Front.Comput.Sci.
DOI

REVIEW ARTICLE

From One to Crowd: A Survey on Crowdsourcing Based Wireless Indoor Localization

Xiaolei ZHOU [1,2], Tao CHEN [2], Deke GUO (✉)[2], Xiaoqiang TENG [2], Bo YUAN [2]

1 Nanjing Telcommunication Technology Research Institute,
National University of Defense Technology, Nanjing, 210089, China
2 Science and Technology on Information Systems Engineering Laboratory,
National University of Defense Technology, Changsha, 410073, China

Abstract Wireless indoor localization has attracted growing research interest in the mobile computing community for the last decade. Various indoor available signals, including the radio frequency, ambient, vision, and motion signals are extensively exploited to derive location estimation in the indoor environments. The physical measurements of those signals, however, are still limited by both the resolution of devices and the spatial-temporal variance of signals. One type of noisy signals, as a complementary to another type signals, can benefit the wireless indoor localization in many ways, since those signals are related in physics and independent in noise. In this article, we survey this new trend of integrating multiple chaotic signals to enable the crowdsourced localization system. Specially, we first present a three-layer framework for crowdsourcing based indoor localization with integrating multiple signals and illustrate the basic methodology for taking use of the available signals. Next, we study the mainstream signals involved in indoor localization approaches, in terms of their characteristics and typical usages. Furthermore, considering the multiple different outputs from different signals, we present the high-level insights to integrate them together to achieve the localizability in different scenarios.

Keywords Wireless indoor localization, Crowdsourcing system, Crowdsensing

Received month dd, yyyy;

E-mail: *guodeke@gmail.com*

1 Introduction

Wireless indoor localization has become one of the most attractive research trends during the last decade. Compared with the outdoor space, the indoor localization is much more challenging, since the widely used GNSS signal (e.g., GPS) can hardly get into the inside of the complicated indoor environments. The indoor environment is therefore termed as the GNSS-denied environment.

The early studies of indoor localization exploit dedicated devices. Roy Want and Andy Hooper [1] propose the first indoor localization system, Active Badge, in 1992. They designed an active infrared badge with a unique ID for each user, which emits pulse-width modulated infrared signals periodically. Those signals can be identified with the pre-deployed infrared receivers to infer the room-level locations. Andy Hooper further proposed Bat [2], a similar system with Active Badge, to exploit wireless and ultrasound signals to realize indoor localization. Inspired by such efforts, researchers have exploited the other signals for localization purpose, including acoustic [3, 4], vision [5], radio frequency [6] and RFID [7]. Those approaches exploit either the *time* or *power* of specific signals for ranging, and then estimate the locations with multiple ranging relations. However, they usually require expensive and *dedicated* devices, which greatly limit their applications in reality. Besides, the deployment of those dedicated devices also requires careful calibration of the indoor environments with *experts*. In a word, the early studies of indoor localization can be described as *dedicated experts* with *dedicated devices*. The recent development on indoor localization is boosted by two milestones, the pervasiveness of mobile devices and the emerging of crowdsourcing.

Pervasiveness of mobile devices relaxes the requirements of indoor localization from the *dedicated devices* to the *pervasive devices*. As the mobile devices, such as the smartphones and various wearable devices, become popular, the physical measurements of indoor available signals become much easier than before. The richness of the build-in sensors on the mobile devices not only eases the physical measurement of the ambiance signals, but also provides a new opportunity to explore those physical measurements of various signals to enable indoor localization. However, those signals are sensitive to the complex indoor environments. For example, the Received Signal Strength Indicator (RSSI)-based approaches suffer from dramatic performance degradation, due to the multipath fading and temporal dynamics. To design an accurate and reliable localization system, we have to carefully analyze the factors that affect the quality of the signals. In this paper, we first discuss the influence of those factors, and then analyze the key features of the widely used signals in the indoor localization and navigation systems.

Emerging of crowdsourcing turns the participants of indoor localization from *experts* to *ordinary individuals*. Crowdsourcing-based indoor localization is a rising trend among the recent efforts in indoor localization. Crowdsourcing has been regarded as a new collaborative paradigm that leverages the power of tens of thousands of participants to accomplish one specific task. Since the calibration of the indoor fingerprinting database, also known as site survey, is usually labor-intensive, the construction of the indoor localization seems not cost efficient in practice. Owing to crowdsourcing, ordinary individuals are able to participate such a process. More importantly, those individuals are also the users that benefit from the constructed localization systems; and hence are willing to share their data with others to achieve a globally localizable indoor environments. The advantage of crowdsourcing greatly encourages the development of the indoor localization and indoor location-based services (ILBS).

In this survey, we take an important philosophy in systemic theory, *order out of chaos* [8], to reconsider the indoor localization. We not only illustrate the state-of-the-art works in indoor wireless localization, but also discuss how to boost a crowdsourcing based indoor localization system. Various signals have been exploited to derive an estimation of locations. However, existing techniques exhibit different limitations. For example, the channel state information (CSI)-based approaches [9, 10] are considered as the best indoor ranging technique, and achieves centimeter-level accuracy [11]. However, the localization accuracy shows significantly long-tail effect for the entire localization accuracy [12], since the strict requirements on sufficient line-of-sight relations. Considering the pervasive requirements and complexity of the indoor spaces, building an indoor localization system means not only to exploit various signals to derive location estimations, but also to integrate those estimations into order. This survey provides a unique perspective to understand the crowdsourcing-based indoor localization systems. The contributions of this survey is summarized as follows:

First, we summarize the signals exploited in current works and analyze their characteristics, in terms of the spatial-temporal variance, device diversity, human factors, etc.

Second, we illustrate both the high-level insights and detailed techniques to integrate various signals to provide the pervasive indoor location service.

Third, we propose the challenging issues to boost crowdsourcing based indoor localization systems.

We expect this survey to act as a guideline for both the researchers and the developers in both the academic and engineering community to design novel crowdsourcing based indoor localization systems. There are several very good surveys for indoor localization [13–16]. Yang et al. review the underlying techniques of radio frequency signal in [13] and motion signal in [14]. Pei et al. present the first work for crowdsourcing-based indoor localization, and emphasize only the fingerprinting-based approaches with opportunistic signals. Adler et al. [16] reviews the high-level statistics results for studies on indoor localization since 2010 to 2014 for their categories. Compared with those works, we provide a thorough and comprehensive review for the crowdsourcing-based indoor localization in terms of the temporal-spatial characteristics of the underlying signals, integration techniques, and the crowdsourcing designs. We also suggest the readers of interest to refer those surveys as complementary of this paper.

The rest of this paper is organized as follows: Section 2 presents an overview of crowdsourcing-based indoor localization in aspects of a three-layer framework and the basic methodology. Corresponding to the three-layer framework, we review the recent developments of indoor localization in aspects of *characteristics of underlying signals*, *integration techniques*, and *crowdsourcing designs*in Section 3, Section 4 and Section 5, respectively. Section 6 summarizes the challenges for crowdsourcing-based indoor localization. Finally, we conclude this paper in Section 7.

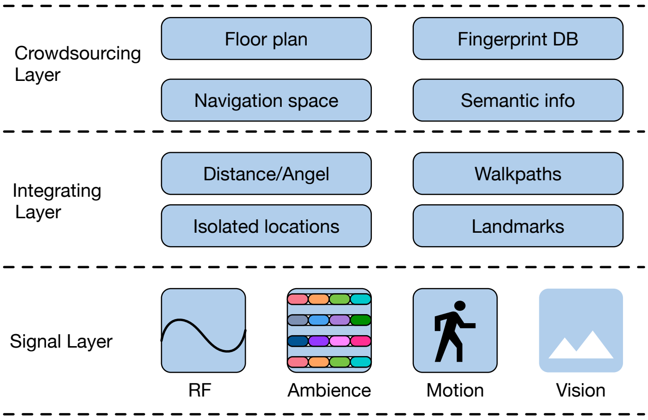


Fig. 1 A three-layer framework for crowdsourcing based indoor wireless localization systems.

2 Overview

In the past decades, so many innovations have been presented to contribute to break the fetters of the complexity in the indoor spaces, and pursuit higher accuracy of localization. This survey focuses on the state-of-the-art works in crowdsourcing based indoor localization. Before we step deeply into those innovative techniques, we first present a three-layer framework to figure out the roles of those amazing techniques in the whole picture of indoor localization. We expect such a framework would help the readers to understand the fundamental problems in the indoor localization domain. We then summarize the basic methodology in current studies.

2.1 Framework

We first present a three-layer framework for the current studies in crowdsourcing based indoor localization, as shown in Fig.1. Such a framework consists of three layers, namely *signal layer*, *integrating layer*, and *crowdsourcing layer*.

Signal layer. The signals are the foundation of building an indoor localization system. A large body of current works focus on the design of the localization approach with proper signals. One of the most challenging issues is the lack of pervasive and stable signal for localization, such as the widely used GPS signal in outdoor localization. It motivates the researches to exploit various available signals to infer the possible locations of any indoor object.

We category the current available signals indoor into four types in the signal layer, namely *radio frequency signal*, *ambient signal*, *motion signal*, and *vision signal*, and analyze the time-spatial characteristics of those signals with different types, as well as their representative usages for localization purpose in Section 3.

Integration layer. The outputs of *signal layer* are distance, angle, isolated locations, walk-paths, and indoor landmarks, with either fingerprinting or ranging techniques. Considering the complexity and diversity of the indoor spaces, it is difficult to derive a pervasive localization approach to achieve a promising localization accuracy with merely one certain signals. The localization techniques with different signals might be complementary to each other; and hence can be integrated to facilitate the indoor localization. We summarize the high-level insight to integrate different signals together for localization in Section 4.

Crowdsourcing layer. A large body of indoor localization approaches are designed in a crowdsourcing manner. The large-scale participations of anonymous users benefit the indoor localization in many ways, including the automatic construction of the indoor floor plans, fingerprinting database and the navigation space with all the walkable paths in certain environments. Besides, mining the large-scale user data is an effective way to discover the semantic information in the surroundings. Mapping the semantic information with the locations would facilitate the intelligent indoor location-based services. We illustrate the current works in crowdsourcing systems in Section 5.

In summary, the chaotic physical measurements of various signals are converted into isolated locations, ranging parameters, walk-paths, etc. Those intermediate location-related outputs are integrated in the integrating layer to achieve their full potential for localization in various scenarios. Finally, the measurements from large-scale anonymous users contribute to construct a pervasive indoor localization system. Owing to the complex indoor environments and the limitations of the devices, those measurements of different signals are with unknown and unpredictable noise. Therefore, the physical measurement at a certain spot from a single user is hardly to achieve localizability. Integrating those signals truly embodies the philosophy in systemic theory: *order out of chaos* [8].

2.2 Methodology

The key to any indoor localization systems is essentially how to convert the chaotic physical measurements of various signals to either locations or location-related parameters. The later one includes *distance/angle*,*walk-paths* and *signatures*. Before going deeply into the detail of any specific approach, we first provide an overview of the current works in the methodology level, and summarize the behind methods for

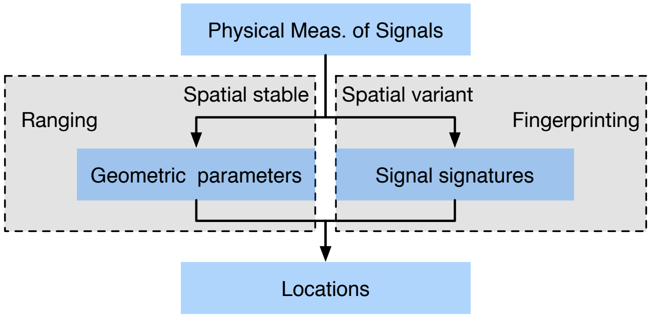


Fig. 2 Methodology of localization.

the innovations in the current literature as *ranging* and *fingerprinting*. We find that most works either directly use one of the two methods, or a combination of them. In general, if the measurement of signal is spatial stable, it is more suitable for the ranging method. Otherwise, it is probably more suitable for the fingerprint method, as shown in Fig. 2

2.2.1 Ranging

With various ranging techniques, the physical measurements of some signals are firstly converted to some geometric parameters (e.g., distance and angle), and finally derive the actual location with different geometric models. With proper geometric models, those parameters are then converted into locations with trilateration or multilateration approaches. The range of geometrics models does not exceed the models utilized in the conventional problem of localization [17], which includes Time of Arrival (ToA), Time difference of Arrival (TDoA), Angle of Arrival (AoA).

To derive a high-accuracy location estimation, however, usually has some strict requirements on the stability of spatial factors. In other words, only those signals that are robust to the *spatial-diversity* can be utilized for ranging. Those physical measurements include 1) *power* of radio frequency signals, 2) *time* of radio frequency signals, 3) *time* of acoustic signals, 4) *walk-paths* with motion signals, and 5) direct measurements of distance or angle with multiple images or point cloud data.

Despite the clear geometric relationships in the underlying models, it is still challenging to derive a high-accuracy estimation of the location [18, 19]. First, the indoor spaces are with rich multipath effects, which makes the physical measurements are sensitive to the spatial diversity. Second, the accuracy of ranging is heavily relied on the resolutions of the devices. For example, the motion signals measured by different devices usually vary a lot. Finally, some ranging techniques require specific infrastructures, such as MIMO devices or dedicated device on receiver side (e.g., Intel 5300 NIC for measuring PHY layer information).

In summary, some of the physical measurements are convertible to the geometric parameters to facilitate the location estimation. We further illustrate the typical ranging techniques with various signals in Section 3.

2.2.2 Fingerprinting

Different with the *spatial-stable* requirements on the ranging-based approaches, Fingerprinting-based approaches are exactly taking use of the *spatial-diversity* of various signals to estimate the locations. The main idea is to survey the surrounding signatures at every location in the areas of interests and then build a fingerprint database. The location is then estimated by mapping the measured fingerprints against the database.

The range of signals of fingerprinting based approaches is much larger than that of the ranging-based approaches. Nearly all kinds of ambient signals and RF signals can serve as fingerprinting. Besides, the combination of different signals further improve the uniqueness of the fingerprinting at each spots. Therefore, researchers have striven to exploit different signatures of the existing devices to enlarge the fingerprinting space.

