Wireless Localization with Spatial-Temporal Robust Fingerprints

DANYANG LI, JINGAO XU, and ZHENG YANG, School of Software and BNRist, Tsinghua University
CHENSHU WU, Department of Electrical & Computer Engineering, University of Maryland, College Park
JIANBO LI, College of Computer Science and Technology, Qingdao University
NICHOLAS D. LANE, University College London and Bell Labs

Indoor localization has gained increasing attention in the era of the Internet of Things. Among various technologies, WiFi fingerprint-based localization has become a mainstream solution. However, RSS fingerprints suffer from critical drawbacks of spatial ambiguity and temporal instability that root in multipath effects and environmental dynamics, which degrade the performance of these systems and therefore impede their wide deployment in the real world. Pioneering works overcome these limitations at the costs of ubiquity as they mostly resort to additional information or extra user constraints. In this article, we present the design and implementation of ViViPlus, an indoor localization system purely based on WiFi fingerprints, which jointly mitigates spatial ambiguity and temporal instability and derives reliable performance without impairing the ubiquity. The key idea is to embrace the spatial awareness of RSS values in a novel form of RSS Spatial Gradient (RSG) matrix for enhanced WiFi fingerprints. We devise techniques for the representation, construction, and localization of the proposed fingerprint form and integrate them all in a practical system. Extensive experiments across 7 months in different environments demonstrate that ViViPlus significantly improves the accuracy in localization scenarios by about 30% to 50% compared with the state-of-the-art approaches.

CCS Concepts: • **Human-centered computing** → **Ubiquitous and mobile computing**; • **Networks** → **Location based services;**

Additional Key Words and Phrases: Indoor localization; RSS fingerprint; spatial gradient

ACM Reference format:
Danyang Li, Jingao Xu, Zheng Yang, Chenshu Wu, Jianbo Li, and Nicholas D. Lane. 2021. Wireless Localization with Spatial-Temporal Robust Fingerprints. *ACM Trans. Sen. Netw.* 18, 1, Article 15 (October 2021), 23 pages.
https://doi.org/10.1145/3488281

A preliminary version of this article appeared in the International Conference on Mobile Ad-hoc and Sensor Systems (IEEE MASS'18).

Authors' addresses: D. Li, J. Xu, and Z. Yang (corresponding author), School of Software and BNRist, Tsinghua University, 100084; emails: lidy19@mails.tsinghua.edu.cn, {xujingao13, hmilyyz}@gmail.com; C. Wu, Department of Electrical & Computer Engineering, University of Maryland, College Park, 20742; email: wucs32@gmail.com; J. Li, College of Computer Science and Technology, Qingdao University, 266071; email: lijianbo@188.com; N. D. Lane, University College London and Bell Labs, WC1E 6BT; email: niclane@acm.org.

1550-4859/2021/10-ART15 $15.00
https://doi.org/10.1145/3488281

1 INTRODUCTION

Wireless indoor localization has attracted significant research interest from both academia and industrial sides due to the popularity of mobile and ubiquitous computing. Among various solutions based on technologies like RFID [30, 38], visual images [47], and inertial sensors [14], WiFi-based localization systems have been widely studied and deployed as one of the most promising solutions for ubiquitous indoor localization [21, 23, 28, 31, 36, 46, 50, 55]. Built upon widely deployed WiFi infrastructure, WiFi-based localization systems usually employ the easily accessible **Received Signal Strengths (RSSs)** as location fingerprints. As a result, these systems are free of extra hardware or dedicated equipment, rendering them especially attractive for commercial and pervasive deployment. In addition to several start-up productions, the RSS fingerprint-based method has been incorporated in the positioning services of great companies like Google, Apple, Cisco, Huawei, and Baidu.

Fingerprint-based approaches generally consist of two stages. In the first training phase, RSS fingerprints are collected with location labels by site survey to form a fingerprint database (a.k.a radio map). Then, during the localization stage, a user is located by matching his/her fingerprint observation against the fingerprint database. Despite extensive research, however, this technology has not yet stepped in the prime time for wide deployment. Leading companies like Google and Baidu only integrate indoor map services in sporadic areas like large malls, airports, and museums. The primary hurdles are twofold: expensive training efforts of site survey and unreliable accuracy. While the former has recently been efficiently addressed by crowdsourcing-based approaches [28, 36, 41, 50, 51], the unstable accuracy still remains a critical drawback to its widespread adoption [25].

Fingerprint-based systems assume that RSS fingerprints are spatially unique and temporally stable for a specific location and could be leveraged as location fingerprints. Due to inherent wireless signal properties, however, RSS values suffer from severe uncertainty, especially in complex indoor environments on account of multipath effects and environmental changes, which dramatically impair spatial uniqueness and temporal stability. The resulting effects on RSS fingerprints are twofold:

- *Spatial ambiguity*: Fingerprints from distinct locations may be similar, and thus fingerprint mismatches would happen among distant locations. Spatial ambiguity is recognized as the root cause of large errors [23, 47] and may lead to large location errors of even up to 10 meters [25].
- *Temporal instability*: Location fingerprints would vary over time, deviating from and therefore failing to match the initially collected ones. Temporal instability further leads to gradually deteriorated performance over time [13, 42, 43].

In a nutshell, both of them degrade the performance of RSS fingerprinting and further prevent it from practical deployment.

Many pioneers have been addressing the above spatial and temporal shortcomings of RSS fingerprints. Major efforts usually overcome the spatial ambiguity by leveraging (1) *user mobility*: fingerprint ambiguity is reduced by user mobility, either by eliminating less likely candidates with inertial sensor hints [14, 34] or by constructing mobile forms of fingerprints that combine subsequent measurements along user movements [33, 53, 54]; (2) *extra ranging*: eliminating spatial ambiguity by geometry constraints gathered from acoustic ranging [23] or spatial images [47]; or (3) *CSI*: to avoid the drawbacks of RSS, some efforts are made to exploit physical layer information such as **Channel State Information (CSI)** for localization. Regarding temporal variations, periodic re-calibration should be applied to the fingerprint database, either by reconstruction or self-adaptation [13, 39, 43]. Despite the performance gains achieved, these methods also cause one or more of the following pains that hamper the superior ubiquity of fingerprint-based systems:

(1) require accurate user mobility, which is difficult to obtain by erroneous smartphone built-in sensors and cannot locate a stationary user; (2) impose constraints on user behaviors of moving or using smartphones; (3) rely on intentional cooperation among multiple users; (4) introduce extra information inputs beyond RSS measurements, such as digital floorplan, inertial data, images, and so forth, some of which are even not readily accessible from smartphones. In addition, existing approaches mostly only deal with one single aspect of either spatial ambiguity or temporal instability.

In this article, we present the design and implementation of ViViPlus, an indoor localization system purely based on WiFi fingerprints, which jointly mitigates spatial ambiguity and temporal instability and derives reliable performance without the limitations mentioned above. The key intuition is to embrace the spatial awareness of RSS values for enhanced WiFi fingerprint representation. In particular, the spatial relationships among the RSSs of one AP at multiple neighboring locations tend to be more robust than individual RSS values from one single location. In ViViPlus, we propose ***RSS Spatial Gradient*** **(RSG)** for fingerprinting, which depicts the RSS differences among a set of selected nearby locations. As a spatially relative form, the RSG matrix can better mitigate the fingerprint ambiguity due to multipath fading and temporal variations due to temporal dynamics like environmental changes, AP power adjustment, and so forth. For example, if one AP adjusts its transmission power, the RSS values observed at several selected adjacent locations would suffer identical or similar changes, keeping the corresponding RSG less affected or unaffected.

ViViPlus's design follows the classical fingerprinting framework and requires no more inputs than any previous RSS fingerprint-based systems. Translating the intuitive idea into a practical system, three challenges reside:

(1) How to represent an effective fingerprint form based on RSG? The key is to formulate a form that is compatible with the traditional RSS fingerprint database. We define an RSG Matrix for each location. Each row in the matrix depicts the RSG profile of one AP among several neighboring locations. To obtain an optimal matrix, we carefully select a subset of neighboring locations to form a reliable RSG profile.
(2) How to construct such an RSG matrix-based fingerprint for a single query? The construction of the RSG matrix needs RSS observations from multiple locations. However, in typical scenarios, a user only reports one single measurement from one location (or at most a sequence of measurements along a path for a moving user). Our trick is to reuse the data of a current candidate location to construct a query RSG matrix. By doing this, the RSG matrix would be similar to the candidates if they come from the same location and largely deviate from each other if not.
(3) How to efficiently leverage the RSG matrix-based fingerprints for accurate localization? Instead of comparing two matrices row by row, we conduct fingerprint comparison in a global way by extracting and matching the matrix features. Furthermore, we design an adversarial learning-based model that can remove the influence of *temporal instability* contained in the RSG matrix and further extract *dynamics-resistant* features.

To evaluate ViViPlus, we conduct comprehensive experiments in multiple different buildings with various conditions. We deploy ViViPlus in real business environments and continuously evaluate the system performance across 7 months. The results demonstrate that ViViPlus achieves reliable performance, with a mean accuracy of 2.13m and a 95th percentile accuracy of 4.91m, outperforming even the best among four comparative approaches by >27% and >20%, respectively. Even 7 months after the fingerprint database is established, the localization success rate of ViViPlus maintains >80%, outperforming other works by more than 15%. Our vision is to replace

previous RSS fingerprints with RSG matrix for WiFi-based location systems in a real-world deployment.

The core contributions are summarized as follows:

(1) We explore spatial awareness of WiFi signals from the perspective of *RSG* for localization. RSG exploits the underlying spatial features of RSSs from nearby locations, which better mitigates the spatial ambiguity and temporal instability than the original RSS fingerprints.
(2) We propose the RSG matrix to formulate a novel fingerprint form based on RSG features. Derived from the pure RSS fingerprint database, the RSG matrix is fully compatible and requires no more information inputs than any existing fingerprint-based systems. We design algorithms for the representation and construction of the proposed RSG matrix for efficient localization.
(3) We implement ViViPlus on smartphones and conduct extensive experiments in real environments for 7 months. Encouraging results demonstrate that ViViPlus achieves remarkable performance gain without the pains of resorting to additional information or restrictions to users.

The rest of the article is organized as follows. An overview is presented in Section 2. We introduce RSG in Section 3 and present the ViViPlus design in Section 4. Implementation and evaluation are conducted in Section 5. We review related works in Section 6 and conclude this work in Section 7.

2 OVERVIEW

2.1 Classical Fingerprinting Framework

The mainstream of existing approaches employ RSS observations at one location as its fingerprints [23, 28, 50, 51]. Such systems typically consist of two phases: In the first *offline training phase*, an RSS fingerprint database is constructed with locationally labeled fingerprints. In the *online localization phase*, location is determined by fingerprint matching.

