, Rice Univ., 2002.

[24] T. Roos, P. Myllymki, H. Tirri, P. Misikangas, and J. Sievnen, "A Probabilistic Approach to WLAN User Location Estimation," *Int'l J. Wireless Information Networks,* vol. 9, pp. 155-164, 2002.

[25] R. Battiti, M. Brunato, and A. Villani, "Statistical Learning Theory for Location Fingerprinting in Wireless LANs," Technical Report DIT-02-0086, Universita di Trento, Dipartimento di Informatica e Telecomunicazioni, Oct. 2002.

[26] B.-C. Liu, K.-H. Lin, and J.-C. Wu, "Analysis of Hyperbolic and Circular Positioning Algorithms Using Stationary Signal-Strength-Difference Measurements in Wireless Communication," *IEEE Trans. Vehicular Technology,* vol. 55, no. 2, pp. 499-509, Mar. 2006.

[27] J. Hightower and G. Borriella, "Location Systems for Ubiquitous Computing," *Computer,* vol. 34, no. 8, pp. 57-66, 2001.

[28] K. Pahlavan, X. Li, and J. Maleka, "Indoor Geolocation Science and Technology," *IEEE Comm. Magazine,* vol. 40, no. 2, pp. 112-118, Feb. 2002.

[29] C.-T. Huang, C.-H. Wu, Y.-N. Lee, and J.-T. Chen, "A Novel Indoor RSS-Based Position Location Algorithm Using Factor Graphs," *IEEE Trans. Wireless Comm.,* vol. 8, no. 6, pp. 3050-3058, June 2009.

[30] Ekahau, http://www.ekahau.com, 2012.

[31] M.B. Kjærgaard, "Automatic Mitigation of Sensor Variations for Signal Strength Based Location Systems," *Proc. Second Int'l Workshop Location and Context Awareness,* 2006.

[32] M.B. Kjærgaard, "Indoor Location Fingerprinting with Heterogeneous Clients," *Pervasive and Mobile Computing,* vol. 7, no. 1, pp. 31-43, Feb. 2011.

[33] "Tcpdump/Libpcap Public Repository," http://www.tcpdump.org, 2012.

[34] BlueZ, "Official Linux Bluetooth Protocol Stack," http://www.bluez.org, 2012.

[35] K. Kaemarungsi and P. Krishnamurthy, "Properties of Indoor Received Signal Strength for WLAN Location Fingerprinting," *Proc. First Ann. Int'l Conf. Mobile and Ubiquitous Systems: Networking and Services (MobiQuitous '04),* pp. 14-23, 2004.

[36] A. Madhavapeddy and A. Tse, "A Study of Bluetooth Propagation Using Accurate Indoor Location Mapping," *Proc. Seventh Int'l Conf. Ubiquitous Computing (UbiComp '05),* pp. 105-122, Sept. 2005.

A.K.M. Mahtab Hossain received the BSc degree in computer science and engineering from the Bangladesh University of Engineering and Technology in 2003, the MEngg degree in computer science from the Asian Institute of Technology (AIT), Thailand, in 2005, and the PhD degree in electrical and computer engineering from the National University of Singapore in 2010. He joined the Internet Education and Research Laboratory (intERLab) of AIT as a research specialist in May 2010. His current research interests include indoor localization, mobile IP, and wireless ad hoc networks.

Yunye Jin received the BEng (Hons) and PhD degrees in electrical and computer engineering from the National University of Singapore in 2007 and 2012, respectively. He joined the Institute for Infocomm Research, A*STAR, Singapore, as a research scientist in 2011. His current research interests include indoor pedestrian tracking and wireless sensor networks.

Wee-Seng Soh received the BEng (Hons) and MEng degrees in electrical engineering from the National University of Singapore (NUS) in 1996 and 1998, respectively. In 1998, he was awarded the Overseas Graduate Scholarship by the National University of Singapore to study at Carnegie Mellon University, Pittsburgh, Pennsylvania, where he received the PhD degree in electrical and computer engineering in 2003. Since 2004, he has been with the Department of Electrical and Computer Engineering, National University of Singapore, where he is currently an assistant professor. Prior to joining NUS, he was a postdoctoral research fellow in the Electrical Engineering and Computer Science Department, University of Michigan. He has served on the technical program committees (TPC) of 20 conferences and also served as a TPC cochair for ICCS 2008 and WUnderNet 2011. He is currently serving as an area editor of *Computer Communications* (Elsevier). His current research interests include wireless networks, underwater networks, and indoor tracking/localization techniques. He is a member of the IEEE.

Hien Nguyen Van received the BEng (Hons) degrees in electrical and computer engineering from the National University of Singapore in 2007. Since 2008, he has been a doctoral student with the Computer Vision Lab (Center for Automation Research) at the University of Maryland, College Park. His current research interests include statistical video analysis, compressed sensing and its applications to pattern recognition.