The major drawback of fingerprinting lies on the labor-intensive calibration process of various signals. The grain size of the fingerprinting database directly affect the localization accuracy.

3 One: Signals

The physical measurements of various signals are the foundation of any localization approach. The lack of a GPS-like global reference signal in the indoor environments makes the indoor localization a challenging issue, until the sensor-rich mobile devices become popular in our daily life. To date, the mainstream smartphones have been integrated with multiple sensors, including WiFi, Bluetooth, GPS, GSM/CDMA Cell, near filed communication (NFC), cameras, microphone, light sensor, accelerometer, gyroscope, magnetometer, barometer, and proximity, etc. Those on-board sensors not only enlarge the sensing ability of the mobile device, but also achieve plentiful physical measurements of various signals to facilitate the computation of the indoor locations.

This section focuses on the signal layer in crowdsourcing based indoor localization systems. We divide those signals into four major categories, *radio frequency signals*, *ambient signals*, *motion signals*, and *vision signals*. In the rest of this section, we focus on the characteristics of those signals and their typical usages for the indoor localization. We summarize the signals utilized in the mainstream indoor localization approaches and thoroughly analyze their characteristics.

3.1 Radio frequency signal

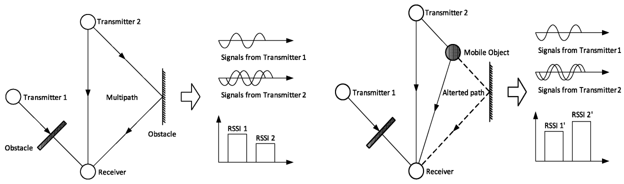
Radio frequency (RF) signals are the electromagnetic waves with frequencies ranging from around 3 kHz to 300 GHz. The RF signals are widely used in radar and wireless communication. Considering the hardware requirements to measure those RF signals, the physical measurements of three types of signals are supported by current mainstream mobile devices, i.e., *WiFi*, *Bluetooth*, *FM Radio*. This section focuses on the physical measurement of those three types of RF signals, as well as their characteristics and typical usages.

3.1.1 RSSI of WiFi signal

Nowadays, WiFi access point (AP) has become one of the most common infrastructures in many indoor spaces, ranging from commercial facilities to offices buildings [10, 20]. On the other hand, the physical measurement of WiFi signals are supported by most of the mainstream off-the-shelf mobile devices. Therefore, the WiFi signals are adopted in a large body of indoor localization systems [21].

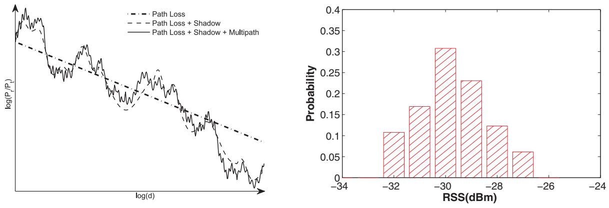
RSSI of WiFi is a signature in the MAC layer that describe the the *power* of WiFi signals during propagation in the MAC layer. As for the WiFi signals, RSSI of the WiFi signal from different APs can be obtained as a MAC layer signature. Currently, the RSSI measurement is supported by 802.11 a/g/n protocols at the receiver side. Despite the easy availability of the WiFi RSSI, the rich multipath fading and temporal dynamics in the indoor environments make such a power signature fragile to the spatial factors. Multipath effect means that the signals reaching the receiving antenna by two or more paths, as shown in Fig. 3. Therefore, the RSSI at some spots are increased due to the constructive, while decreased at some other spots due to destructive. Besides, the RSSI of WiFi signals are easily affected by the device-variance and other human factors, such as the attitude of device and the human movements in the surroundings. Fig. 4 shows the spatial and temporal variance of WiFi RSS. In the recent literatures, the RSSI of WiFi signal is exploited in both ranging-based and fingerprinting based approaches.

Ranging. The RSSI of WiFi signal is a power signature; and hence can be converted to distance with the prevalent Log-normal Distance Path Loss (LDPL) model [22] as Equation (1).



(a) Destructive. (b) Constructive.

Fig. 3 Multipath effect [23].



(a) Spatial variations. (b) Temporal variance.

Fig. 4 Spatial and temporal variance of WiFi RSSI [13].

$$PL(d) = PL(d_0) + 10n\lg(\frac{d}{d_0} + X_\sigma), \tag{1}$$

where $PL(d)$ denotes the measured path loss at distance d, $PL(d_0)$ denotes the average path loss at the reference point with distance d_0 to the source as a prior, n is the path loss exponent, and X_σ is a zero-mean normal random noise.

Actually, the path loss stems from the dissipation of transmission power in the propagation channel, while shadowing results from the obstacles that attenuate signal power through absorption, reflection, scattering, and diffraction.

The LDPL model relates the RSSI to the propagation distance from the reference point. When one is in the indoor environments, the device obtains the RSSIs from multiple APs. Such a scenario is usual in our daily experience. When the locations of those APs are known, with distances computed by LDPL model, the user's location can be estimated by multilateration. Due to the rich multipath fading and the high temporal dynamics in complex indoor environment, the parameters in the LDPL model differ a lot. Before converting the RSSI to the distance of propagation path, a calibration phase is required to estimate those parameters. Such a phase usually involves large-scale human-intensive measurements at different reference points.

To reduce the human efforts during such a phase, most works exploit the relations between the RSS of multiple APs to enable the dynamic and automatic estimation of those parameters. Gwon and Jain propose TIX [24], to modify the protocols in the commercial APs to make any AP able to measure the RSSI from the neighboring APs. Those RSSIs and distances achieve distance-related liner functions. Thus, the RSS at any spot in the area of interest is able to be estimated. Lim et al. [25] utilize a unique technique, truncated singular value decomposition to construct the mapping between the signal and distance. Ji et al. [26] propose a novel approach to dynamically construct the indoor signal map.

Those works either require the APs' locations known as priori knowledge, or the human intervention to estimate the parameters for the specific indoor environments. Chintalapudi et al. propose EZ [27] to break those limitations and realize the RSSI-based indoor localization without any priori knowledge and extra devices. The main idea of EZ is that when sufficient distances between APs and the mobile devices, their relative locations are able to be estimated. If any three global locations of those mobile devices are available, the actual locations of the other mobile devices can be therefore uniquely determined. EZ [27] is considered the first work that realize indoor localization without any human intervention. However, it still requires the APs are deployed with high density and the occasionally obtained accurate locations of the mobile devices via GPS signals. Abhishek et al. propose WiGEM [28], a learning-based approach to automatically estimate the parameters in the propagation model with Gaussian Mixture Model(GMM).

In a word, ranging with either RSSI of WiFi signal lack of accuracy guarantee. The localization error in those approaches ranges from 2 to 10 meters, which depends on the actual settings of various indoor environments. The root reason is that the MAC layer power signature of WiFi is affected by both the spatial and temporal factors. A finer grained ranging technique with the PHY layer measurements of WiFi signals as *channel response* will be illustrated in Section 2. We now discuss how to utilize the spatial variance of the WiFi signals to achieve indoor localization.

Fingerprinting Fingerprinting-based approaches utilize the spatial variance of signals to extract a unique signature for each spot. When there is a localization request, one can estimate his/her location by finding a best match in the fingerprinting database. Such an idea is first presented by Bahl et al. with RADAR [29] system, which is considered as the first work in the fingerprinting-based indoor localization approach as well. Since then, most of the fingerprinting-based approaches are either variants or improvements of the RADAR system. Horus [30] approximates the fingerprint at each spot by a parametric distribution to improve the performance of fingerprint matching.

The defects of the WiFi RSSI fingerprinting based approach are significant. First, the spatial variance of WiFi RSSI does not always achieve distinction among different spots. Owing to the complexity of the wireless propagation in the indoor environments, the fingerprints with WiFi RSSI at different spots might be similar, and further lead to the false positive in the matching, or mis-matching for the fingerprinting-based approach. Besides, construction of the indoor fingerprints database usually requires large-scale and labor-sensitive. A more strict requirement on the localization accuracy in the fingerprinting-based system requires a finer grained calibration of the indoor RSS map, which definitely enlarge the efforts and overhead of the fingerprint database.

Both of the two challenges can not be directly addressed with the WiFi signal itself. The current innovations integrate multiple signals with either independent or identical independent noise to improve the distinction of the fingerprints, and integrate with other dimensional measurements to eliminate the human-intervention for fingerprint database construction. For the ease of presentation, we just discuss the fingerprinting with merely WiFi RSSI here, and leave the illustration of the detailed solutions for those two important issues in the integration layer (See Section 4).

3.1.2 CSI of WiFi signal

Multipath effect is the major error source in the power-based ranging. To eliminate the multipath effect of WiFi signal, the channel response at PHY layer information is exploited. In the early study, the physical measurements on the wireless channel response require dedicated professional equipment. As the 802.11 a/g/n standard, channel response can be partially extracted from off-the-shelf Orthogonal Frequency Division Multiplexing (OFDM) receivers in the format of Channel State Information (CSI). It reveals a set of channel measurements depicting the amplitudes and phases of every sub-carrier in time domain.

Compared to RSSI of WiFi in the MAC-layer, the PHY layer power feature, channel response, is able to discriminate multipath characteristics. To help the readers well understand channel response, we briefly explain the two components in CSI, i.e.,channel frequency response (CFR) and channel impulse response (CIR).

CFR is a description of the channel performance at each subcarrier in the frequency domain. The OFDM communication transmits signals across orthogonal subcarriers at different frequencies. Each symbol $X(f)$ is modulated on a different subcarrier f, and the quality of the received symbol $Y(f)$ will depend on the channel $H(f)$:

$$Y(f) = H(f)X(f). \tag{2}$$

Vector $\mathbf{H} = (H(f))_{f=1,\ldots,F}$ is the channel frequency response (CFR).

CIR is a temporal linear filter to characterize the multipath for the wireless channel in the time domain. Suppose there are N path from the transmitter to the receiver, and the amplitude, phase, and time delay of the i^{th} path are denoted as a_i, θ_i and τ_i, respectively. Under the time-invariant assumption, CIR $h(\tau)$ is denoted as:

$$h(\tau) = \sum_{i=1}^{N} a_i e^{-j\theta_i} \sigma(\tau - \tau_i), \tag{3}$$

where $\sigma(\tau)$ is the Dirac delta function. Given infinite bandwidth, CFR is the Fourier transform of CIR.

There are two ways to utilize the power-based ranging with CIR. One is to extract the multipath factors [31].

$$P_r(d) = (\frac{\lambda}{4\pi d})^2 P_t G_t G_r. \tag{4}$$

Ranging with CIR. CSI based ranging is currently the most reliable distance measurements in the indoor environments. There are two ways to derive the distance form CIR. The first one is to utilize the Friis equation to simultaneously extract all the multipath sections [31]. The other one is to extract the Line-of-Sight (LoS) path to estimate the propagation distance [32]. The accuracy of the CIR based ranging depends on the time resolution of the multipath, which is equivalent to the system bandwidth.

Fingerprinting with CSI of WiFi. Since the CSI of WiFi signal achieves better spatial distinction and temporal stability than RSSI of WiFi, it is more suitable to serve as fingerprints for indoor localization. Both CIR [9,33] and CFR [34,35] have been used in fingerprinting. The accuracy of the CSI based fingerprinting approaches achieves sub-meter level accuracy, which overweights the RSS based approach.

Due to the page limits, we omit the details of the CSI based ranging and fingerprinting. We suggested the interested readers to [13], which provides a thorough survey for the CSI based indoor localization approaches.

3.1.3 Bluetooth signal

Bluetooth becomes popular for short distance communication. To date, the Bluetooth module has been integrated into the mainstream smartphones and wearable devices. Similar to the WiFi signal, Bluetooth operates at 2.4GHz. The transmitting power of Bluetooth, however, is much lower than WiFi; and hence prolong the life time of battery. Similar to the WiFi signals, Bluetooth signals are easily affected by the multipath effect and the temporal dynamics as well. The limited coverage Bluetooth signal makes it not suitable to serve as fingerprints in large-scale indoor environments. Therefore, most of the current works take the signal strength of Bluetooth as the coarse-grained ranging, rather than fingerprinting [36, 37].

Antti et al. [36] propose the first Bluetooth based indoor localization system. The RSS of Bluetooth is converted to distance with LDPL model. The 3D position is achieved with those distances with the extended kalman filter (EKF). The experimental results show that the mean localization accuracy is 3.76 meters. While Mortaza et al. [37] propose the first Bluetooth fingerprinting-based approach for indoor localization. The inquiry response rate, which is the percentage of inquiry responses to total inquiries, servers as fingerprints for localization. However, such an approach is still challenged by the limited coverage of Bluetooth devices.

Recent works focus on the utilization of the Bluetooth low energy devices such as iBeacon, which are emerging as the new generation infrastructure in many indoor scenarios. Those devices are able to transmit a universally unique identifier, UUID, as well as the message to the nearby devices. They are cheap in expense and with longer lifetime than the WiFi infrastructure. Lin et al. [38] used an RSS-based algorithm to locate a mobile device in the indoor environments using the Bluetooth sensors.

3.1.4 FM broadcast signals

Frequency modulation (FM) broadcast signal is pervasively available across the world, which is still popular to provide the public with high-fidelity sound over broadcast radio, including news, musics, traffic messages, and advertisements.

FM signals operate at the frequency range of 88-108MHz, which makes them less susceptible to the presence and orientation of humans and small-scale furnitures. Furthermore, FM signals are significantly stronger than WiFi signals in the sense that they can easily cover areas of hundreds of kilometers, while achieving good indoors penetration. From the infrastructure point of view, there are thousands of commercial and amateur FM signals being broadcasted continuously across the world, eliminating the need for deploying any custom infrastructure. Also, most mobile devices, even the lower-end ones, are equipped with FM radio receivers that are lower power and less costly compared to the WiFi receivers.

Most of the current works take the signal strength of the

Table 1 Comparison of radio frequency signals.