Formally, the area of interest is sampled as a discrete location space $\boldsymbol{L} = \{\boldsymbol{l}_1, \boldsymbol{l}_2, \ldots, \boldsymbol{l}_n\}$, where n is the amount of sample locations. For each location $\boldsymbol{l}_i$ with coordinates (x_i, y_i), a corresponding fingerprint is denoted as $\boldsymbol{f}_i = \{f_{i1}, f_{i2}, \ldots, f_{im}\}$, $1 \le j \le m$, where f_{ij} denotes the RSS value (or the RSS distribution in probabilistic algorithms [55]) of the jth AP and m is the total number of APs in the targeted location space. All fingerprints form a fingerprint space $\boldsymbol{F} = \{\boldsymbol{f}_1, \boldsymbol{f}_2, \ldots, \boldsymbol{f}_n\}$, corresponding to the location space $\boldsymbol{L}$. The radio map consisting of $<\boldsymbol{l}_i, \boldsymbol{f}_i>$ terms is usually constructed either manually by site survey or automatically by crowdsourcing [28, 50]. Then location is estimated by retrieving the best matches of a query fingerprint $\boldsymbol{f}_q$ against $\boldsymbol{F}$, using some specific fingerprint similarity measures such as Euclidean distance:

$$\arg\min_{1 \le i \le n} \sqrt{\sum_{1 \le j \le m} (f_{ij} - f_{qj})^2}. \tag{1}$$

2.2 ViViPlus Overview

The design of ViViPlus follows the classical fingerprint framework, with no more inputs than any existing fingerprint-based systems. By doing this, we retain the elegant ubiquity of WiFi fingerprinting. ViViPlus contains three unique modules, i.e., RSG matrix database construction, query RSG matrix construction, and RSG matrix-based location estimation, as illustrated in Figure 1.

The RSG matrix database is constructed during the offline phase based on a conventional RSS fingerprint database, with neither mobility nor other information requirements. The RSG matrix

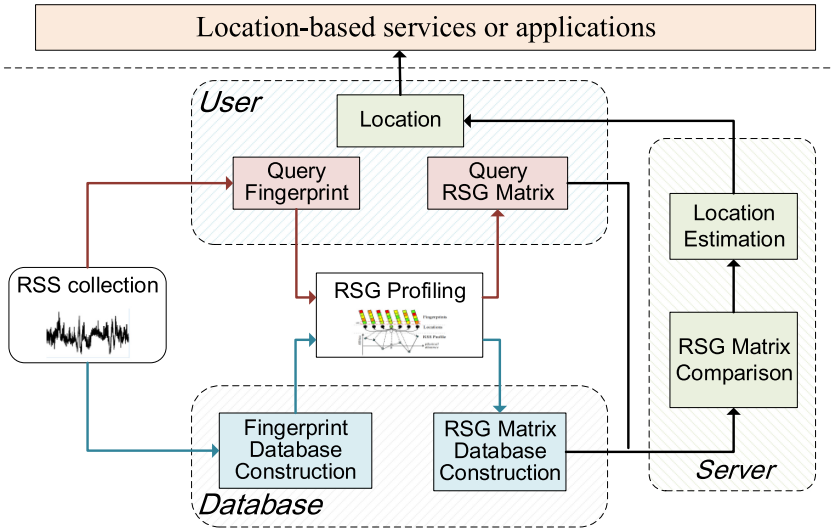


Fig. 1. Overview of ViViPlus workflow.

for one reference location is formed upon the RSS fingerprints from itself and its multiple neighboring locations.

In the localization stage, a user reports a query RSS fingerprint in the identical form as classical RSS fingerprints. Due to lack of location information and thus no neighboring locations, the query is then transformed into an RSG matrix depending on the information of a current candidate location. Specifically, when matching a user query $\boldsymbol{f}_Q$ against a candidate reference location $\boldsymbol{l}_C$, we construct an RSG matrix $G_{Q(C)}$ tailored to this candidate location by reusing its RSG information (nearby locations and corresponding fingerprints). Therefore, the query RSG matrix is customized for each candidate location. The ultimate location estimate of query $\boldsymbol{f}_Q$ is then determined by the top k locations that output the largest similarity between their RSG matrices.

The proposed ViViPlus is computation efficient. The relatively complex RSG matrix construction is a one-time effort during the offline phase. During the online phase, the extra computation costs compared to traditional RSS fingerprint-based approaches lie in the query RSG matrix construction, which can be fortunately computed in constant time. For practicality, we also implement an AP selection procedure to include only good-quality APs (scored by AP information entropy [5]) for localization. To further reduce the computation costs, in practice, we do not need to try every reference location but can efficiently shrink the search space by looking at common APs.

3 SPATIAL AWARENESS OF RSS FINGERPRINTS

3.1 Limitations of RSS Fingerprint

While spatial uniqueness and temporal stability are two fundamental assumptions of RSS fingerprinting, these two properties do not always hold in practice.

On one hand, multipath effects of wireless signals render similar RSSs over different locations of the same AP, resulting in the *spatial ambiguity* of RSS fingerprints, which means that fingerprints from distinct (and distant) locations may be similar to each other. Figure 2(a) shows the self-similarity matrix, using Euclidean distance, among each pair of fingerprints from 72 locations in a classroom. As seen, the most similar fingerprints for each query do not always appear at the true locations (on the diagonal line). In contrast, distant locations may possess more similar

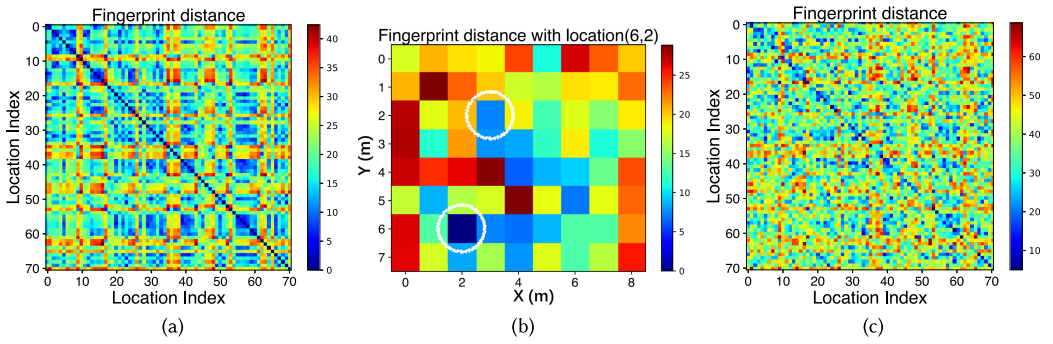


Fig. 2. Fingerprint spatial ambiguity and temporal instability: (a) Fingerprint distance matrix of 70 different locations in a room. While most of the similar fingerprints appear at the true location (on diagonal line), the top k most similar fingerprints may be far away. (b) Fingerprint distances of a specific location (6, 2) to all other locations. (c) Re-collected fingerprint distance matrix against those used in (a). For some queries, significant parts of the best-matched locations are not true locations. We index the 70 locations in an 8 m $\times$ 9 m room with an S-shaped snakelike manner in (a) and (c).

fingerprints. Taking location (6, 2) as an example, we calculate the fingerprint similarity to all 70 locations in the room and depict the results in Figure 2(b). The second most similar fingerprint appears at the location (2, 3), which is about 4m away from the true location (note that the width and the length of the classroom are only about 9 m). Spatial ambiguity is recognized as the root cause of large location errors in WiFi localization [23, 47].

On the other hand, RSS is sensitive to uncertain environmental dynamics due to severe multipath effects in complex indoor environments [21, 43, 55]. RSS variations may induce temporal instability; i.e., location fingerprints would vary over time, deviating from and therefore failing to match the initially collected ones. As shown in Figure 2(c), we re-collected the fingerprints of all locations in the same room on a different day and calculated the similarity with those previously collected. Compared with Figure 2(a), more locations cannot be correctly located using the newly collected fingerprints, which indicates that they have deviated from the original version due to temporal changes. As a result, temporal instability would gradually degrade the localization performance over time, especially during long-term deployment.

In complex indoor environments, spatial ambiguity and temporal instability are even severe yet inevitable due to multipath fading and temporal dynamics. Consequently, they become the major obstacle behind the limited accuracy and reliability of WiFi localization based on RSS fingerprints.

3.2 RSS Spatial Gradient

In contrast to previous WiFi fingerprinting that mainly employs RSS vectors as fingerprints, we propose RSG to explore and exploit spatial features of RSS fingerprints. The key insight is that certain spatial relationships among RSSs from multiple adjacent locations keep relatively stable, although their individual RSS might be altered by signal distortions. As shown in Figure 3, RSS differences between two neighboring locations are more stable than individual RSS values (with about 80% improvement), regardless of the AP's signal strengths. Therefore, we can seek a set of neighboring locations to form an RSG matrix as more favorable fingerprints for our advantages of building a more accurate and reliable fingerprinting scheme. Note that we do not assume that RSS observations are similar among neighboring locations as we explore the stability of RSS differences among them. However, we neither require that every pair of locations hold stable RSS differences.

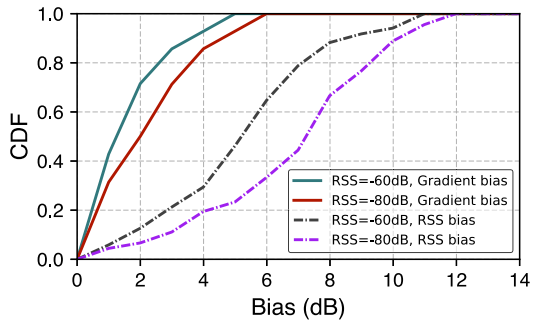


Fig. 3. We use the identical devices to collect RSS fingerprints simultaneously at different locations. RSS differences between neighboring locations tend to be more stable than individual RSS values regardless of the signal strength in different locations.

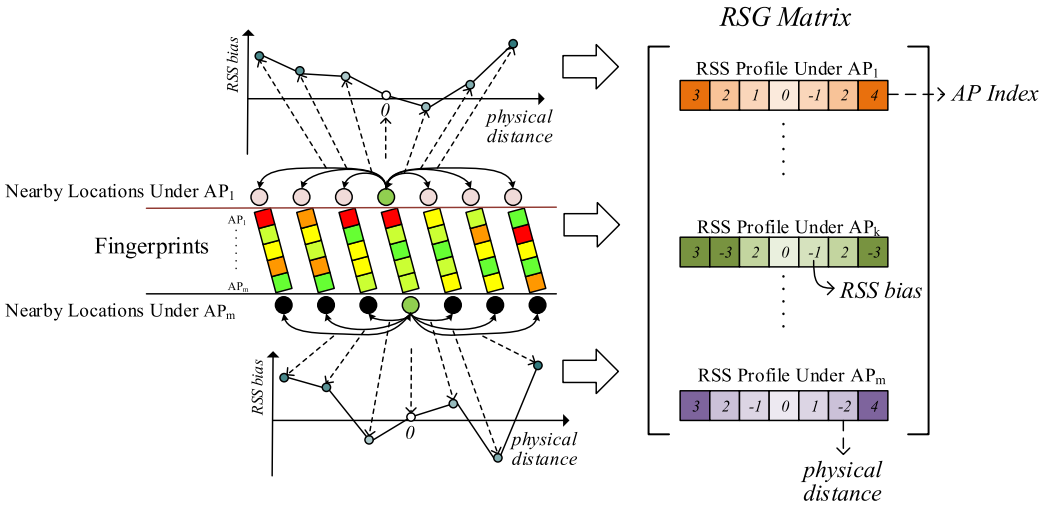


Fig. 4. Illustration of RSG matrix (of the central green sample location).

As demonstrated in the following, we only need to select a subset of neighboring locations that possess more stable RSS spatial relationships.

Benefiting multiple RSSs across different spatial locations, the resulting profiles also would be more distinguishable than a single RSS fingerprint. Hence, we explore and exploit such RSS spatial gradient as a more favorable feature than the widely adopted original RSS fingerprints for location mapping.

3.2.1 RSG Matrix Specification. An RSG matrix for a specific location depicts the RSS differences of every AP between itself and its neighboring locations, as shown in Figure 4. Specifically, for a location $\boldsymbol{l}_i$, the RSG matrix is defined as

$$G_i = (\vec{g_{i1}},\ \vec{g_{i2}},\ \ldots,\ \vec{g_{im}})^{\mathrm{T}}, \tag{2}$$

where m is the total number of APs that are selected for location $\boldsymbol{l}_i$. $\vec{g_{ik}}$ is a series of RSS differences of AP k between $\boldsymbol{l}_i$ and its $2r+1$ neighboring locations, which is defined as

$$\vec{g_{ik}} = \{<d(\boldsymbol{l}_i, \boldsymbol{l}_j), \phi(f_{ik}, f_{jk})>, i-r \le j \le i+r\}, \tag{3}$$

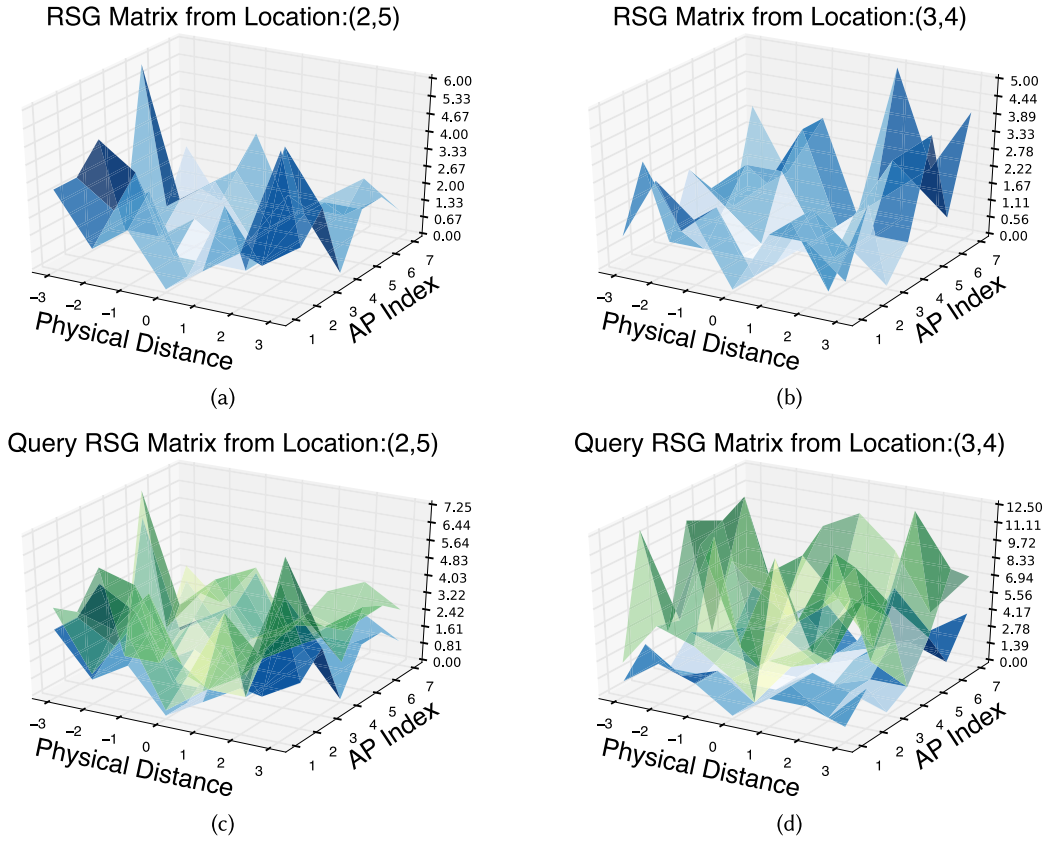


Fig. 5. Illustrations of RSG matrices at different locations, and we use negative and positive values to denote the relative physical distance between target location and reference location in different directions: (a) RSG matrix from location (2, 5). (b) RSG matrix from location (3, 4). (c) Query RSG matrix from (2, 5) generated upon the RSG matrix of candidate location (2, 5), which is similar to (a). (d) Query RSG matrix from (2, 5) built upon a different location (3, 4), which is significantly different from (b).

where $\phi(f_{ik}, f_{jk})$ is the RSS difference of AP k between location $\boldsymbol{l}_i$ and its neighboring location $\boldsymbol{l}_j$ (i.e., the respective kth item in their corresponding fingerprints $\boldsymbol{f}_i$ and $\boldsymbol{f}_j$). In the section, we simply calculate

$$\phi(f_{ik}, f_{jk}) = f_{ik} - f_{jk}, \tag{4}$$

and in Section 4.1.2 we will present a discrete definition to enhance the robustness. $d(\boldsymbol{l}_i, \boldsymbol{l}_j)$ denotes the physical distance between $\boldsymbol{l}_i$ and $\boldsymbol{l}_j$. The $2r + 1$ neighboring locations, i.e., $\boldsymbol{l}_{i-1}$ to $\boldsymbol{l}_{i-r}$ and $\boldsymbol{l}_{i+1}$ to $\boldsymbol{l}_{i+r}$, are selected from the surrounding subspace of $\boldsymbol{l}_i$ and ordered in physical distance to the current location $\boldsymbol{l}_i$. As a result, the RSG matrix G_i for a reference location $\boldsymbol{l}_i$ is a $m \times (2r+1)$ matrix, where the r + 1th column is all zeros (These will not necessarily be zeros when profiling a query fingerprint, as detailed later in Section 4). Figure 5(a) and 5(b) show two illustrative RSG matrices for two different locations, respectively.

3.2.2 RSG Matrix Superiority. We qualitatively analyze the advantages of the proposed RSG matrix to traditional RSS fingerprints regarding temporal stability and spatial uniqueness.

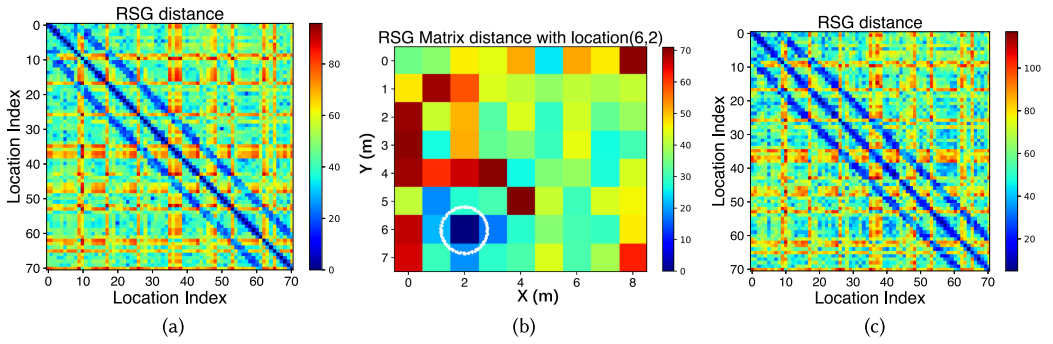


Fig. 6. RSG matrix mitigates spatial ambiguity and temporal instability: (a) Distance matrix of RSG matrix of a room. (b) RSG matrix distance for a specific location (6, 2). (c) Distance matrix of RSG matrix of re-collected fingerprint against the previous data used in (a). In both (a) and (c), the best-matched locations for each query almost always appear at close locations to the true locations (on the diagonal line and its two parallel lines).

First, the RSG matrix exploits the spatially differential RSSs among a set of nearby locations, which turns out to be more stable against temporal dynamics than previous RSS fingerprints formed by absolute RSS values observed at a single location [33, 43, 53]. Specifically, let f^1_{ik} and f^1_{jk} denote the RSS values of AP_k at two nearby locations $\boldsymbol{l}_i$ and $\boldsymbol{l}_j$ at time t_1, and f^2_{ik} and f^2_{jk} denote those measured at time t_2. Accounting for that temporal dynamics in the environments would be similar to neighboring locations, the RSS differences over time would also be likely similar, i.e., $|f^1_{ik} - f^1_{jk}| \approx |f^2_{ik} - f^2_{jk}|$, while their respective RSS value changes $|f^1_{ik} - f^2_{ik}|$ and $|f^1_{jk} - f^2_{jk}|$ could be significantly large, as shown in Figure 3.

Second, the RSG matrix is also more distinctively in space. The reasons are twofolds: (1) different from traditional RSS fingerprints that rely on the measurements solely from one single location, the RSG matrix synthesizes more information from multiple locations for location distinction and naturally possesses better spatial resolution; (2) as will be depicted in Section 4, our novel scheme for RSG matrix generation and comparison further improves the spatial uniqueness. To be brief, the RSG matrix for a user query is constructed upon the fingerprints of a candidate reference location and its selected neighbors. Therefore, the query RSG matrices of an identical query vary upon different candidate locations. Hence, the constructed RSG matrix would be similar to the reference matrix if the query comes from the same location as the reference one (Figure 5(a) and 5(c)); otherwise, they will be significantly different (Figure 5(b) and 5(d)). For two distant locations that suffer from spatial ambiguity under traditional RSS fingerprints, their RSG matrices would still be distinctive since they are unlikely to hold consistent gradients with the neighboring fingerprints even though their fingerprints are similar.

Figure 6 illustrates the advanced results for the same dataset as in Figure 2, yet using RSG matrix as fingerprints. Comparing Figure 2(a) with Figure 6(a) and Figure 2(b) with Figure 6(b), the spatial ambiguity is clearly reduced by using the RSG matrix. If we compare Figure 2(c) with Figure 6(c), one can also observe that the performance is well retained regarding temporal dynamics. Specifically, in both Figure 6(a) and 6(c), the three best-matched RSG matrices almost always appear at the three closest locations, as indicated in the three diagonal lines in the figures (note that the three lines are not close together because we index the 70 locations in an 8 m $\times$ 9 m room with an S-shaped snakelike manner).