Signals	WiFi RSSI	WiFi-CSI	Bluetooth RSSI	FM radio
Spatial-variance	high	high	high	high
Time-variance	high	low	low	low
Device-variance	high	low	high	low
Human factor	high	high	high	low
Frequency	2.4GHz	2.4GHZ	2.4GHz	800MHz
Energy-consumption	high	high	low	low
Ranging accuracy	low	high	low	low
Fingerprinting	low	high	low	high

FM signals as a power signature to design fingerprinting-based approaches [39–42]. Matic et al. [39] demonstrate the feasibility of taking the p,ower of FM radio as fingerprint for indoor localization, and prove that the FM radio achieves similar accuracy with the WiFi signal. Bedsides, combining the RSS of WiFi and FM signal together as fingerprints can further improve the overall localization accuracy, since the noise and multipath for those signals are independent. With a similar insight, Chen et al. [41] exploit the PHY layer information (SNR, multipath) of FM signals to augment the robustness of fingerprints. Due to the characteristics of lower frequency and longer wavelength of FM radio, the experimental results show that the overall localization accuracy has been improved about 83% than the approach with merely WiFi signal. Although those works achieve significant accuracy, the overhead of constructing the FM radio map is non-trivia.

To reduce such an overhead, Yoon et al. [42] propose a novel approach to automatic construct the fingerprint database. The path loss of the FM radio signals vary a lot in the outdoor and indoor environments. They model the propagation of FM signals in indoor environments with large-scale observations; and hence can directly compute the signal strength of FM signal from different tower at any indoor spot. The experimental results show that their approach achieves an accuracy of 6 meters on average.

Despite those innovations, the FM radio is still not in mainstream for indoor localization. The root reason is that the physical measurement of the RSS of FM radio is still not directly supported by the mobile devices. We notice that, the RSS of FM radio in those works are measured by the dedicated devices, such as the USRP device or specific FM radio receiver. As for the widely used smartphones, there lack of API for developers to directly measure the FM radio. In 2011, Sony published FM radio module for the Xperia phones as an open source, while the Communications Research Centre in Canada developed an Android library to access the FM-RDS capabilities of the Samsung Galaxy S. However, these two models of smartphones are outdated right now, and there is no any further update for those two projects. We expect there would be some updated liberality for the FM radio measurements in the future, to facilitate the real application of FM signal in various scenarios. We summarize the characteristics of four physical measurements of different RF signal in Table 1.

3.2 Ambient signals

To push the limit of the radio frequency signals, the researchers have taken the pervasive ambient signals in the indoor environments into consideration for indoor localization. The scope of the ambient signals includes the acoustic, geo-magnetic field, and the visible light. The physical measurements of those signals are supported by the mainstream mobile devices. With the ambient signals, we can either obtain more accurate ranging, or more plentiful signatures of the spatial variance; and hence can greatly improve the robustness of indoor localization to the temporal dynamics.

3.2.1 Acoustic signal

Several localization systems extended the scope of the signals utilized in the indoor localization into the acoustic signals. Compared with the RF signals, the acoustics are far more omnipresent in the indoor environments. Actually, the community has contributed to utilize the acoustics to construct the indoor localization systems. The first work can be traced back to the approach proposed in [43], which use the pre-deployed microphones to detect the presence of the people. Similarly, [44] designed a ultrasonic sonar sensing system to silently localize the pedestrian. However, those works require expensive dedicated devices, which limits the large-scale application in reality. Recently, the researchers have dived more deeply to exploit the characteristics of the acoustics in the indoor spaces. The most significant characteristics of

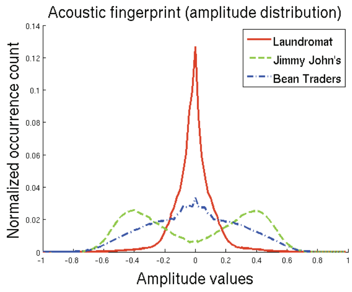


Fig. 5 Spatial distinciton of acoustic fingerprint [47].

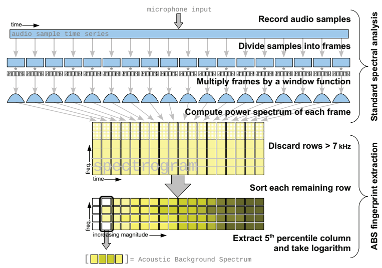


Fig. 6 Acoustic Background Spectrum fingerprint extraction [48].

indoor acoustic signals are spatial variance, and low-velocity in propagation.

Ranging with low-speed signal. Although the acoustic signals are challenged by the multipath effect as the radio frequency signals, it is still able to be used for accurate ranging. The propagation speed of the acoustic signal is much lower in than the RF signals in the air, i.e., 340 meters per second.

Based on the low-velocity, Peng et al. propose a high accuracy Time of Arrival (TOA) ranging system, BeepBeep, in [45] to measure the distance between a pair of smartphones. BeepBeep makes each device emit a specially-designed acoustic signal (i.e., Beep) by its speaker, and records two Beeps. One is from itself, while the other one is from its peer. As such, the time-of-flight between the pair of devices can be therefore derived in a synchronization-free way. Instead of using the local timestamps at each device, the time variables are computed by counting the number of samples between the two Beeps in the local records, since the sampling rate of acoustic is constant. As a result, BeepBeep improves the TOA estimation accuracy to the sampling rate granularity. That is, one to two centimeters accuracy for more than ten meters between the two devices.

Lim et al. [25] utilize the ray tracing and simulated annealing algorithm to estimate the parameters in the sound radio propagation model, and dynamically construct the indoor signal map. Guoguo to localize the pedestrian in [46], which deploys the anchor nodes at predefined locations, and then localize the accessed human with device-anchor distance. The prototype of Guoguo shows centimeter-level localization accuracy in various indoor environments.

Fingerprinting with spatial variance. The received signal is indeed a combination of the signal from the persistent acoustic source and the impulse response of the room. Therefore, the received acoustic signal is easily affected by the geometry of the room and the furnishings inside. Besides, the complexity and variety of the indoor environments make the spatial significant. Based on this, the researchers demonstrate the feasibly to fingerprinting the acoustic signals to localize the users at different spots.

Azizyan et al. propose SurroundSense in [47] to fingerprint multiple signals, including WiFi, sound, acceleration, color and light to infer the store-level location of the users'. The behind intuition is that the services provided by neighboring stores are generally different. Therefore, SurroundSense takes the amplitude distribution of the sampled sound signals (See Fig. 5), together with the physical measurements obtained by the other build-in sensors as a distinguishable ambient fingerprint. Tarizia et al. take one step further to discover the localization only with the acoustic signals in [48]. They propose a new ambient sound fingerprint, Acoustic Background Spectrum (ABS), to provide room-level localization in more general scenarios. In stead of using the amplitude of audio signal directly, ABS extracts the signal features in the frequency domain to label different rooms (See Fig. 6). The trace-driven simulations show that, nearly 92% neighboring rooms can be distinguished, even if their room type is similar. However, the accuracy of ABS-based localization is only 69% in the experiment.

In summary, using the acoustic signals as the ambient fingerprint can only provides coarse-grained localization accuracy, due to the influence of noise, similarity, temporal dynamics, and the other environmental factors. However, it can be utilized as a complementary of the other localization techniques to differ the neighboring rooms.

3.2.2 Geo-magnetic

The geo-magnetic filed is generated by the motion of molten iron alloys in the Earth's outer core. It is pervasively exiting while remains undisturbed in most of the outdoor environments on the earth. The behind reason is that the disturbances of the outdoor field do not change quickly. Based on this property, the compass, which is one of the most great inventions in the human history, has been utilized world-wide to estimate the direction in a brunch of navigation applications [49].

However, when the scenario turns to the indoor environment, the geo-magnetic seems much easier to be disturbed by the surroundings. For example, the indoor field may change even on a scale of a few centimeters or less. The root reason is that the physical measurement of geo-magnetic filed is easy to be affected by the floors, walls and physical objects inside, such as furnitures and machines. In fact, the metal material in the buildings is considered as the dominate factor that disturbs the indoor magnetic field. Actually, the reinforcing steal is the fundamental material in the modern architecture. Besides, a large amount of the furnitures and machines are made of metal. Angermann et al. reveal the characterization of geo-magnetic filed in the indoor environments thoroughly in [50] based on extensive experiments with a specific designed device. They have demonstrated the spatial variation of the geo-magnetic field. Even single samples of the geo-magnetic field allow to resolve the location with sub-centimeter accuracy with no ambiguity on a segment of length one meter. Such a fact is the key proof for using the geo-magnetic filed as location-dependent fingerprint in the indoor environments.

Based on the spatial variation of the geo-magnetic filed, a large body of works have contributed to construct geo-magnetic fingerprint floor plan for localization purpose. The authors in [51] fingerprint the ambient filed along the straight pathways. [52] introduced a pedestrian navigation system with human motion recognition, although a pre-established magnetic field map is equipped to correct the severe disturbance of indoor direction sensing. Besides, magnetic signatures are leveraged in [53] to identify locations and rooms. Although mobile phones are used to measure magnetic field intensity, the system relies on pillars and offers only room-level positioning accuracy. Grand et al. [54] proposed a lightweight magnetic map construction method and used an online particle filter to estimate the location of a hand-held device. Similarly, a particle-filtering-based engine was designed in [21] to localize and track users with a given geomagnetic fingerprint map.

Shu et al. propose FollowMe in [55], which utilizes the features of geomagnetic field without customized hardware and avoids the time-consuming map construction process.They have demonstrated that the uniqueness of the geo-magnetic filed measurements along pathways in the indoor environments. In contrast to the other works with the geo-magnetic filed measurements, FoolowMe dose not convert the measurements to the actual positions.

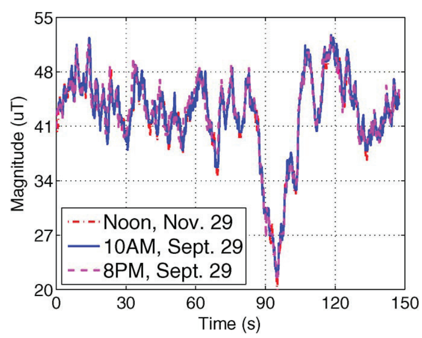


Fig. 7 Temporal variance of geo-magnetic filed [55].

3.2.3 Visible light

Lights are extensive existing in most of indoor environments. Dislike the RF or acoustic signals, light is unable to penetrate the walls or furnitures, and hence better reflects the structures and furnishing of the indoor spaces. However, the physical measurement of the strength of the conventional light sources is coarse-grained. For a long time, the light servers merely as a complementary to the other signals for coarse-grained indoor localization. For example, SurroundSense [47] takes the light as part of the fingerprints for room-level localization, while IODector [56] utilizes the light as the evidence to determine whether in the indoor or outdoor environments.

The emerging Light-emitting Diode (LED), as well as the Visible Light Communication (VLC) techniques [57, 58], provide a new opportunity for the indoor localization design. LED shows its potential to replace the conventional bulbs, and offers a revolutionary lighting technology with longer lifetime, better lighting efficiency, energy saving and environment preservation [59]. VLC takes the instantaneous on and off state of LEDs to embed digital information in the light to achieve high speed data transmission. Note that, such a process is not visible for human eyes.

Several researchers propose indoor localization systems using the VLC techniques which are expected more accurate than those based on other techniques (e.g., WiFi, FM, Bluetooth) [59–65]. Li et al. [59] exploit LED beacons and a custom light sensor that plugs into a smartphone's audio port to locate the mobile device in the indoor space using BFSK. The proposed system required users to perform gestures and offered half-meter accuracy. Kuo et al. used a camera embedded in a mobile device to estimate the location and orientation of the mobile device in the indoor space. In Luxapose system, the Angle-of-Arrival (AoA) method was used for triangulation to localize the receiver device. The proposed system achieved localization accuracy of about 0.1 meters. Similar to Luxapose, Rajagopal et al. [62]

Table 2 The details of Android mobile devices used in our experiments.

	WiFi	GPS	3G/4G	Camera	Resolution	RAM	Processor
MI	√	√	√	√	20 million pixels	2GB	Nvidia Tegra4
Nubia Z5	√	√	√	√	13 million pixels	2GB	Qualcomm APQ
Meizu	√	√	√	√	12 million pixels	2GB	Helio X20

Table 3 keypoints number in images by different devices.

	Mi3	Meizu	Nubia Z5
Number of keypoints	133	119	316

used cameras to receive from multiple luminaires each of which creates a visual landmark. Lee et al. [66] proposed RollingLight system to allow a light to deliver information to diverse rolling-shutter cameras by overcoming the diversity. This diversity is caused by the heterogeneous sampling rates of the phones. Different above systems, Tian et al. [63] proposed DarkLight, a new VLC primitive that allows light-based communication to be sustained even when LEDs emit extremely-low luminance by encoding data into ultra-short, imperceptible light pulses. Yang et al. proposed PIXEL [64], the polarization-based modulation, which is flicker-free, to enable a low pulse rate VLC.

3.3 Vision signal

Another type of signals widely used in indoor localization is the vision signals. As *camera* becomes one the major components in mainstream smartphones. Some models of smartphones are even equipped with more than one camera to improve the quality of images. On the other hand, image sharing takes a great part of the on-device applications in our daily activities. The availability of the images in the indoor environments has been greatly guaranteed.

Compared to those RF signals and ambient signals, the vision signals contains ever plentiful information. Through the vision signals, one can not only extract the physical structures of the indoor environments, but also obtain rich semantic information, such as points of interest (PoI), names of shops, concurrence of multiple entities. Fingerprint those semantic information greatly improves the robustness to the temporal dynamics for the fingerprinting-based indoor localization systems. Besides, integrating multiple images might help to derive distance relations, which contribute to the ranging-based indoor localization.

This paper reviews the vision-based indoor localization in aspects of the image-based indoor localization and the point cloud based indoor localization.