The above superior properties lay a solid foundation for location distinction. In the following, we design and implement a full functional localization system based on the proposed RSG.

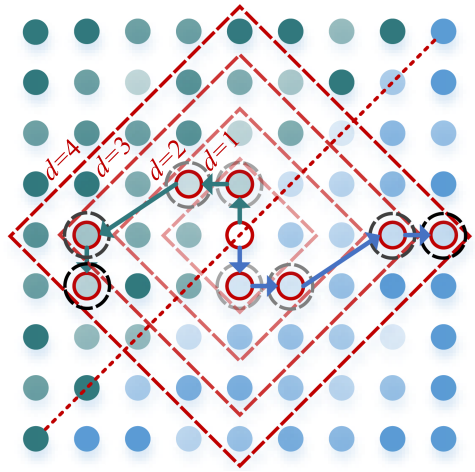


Fig. 7. Illustration of representative neighboring locations selection. Circles with darker color indicate larger RSS difference variance to the targeted point marked by empty circles.

Unless otherwise mentioned, we still refer fingerprints to traditional RSS fingerprints in the following.

4 VIVIPLUS DESIGN

In this section, we present the design and implementation of the ViViPlus system that exploits the RSG matrix as fingerprints.

4.1 Realization of RSS Spatial Gradient

The RSG matrix for each reference location is built upon the original RSS fingerprint radio map. In this section, we first discuss how to form a good RSG matrix for each reference location. Then we consider discrete the RSG matrix to increase the robustness of our system. Finally, we present how to profile a query fingerprint based on a reference location's RSG.

Since an RSG matrix is a collection of the RSG of multiple APs (with each column corresponding to one AP), we interpret on a single AP basic. The RSG matrix construction is then a repeat over multiple selected APs.

4.1.1 Profiling a Reference Location. For a reference location $\boldsymbol{l}_i$ with an average fingerprint $\boldsymbol{f}_i = \{f_{ik}, 1 \leq k \leq m\}$, we consider all its neighboring locations within r sample points. Our goal is to select $c = 2r$ locations among them as the neighbors for RSG construction for each AP. Recalling that the key insight of RSG is to exploit stable spatial features, we achieve this by selecting a subset of locations that produce the most stable RSG for each AP.

Previously, we mentioned the representative fingerprint $\boldsymbol{f}_i$ for a location $\boldsymbol{l}_i$ as the average one. In practice, the fingerprint database generally stores all raw fingerprint measurements for each location [40, 55]. Therefore, there will be multiple RSS records of each AP for each location. Therefore, we leverage all available RSS observations to optimize the spatial stability in RSG difference variances.

Specifically, for location $\boldsymbol{l}_i$, denote its averaged RSS for the kth AP as f_{ik}. Suppose there are p RSS records, i.e., $\{f_{jk}^{(x)}, 1 \leq x \leq p\}$, for a location $\boldsymbol{l}_j$ in its neighboring set. As shown in Figure 7, we calculate every RSS difference $\phi(f_{ik}, f_{jk}^{(x)}), 1 \leq x \leq p$ and derive the corresponding variance.

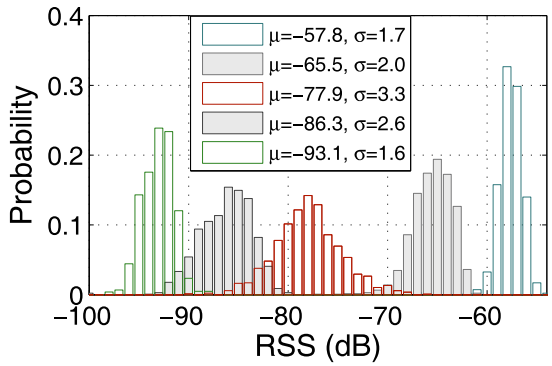


Fig. 8. Illustration of Gaussian-like RSS distribution over time.

To guarantee sufficient space coverage of the selected neighbors, we increase the distance d step by step from 1 to r and select the two locations with the smallest RSS difference variances for each step. In practice, this policy is equivalent to selecting the two locations with the smallest RSS variances. By doing this, we obtain $c = 2r$ neighboring locations in total that cover physical distances from 1 to r sample points. For each selected location $\boldsymbol{l}_j$, we calculate the RSS difference as

$$\phi(f_{ik}, f_{jk}) = f_{ik} - f_{jk}, \tag{5}$$

where $f_{jk} = \frac{1}{p}\sum_{t=1}^{p} f_{jk}^{(t)}$. Afterward, we generate the RSG for current location $\boldsymbol{l}_i$ by dividing the selected neighbors into two parts with identical sizes and ordering them by physical distances to $\boldsymbol{l}_i$, resulting in the RSG profile

$$\vec{g_{ik}} = \{< d(\boldsymbol{l}_i, \boldsymbol{l}_j), \phi(f_{ik}, f_{jk}) >, i - r \le j \le i + r\} \tag{6}$$

of the kth AP as in Figure 4. Repeating the above procedure for each AP in $\boldsymbol{f}_i$, we obtain the RSG matrix whose kth column corresponds to the RSG profile of the kth AP.

The above profiling procedure is similar to that in [40], which optimizes the stability in terms of entire fingerprint similarity. Differently, ViViPlus considers spatial stability based on the RSS gradient for the RSG matrix.

The complexity of profiling each sample location is $O(N_r(m + 1))$, where N_r is the number of neighboring sampling locations within the range r and m is the number of raw fingerprint records of each sample location. Taking the grid topology in Figure 7 as an example, the complexity is $O(\frac{4}{3}(4^r - 1)(m + 1))$. Considering that r is generally a small constant number (e.g., <5), the complexity is then linear to m.

4.1.2 Discretization of RSG matrix. In practice, there might be uncertainties even when we use the averaged RSS for calculation of RSS difference. To improve the robustness, we further discretize the RSG matrix. After discretization, each RSG matrix element reflects the strength relationship of the RSSs between the current location and its nearby locations, which is more stable and reliable than the absolute RSS difference.

As shown in Figure 8, multiple RSS measurements from a specific AP over time typically follow a Gaussian distribution [5, 21, 55], and we adopt the discretization method based on t-test as in [33]. $t-$Test is a statistical hypothesis test that determines whether the null hypothesis (no difference between sample values) is rejected or accepted.

Let $R_{ik} = \{f_{ik}^{(x)}\}$ and $R_{jk} = \{f_{jk}^{(x)}\}, 1 \le x \le p$ denote two sets of RSS samples from AP k at location $\boldsymbol{l}_i$ and $\boldsymbol{l}_j$. The mean and standard variance of R_{ik} and R_{jk} are denoted by (μ_{ik}, σ_{ik}) and

(μ_{jk}, σ_{jk}), respectively. Then we test the following two-side hypothesis:

$$\begin{cases} H_0: & \mu_{ik} = \mu_{jk} \\ H_1: & \mu_{ik} \neq \mu_{jk}. \end{cases} \tag{7}$$

The t-statistic can be calculated as

$$t^k = \frac{\mu_{ik} - \mu_{jk}}{\sqrt{\frac{\sigma_{ik}^2}{n} + \frac{\sigma_{jk}^2}{n}}}. \tag{8}$$

Therefore, we can compute the cumulative distribution for t distribution at values in t^k, and we set the signification level $\alpha = 0.1$. The null hypothesis H_0 can be rejected with confidence (i.e. $\mu_{ik} \neq \mu_{jk}$) if a cumulative density is greater than $1 - \alpha$, and accepted (i.e. $\mu_{ik} = \mu_{jk}$) otherwise.

Based on the significance test, we compute the RSS gradient (i.e., the discrete RSS difference) $\phi(f_{ik}, f_{jk})$ as

$$\phi(f_{ik}, f_{jk}) = \begin{cases} 1, & H_0 \text{ is accepted} \\ 2, & H_0 \text{ is rejected and } \mu_{ik} \geq \mu_{jk} \\ 0, & H_0 \text{ is rejected and } \mu_{ik} < \mu_{jk}. \end{cases} \tag{9}$$

The improvement by the disretized RSS gradient would be compared with the non-discrete RSS difference in the evaluation in Section 5.2.3.

4.1.3 Profiling a Query Fingerprint. In ViViPlus, a user queries his/her location by reporting an RSS fingerprint. We then construct an RSG matrix based on a single query fingerprint. The key intuition is to reuse the data in the RSG matrix database for reference locations to generate the query RSG matrix. By doing this, a ViViPlus user is not required to report any mobility data, only an RSS fingerprint, as if he/she is using a conventional RSS fingerprint-based system.

Consider that we match a query RSS fingerprint $\boldsymbol{f}_Q$ against a candidate reference location $\boldsymbol{l}_C$. We then generate the query RSG matrix by substituting the query fingerprint as the representative fingerprint of $\boldsymbol{l}_C$. Specifically, suppose

$$\vec{g_{Ck}} = \{<d(\boldsymbol{l}_C, \boldsymbol{l}_j), \phi(f_{Ck}, f_{jk})>, j \in N(\boldsymbol{l}_C)\} \tag{10}$$

is the RSG for the kth AP in the reference RSG matrix G_C for the current location, where $N(\boldsymbol{l}_C)$ denotes the selected neighbors. The corresponding column in the query RSG matrix G_Q is calculated as

$$\vec{g_{Qk}} = \{<d(\boldsymbol{l}_C, \boldsymbol{l}_j), \phi(f_{Qk}, f_{jk})>, j \in N(\boldsymbol{l}_C)\}, \tag{11}$$

where we replace f_{Ck} in $\vec{g_{Ck}}$ and keep everything else (including $\boldsymbol{l}_C$) unchanged. Repeat this for all m APs, and we derive a query RSG matrix:

$$G_Q = (\vec{g_{Q1}}, \vec{g_{Q2}}, \ldots, \vec{g_{Qm}})^{\mathrm{T}}. \tag{12}$$

Generally, the resulted query RSG matrix will be similar to the reference one if the query fingerprint is from the same or close location with $\boldsymbol{l}_C$. Otherwise, the two matrices would deviate from each other significantly because two different locations are unlikely to share consistent RSS spatial gradients over $N(\boldsymbol{l}_C)$ locations. This reference-data-based query matrix generation is a unique design, which enables ViViPlus to deal with a single query fingerprint from one user without any additional information.

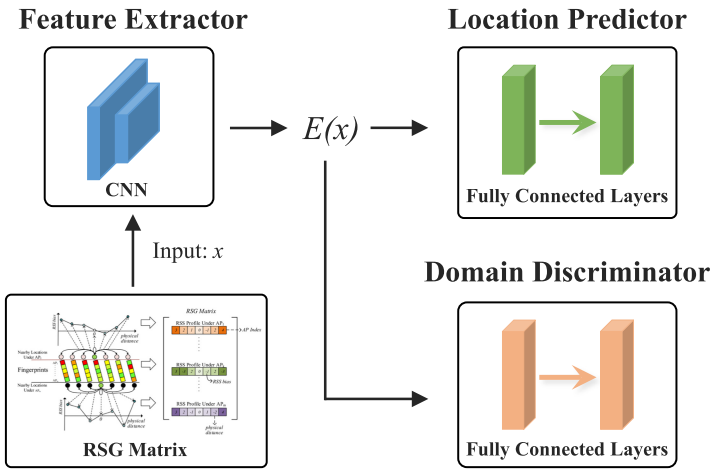


Fig. 9. Overview of domain adversarial learning-based model.

4.2 Localization with RSG Matrices

Different approaches can be used for RSG matrix localization. We consider a **domain adversarial neural network (DANN)**-based method, an eigenvector-based method, and a SIFT-based method in ViViPlus.

4.2.1 DANN-based Method. In order to achieve accurate and robust localization performance, we leverage the DANN to further extract the dynamics-resistant feature based on the generated RSG matrix. An overview of the proposed domain adversarial learning model is shown in Figure 9. The input RSG matrix is first transformed into latent representation $E(x)$ by the **feature extractor**. Using the learned feature representations, the **location predictor** is leveraged to maximize the localization accuracy and obtain the location predictions. To remove domain-specific features, a **domain discriminator** is designed to label each domain (specifically, to identify *when the fingerprints are collected*). The feature extractor tries its best to cheat the domain discriminator and at the same time boost the accuracy of the localization results, which is termed as a minimax game [20, 56]. Through this minimax game, the feature extractor can finally learn the domain-independent features $E(x)$ for all RSG matrices. During the localization stage, multiple query RSG matrices representing different reference locations generated from the original query fingerprints are predicted by the well-trained model, respectively, and we take the one with the highest prediction probability to its real location as the localization result.

4.2.2 Eigenvector-based Method. As the fingerprints involved in ViViPlus are in a matrix form, we intend to devise matrix features for fingerprint matching. Specifically, we exploit eigenvectors and eigenvalues, which are the most common and useful characteristics of a matrix that reflect the spatial feature and dimensionality of the specific matrix. Under this context, an RSG matrix G can also be expressed as a set of $\{<\lambda_1, \vec{\mu_1}>, <\lambda_2, \vec{\mu_2}>, \ldots, <\lambda_m, \vec{\mu_m}>\}$, where $\vec{\mu_1}, \vec{\mu_2}, \ldots, \vec{\mu_m}$ denotes the m eigenvectors under eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_m$. We calculate a representative vector $\vec{v}$ as $\lambda_1\mu_1 + \lambda_2\mu_2 + \ldots + \lambda_m\mu_m$ for each RSG matrix G_i for location l_i. Then we only need to calculate the similarity of the representative vector $\vec{v_i}$ for localization. The eigenvector method depicts the spatial feature of a matrix in a way more or less like individually considering each column of the RSG matrix that represents the RSS gradients under each AP, which makes the eigenvector method a more practical significance method.

Table 1. Data Collection in Different Scenarios

#	Building Type (Areas)	Size (m^2)	Density	Devices	#Region	#Samples	Duration
1	Academic (Public areas)	600	1 m × 1 m	Nexus 5/7, two Nexus 6p	13	72K	2 weeks
2	Office (Whole floor)	1,500	2 m × 2 m	Two Nexus S	20	65K	2 weeks
3	Classroom (Public areas)	2,000	1.2 m × 1.2 m	Nexus 7, two Nexus 6p	18	96K	2 weeks
4	Shopping mall (Public areas)	2,130	————	HUAWEI P9, imoo Z5/Z6	30	288K	7 months

4.2.3 SIFT-like Method. As shown in Figure 5, an RSG matrix can also be treated as an image with only one channel. Thus, we can also apply developed computer vision techniques such as SIFT [24] for our matrix matching. The SIFT first extracts keypoints from the one channel RSG image and appends detailed descriptors to each keypoint. The keypoint with descriptor in the reference RSG image can be expressed as $R_i = (r_{i1}, r_{i2}, \ldots, r_{in})$, and $S_i = (s_{i1}, s_{i2}, \ldots, s_{in})$ in the query RSG image, where n is the dimension of the descriptor vector. The similarity metric between any two descriptors can be expressed as $d(R_i, S_i) = \sqrt{\sum_{j=1}^{n}(r_{ij} - s_{ij})^2}$. The reference RSG image with the highest overall similarity is used as the matching result for the query RSG image. The SIFT-like method is a more distinctive algorithm that treats the RSG matrix as an image and focuses on key elements of the matrix.

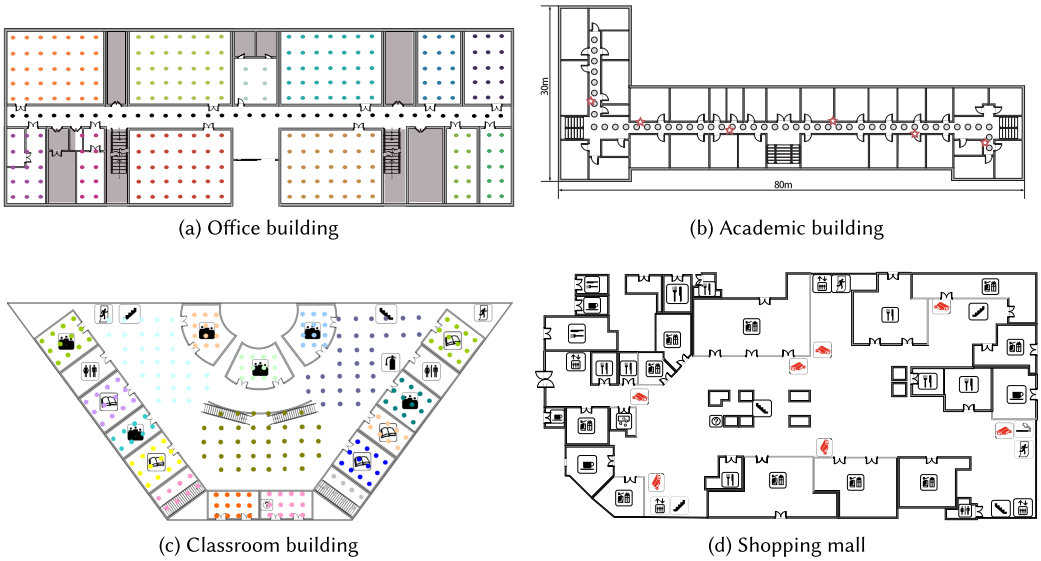
5 IMPLEMENTATIONS AND EVALUATION

5.1 Experiment Methodology

We have implemented ViViPlus on Android platforms and conduct experiments using seven different devices over various scenarios. And we deploy ViViPlus in real business environments and continuously evaluate the system performance across 7 months. In this section, we first introduce the experimental settings and then present a detailed evaluation.

5.1.1 System Implementation & Workflow. We implemented the client of ViViPlus on the Android mobile platforms and the server on the Linux OS. In brief, the clients collect raw RSS data and upload it to the server. The server-side constructs the RSG matrix based on the query fingerprint and compares the query RSG matrix with the pre-constructed database for localization. Finally, the server returns the localization result to the client for location-based services or applications.

5.1.2 Experimental Scenarios & Datasets. We conduct our extensive experiments in four different buildings, as shown in Figure 10. The four buildings have different environmental conditions. In particular, the classroom building and shopping mall are much more crowded than the academic building and school office building. The data collection details in each building are summarized in Table 1. We use different sampling densities for different scenarios that have different requirements on localization accuracy. Dense sampling can improve localization accuracy but also has a greater burden of site survey. However, sampling density does not affect the RSG profiling due to the flexible reference location selection strategy, as described in Section 4.1. The training samples are collected once at the beginning, while test samples are collected multiple times at intervals. When collecting fingerprints for training, we employ a typical sampling frequency of around 1 Hz. We employ seven phones of five different models manufactured by different companies for data collection, which are equipped with different WiFi chips. We additionally employ two kinds of smartphones, imoo Z5 and imoo Z6, to collect query fingerprints during the localization stage.



(a) Office building (b) Academic building

(c) Classroom building (d) Shopping mall

Fig. 10. Experimental areas.

5.1.3 Comparative Methods. To extensively evaluate the performance of ViViPlus, we additionally implement five different state-of-the-art approaches for comparison, which have been proposed to enhance the primary RSS fingerprinting. The five methods are (1) **Horus** [55]: a classical probabilistic algorithm; (2) **TW-KNN** [11]: it applies an iterative, recursive weighted average filter to form temporally weighted RSSs as fingerprints; (3) **GIFT** [33]: a binary metric of differential RSSs at two adjacent locations along a moving trace is exploited—as GIFT is designed for mobile traces, we combine queries from two adjacent locations as one for GIFT in the localization experiments and implement normal GIFT for tracking experiments; (4) **ViVi** [40]: a most related system that puts forward a similar concept of *fingerprint spatial gradient*; (5) **Magicol** [32]: a mobile tracking system using a Particle Filter to fuse traditional WiFi fingerprints and magnetic signals.

In our experiments, we compare our system ViViPlus with Horus, TW-KNN, GIFT (modified), and ViVi. In tracking experiments, we integrate ViViPlus with a **Particle Filter (PF)** [49] and compare it with GIFT (original), Magicol, and Horus (with a PF). We mainly aim to show the advantages of the proposed RSG over RSS fingerprints. Therefore, we focus on the relative accuracy improvement rather than the absolute accuracy achieved by ViViPlus. Hence, we mainly implement the core fingerprinting and matching components of the above systems. For example, we omit the clustering step for fast localization in Horus [55]. Then we apply identical preprocessing steps (e.g., AP selection) to all methods.

5.1.4 Evaluation Metrics. Similar to existing works, we adopt two methods to evaluate the localization performance: (1) Distance-level localization bias, which is a fine-grained evaluating indicator. The Euler distance between localization result and ground truth is defined as localization bias. The average location error and the 95th percentile error are adopted as major performance metrics in this method. (2) Region-level localization success rate, which is a relatively coarse-grained but intuitive and meaningful indicator. We count the rate that a system locates a user in the correct room or region we segmented. Both of the two indicators are leveraged in experiments in the office, academic, and classroom buildings. And only the localization success rate is used in the shopping mall because of the large fingerprints sampling density and irregular shape of public areas.

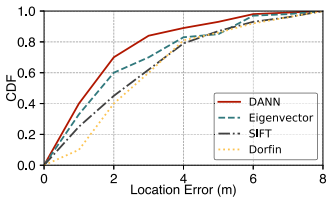


Fig. 11. Localization methods with RSG.