3.3.1 Image

The images in the context of indoor localization is directly captured by the smartphones. One of the typical usage for the image-based indoor localization is to take the images as fingerprints for the indoor landmarks. Before we further review the image-based localization approaches, the characteristics of the image signal is studied through a real experiments. We captured three images with three different models of smartphones simultaneously at the same position. The models of smartphones in the experiments are listed in Table 2. We take the Scale-Invariant Feature Transform (SIFT) algorithm in [67] to evaluate the similarity of any two images. It can be seen from Table 3 that the number of keypoints of images are significantly varies a lot for different devices. The above experiment shows that the images are sensitive to the device diversity. Besides, the environmental dynamics, such as illumination variation, also affects the quality of the captured images. Therefore, the images can not be directly used as a fingerprints for localization purpose.

To overcome the above challenging issues, a large body of innovations are proposed in the last decades. Instead of directly fingerprinting those images captured by the smartphones, researchers focuses on fingerprinting the extracted semantic informations as the signatures for the positions. Wang et al. [68] propose a crowdsourcing framework to integrate the RF signals (WiFi and GSM signal) and the images to construct the fingerprint database. With the proper selection of the searching space, such an approach overweights the traditional image-based fingerprinting approaches. Gao et al. [69] propose Sextant by leveraging environmental reference objects, such as store logos, to locate a mobile device in the indoor environments with triangulation. Since the reference objects seldom move, Sextant avoids extensive human efforts in obtaining and maintaining RF signatures in mainstream indoor localization technologies. Xu et al. [70] take images as a complementary of the WiFi based indoor localization to improve the overall accuracy. Wang et al. [71] take images and semantics (e.g., store names) to localize mobile devices in the indoor commercial scenario.

Besides, some other works integrate the images with the RF signals and ambient signals to comprehensively characterize the indoor spots. Chon et al. [72] focus on characterizing the indoor places with the opportunistic photos and audio clip from the users' side in a crowdsensing manner. They propose a novel framework, CrowdSense@Place (CSP), to employ five different images and audio classifiers to abstract the characteristics of different places. In general, such an approach actually exploits three different kinds of signals, the vision signal (images), acoustic signal (speech and ambience sound), and the radio frequency (Wifi and GSM).

Several other researchers focus on compressing of the original images. With uploading the compressed information of the captured images as a query, the backend server returns the corresponding locations to the user. The utilized compressing techniques include JPEG compressed query image [73], compressing the extracted query features [74], quantizing the query features into a bag of visual words and sending only the respective indices [75], and the signatures by exploiting the similarities of spatially neighboring reference images [76].

3.3.2 Point cloud

Compared with those indoor RF signals and ambient signals, the indoor structural information is much more stable and robust to the temporal dynamics. However, the measurement of such structural information requires dedicated devices. Thanks to the rapid development of the pervasive and ubiquitous computing, more and more hand-held devices have equipped with more than one camera. Those new models of the mobile devices not only improve the quality of the captured images, but also provide an opportunity to obtain the point cloud of the indoor environments. The point cloud greatly ease the burden for extract the indoor structural information.

A point cloud is a set of vertices in a three-dimensional coordinate system with X, Y, and Z values (as shown in Fig. 8). A variety of methods have been proposed to obtain point clouds. One of the most representative techniques is the Structure from Motion (SfM) [77], where point clouds are derived from images and 3D scanners (e.g., Google Project Tango kit, Kinect, Xtion Pro Live, and Lidar). The features extracted from point clouds are usually not affected by variations in scale, rotation and illumination [78]. Therefore, the point cloud is robust to environment factors. Despite the above benefits, the usages of point clouds face several major challenges [79]. First, the layouts of indoor environments are complicated and the 3D geometry of objects is with messy surroundings and substantial variation between parts. Therefore, point clouds are sensitive to layouts of indoor

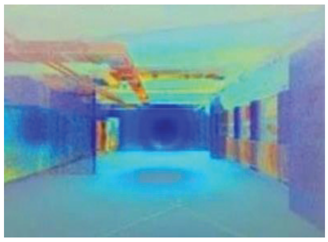


Fig. 8 An example of a point cloud where different colors represent different depths between the point in a point cloud and the 3D scanner.

environment changing. Second, point clouds captured by scanning devices are often noisy, may be distorted, and can have large gaps. To help the readers better understand the characteristics of the point clouds, we conduct real experiments to investigate the influence of the environmental factors to the indoor point cloud, in terms of device diversity, illumination variation, and different hand poses.

We ask 10 volunteers (different in genders, heights and weights) to conduct the experiments. 7889 point clouds along 288 walking traces are collected in our office building during a three-month period with two different devices(i.e., Google project tango tablet and the Xtion PRO Live camera). We take the number of the matched keypoints to measure the similarity of the difference of the two point clouds. The detection of the keypoints is based on the SIFT algorithm in [67].

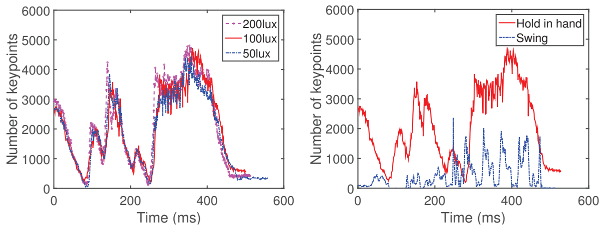
We first investigate the influence of illumination on point clouds. The point clouds are collected under three different illumination conditions (i.e., 200lux, 100lux, and 50lux) along the same path by the same volunteer. The experimental results are shown in Fig. 9(a). The number of keypoints of the collected point clouds is slightly changing among different illumination conditions, and such a changing is negligibly slight compared with images. Guo et al. demonstrate the similar performance for the point clouds in [78]. Therefore, the point clouds are robust to illumination variation.

To study the influence of device diversity on the point clouds, we collect the point clouds with four different devices along the same trajectory in our office building. The results are shown in Table 4. With different devices, the number of the detected keypoints significantly changes. Therefore, the quality of the point clouds is sensitive to the device diversity.

Finally, we evaluate the performance of the point clouds with different hand poses of the hand-held devices. We take two poses, i.e. hold in hand and swing in hand, in our experiments. A volunteer is asked to walk along a trajectory with the same device in our office building twice with the above two hand poses. The experimental results are shown in

Table 4 The number of points of the point cloud acquired by different 3D scanners.

	Velodyne Lidar	Kinect	Xtion Pro Live	Google Tango Device
Nomber of points	700, 000	300, 000	300, 000	10, 000



(a) Different illumination conditions. (b) Different hand poses.

Fig. 9 The number of keypoint detected in point clouds acquired under different illumination conditions and hand poses.

Fig. 9(b). As we can directly learn form the experimental, different hand poses lead to different numbers of detected keypoints.

In summary, the point clouds are robust to the illumination conditions, but sensitive to the device diversity and hand poses. Such characteristics make the point clouds based indoor localization still a very challenging issue in practice, especially for the crowdsourcing based solutions. The root reason is that the point clouds derived form the crowdsourcing data lack of quality guarantee. However, the interest of utilizing the point clouds in the context of indoor localization is ever growing in the last couple of years. We then review the current progress for the point clouds based indoor localization.

3.3.3 Indoor Localization with Point Clouds

The utilization of point clouds in the context of indoor localization can be roughly divided into two categories, i.e., Simultaneous localization and mapping (SLAM)-based and 3D maps-based.

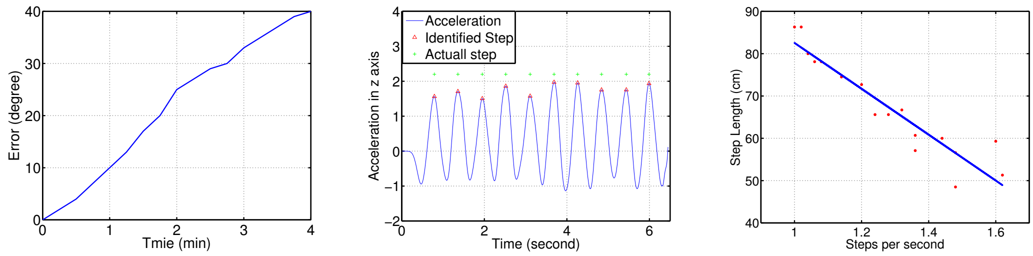
SLAM-based approaches. Traditionally, the major objective of SLAM is to reconstruct maps in an unexplored environment using different sensors (e.g., laser sensors, cameras, and odemetry). As a brand new and information-rich signal, point clouds provide SLAM a powerful techniques to characterize the indoor environments. Henry et al. [80] present one of the first full SLAM systems with merely point clouds. The propose a novel joint optimization algorithm that combines visual features and shape-based alignment to achieve this. Endres et al. [81] optimize the 3D pose graph using the g2o [82] framework. The authors in [82] also incorporate geometric error to allow tracking through scenes with little texture. LSD-SLAM [83] utilize direct image alignment together with filtering-based estimation of semi-dense depth maps as originally proposed in [84]. We propose a point cloud based navigation approach in [85]. Those innovations greatly expand the usages of the point clouds for localization purpose.

3D maps-based approaches. 3D maps-based indoor localization systems aim to localize a object by constructing 3D maps with point clouds. 3D mapping techniques have been developed using range scans in the robotics and computer vision communities [86–88]. KinectFusion [89] use surface reconstruction based method on a voxel grid containing the truncated signed distance [90] to the surface. Keller et al. [91] propose a real-time dense reconstruction method with equivalent quality to existing online methods, but with support for additional spatial scale and robustness in dynamic scenes. Al-Nuaimi et al. [92] propose an indoor location retrieval method as a part-in-whole matching problem of KinectFusion (KinFu) query scans in large-scale target indoor point clouds. Such an approach achieves an average localization accuracy of 6cm. Winterhalter et al. [93] propose a novel approach that globally localizes a user in a given 2D floor plan with a Google Tango project device. They also accurately track the user in such an environment using particle filters by estimating the 6 DoF pose of the device. Such an approach achieved 0.2 – 0.5m localization errors.

3.4 Motion signals

The MEMS inertial motion unit (IMU) is embedded in nearly all the mainstream mobile devices. Even in the pervasive wearable device, such as the smart bracelet, IMU is one of the standard components. Typically, IMU consists of accelerator, magnetometer, and gyroscope, which can measure the movements of the devices with the acceleration, geo-magnetic filed, and the angular velocity, respectively. With those physical measurements, one's movement can be therefore tracked with dead reckoning.

The physical measurements with those sensors, however, are with non-negligible noise. Besides, different models of IMUs vary a lot in both accuracy and resolution. More seriously, those two aspects of errors will accumulate during the process of dead reckoning, and finally lead to an inaccurate location estimation. we first give a brief illustration of the pedestrian dead reckoning, and then discuss how to deal with the errors in dead reckoning.



(a) Error in direction estimation with gyroscope. (b) Step detection accuracy. (c) Step Length vs. Frequency.

Fig. 10 Error analysis in dead reckoning.

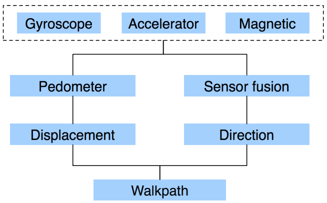


Fig. 11 Pedestrian dead reckoning.

3.4.1 Pedestrian dead reckoning: convert motions to walk-paths

Dead reckoning is a widely used as a complementary navigation approach to GNSS in marine, air, and automotive for decades. It estimates one's current location by using a previously determined position and estimated speeds and heading over elapsed time; and hence requires no global localizing signals. Different with the conventional dead reckoning, the dead reckoning on mobile devices is usually pedestrian based. The reason is that the large noise in acceleration measurements from the build-in sensors can not be directly integrated to estimate the velocity.

We show a typical pedestrian dead reckoning system in two-dimensional floor plan in Equation 5.

$$P_{k+1} = \begin{pmatrix} x_{k+1} \\ y_{k+1} \end{pmatrix} = \begin{pmatrix} x_k + l_k cos\theta_k \\ y_k + l_k sin\theta_k, \end{pmatrix} \tag{5}$$

where $P_k = (x_k, y_k)^T$ denotes the coordination at the k^{th} step, l_k and θ_k denote the estimations of the step length and walking orientation of user at the k^{th} step in the horizontal plane of earth frame (i.e., East-North-Sky frame), respectively. Note that, the above equation is in the earth's frame of reference, rather than the devices'. Therefore, the device's attitude is required to transit from its local frame of reference to the earth's. The transition from the local frame to the global frame of reference is beyond the scope of this paper. More detailed information can be found in [94].

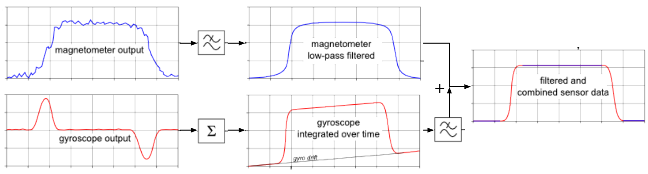


Fig. 12 Illustrative example of Sensor fusion.

3.4.2 Handling the error in dead reckoning

The user's location at the k^{th} step is computed iteratively with two types of inputs, including the location at the previous step and new coming measurements of inertia sensors. However, the location estimation suffers a localization error, due to the noisy measurements. More seriously, such an error will be accumulated over time. We further extract three major error sources from the pedestrian dead reckoning system, i.e., the *step detection*, *orientation*, and *stride length*. To investigate the influence of such factors, we conduct real experiment within a building. The experimental results are shown in Fig. 10. We ask a volunteer to walk with a smartphone steady at hand. To record the ground truth of the volunteers' trajectory, we paste tracer (some special costing) on the shoes of those volunteers' to show their footprints on the floor, and measure the real positions of volunteers with a tapeline.

We first evaluate the error in estimation of rotation with gyroscope measurements. As shown in Fig. 10(a), such an error is approximately linear to time. Similar observations can be found in the recent works [11, 94]. The gradient for error accumulation is device-dependent, which is about 10 degree per minute in our experiments. The user's walking direction is calculated by adding the rotation during one step to the direction at the previous step. Therefore, the error in the rotation leads to an accumulative error in the direction estimation over time.

The user's walking step is detected by finding the peaks in the magnitude of the readings of the triaxial acceleration. To improve the detection accuracy, we utilize the finite impulse response (FIR) filter to filter out both the high frequency

Table 5 Comparison of wireless indoor localization technologies.

Signals	Technique	Range	Positoning Error (m)	Complexity	Robustness	Cost	Infrastructure
WiFi	Ranging/ Fingerprinting	Indoor	1-7	Medium	Depend on the positioning algorithm and fingerprinting database	Medium	WiFi infrastructure
GSM	Ranging/ Fingerprinting	Indoor/ Outdoor	5	Medium	Depend on the positioning algorithm and fingerprinting database	Medium	Base station
Bluetooth	Ranging/ Fingerprinting	Indoor	2-5	Medium	Performance is sensitive to obstacles	Medium	Bluetooth tags
FM broadcast	Ranging/ Fingerprinting	Indoor/ Outdoor	<2	Medium	Depend on the positioning algorithm and fingerprinting database	Medium	Base station
Acoustic	Ranging/ Fingerprinting	Indoor/ Outdoor	0.4	Low	Performance is sensitive to sound noisy	Low	Acoustic sensors
Geo-magnetic	Ranging/ Fingerprinting	Indoor	<1	Medium	Performance is sensitive to metallic objects	Medium	Magnetic compass
Visible light	Ranging	Room	< 0.35	Low	Restriction on the number of LEDs	Low	LED lighting
Image	Image match	Room	1-4	High	Performance is sensitive to scale, rotation, and illumination	High	Infrastructure-free
Point cloud	Ranging/ Match	Several meters	<2	High	Restriction on the number of points	High	Infrastructure-free
Motion	Dead reckoning	Indoor	2-6	Low	Restriction on the number of landmarks	Low	Infrastructure-free

and low frequency noises in the acceleration measurements. Fig. 10(b) shows the step detection accuracy of a 10-step walk, where the green crosses and red triangles represent the actual step and identified step from the magnitude of triaxial acceleration measurements. Recent work [95] has also demonstrated the high accuracy of step detection, regardless of different carrying modes.

Users usually differ in their step lengths, due to the difference in genders, heights, weights and other personal factors. It is difficult to accurately measure the step length for each user via current build-in sensors. Although the error in the step length estimation is small, it also accumulates over time. Recent work [96] suggests that the step length is linear to the frequency of walking, i.e., $stridelength = a \cdot f + b$, where f denotes the step frequency and a, b are two user-dependent parameters. Such assumption is convinced by our experimental results. We ask a volunteer to walk naturally along the same straight line with different step frequencies. Fig. 12 plots the experimental results, where the red dots are the mean values of step length under different step frequencies, and the blue straight line shows the fitted linear relationship between the step length and the step frequency.

We show a typical pedestrian dead reckoning system in Fig. 11. The motions of the users is measured by the onboard sensors, includes gyroscope, accelerator and magnetometer. The step is detected by the pedometer. Together with the stride length estimation, the displacement of each step can be computed. On the other hand, the raw measurements are fused with proper filters, such as Kalman filter [97, 98] and complementary filter [99] to calibrate the errors to estimation the walking direction. The process of sensor fusion is shown in Fig.12. The errors in the magnetometer and gyroscope are independent, fusing them together can eliminate both the drifts and jitters in the raw measurements. With the computed displacement and direction, the motion signals are converted to the walk-path of users.

Due to the page limits, we only take the representative works for the optimization of the error in dead reckoning. We refer the interested readers to [14] for a detailed review on the error optimization techniques for dead reckoning.

3.5 Summary

In this section, we illustrate four types of signals which are frequently involved in the efforts on indoor localization. We compare the representative localization techniques for each kind of signal mentioned in Table 5, in terms of methodology, application range, accuracy, complexity, robustness to the temporal-spatial dynamics, cost, and the requirements on the infrastructures. We believe this comparison might be helpful for the readers to rapidly get an overview for the performance and limitation of the usage of different indoor available signals.

4 Integration layer

We illustrate the representative approaches for integrating multiple measurements of signals in this section.

The outputs of the signal layer are spot-dependent features, ranging measurements, walk-paths, and isolated locations. Most of the mainstream studies focus on the signal layer. A large body of innovations are proposed to exploit a certain kind of signals to design different localization approaches. To date, nearly all the signals available in current indoor environments have been exploited to support the location inference.

Building a good indoor localization system, however, is not as simple as computing the location of an objective with some of the available signals. The complexity of the indoor environment, as well as the available infrastructures, greatly limits the application of the indoor localization techniques. On the other hand, the outputs derived from different types of signals might be correlated. Besides, the noises in the measurements of different signals are usually independent. Thus, integrating those signals and outputs can improve the localizability. With the help of integrating, some spots, which are difficult to derive a promising location estimation, can be accurately localized.

Generally, the objectives of signal integration is three-fold. **First**, integrating the measurements of different signals from the a single user help us to fully exploit the utilization of the available signals. As different signals measured at the same spot, they are correlated with the locations. **Second**, integrating the measurements of the same signals from different users help us extend the scope of localization. As the users walk in the indoor environments, the measurements of a certain kind signal at different spot are correlated with the walk-paths. **Finally**, integrating the correlated and different signals to improve the localization accuracy. The noises in the measurements of different signals are either independent or independent and identically distributed. The actual independence of different signals depends on the characteristics of each signal. Integrating those different signals with proper techniques can eliminate the influence of measurement noise.

In a word, the signals in chaos are turned into order in the integration layer; hence is able to support the construction of the crowdsourcing system. This section summarizes five representative high-level insights for integrating those signals, as well as the innovative techniques in the current literatures.

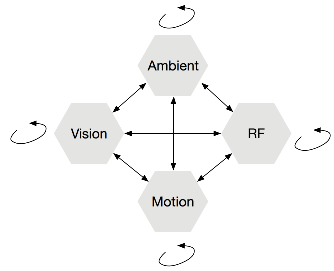


Fig. 13 Integrating various types of signals.

4.1 Integrating multiple independent signals as fingerprint

Generally, any signal with spatial-variance is able to serve as fingerprint for localization. However, the ambiguity of fingerprint leads to false positive matching for the location estimation. The major reason lies on the temporal dynamics and the noise in the physical measurements by different devices. Combining the measurements of independent signals greatly improves the ambiguity of the fingerprints, since the influences of the temporal and devices factors on different signals vary a lot. For example, for the case the fingerprints constructed by merely RSS of WiFi signal at two separated positions, they are less likely to be similar for the fingerprints constructed by the measurements on geo-magnetic filed. The spatial variance of the RSS of WiFi is shaped by the multipath fading and temporal dynamics, while that of geo-magnetic filed is shaped by the distortion of the iron and electronic devices and furnitures. Therefore, the fingerprinting-based approaches strive to integrate more available and independent signals to improve the spatial distinction of fingerprints.

As illustrated in Section 3, the signals that can be utilized as fingerprints include RSS of WiFi, CSI of WiFi, Bluetooth, FM radio, acoustic, light, images, and etc. Chen et al. [41] integrate the RSS of WiFi and FM radio together as fingerprints, which improves the localization with 83% than merely with WiFi signal. SurroundSense [47] integrate the measurements of sound, light, color of images, acceleration, WiFi together to realize a room-level localization, and achieves an average accuracy of 87% in the experiments. More utilization of the other signals combinations as fingerprints still requires the contribution form the interested readers in the futures, especially considering the signals with finer grained spatial distinction, such as the images and point cloud.

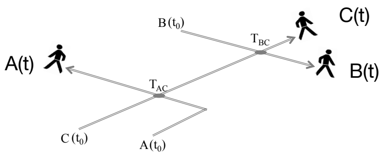


Fig. 14 An illustrative example of Escort system [100].

4.2 Integrating multiple walk-paths

One of the most widely used signal integration is to merge multiple walk-paths of independent users. Those walk-paths are derived from the dead reckoning with users' motion signals. When users walk in the indoor environments, their walk-paths are actually parts of the floor plan. Under different scenarios, integrating multiple walk-paths achieves the full potential of the current available signals.

Escort [100] is the first work that present the idea of merging multiple walk-paths. It takes use of the *encounters* (intersections) during the walk-paths to facilitate the application of finding the other person. Fig.14 shows an illustrative example for *Escort*. Three different users, denoted as A, B and C, locate at different and unknown spots at time t_0, and walk independently. Their walk path are shown as three lines in the figure. T_{AC} and T_{BC} are the intersections of the walk-paths of user A, C, and user B,C, respectively. When A wants to navigate to B, *Escort* plans a route, composed of segments between distinct intersections. For instance, if A has met C in the past, and C has met B recently, a walkable path from A's current position to B's position can be achieved by joining $(A(t), T_{AC})$, (T_{AC}, T_{BC}), and $(T_{BC}, b(t))$ together. *Escort* employs acoustic signal to identify the "encounters" during the walk, as well as a beacon to correct the accumulative error in dead reckoning.

Owing to walk-paths merging, it is no need to localize each user in real time, as well as the indoor floor plan, to navigate to the other person. The limitations of *Escort* are significant. It requires the device to predictably emit a acoustic signal to identify the encounters; hence leads to an additional modification at the users' devices. Besides, *escort* requires extra hardwares, i.e., beacons, as the anchor nodes to correct the error in dead reckoning.

Wu et al. [101] remove both of the above two requirements for path merging. They transform the indoor floor plan to a *stress-free* floor plan. A mesh of grids are constructed within the indoor floor plane, and intersecting locations of such a mesh is sampled. The distance between any pair of samples is then calculated. With the Multi-model scaling (MDS) approach, all the samples in to a d-dimension Euclidean space. In such a stress-free floor plan, the distance between

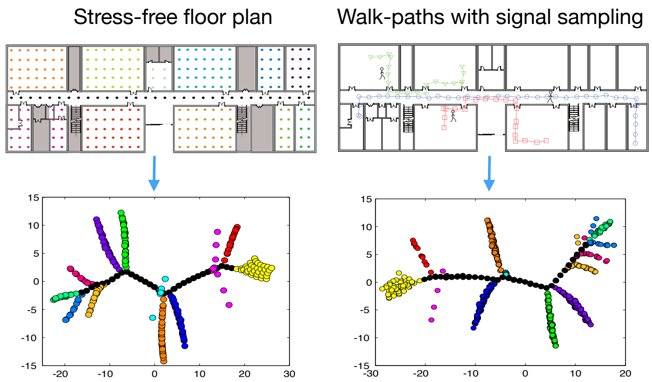
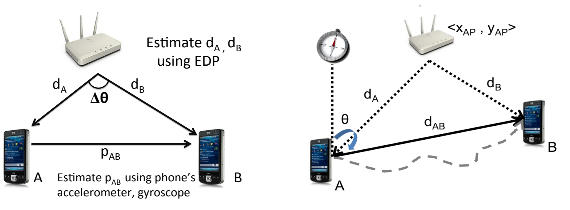


Fig. 15 Mapping the *stress-free floor plane* and the fingerprinting space [101].

any pair of samples is the actual walk distance between the corresponding locations in the floor plan. On the other hand, the users contribute their walk-paths, as well as the fingerprints achieved by measurements of WiFi RSSI along their walk-paths. The distance of between two fingerprint at different walk-paths is utilized to identify the overlapped part of the paths. The walk-paths and fingerprints form the fingerprint space, as shown in the right part of Fig.15. The stress-free floor plane and the fingerprint space is mapped with MDS approach. Note that, the computation of the walkable distance in the stress-free floor plan is accurate, while the walking distance derived from the walk-paths is with accumulative error. MDS is a error-tolerant approach that can gracefully eliminate the influence of the measurement errors.

Based on the above two works, we can derive two key issues for walk-paths merging: 1) accurate identification of the overlapped part of two different walk-paths, and 2) correct the accumulative error during the dead reckoning. *Escort* utilizes the acoustic signals to identify the "encounters" for the first issue, while employs anchor beacons to correct the error. LiFS [101] exploit the distance of features to address the first issue, while takes the benefits of MDS to eliminate the influence of errors.

Inspired by the above efforts, more works integrate the walk-paths to achieve different objectives. Walkie-Markie [102] takes use of walk-paths to automatically construct the digital indoor floor plane. It exploits the peak of the WiFi RSS along the walk-paths as the possible merging points and propose an spring algorithm to correct the error. Unloc [103] clusters the signals (WiFi RSS, Bluetooth, Magnetic filed) features along the walk-paths as the virtual landmarks to provide pervasive dead reckoning service. It is probably one of the best dead reckoning approaches in the literature. Wu et al. [104] take use of opportunistic static devices (phones and



(a) CUPID [107]. (b) SAIL [108].

Fig. 16 Integrating ranging and walk-paths.

pads) to automatically update the indoor fingerprint database.

4.3 Integrating walk-paths and vision

The vision signal contains more plentiful informations, such as the keypoints, semantic information, indoor entities of interest, etc. Integrating the walk-paths and vision signals, greatly extends the scope of the localization.

Travi-Navi [105], a vision-guided navigation system, enables a user to easily bootstrap and deploy indoor navigation services without building the entire localization system. Specifically, Trivi-Navi used the images of pathway to track the user's walking progress and the particle filtering method as the navigation engine. Besides, it provided multiple navigation traces to multiple destinations by automatically finding shortcuts, weaving the traces to form a holistic trace, and guiding users to destinations with minimum detours. However, Travi-Navi still requires particular infrastructures (e.g., WiFi access point) for practical use. Second, users can be guided from only a limited number of predefined locations, such as the building entrances.

Jigsaw [106] focuses on integrating the walk-paths between the identified landmarks to reconstruct the indoor floor plan. On one hand, it extracts the position, size and orientation information of individual landmark objects from images taken by users. On the other hand, it derives the spatial relation between adjacent landmark objects from the motion data. Finally, the coordinates and orientations of these objects on an initial floor plan are computed. By combining user mobility traces and locations where images are taken, it produces complete floor plans with hallway connectivity, room sizes and shapes. Experiments show that the localization of landmarks achieves 1∼2 meters, and all the corridors are correctly constructed in the digital floor plan.