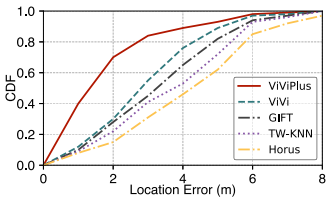


Fig. 12. Different method comparison.

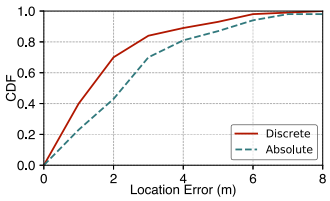


Fig. 13. Different RSG matrix.

5.2 Overall Performance

5.2.1 Different Localization Methods with RSG Matrix. We first explore the best RSG matrix comparing methods. We integrate the results of different phones from all experimental areas for evaluation. In addition to the methods that leverage global features of a matrix for localization (see Section 4.2), we further incorporate the methods in DorFin [44] and implement an additional method that compares RSG matrices row by row. As shown in Figure 11, the DANN-based method achieves the best accuracy, yielding an average accuracy of 2.13 m and a 95th error of 4.91 m. The eigenvector-based method achieves an average accuracy of 2.5 m and a 95th error of 5.61 m. The SIFT-like method achieves an average accuracy of 3.12 m and a 95th error of 6.05 m. The Dorfin-based method achieves an average accuracy of 3.3 m and a 95th error of 6.12 m. Considering the localization accuracy, we use the DANN-based method for following evaluation.

5.2.2 Overall Performance Comparison. The performances of the proposed ViViPlus, as well as the four state-of-the-art and comparative approaches, are depicted in Figure 12. As seen, ViViPlus achieves the best performance among all. The average accuracy outperforms Horus by 50.7%, TW-KNN by 44.2%, and GIFT by 38.9%, and exceeds ViVi by 31.6%. The 95th percentile accuracy outperforms the four comparative approaches by 40.4%, 33.8%, 30.8%, and 22.5%, respectively. The results demonstrate that ViViPlus achieves remarkable performance gains based on only RSS fingerprints without the pains introducing extra information or constraints. Further performance gains by additional information like sensor hints can easily be incorporated in ViViPlus. For example, we implement particle filter in ViViPlus and evaluate it for tracking in Section 5.4.

5.2.3 Absolute vs. Discrete RSG Matrix. We evaluate the benefit of RSG matrix discretization. Evidently in Figure 13, RSG matrix discretization significantly improves the robustness of ViViPlus. While the 95th percentile errors are similar, the average accuracy is improved by 20.9%; it decreased from 2.61m to 2.13m by discretizing the RSG matrix.

5.2.4 Performance with Different Conditions. To examine the robustness and practicability of ViViPlus, we invite three users to evaluate it in different buildings with different devices. As shown in Figure 14, ViViPlus achieves consistently delightful accuracy of 2.01 m, 2.24 m, and 2.43 m in the academic, office, and classroom buildings that suffer from different crowd levels and wireless environments. And as depicted in Figure 15, ViViPlus yields an average localization success rate of 96.2% in the office building, 94.8% in the academic building, 94.1% in the classroom, and 92.4% in the shopping mall. Furthermore, as shown in Figure 16, the average accuracies when using Nexus 6p, Nexus 5, and Nexus 7 pad are 2.05 m, 2.22 m, and 2.35 m, respectively, while the fingerprint databases are not necessarily constructed using the same models of phones (see Table 1). The results demonstrate that ViViPlus achieves similar performance regardless of the environments and devices.

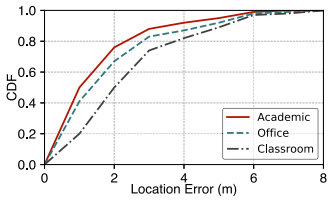


Fig. 14. Different areas (location error).

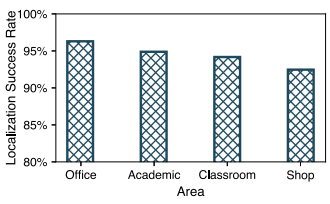


Fig. 15. Different areas (success rate).

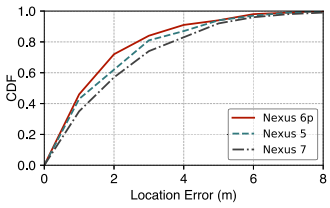


Fig. 16. Different devices.

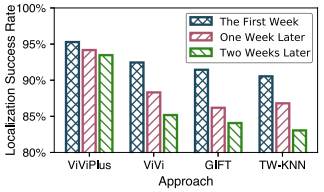


Fig. 17. Comparison in temporal robustness.

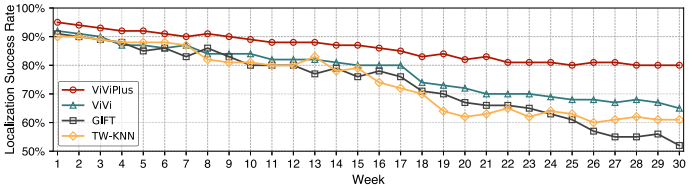


Fig. 18. Comparison in long-term accuracy.

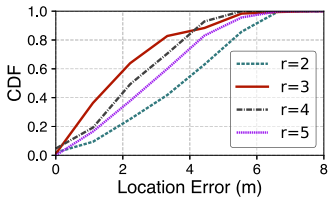


Fig. 19. Impacts of r.

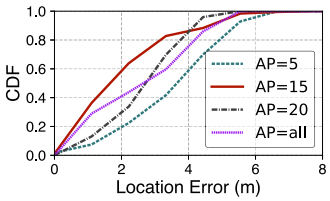


Fig. 20. Impacts of AP number.

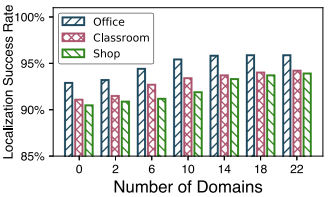


Fig. 21. Impact of domain number.

5.2.5 Performance Comparison with Different Time Interval. We also examine localization robustness in terms of temporal stability. We recollect fingerprints during the different time intervals in all four scenarios and use the original database for localization. Figure 17 depicts the performance of ViViPlus as well as the three approaches. ViViPlus yields a similar success rate of more than 93.4% even after 2 weeks. Compared with related works where the success rates decrease more than 7%, the decrease in success rate for ViViPlus is within 2%. The results demonstrate that RSG matrices-based localization can extract robust features to resist the fingerprint temporal instability caused by wireless signal fluctuation.

5.2.6 Long-term Performance Comparison. We finally compare ViViPlus with related systems to verify the performance of the proposed method in long-term accuracy. We continuously collect RSS fingerprints and evaluate the performance of systems over 7 months. Figure 18 records performance variation during such a long period. As shown, ViViPlus achieves the best performance among all comparative systems at any given time. Even after 7 months, the localization success rate of ViViPlus still maintains 80%, which only decreases by 15% compared with other systems by at least 27%. It is worth mentioning that in the 18th week, due to the re-decoration of the shopping mall, the performance of three comparative works reduces sharply. However, there is merely a negligible effect on the performance of the proposed ViViPlus. The above remarkable results demonstrate that the proposed ViViPlus is qualified for locating in dynamic environments, and the proposed RSG matrices and DANN-based localization are able to maintain the system's performance for a long time.

5.3 Impact of Parameters

5.3.1 Impacts of Neighbor Number. In the above experiments, we use six neighbors ($r = 3$) to generate RSG matrices. Now we examine the impacts of r ranging from 2 to 5. As shown in Figure 19, when r increases from 2 to 3, the average location errors decrease from 3.84 m to 2.5 m. When r further increases to 5, however, the location errors increase to 2.96 m ($r = 4$) and 3.35 m ($r = 5$). The results indicate that the spatial stability among neighboring RSSs only holds within a certain space range. Thus, if using two distant neighbors (too large r), RSS observations from distant locations will also be involved, which will degrade the stableness of the RSG matrix.

5.3.2 Impacts of AP Number. We also examine the impacts of AP numbers, which determine the size of the resulting RSG matrix. We evaluated the performance by randomly choosing 5, 15, 20, and all APs without filtering. As shown in Figure 20, ViViPlus achieves the best performance when using 15 APs. The average accuracy increases by 53.2%, 24%, and 20.4% compared with using 5, 10, and all APs. Also, note that the larger AP number we select, the higher system latency will result. Thus, in ViViPlus, we use 15 APs by default.

5.3.3 Impacts of Domain Number. The localization performance of the DANN-based method depends on the number of source domains in the model training [16]. In this experiment, we divide the datasets into two disjoint sets as the source and target domains. There are 22 different time periods (when collected) as the source domains and 15 different time periods as the target domains. Here, we refer to the domains with and without label information as the source and target domain. Figure 21 shows the accuracy of ViViPlus by using the DANN with the different number of source domains. We can observe that the proposed ViViPlus can achieve better accuracy when the number of source domains increased, and achieves the best performance when using more than 14 source domains in all three scenarios. Therefore, in practical applications, an appropriate increase in the number of source domains is necessary to achieve the best performance.

5.4 ViViPlus in Mobile Tracking

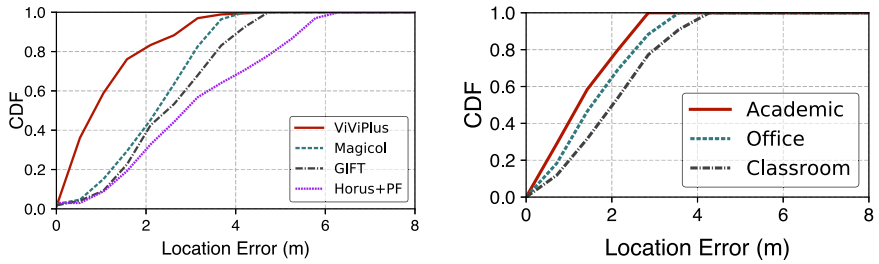
ViViPlus achieves reliable localization with no more inputs than traditional RSS fingerprints. Nevertheless, extra information like inertially sensed user mobility [51] and mature methods like Particle Filter [14] can be easily incorporated in ViViPlus to further improve the performance, especially for continuous tracking. In the following, we integrate a Particle Filter module in ViViPlus and evaluate it for real-time tracking.

Environment Setting: We conduct tracking experiments in the same environments, as shown in Table 1, and collect RSS measurements while walking. We respectively collect 100, 150, and 250 traces in the academic, office, and classroom buildings with more than 14,000 location queries. The traces are collected over different times, with lengths varying from 10 m to 25 m. Figure 22 shows two example areas, one of which is in the corridors of the academic building and the other in the classroom building when there are sometimes many students walking by and sometimes nobody around.