4.4 Integrating walk-paths and ranging

The CSI-based ranging technique is the most promising ways to achieve high-accurate indoor localization, which has also improved the localization accuracy to centimeter level. Typically, it requires sufficient relations in distance or angle to achieve localizability with ranging. Those approaches usually requires the densely deployment of the dedicated hardwares. Besides, the accuracy of ranging, however, heavily relies on the condition of LoS paths. The walls, furnitures, and people are all the obstacles for the WiFi signals. Therefore, it is still changeling to realize a high-accuracy ranging-based indoor localization system, even with the help of the fine-grained PHY-layer channel information.

To push such a limit, the researchers have integrated the ranging techniques with the short-distance walk-paths. Although the pedestrian dead reckoning produces unorganizable error, the estimation of short-distance walk-paths is relative accurate. Sen et al. propose CUPID [107] to demonstrate the feasibility of integrating the ranging and short-distance walk-path together to identify the direct path (LoS path). The basic model of CUPID is shown in Fig. 16(a). As the user walks from location A to B, CUPID tracks the change in angle of the direct paths ($\Delta\theta$). The AP uses the energy of the direct path to estimate the distance of the mobile client, d_A and d_B at A and B respectively. It estimates the user's displacement between location A, and B(p_{AB}) by dead reckoning. Since d_A, d_B and p_{AB} are known, the actual angle change of the direct path can be estimated with AoA estimation.

Following the above work, Mariakakis and Sen further propose SAIL [108], a single AP based indoor localization system based on the triangle depicted in Fig.16(b). Similar to CUPID, the direct distance of d_A, d_B is estimated with CSI-based ranging, while d_{AB} is estimated by dead reckoning. Therefore, the triangle with d_A, d_B and d_{AB} is rotated around the AP in any direction and will still satisfy the side length constraints. The compass heading during the walk-path from A to B, help to determine the orientation of the triangle in the 2-D plane, and finally determine the user's location. Inspired by the above efforts, SpyLoc [109] leverages the benefits of both the dead reckoning and the ranging schemes to build a practical localization system. Given the accumulative errors of the inertial sensors to track user's movement over the time, SpyLoc uses the RF-Beep ranging scheme to calibrate this error in order to improve the localization accuracy. Unlike ranging-based or RF-based localization schemes that require multiple reference points (e.g., access points), using the dead reckoning in SpyLoc reduces the number of needed reference points to at least one reference point to locate and track user's movement accurately. This low dependency on ranging scheme and the elimination of any calibration make SpyLoc a light-weight system and practically suitable for mobile users.

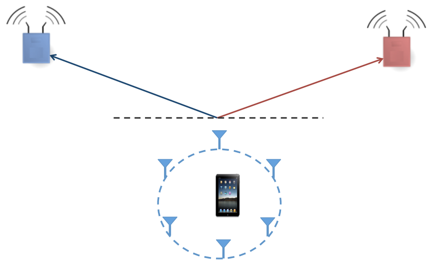


Fig. 17 Ubicarse [11] : Integrating multiple ranging with SAR.

4.5 Integrating multiple ranging

The ranging-based localization approaches usually requires trilateration or multilateration to estimate the position of the object. The insufficient ranging relations in the indoor environments limit the practice of ranging-based localization approaches. More relations usually require deploying more hardware for localization. Kumar et al. propose a novel approach in Ubicarse [11], which integrate multiple ranging at a single spot to achieve localizability. It indeed emulates a large antenna array with Synthetic Aperture Radar (SAR). For a device equipped with two antennas, when one twist it about its vertical axis, the relative trajectory of the two antennas performs a perfect circle, as shown in Fig. 17. Even if the twisting involves unknown trajectories, the distance between these antennas is fixed independent of how the user moves the device. Thus, whenever the user translates the device, the relative position vector of both its antennas remains the same. In contrast, as the device rotates, the relative position vector of the antennas, also rotates. Ubicarse achieves a median error of 39 centimeter in 3-D localization, which is regarded as the state-of-the-art for indoor localization without any specialized infrastructure or fingerprinting.

5 Crowdsourcing Layer

The concept of mobile crowdsourcing, or crowdsensing, is more and more popular [110–112]. Crowdsourcing is a technology with the potential to revolutionize large-scale data gathering in an extremely cost-effective manner [113]. It represents the act of a company or institution taking a function once performed by employees and outsourcing it to an undefined network of people in the form of an open call. This can take the form of peer-production, but is also often undertaken by sole individuals. The crucial prerequisite is the use of the open call format and the wide network of potential laborers. Crowdsourcing has developed to be an online, distributed problem-solving and production model gradually. With time elapsing, the power of crowdsourcing has been confirmed by many successful and profitable cases in a variety of industries such as Threadless, iStockphoto, InnoCentive and so on [114]. For example, The Carwow[1] platform attracts a lot of customs and businesses to build a platform for buying new cars from franchise dealers. Users can choose the car they would like to buy, along with the various specifications and features, then receive offers directly from dealers in the Carwow platform. Meanwhile, this platform also provides information to prospective purchasers and gives feedback on rogue dealers.

Nowadays, the development of wireless and embedded technology has developed smartphones and wearable devices with powerful computing, communicating, and sensing capability. Those mobile devices have acted as an increasingly important information interface between humans and environments. Therefore, mobile crowdsensing has become a promising technology. Many researchers used mobile crwodsensing technology to propose indoor localization systems in order to reduce the time-consuming, labor-intensive, and costs, such as LiFS [115], Walkie-Markie [102]. In this paper, we survey indoor localization systems based on mobile crowdsensing from three categories, including fingerprinting-based indoor localization methods, indoor floorplans construction, and self-deployable indoor navigation systems.

5.1 Construction of fingerprinting database without active intervention

WiFi-based fingerprinting indoor localization method has been one of the most attractive and promising techniques for ubiquitous applications. Typically, the fingerprinting based localization approaches can be divided into two major phases, i.e., off-line phase and on-line phase. During the off-line phase, also known as site survey, a fingerprint database is constructed with site survey. in which mobile devices record the RSS fingerprints (e.g., WiFi signal strengths from multiple Access Points, APs) at every location of an indoor space. In the on-line phase, when a user sends a location query with his current RSS fingerprint, localization algorithms retrieve the fingerprint database and return the matched fingerprints as well as the corresponding locations [115]. It can be seen from the offline phase that the site survey process is time-consuming, labor-intensive, and vulnerable to environmental dynamics. To solve this problem, Zheng et al. [115] proposed LiFS, which first exploited user motions from mobile phones to remove the site survey process of traditional WiFi fingerprinting method, while at the same time, achieve competitive localization accuracy via mobile crowdsensing.

[1] https://www.carwow.co.uk/

Specifically, LiFS system mapped the WiFi fingerprints into the special locations in the indoor floorplans by collecting and processing moving paths with recording RSS fingerprints of users using the Multidimensional scaling (MDS) technology. Furthermore, Rai et al. [116] used the inertial sensors (e.g., accelerometer, compass, gyroscope) presented in the mobile devices to track users as they traverse an indoor environment, while simultaneously performing WiFi scans to construct the fingerprint database via mobile crowdsensing. Chang et al. [117] and Wilk et al. [118] both focused on the radio map update problem and proposed systems that can maintain and update the radio map by utilization of the crowdsourced WiFi fingerprints. Wu et al. [104] proposed AcMu, an automatic and continuous radio map self-updating service for WiFi fingerprinting based methods that exploits the static behaviors of mobile devices.

5.2 Automatic construction of indoor floor plans

The digital indoor floorplans are the key components of indoor localization systems, such as fingerprintging-based, dead reckoning-based, and VLC-based methods. However, the indoor equivalent floorplans are currently very limited due in part to the following reasons: (1) buildings owners may not allow sharing of their floorplans in public for privacy reason and (3) manual creation of indoor floorplans requires slow, labor-intensive tasks [119]. Therefore, many researchers proposed systems to construct digital indoor floorplans with low-cost via mobile crowdsensing, including CrowdInside [120], MapGENIE [121], Walkie-Markie [102], Jigsaw [106], spyloc [109], CrowdMap [122], and the work presented in [123].

Jiang et. al. [123] constructed indoor floorplans by detecting similarities in WiFi signatures between different rooms and hallway segments to find their adjacency and combining inertial data to obtain hallway lengths and orientations. Shen et al. [102] proposed Walkie-Markie system, which leveraged the trend of WiFi signal strength reverses direction as anchor points and combined inertial data to construct the indoor floorplans. Alzantot et al. [120] and Philipp et al. [121] proposed CrowdInside and MapGenie systems, respectively, which only used walking traces of users to construct the indoor floorplans. In Recent works, Gao et al. [109] and Chen et al. [122] proposed JigSaw and CrowMap systems, respectively, which combined the walking traces of users and responding the vision information (e.g., images and videos) captured by mobile devices along walking traces to construct accurate indoor floorplans.

5.3 Construction of self-deployable indoor navigation systems

Self-deployable indoor navigation systems can be easily bootstrapped and deployed for motivated users via mobile crwodsensing. Compared conventional indoor navigation methods, self-deployable indoor navigation systems donot rely on comprehensive indoor localization systems or even floor maps. Recent work demonstrate that such systems are feasible in practice applications, including Travi-Navi [105], FollowMe [55], and iMoon [124].

Zheng et al. [105] proposed Travi-Navi system, which is a vision-guided navigation system, enables a user to easily bootstrap and deploy indoor navigation services without building the entire localization system. Specifically, Trivi-Navi used the images of pathway to track the user's walking progress and the particle filtering method as the navigation engine. Besides, it provided multiple navigation traces to multiple destinations by automatically finding shortcuts, weaving the traces to form a holistic trace, and guiding users to destinations with minimum detours. However, Travi-Navi still requires particular infrastructures (e.g., WiFi access point) for practical use. Shu et al. [55] proposed FollowMe system, used the geomagnetic field to guide a user by providing the "scent" left by the leaders or previous travelers without infrastructure. However, users can be guided from only a limited number of predefined locations (e.g., the building entrances) in those systems. Therefore, Dong et al. [124] proposed iMoon system, which built 3D models of indoor environment from crowdsourced 2D photos using Structure from Motion (SfM) algorithm, and compiles a navigation mesh from the generated 3D models. Zhou et al. propose MagSpider [125], a unique localization-free indoor navigational map construction approach. MagSpider integrated the isolated trajectories with temporal-spatial characteristics to construct the logical indoor map. It only utilizes the unlabeled and noisy measurements by inertial sensors on the smartphones to construct the navigation space; and hence can avoid the labor-intensive indoor floor plan calibration.

5.4 Automatic extraction of semantic information

Indoor environments contains plentiful semantic information, includes position of interest, names of shops or rooms, entities, etc. On the other hand, the smartphone becomes one of the major interaction media between the user and the indoor environments. Mining the interaction to derive the semantic information in the indoor environments is the new open research area for indoor localization.

Recently, several systems have been initially proposed to manually label or learn semantics for objects in an indoor floor plan [126–129]. Gunther et al. [127] generated a semantic map of an indoor environment using a set of 3D point clouds captured by a Kinect amounted on a mobile robot. The proposed system reconstructs the surfaces, detects different types of furnitures and estimates their poses. Zhao et al. [126] jointly inferred the semantic object category and structural class for each point of the global map. SemSense [128] requires each user to actively assign a semantic name to a physical location during the normal check-in operations. TransitLabel [129] was developed to recognize users' activities in transit stations and to infer the functionalities around the physical areas of users.

5.5 Summary

In summary, the crowdsourcing-based indoor localization systems are designed for different purposes. The insight behind those innovations is how to utilize the usersâĂŹ typical interaction with the mobile device to achieve non-conscious participation. Therefore, the crowdsourcing-based approaches usually require proper and innovative integration techniques to integrate more than one type of signals to exhibit the full potential in certain scenarios. Expect for those integration techniques, the underlying techniques for both crowdsourcing-based approaches and non-crowdsourcing-based approaches are exactly the same. Compared with those non-crowdsourcing-based approaches, the crowdsourcing-based approaches greatly reduce the labor cost, but require a longer time for deployment. Besides, since the collected data through crowdsourcing lack of quality guarantee and might be sensitive to the temporal-spatial dynamics and device diversity, the overall localization accuracy with the crowdsourcing-based approach is relative low.

6 Challenges for crowdsourcing-based indoor localization systems

Although mangy works claim that they are suitable to be implemented in a crowdsourcing manner. We find that at least two aspects of challenges are still not taken into consideration for the current studies.

6.1 Incentive mechanism in crowdsourcing-based indoor localization System

Although the concept of mobile crowdsourcing has already shown its promising future in many real applications, the participation of the publics is crucial to any crowdsourcing-based system. The costs of the crowdsensing task, such as the battery, data rates, computation resources, are all paid on the users' side. Besides, the concerns on the privacy might further limit the purpose for participation. Those issues directly lead to the research on incentive mechanisms [130–135]. Yang et al. [130] and Zhang et al. [131] focus on different models utilized in the design of incentive mechanism, while Zhao et al. [132] propose a unique framework for the online incentive mechanizsm. Besides, Luo et al. [134] introduce Tulllock contest to the design of incentive mechanism. Due to the page limitation, this paper can not cover all the recent advances in the incentive mechanisms for mobile crowdsourcing systems, the interested readers are suggested to refer [136] for more thorough reviews.

Most of the current innovations on incentive mechanisms focus on the general crowdsourcing systems, but neglect the special characteristics in the context of localization with various different participations. The localization-related incentive mechanism is still a open research topic. We believe that the efforts on such a topic will greatly improve the feasibility and performance of the crowdsourcing-based indoor localization system.

6.2 Quality control of crowdsourced data

As a new problem-solving paradigm, crowdsourcing changes the fundamental way of data gathering, and ease the burden of calibration the reference for localization purpose. Indoor localization usually has strict requirements for the accuracy. However, the gathered data through crowdsourcing lacks of quality guarantee. On one hand, the raw measurements of the underlying signals are challenged by the noise, temporal-spatial dynamics, devices diversity and other human factors. On the other hand, the data contributed by the mal-participants might be harmful to entire system, which might event caused unrecoverable damages to the system. Finding the mal-participants among tens of thousands of ordinary and anonymous users is very challenging. Besides, proper design of feedback mechanism for the crowdsourcing-based indoor localization system is essential for the quality control of the but neglected problems in the current literature.

To the best of our knowledge, the dedicated study on those two issues for the crowdsourcing-based indoor localization system is still an open area in the mainstream publications. We'd like to encourage the readers of interest focus on those challenging issues to further push the limit of indoor localization.

7 Conclusion

Despite the innovations in the wireless indoor localization recently, a killer application like GPS for the outdoor localization is still far from reality. The root reason is that the indoor environments are complex, which are with rich spatial-temporal variance and dynamics. The limitation of the pervasive devices, as well as the human factors, makes that the current work with a certain kind of signals are not robust. In this survey, we review the current innovations which take use multiple signals as complementary to each other, to facilitate the construction of crowdsourcing-based indoor localization system. We notice that, the physical measurements of different types of signals vary a lot in characteristics, such as propagation, spatial variance, noise distribution, temporal dynamics. However, those signals might be either related in physical or independent in noise. Integrating those chaotic signals achieves the objectives that are impossible for any single kind of signals.

In this article, we systemically studied the signals involved in the mainstream indoor localization approaches, in terms of the characteristics, spatial-temporal variance and dynamics. Considering the limitation of each single kind of signal, we emphasize the integrating approach to fully achieve the potential of the indoor available signals. Despite pioneer efforts in integrating multiple signals, the realm is still in its infancy and continues to develop from diverse perspectives:

First, integrating multiple signals to improve the worst-case accuracy in the indoor localization. Recently, the CSI ranging-based approach exhibits its promising ability for improving the localization accuracy. However, the long-tail effect still limits the large-scale application in practice. One of the dominant reasons is that the insufficient LoS path coverage. For example, the corner of a room can hardly reach the LoS path from the APs. Wu et al. [137] propose an novel approach for real-time LoS path indemnification for WiFi signal. Can we utilize the other signals as a complementary when LoS path is unavailable? We have already shown the powerful ability of integrating walk-paths and ranging. We believe that such a direction is worth attempting with various independent signals, to improve the overall localization accuracy.

Second, integrating the semantic informations with the physical measurements of signals to enable the context-aware location-based service. Most of the current works in indoor localization usually follow the ideas in the outdoor space, and strive to derive an accurate coordination of the people or entity. Actually, the indoor environments contains plentiful semantic information, which is probably more valuable than merely a blank and cold coordination. On the other hand, the smartphones have become one of the major interaction tools between users and the surroundings. Mining those interactions can help to automatically obtain the semantic information of interest, and further facilitate the intelligent indoor location based service. We believe it will become one of the heated topic in the future.

Third, integrating the physical measurement on different devices to enable cooperative localization. Thanks to the rapid development of pervasive computing, the range of the pervasive devices are continuously extended. the smart watches, smart bracelets, smart glasses, VR devices, etc. have been emerging in our life. Most of the current works still takes the smartphones as the platform to localization. The problem of cooperative location with multiple pervasive devices has been seldom noticed. On the other hand, smart glasses provide with plentiful vision signals in the first-person view, while the smart watches and bracelets can help to more preciously shape the motion of users'. We believe integrating the physical measurements on different devices can help to not only improve the accuracy of localization, but also realize fine-grained interaction discovery. The later one would contribute to obtain the semantic information, and leaves largely open and attractive research opportunities.

Acknowledgment

The authors would like to thank the anonymous reviewers for their valuable comments.

This work is partly supported by the National Natural Science Foundation for Outstanding Excellent young scholars of China under Grant No. 61422214, National Basic Research Program (973 program) under Grant No. 2014CB347800, the Program for New Century Excellent Talents in University, the Hunan Provincial Natural Science Fund for Distinguished Young Scholars under Grant No. 2016JJ1002, the Research Funding of NUDT under Grant Nos. JQ14-05-02 and ZDYYJCYJ20140601.

References

1. Want R, Hopper A, Falcão V, Gibbons J. The active badge location system. ACM Transactions on Information Systems (TOIS), 1992, 10(1): 91–102
2. Harter A, Hopper A, Steggles P, Ward A, Webster P. The anatomy of a context-aware application. Wireless Networks, 2002, 8(2/3): 187–197
3. Ward A, Jones A, Hopper A P C I. A new location technique for the active office. Personal Communications, IEEE, 1997, 4(5)
4. Priyantha N B, Chakraborty A, Balakrishnan H. The Cricket location-support system. In: Proceedings of ACM Annual International Conference on Mobile Computing and Networking. 2000, 32–43
5. Williams A, Ganesan D, Hanson A. Aging in place: fall detection and localization in a distributed smart camera network. In: Proceedings of ACM International Conference on Multimedia. 2007, 892–901
6. Krishnan P, Krishnakumar A S, Ju W H, Mallows C, Gamt S N. A system for LEASE: location estimation assisted by stationary emitters for indoor RF wireless networks. In: Proceedings of IEEE International Conference on Computer Communications. 2004, 1001–1011
7. Ni L M, Liu Y H, Lau Y C, Patil A P. LANDMARC: indoor location sensing using active RFID. Wireless Networks, 2004, 10(6)
8. Holland J H. Emergence: from chaos to order. Addison-Wesley Longman Publishing Co., Inc., 1998
9. Jin Y Y, Soh W, Wong W. Indoor localization with channel impulse response based fingerprint and nonparametric regression. IEEE Transactions on Wireless Communications, 2010, 9(3): 1120–1127
10. Wu K S, Li H C, Wang L, Yi Y W, Liu Y H, Chen D H, Luo X N, Zhang Q, Ni L M. hjam: Attachment transmission in wlans. IEEE Transactions on Mobile Computing, 2013, 12(12): 2334–2345
11. Kumar S, Gil S, Katabi D, Rus D. Accurate indoor localization with zero start-up cost. In: Proceedings of ACM Annual International Conference on Mobile Computing and Networking. 2014, 483–494
12. Xiao J, Yi Y W, Wang L, Li H C, Zhou Z M, Wu K S, Ni L M. Nomloc: Calibration-free indoor localization with nomadic access points. In: Proceedings ofIEEE International Conference on Distributed Systems. 2014, 587–596
13. Yang Z, Zhou Z M, Liu Y H. From RSSI to CSI: indoor localization via channel response. ACM Computing Survey, 2013, 46(2): 25:1–25:32
14. Yang Z, Wu C S, Zhou Z M, Zhang X L, Wang X, Liu Y H. Mobility increases localizability: a survey on wireless indoor localization using inertial sensors. ACM Computing Surveys, 2015, 47(3): 54:1–54:34
15. Pei L, Zhang M, Zou D P, Chen R Z, Chen Y W. A survey of crowd sensing opportunistic signals for indoor localization. Mobile Information Systems, 2016, 2016: 4041291:1–4041291:16
16. Adler S, Schmitt S, Wolter K, Kyas M. A survey of experimental evaluation in indoor localization research. In: 2015 International Conference on Indoor Positioning and Indoor Navigation. 2015, 1–10
17. Liu Y H, Yang Z, Wang X P, Jian L R. Location, localization, and localizability. Journal of Computer Science and Technology, 2010, 25(2): 274–297
18. Chen T, Yang Z, Liu Y H, Guo D, Luo X S. Localization-oriented network adjustment in wireless ad hoc and sensor networks. IEEE Transactions on Parallel and Distributed Systems, 2014, 25(1): 146–155
19. Chen T, Guo D, Yang Z, Chen H H, Luo X S. Improving the efficiency of localization-oriented network adjustment in wireless sensor networks. IEEE Communications Letters, 2011, 15(9): 983–985
20. Wu K S, Tan H Y, Liu Y H, Zhang J, Zhang Q, Ni L M. Side channel: Bits over interference. IEEE Transactions Mobile on Computing, 2012, 11(8): 1317–1330
21. Francisco J A, Martin R P. A method of characterizing radio signal space for wireless device localization. Tsinghua Science and Technology, 2015, 20(4): 385–408
22. Seidel R T SS. Y. 914 mhz path loss prediction models for indoor wireless communications in multifloored buildings. IEEE Transactions on Antennas and Propagation, 1992, 40(2): 207–217
23. Zhou Z M, Yang Z, Wu C S, Shangguan L F, Cai H B, Liu Y H, Ni L M. Wifi-based indoor line-of-sight identification. IEEE Transactions on Wireless Communications, 2015, 14(11): 6125–6136
24. Gwon Y, Jain R. Error characteristics and calibration-free techniques for wireless lan-based location estimation. In: Proceedings of ACM International Symposium on Mobility Management and Wireless Access. 2004, 2–9
25. Lim H, Kung L C, Hou J C, Luo H. Zero-configuration indoor localization over ieee 802.11 wireless infrastructure. Wireless Networks, 2010, 16(2): 405–420
26. Ji Y, Biaz S, Pandey S, Agrawal P. Ariadne: a dynamic indoor signal map construction and localization system. In: Proceedings of ACM International Conference on Mobile Systems, Applications, and Services. 2006, 151–164
27. Chintalapudi K, Padmanabha Iyer A, Padmanabhan V N. Indoor localization without the pain. In: Proceedings of ACM Annual International Conference on Mobile Computing and Networking. 2010, 173–184
28. Goswami A, Ortiz L E, Das S R. Wigem: A learning-based approach for indoor localization. In: Proceedings of ACM Conference on Emerging Networking Experiments and Technologies. 2011, 3
29. Bahl P, Padmanabhan V N. RADAR: an in-building RF-based user location and tracking system. In: Proceedings of IEEE International Conference on Computer Communications. 2010, 775–784
30. Youssef M, Agrawala A. The horus wlan location determination system. In: Proceedings of ACM International Conference on Mobile Systems, Applications, and Services. 2005, 205–218
31. Zhang D, Liu Y H, Guo X N, Gao M, Ni L M. On distinguishing the multiple radio paths in rss-based ranging. In: Proceedings of IEEE International Conference on Computer Communications. 2012, 2201–2209
32. Wu K S, Xiao J, Yi Y W, Gao M, Ni L M. FILA: fine-grained indoor localization. In: Proceedings of IEEE International Conference on

Computer Communications. 2012, 2210–2218

33. Nerguizian C, Despins C L, Affes S. Geolocation in mines with an impulse response fingerprinting technique and neural networks. IEEE Transactions on Wireless Communications, 2006, 5(3): 603–611
34. Sen S, Radunovic B, Choudhury R R, Minka T. You are facing the mona lisa: spot localization using phy layer information. In: Proceedings of 10th ACM International Conference on Mobile Systems, Applications, and Services. 2012, 183–196
35. Xiao J, Wu K S, Yi Y W, Ni L M. FIFS: fine-grained indoor fingerprinting system. In: Proceedings of International Conference on Computer Communications and Networks. 2012, 1–7
36. Kotanen A, Hännikäinen M, Leppäkoski H, Hämäläinen T. Experiments on local positioning with bluetooth. In: Proceedings of International Symposium on Information Technology. 2003, 297–303
37. Bargh M S, Groote d R. Indoor localization based on response rate of bluetooth inquiries. In: Proceedings of the ACM International Workshop on Mobile Entity Localization and Tracking in GPS-less Environments. 2008, 49–54
38. Lin X, Ho T, Fang C, Yen Z, Yang B, Lai F. A mobile indoor positioning system based on ibeacon technology. In: Proceedings of International Conference of the IEEE Engineering in Medicine and Biology Society. 2015, 4970–4973
39. Matic A, Papliatseyeu A, Osmani V, Mayora-Ibarra O. Tuning to your position: Fm radio based indoor localization with spontaneous recalibration. In: Proceedings of IEEE International Conference on Pervasive Computing. 2010, 153–161
40. Moghtadaiee V, Dempster A G, Lim S. Indoor localization using fm radio signals: A fingerprinting approach. In: Proceedings of IEEE International Conference on Indoor Positioning and Indoor Navigation. 2011, 1–7
41. Chen Y, Lymberopoulos D, Liu J, Priyantha B. Fm-based indoor localization. In: Proceedings of 10th ACM International Conference on Mobile Systems, Applications, and Services. 2012, 169–182
42. Yoon S, Lee K, Rhee I. FM-based indoor localization via automatic fingerprint DB construction and matching. In: Proceedings of ACM International Conference on Mobile Systems, Applications, and Services. 2013, 207–220
43. Scott J, Dragovic B. Audio location: Accurate low-cost location sensing. In: International Conference on Pervasive Computing. 2005, 1–18
44. Tarzia S P, Dick R P, Dinda P A, Memik G. Sonar-based measurement of user presence and attention. In: Proceedings of ACM International Joint Conference on Pervasive and Ubiquitous Computing. 2009, 89–92
45. Peng C Y, Shen G B, Zhang Y G, Li Y L, Tan K. Beepbeep: a high accuracy acoustic ranging system using cots mobile devices. In: Proceedings of the 5th international conference on Embedded networked sensor systems. 2007, 1–14
46. Liu K K, Liu X X, Li X L. Guoguo: enabling fine-grained indoor localization via smartphone. In: Proceedings of ACM International Conference on Mobile Systems, Applications, and Services. 2013, 235–248
47. Azizyan M, Constandache I, Roy Choudhury R. Surroundsense: mobile phone localization via ambience fingerprinting. In: Proceedings of ACM Annual International Conference on Mobile Computing and Networking. 2009, 261–272
48. Tarzia S P, Dinda P A, Dick R P, Memik G. Indoor localization without infrastructure using the acoustic background spectrum. In: Proceedings of 9th ACM International Conference on Mobile Systems, Applications, and Services. 2011, 155–168
49. Lin Q, Guo Y. Accurate indoor navigation system using human-item spatial relation. Tsinghua Science and Technology, 2016, 21(5): 521–537
50. Angermann M, Frassl M, Doniec M, Julian B J, Robertson P. Characterization of the indoor magnetic field for applications in localization and mapping. In: Proceedings of IEEE International Conference on Indoor Positioning and Indoor Navigation. 2012, 1–9
51. Haverinen J, Kemppainen A. A global self-localization technique utilizing local anomalies of the ambient magnetic field. In: Proceedings of IEEE International Conference on Robotics and Automation. 2009, 3142–3147
52. Glanzer G, Walder U. Self-contained indoor pedestrian navigation by means of human motion analysis and magnetic field mapping. In: Proceedings of IEEE Wireless Communications and Networking Conference. 2010, 303–307
53. Gozick B, Subbu K P, Dantu R, Maeshiro T. Magnetic maps for indoor navigation. IEEE Transactions on Instrumentation and Measurement, 2011, 60(12): 3883–3891
54. Le Grand E, Thrun S. 3-axis magnetic field mapping and fusion for indoor localization. In: Proceedings of IEEE Conference on Multisensor Fusion and Integration System. 2012, 358–364
55. Shu Y C, Shin K G, He T, Chen J M. Last-mile navigation using smartphones. In: Proceedings of ACM Annual International Conference on Mobile Computing and Networking. 2015, 512–524
56. Zhou P F, Zheng Y Q, Li Z J, Li M, Shen G B. IODetector: a generic service for indoor outdoor detection. In: Proceedings of ACM Conference on Embedded Networked Sensor Systems. 2012
57. Tsonev D, Videv S, Haas H. In: Dingel B, Tsukamoto K, eds. Light fidelity (Li-Fi): towards all-optical networking. SPIE, 2013
58. Burchardt H, Serafimovski N, Tsonev D, Videv S, Haas H. VLC: beyond point-to-point communication. IEEE Communications Magazine, 2014, 52(7): 98–105
59. Li L, Hu P, Peng C Y, Shen G B, Zhao F. Epsilon: A visible light based positioning system. In: Proceedings of USENIX Symposium on Networked Systems Design and Implementation. 2014, 331–343
60. Li T X, An C, Tian Z, Campbell A T, Zhou X. Human sensing using visible light communication. In: Proceedings of ACM Annual International Conference on Mobile Computing and Networking. 2015, 331–344
61. Kuo Y, Pannuto P, Hsiao K, Dutta P. Luxapose: indoor positioning with mobile phones and visible light. In: Proceedings of ACM Annual International Conference on Mobile Computing and Networking. 2014, 447–458
62. Rajagopal N, Lazik P, Rowe A. Visual light landmarks for mobile