Tracking Performance: To compare the performance of ViViPlus for tracking, we also implement Magicol (WiFi plus magnetism sensor with PF) in addition to the aforementioned GIFT (with PF) and Horus (with PF). As shown in Figure 23(a), ViViPlus outperforms Horus by 55.1%, GIFT by 34.5%, and Magicol by 29.6% in terms of mean accuracy. The 95th percentile accuracy outperforms the three comparative approaches by 50.2%, 32.3%, and 24.5%, respectively. ViViPlus still yields better accuracy even compared to Magicol, which employs WiFi together with magnetism signals. Moreover, sensor data like magnetism can also be easily integrated into ViViPlus.

(a) Empty office building (b) Crowded classroom building

Fig. 22. Tracking environments.



(a) Different methods (b) Different locations

Fig. 23. Result of mobile tracking.

The tracking performance of ViViPlus in different buildings is compared in Figure 23(b). The average accuracy in the three buildings is 1.76 m, 1.93 m, and 2.12 m, with a minor bias within 0.4 m. The results demonstrate consistent reliable performance of ViViPlus for deployment over different areas.

System Latency: We also evaluate the latency of ViViPlus in real-time mobile tracking. We randomly select more than 4,000 queries to calculate the average computational latency using a laptop with an i5-5300 core. The average computational latency of ViViPlus (RSG matrix with DANN) is 0.15 s, which can meet the latency requirements in real application scenarios. And as a comparison, the average latency of GIFT, Magicol, and Horus are 0.19 s, 0.25 s, and 0.14 s, respectively.

6 RELATED WORKS

ViViPlus is closely related to several related works in the literature of indoor localization.

Easing Deployment: Site survey has been a major bottleneck for fingerprint-based localization, which is time-consuming and labor-intensive. Among various research efforts, a recent crowdsourcing-based approach sheds promising light in easing the site survey costs [28, 36, 41, 50]. Both [7] and [8] propose zero-calibration systems for realistic environments. OIL [26] facilitates rapid coverage by building an organic surveying system. **Simultaneous Localization and Mapping (SLAM)** techniques are incorporated to avoid the training costs, which result in a set of technique advancements including WiFiSLAM [10], GraphSLAM [15], and SemanticSLAM [1]. In addition to radio maps, pioneer works including [9, 17, 31] further consider automatic construction of floorplans, which also stimulates the practical deployment of WiFi-based localization significantly. Specifically, Walkie-Markie [31] exploits RSS trends

along pathways for the reconstruction of indoor pathways. Targeting a different perspective of accuracy for fingerprint-based localization, ViViPlus is orthogonal with these works and can be put together to enable practical indoor positioning services.

Using Extra Hints: Many works attempt to improve the accuracy of WiFi fingerprinting by leveraging additional information. Ranging via acoustic signals [23] or WiFi Direct [18] among multiple devices is introduced to alleviate fingerprint ambiguity. Fusing inertial sensor data also attracts extensive studies. SurroundSense [2] integrates various sensor hints as multi-modal fingerprints for localization. More commonly, motion information is fused to provide relative locations to improve fingerprint-based localization [51]. [14] and [28] utilize geometric constraints imposed by both mobility information and digital floorplan. Most recently, several works exist that utilize a fusion of camera and mobile sensors with a wide variety of applications. Argus [47] makes use of visual images to obtain extra position constraints for fingerprinting. ClickLoc [48] leverages sensor-enriched photos and enables user localization with a single photo of the surrounding **place of interest (POI)** with high accuracy and short delay. PHADE [4] relies on surveillance cameras to view users' motion patterns and compare the motion with the trajectory calculated from IMU sensors on the user's phone to identify each user and track them. While remarkable accuracy is gained by these approaches, they generally rely on additional information from extra sensors, multiple devices, or many participatory users. On the contrary, ViViPlus improves WiFi fingerprinting without any additional constraints on system inputs.

Physical Layer CSI: Recently, CSI has been leveraged for precise localization [19, 22, 27, 35, 37, 52]. With higher resolution to multipath fading, CSI-based systems can yield decimeter-level accuracy [19, 35]. PinLoc [29] uses CSI as fingerprints to improve localization accuracy, but with significant deployment costs. FILA [45] extracts the line-of-sight signal from CSI for accurate ranging. SpotFi [19] achieves decimeter-level location accuracy by accurately computing the **angle of arrival (AoA)** of multipath components using CSI. Chronus [35] splices multiple WiFi channel and achieves decimeter-level localization accuracy with a single WiFi device. CSI is also exploited for passive localization and tracking [22, 37]. LiFS [37] leverages the shadowing effect caused by the person's blocking line-of-sight paths of WiFi links to achieve passive localization. DynamicMusic [22] leverages the incoherence between the signal reflected from the moving person and the static signal to separate the former signal and calculate its AoA for tracking. Despite its high precision, CSI is not readily available on commercial smartphones. Therefore, these systems rely on customized hardware or specialized WiFi Network Interface Cards like Intel 5300, which largely limits the ubiquity for deployment.

Spatial Awareness: Some recent innovations also explore RSS spatial awareness for enhanced fingerprinting [6, 31, 33]. Early efforts such as the well-known Horus system [55] employ advanced probabilistic methods to enhance the first fingerprinting system RADAR [3]. Recently, WalkieMarkie [31] is the first to explore RSS changing trends along pathways for floorplan construction. INTRI [12] combines RSS contours with traditional fingerprinting for better accuracy. GIFT [33] defines a binary differential RSS, i.e., the difference of RSSs from two continuous locations, as a replacement of absolute RSSs to deal with signal variations. While these works inspire the design of ViViPlus, they rely on user movements and are only applicable to continuous tracking. RSSRatio [6] leverages differential RSS on two antennas on MIMO systems, which is not suitable for commodity smartphones. ViVi [40] is the most related work to ViViPlus. It employs *fingerprint spatial gradient* for better fingerprinting. Compared with the *fingerprint spatial gradient* used by ViVi, ViViPlus exploits an intrinsic *RSS spatial gradient* to portray the spatial features of neighboring locations in a fine-grained manner, which achieves better performance regarding spatial ambiguity and temporal instability.

7 CONCLUSION

In this article, we present the design and implementation of ViViPlus, an indoor localization system purely based on WiFi fingerprints, which jointly mitigates spatial ambiguity and temporal instability and derives reliable performance without impairing the ubiquity. ViViPlus exploits the spatial awareness of RSS values by formulating an RSG matrix as enhanced WiFi fingerprints. We devise techniques for the representation, construction, and localization of the proposed RSG matrix and integrate ViViPlus as a fully practical system that requires no more inputs than any previous RSS fingerprint-based systems. We conduct extensive experiments across 7 months in different buildings and implement five other systems for comparison. The results demonstrate that ViViPlus significantly improves the accuracy in both localization and tracking scenarios by about 30% to 50% compared with the state-of-the-art approaches.

REFERENCES

[1] H. Abdelnasser, R. Mohamed, A. Elgohary, M. F. Alzantot, H. Wang, S. Sen, R. R. Choudhury, and M. Youssef. 2016. SemanticSLAM: Using environment landmarks for unsupervised indoor localization. *IEEE Transactions on Mobile Computing* 15, 7 (July 2016), 1770–1782.

[2] Martin Azizyan, Ionut Constandache, and Romit Roy Choudhury. 2009. SurroundSense: Mobile phone localization via ambience fingerprinting. In *Proceedings of the 15th Annual International Conference on Mobile Computing and Networking*. ACM, 261–272.

[3] Paramvir Bahl, Venkata N. Padmanabhan, et al. 2000. RADAR: An in-building RF-based user location and tracking system. In *IEEE Infocom*, Vol. 2. IEEE, 775–784.

[4] Siyuan Cao and He Wang. 2018. Enabling public cameras to talk to the public. *PACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 2, 2 (2018), 1–20.

[5] Yiqiang Chen, Qiang Yang, Jie Yin, and Xiaoyong Chai. 2006. Power-efficient access-point selection for indoor location estimation. *IEEE Transactions on Knowledge and Data Engineering* 18, 7 (2006), 877–888.

[6] Wei Cheng, Kefeng Tan, Victor Omwando, Jindan Zhu, and Prasant Mohapatra. 2013. RSS-Ratio for enhancing performance of RSS-based applications. In *2013 Proceedings IEEE INFOCOM*. IEEE, 3075–3083.

[7] Krishna Chintalapudi, Anand Padmanabha Iyer, and Venkata N. Padmanabhan. 2010. Indoor localization without the pain. In *Proceedings of the 16th Annual International Conference on Mobile Computing and Networking*. ACM, 173–184.

[8] Rizanne Elbakly and Moustafa Youssef. 2016. A robust zero-calibration RF-based localization system for realistic environments. In *2016 13th Annual IEEE International Conference on Sensing, Communication, and Networking (SECON'16)*. IEEE, 1–9.

[9] A. Eleryan, M. Elsabagh, and M. Youssef. 2011. Synthetic generation of radio maps for device-free passive localization. In *IEEE GLOBECOM*.

[10] Brian Ferris, Dieter Fox, and Neil D. Lawrence. 2007. WiFi-SLAM using Gaussian process latent variable models.. In *IJCAI*, Vol. 7. 2480–2485.

[11] D. Han, S. Jung, M. Lee, and G. Yoon. 2014. Building a practical Wi-Fi-based indoor navigation system. *IEEE Pervasive Computing* 13, 2 (Apr. 2014), 72–79.

[12] Suining He, Tianyang Hu, and S-H. Gary Chan. 2015. Contour-based trilateration for indoor fingerprinting localization. In *Proceedings of the 13th ACM Conference on Embedded Networked Sensor Systems*. ACM, 225–238.

[13] Suining He, Wenbin Lin, and S-H. Gary Chan. 2017. Indoor localization and automatic fingerprint update with altered AP signals. *IEEE Transactions on Mobile Computing* 16, 7 (2017), 1897–1910.

[14] Sebastian Hilsenbeck, Dmytro Bobkov, Georg Schroth, Robert Huitl, and Eckehard Steinbach. 2014. Graph-based data fusion of pedometer and WiFi measurements for mobile indoor positioning. In *Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing*. ACM, 147–158.

[15] Joseph Huang, David Millman, Morgan Quigley, David Stavens, Sebastian Thrun, and Alok Aggarwal. 2011. Efficient, generalized indoor wifi Graphslam. In *2011 IEEE International Conference on Robotics and Automation*. IEEE, 1038–1043.

[16] Wenjun Jiang, Chenglin Miao, Fenglong Ma, Shuochao Yao, Yaqing Wang, Ye Yuan, Hongfei Xue, Chen Song, Xin Ma, Dimitrios Koutsonikolas, et al. 2018. Towards environment independent device free human activity recognition. In *Proceedings of the 24th Annual International Conference on Mobile Computing and Networking*. 289–304.

[17] Yifei Jiang, Yun Xiang, Xin Pan, Kun Li, Qin Lv, Robert P. Dick, Li Shang, and Michael Hannigan. 2013. Hallway based automatic indoor floorplan construction using room fingerprints. In *Proceedings of the 2013 ACM International Joint Conference on Pervasive and Ubiquitous Computing*. ACM, 315–324.

[18] Junghyun Jun, Yu Gu, Long Cheng, Banghui Lu, Jun Sun, Ting Zhu, and Jianwei Niu. 2013. Social-Loc: Improving indoor localization with social sensing. In *Proceedings of the 11th ACM Conference on Embedded Networked Sensor Systems*. ACM, 14.

[19] Manikanta Kotaru, Kiran Joshi, Dinesh Bharadia, and Sachin Katti. 2015. Spotfi: Decimeter level localization using wifi. In *ACM SIGCOMM Computer Communication Review*, Vol. 45. ACM, 269–282.

[20] Danyang Li, Jingao Xu, Zheng Yang, Yumeng Lu, Qian Zhang, and Xinglin Zhang. 2021. Train once, locate anytime for anyone: Adversarial learning based wireless localization. In *IEEE INFOCOM 2021-IEEE Conference on Computer Communications*.

[21] Liqun Li, Guobin Shen, Chunshui Zhao, Thomas Moscibroda, Jyh-Han Lin, and Feng Zhao. 2014. Experiencing and handling the diversity in data density and environmental locality in an indoor positioning service. In *Proceedings of the 20th Annual International Conference on Mobile Computing and Networking*. ACM, 459–470.

[22] Xiang Li, Shengjie Li, Daqing Zhang, Jie Xiong, Yasha Wang, and Hong Mei. 2016. Dynamic-music: accurate device-free indoor localization. In *Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing*. ACM, 196–207.

[23] Hongbo Liu, Yu Gan, Jie Yang, Simon Sidhom, Yan Wang, Yingying Chen, and Fan Ye. 2012. Push the limit of WiFi based localization for smartphones. In *Proceedings of the 18th Annual International Conference on Mobile Computing and Networking*. ACM, 305–316.

[24] David G. Lowe. 2004. Distinctive image features from scale-invariant keypoints. *International Journal of Computer Vision* 60, 2 (2004), 91–110.

[25] Dimitrios Lymberopoulos, Jie Liu, Xue Yang, Romit Roy Choudhury, Vlado Handziski, and Souvik Sen. 2015. A realistic evaluation and comparison of indoor location technologies: Experiences and lessons learned. In *Proceedings of the 14th International Conference on Information Processing in Sensor Networks*. ACM, 178–189.

[26] Jun-geun Park, Ben Charrow, Dorothy Curtis, Jonathan Battat, Einat Minkov, Jamey Hicks, Seth Teller, and Jonathan Ledlie. 2010. Growing an organic indoor location system. In *Proceedings of the 8th International Conference on Mobile Systems, Applications, and Services*. ACM, 271–284.

[27] Kun Qian, Chenshu Wu, Yi Zhang, Guidong Zhang, Zheng Yang, and Yunhao Liu. 2018. Widar2. 0: Passive human tracking with a single wi-fi link. In *Proceedings of the 16th Annual International Conference on Mobile Systems, Applications, and Services*. ACM, 350–361.

[28] Anshul Rai, Krishna Kant Chintalapudi, Venkata N. Padmanabhan, and Rijurekha Sen. 2012. Zee: Zero-effort crowdsourcing for indoor localization. In *Proceedings of the 18th Annual International Conference on Mobile Computing and Networking*. ACM, 293–304.

[29] Souvik Sen, Božidar Radunovic, Romit Roy Choudhury, and Tom Minka. 2012. You are facing the Mona Lisa: Spot localization using PHY layer information. In *Proceedings of the 10th International Conference on Mobile Systems, Applications, and Services*. ACM, 183–196.

[30] Longfei Shangguan, Zheng Yang, Alex X. Liu, Zimu Zhou, and Yunhao Liu. 2017. STPP: Spatial-temporal phase profiling-based method for relative RFID tag localization. *IEEE/ACM Transactions on Networking* 25, 1 (2017), 596–609.

[31] Guobin Shen, Zhuo Chen, Peichao Zhang, Thomas Moscibroda, and Yongguang Zhang. 2013. Walkie-Markie: Indoor pathway mapping made easy. In *Presented as Part of the 10th {USENIX} Symposium on Networked Systems Design and Implementation ({NSDI}'13)*. 85–98.

[32] Yuanchao Shu, Cheng Bo, Guobin Shen, Chunshui Zhao, Liqun Li, and Feng Zhao. 2015. Magicol: Indoor localization using pervasive magnetic field and opportunistic WiFi sensing. *IEEE Journal on Selected Areas in Communications* 33, 7 (2015), 1443–1457.

[33] Yuanchao Shu, Yinghua Huang, Jiaqi Zhang, Philippe Coué, Peng Cheng, Jiming Chen, and Kang G. Shin. 2016. Gradient-based fingerprinting for indoor localization and tracking. *IEEE Transactions on Industrial Electronics* 63, 4 (2016), 2424–2433.

[34] Wei Sun, Junliang Liu, Chenshu Wu, Zheng Yang, Xinglin Zhang, and Yunhao Liu. 2013. MoLoc: On distinguishing fingerprint twins. In *IEEE ICDCS*.

[35] Deepak Vasisht, Swarun Kumar, and Dina Katabi. 2016. Decimeter-level localization with a single WiFi access point. In *13th {USENIX} Symposium on Networked Systems Design and Implementation ({NSDI}'16)*. 165–178.

[36] He Wang, Souvik Sen, Ahmed Elgohary, Moustafa Farid, Moustafa Youssef, and Romit Roy Choudhury. 2012. No need to war-drive: Unsupervised indoor localization. In *Proceedings of the 10th International Conference on Mobile Systems, Applications, and Services*. ACM, 197–210.

[37] Ju Wang, Hongbo Jiang, Jie Xiong, Kyle Jamieson, Xiaojiang Chen, Dingyi Fang, and Binbin Xie. 2016. LiFS: low human-effort, device-free localization with fine-grained subcarrier information. In *Proceedings of the 22nd Annual International Conference on Mobile Computing and Networking*. ACM, 243–256.

[38] Jue Wang and Dina Katabi. 2013. Dude, where's my card? RFID positioning that works with multipath and non-line of sight. In *ACM SIGCOMM Computer Communication Review*, Vol. 43. ACM, 51–62.
[39] Yutian Wen, Xiaohua Tian, Xinbing Wang, and Songwu Lu. 2015. Fundamental limits of RSS fingerprinting based indoor localization. In *2015 IEEE Conference on Computer Communications (INFOCOM'15)*. IEEE, 2479–2487.
[40] Chenshu Wu, Jingao Xu, Zheng Yang, Nicholas D. Lane, and Zuwei Yin. 2017. Gain without pain: Accurate WiFi-based localization using fingerprint spatial gradient. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 1, 2 (2017), 29.
[41] Chenshu Wu, Zheng Yang, and Yunhao Liu. 2015. Smartphones based crowdsourcing for indoor localization. *IEEE Transactions on Mobile Computing* 14, 2 (2015), 444–457.
[42] Chenshu Wu, Zheng Yang, and Chaowei Xiao. 2018. Automatic radio map adaptation for indoor localization using smartphones. *IEEE Transactions on Mobile Computing* 17, 3 (2018), 517–528.
[43] Chenshu Wu, Zheng Yang, Chaowei Xiao, Chaofan Yang, Yunhao Liu, and Mingyan Liu. 2015. Static power of mobile devices: Self-updating radio maps for wireless indoor localization. In *2015 IEEE Conference on Computer Communications (INFOCOM'15)*. IEEE, 2497–2505.
[44] Chenshu Wu, Zheng Yang, Zimu Zhou, Yunhao Liu, and Mingyan Liu. 2017. Mitigating large errors in WiFi-based indoor localization for smartphones. *IEEE Transactions on Vehicular Technology* 66, 7 (2017), 6246–6257.
[45] Kaishun Wu, Jiang Xiao, Youwen Yi, Min Gao, and Lionel M. Ni. 2012. Fila: Fine-grained indoor localization. In *2012 Proceedings IEEE INFOCOM*. IEEE, 2210–2218.
[46] Yaxiong Xie, Zhenjiang Li, and Mo Li. 2018. Precise power delay profiling with commodity Wi-Fi. *IEEE Transactions on Mobile Computing* 18, 6 (2018), 1342–1355.
[47] Han Xu, Zheng Yang, Zimu Zhou, Longfei Shangguan, Ke Yi, and Yunhao Liu. 2015. Enhancing wifi-based localization with visual clues. In *Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing (UbiComp'15)*. Association for Computing Machinery, New York, NY, USA, 963–974.
[48] Han Xu, Zheng Yang, Zimu Zhou, Longfei Shangguan, Ke Yi, and Yunhao Liu. 2016. Indoor localization via multi-modal sensing on smartphones. In *Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing*. ACM, 208–219.
[49] Jingao Xu, Hengjie Chen, Kun Qian, Erqun Dong, Min Sun, Chenshu Wu, Li Zhang, and Zheng Yang. 2019. iVR: Integrated vision and radio localization with zero human effort. In *PACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 3, 3 (2019), 1–22.
[50] Zheng Yang, Chenshu Wu, and Yunhao Liu. 2012. Locating in fingerprint space: wireless indoor localization with little human intervention. In *Proceedings of the 18th Annual International Conference on Mobile Computing and Networking*. ACM, 269–280.
[51] Zheng Yang, Chenshu Wu, Zimu Zhou, Xinglin Zhang, Xu Wang, and Yunhao Liu. 2015. Mobility increases localizability: A survey on Wwireless indoor localization using inertial sensors. *Computing Surveys* 47, 3 (April 2015), Article 54, 34 pages.
[52] Zheng Yang, Zimu Zhou, and Yunhao Liu. 2013. From RSSI to CSI: Indoor localization via channel response. *ACM Computing Surveys (CSUR)* 46, 2 (2013), 25.
[53] Xuehan Ye, Yongcai Wang, Wei Hu, Lei Song, Zhaoquan Gu, and Deying Li. 2016. WarpMap: Accurate and efficient indoor location by dynamic warping in sequence-type radio-map. In *2016 13th Annual IEEE International Conference on Sensing, Communication, and Networking (SECON'16)*. IEEE, 1–9.
[54] Zuwei Yin, Chenshu Wu, Zheng Yang, and Yunhao Liu. 2017. Peer-to-peer indoor navigation using smartphones. *IEEE Journal on Selected Areas in Communications* 35, 5 (2017), 1141–1153.
[55] Moustafa Youssef and Ashok Agrawala. 2008. The Horus location determination system. *Wireless Networks* 14, 3 (2008), 357–374.
[56] Mingmin Zhao, Shichao Yue, Dina Katabi, Tommi S. Jaakkola, and Matt T. Bianchi. 2017. Learning sleep stages from radio signals: A conditional adversarial architecture. In *International Conference on Machine Learning*.

Received April 2021; revised August 2021; accepted August 2021