devices. In: Proceedings of International Symposium on Information Processing in Sensor Networks. 2014, 249–260
63. Tian Z, Wright K, Zhou X. The darklight rises: visible light communication in the dark. In: Proceedings of Annual International Conference on Mobile Computing and Networking. 2016, 2–15
64. Yang Z C, Wang Z Y, Zhang J S, Huang C Y, Zhang Q. Wearables can afford: Light-weight indoor positioning with visible light. In: Proceedings of 13th ACM International Conference on Mobile Systems, Applications, and Services. 2015, 317–330
65. Zhang C, Tabor J, Zhang J L, Zhang X Y. Extending mobile interaction through near-field visible light sensing. In: Proceedings of ACM Annual International Conference on Mobile Computing and Networking. 2015, 345–357
66. Lee H, Lin H, Wei Y, Wu H, Tsai H, Lin K C. Rollinglight: Enabling line-of-sight light-to-camera communications. In: Proceedings of 13th ACM International Conference on Mobile Systems, Applications, and Services. 2015, 167–180
67. Delponte E, Isgro F, Odone F, Verri A. SVD-matching using SIFT features. Graphical models, Elsevier, 2006, 68(5): 415–431
68. Wang H, Zhao D, Ma H D, Xu H Y, Hou X B. Crowdsourcing based mobile location recognition with richer fingerprints from smartphone sensors. In: Proceedings of IEEE International Conference on Parallel and Distributed Systems. 156–163
69. Gao R P, Tian Y, Ye F, Luo G J, Bian K G, Wang Y Z, Wang T, Li X M. Sextant: Towards ubiquitous indoor localization service by photo-taking of the environment. IEEE Transactions on Mobile Computing, 2016, 15(2): 460–474
70. Xu H, Yang Z, Zhou Z M, Shangguan L F, Yi K, Liu Y H. Enhancing wifi-based localization with visual clues. In: Proceedings of ACM International Joint Conference on Pervasive and Ubiquitous Computing. 2015, 963–974
71. Wang S L, Fidler S, Urtasun R. Lost shopping! monocular localization in large indoor spaces. In: Proceedings of IEEE International Conference on Computer Vision. 2015, 2695–2703
72. Chon Y, Lane N D, Li F, Cha H, Zhao F. Automatically characterizing places with opportunistic crowdsensing using smartphones. In: Proceedings of ACM International Joint Conference on Pervasive and Ubiquitous Computing. 2012, 481–490
73. Chao J S, Steinbach E G. Preserving SIFT features in jpeg-encoded images. In: Proceedings of 18th IEEE International Conference on Image Processing. 2011, 301–304
74. Chandrasekhar V, Takacs G, Chen D M, Tsai S S, Grzeszczuk R, Girod B. Chog: Compressed histogram of gradients A low bit-rate feature descriptor. In: Proceedings of IEEE Conference on Computer Vision and Pattern Recognition. 2009, 2504–2511
75. Chen D M, Tsai S S, Chandrasekhar V, Takacs G, Singh J P, Girod B. Tree histogram coding for mobile image matching. In: Proceedings of IEEE Data Compression Conference. 2009, 143–152
76. Opdenbosch D V, Schroth G, Huitl R, Hilsenbeck S, Garcea A, Steinbach E G. Camera-based indoor positioning using scalable streaming of compressed binary image signatures. In: Proceedings of IEEE International Conference on Image Processing. 2014, 2804–2808
77. Wu C. Towards linear-time incremental structure from motion. In: Proceedings of IEEE International Conference on 3D Vision. 2013, 127–134
78. Guo Y, Bennamoun M, Sohel F, Lu M, Wang J. 3D object recognition in cluttered scenes with local surface features: A survey. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2014, 36(11): 2270–2287
79. Chen K, Lai Y, Wu Y X, Martin R R, Hu S M. Automatic semantic modeling of indoor scenes from low-quality rgb-d data using contextual information. ACM Transactions on Graphics, 2014, 33(6)
80. Henry P, Krainin M, Herbst E, Ren X, Fox D. Rgb-d mapping: Using kinect-style depth cameras for dense 3d modeling of indoor environments. The International Journal of Robotics Research, 2012, 31(5): 647–663
81. Endres F, Hess J, Engelhard N, Sturm J, Cremers D, Burgard W. An evaluation of the rgb-d slam system. In: Proceedings of IEEE International Conference on Robotics and Automation. 2012, 1691–1696
82. Kümmerle R, Grisetti G, Strasdat H, Konolige K, Burgard W. g 2 o: A general framework for graph optimization. In: Proceedings of IEEE International Conference on Robotics and Automation. 2011, 3607–3613
83. Engel J, Schöps T, Cremers D. Lsd-slam: Large-scale direct monocular slam. In: Proceedings of European Conference on Computer Vision. 2014, 834–849
84. Engel J, Sturm J, Cremers D. Semi-dense visual odometry for a monocular camera. In: Proceedings of IEEE International Conference on Computer Vision. 2013, 1449–1456
85. Teng X Q, Guo D K, Guo Y L, Zhou X L, Ding Z L, Liu Z. Ionavi: An indoor-outdoor navigation service via mobile crowdsensing. ACM Transactions on Sensor Networks, 2017, 13(2): 12:1–12:28
86. Thrun S, Burgard W, Fox D. A real-time algorithm for mobile robot mapping with applications to multi-robot and 3d mapping. In: Proceedings of IEEE International Conference on Robotics and Automation. 2000, 321–328
87. Triebel R, Burgard W. Improving simultaneous mapping and localization in 3d using global constraints. In: Proceedings of The Twentieth National Conference on Artificial Intelligence and the Seventeenth Innovative Applications of Artificial Intelligence Conference. 2005, 1330
88. May S, Droeschel D, Holz D, Fuchs S, Malis E, Nüchter A, Hertzberg J. Three-dimensional mapping with time-of-flight cameras. Journal of Field Robotics, 2009, 26(11-12): 934–965
89. Newcombe R A, Izadi S, Hilliges O, Molyneaux D, Kim D, Davison A J, Kohi P, Shotton J, Hodges S, Fitzgibbon A. Kinectfusion: Real-time dense surface mapping and tracking. In: Proceedings of IEEE International Symposium on Mixed and Augmented Reality. 2011, 127–136
90. Curless B, Levoy M. A volumetric method for building complex models from range images. In: Proceedings of IEEE International Symposium on Mixed and Augmented Reality. 1996, 303–312

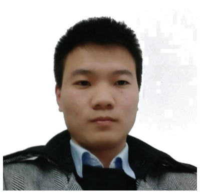
91. Keller M, Lefloch D, Lambers M, Izadi S, Weyrich T, Kolb A. Real-time 3d reconstruction in dynamic scenes using point-based fusion. In: Proceedings of IEEE International Conference on 3D Vision. 2013, 1–8
92. Al-Nuaimi A, Piccolrovazzi M, Gedikli S, Steinbach E, Schroth G. Indoor location retrieval using shape matching of kinectfusion scans to large-scale indoor point clouds. In: Proceedings of 2015 Eurographics Workshop on 3D Object Retrieval. 2015, 31–38
93. Winterhalter W, Fleckenstein F, Steder B, Spinello L, Burgard W. Accurate indoor localization for rgb-d smartphones and tablets given 2d floor plans. In: Proceedings of IEEE International Conference on Intelligent Robots and Systems. 2015, 3138–3143
94. Zhou P F, Li M, Shen G B. Use it free. In: Proceedings of ACM International Conference on Mobile Systems, Applications, and Services. 2014, 605–616
95. Li F, Zhao C S, Ding G Z, Gong J, Liu C X, Zhao F. A reliable and accurate indoor localization method using phone inertial sensors. In: Proceedings of ACM Conference on Ubiquitous Computing. 2012, 421–430
96. Cho D K, Mun M, Lee U, Kaiser W J, Gerla M. Autogait: A mobile platform that accurately estimates the distance walked. In: Proceedings of IEEE International Conference on Pervasive Computing. 2010, 116–124
97. Foxlin E. Pedestrian tracking with shoe-mounted inertial sensors. IEEE Computer Graphics and Applications, 2005, 25(6): 38–46
98. Bertozzi M, Broggi A, Fascioli A, Tibaldi A, Chapuis R, Chausse F. Pedestrian localization and tracking system with kalman filtering. In: Proceedings of IEEE Intelligent Vehicles Symposium. 2004, 584–589
99. Roy N, Wang H, Choudhury R R. I am a smartphone and i can tell my user's walking direction. In: Proceedings of ACM International Conference on Mobile Systems, Applications, and Services. 2014, 329–342
100. Constandache I, Bao X, Azizyan M, Choudhury R R. Did you see bob?: human localization using mobile phones. In: Proceedings of ACM International Conference on Mobile Systems, Applications, and Services. 2010
101. Wu C, Yang Z, Liu Y. Smartphones based crowdsourcing for indoor localization. IEEE Transactions on Mobile Computing, 2015, 14(2): 444–457
102. Shen G B, Chen Z, Zhang P C, Moscibroda T, Zhang Y G. Walkie-markie: Indoor pathway mapping made easy. In: Proceedings of USENIX Symposium on Networked Systems Design and Implementation. 2013, 85–98
103. Wang H, Sen S, Elgohary A, Farid M, Youssef M, Choudhury R R. No need to war-drive: unsupervised indoor localization. In: Proceedings of ACM International Conference on Mobile Systems, Applications, and Services. 2012
104. Wu C S, Yang Z, Xiao C W, Yang C F, Liu Y H, Liu M Y. Static power of mobile devices: Self-updating radio maps for wireless indoor localization. In: Proceedings of IEEE International Conference on Computer Communications. 2015, 2497–2505
105. Zheng Y Q, Shen G B, Li L Q, Zhao C S, Li M, Zhao F. Travi-navi: Self-deployable indoor navigation system. In: Proceedings of ACM Annual International Conference on Mobile Computing and Networking. 2014, 471–482
106. Gao R, Zhao M, Ye T, Ye F, Wang Y, Bian K, Wang T, Li X. Jigsaw: Indoor floor plan reconstruction via mobile crowdsensing. In: Proceedings of International Conference on Mobile computing and networking. 2014, 249–260
107. Sen S, Lee J, Kim K, Congdon P. Avoiding multipath to revive inbuilding wifi localization. In: Proceedings of ACM International Conference on Mobile Systems, Applications, and Services. 2013, 249–262
108. Mariakakis A T, Sen S, Lee J, Kim K H. Sail: single access point-based indoor localization. In: Proceedings of ACM International Conference on Mobile Systems, Applications, and Services. 2014, 315–328
109. Uddin M, Nadeem T. Spyloc: a light weight localization system for smartphones. In: Proceedings of IEEE International Conference on Sensing, Communication and Networking, Singapore, June. 2014, 72–80
110. Ganti R K, Ye F, Lei H. Mobile crowdsensing: current state and future challenges. IEEE Communications Magazine, 2011, 49(11): 32–39
111. Ma H D, Zhao D, Yuan P Y. Opportunities in mobile crowd sensing. IEEE Communications Magazine, 2014, 52(8): 29–35
112. Guo B, Wang Z, Yu Z W, Wang Y, Yen N Y, Huang R H, Zhou X S. Mobile crowd sensing and computing: The review of an emerging human-powered sensing paradigm. ACM Computing Surv., 2015, 48(1): 7:1–7:31
113. Chen S, Li M Y, Ren K, Fu X W, Qiao C M. Rise of the indoor crowd: Reconstruction of building interior view via mobile crowdsourcing. In: Proceedings of ACM Conference on Embedded Networked Sensor Systems. 2015, 59–71
114. Guazzini A, Vilone D, Donati C, Nardi A, Levnajic Z. Modeling crowdsourcing as collective problem solving. Scientific Reports, 2015, 5: 16557–16557
115. Yang Z, Wu C, Liu Y. Locating in fingerprint space: wireless indoor localization with little human intervention. In: Proceedings of ACM Annual International Conference on Mobile Computing and Networking. 2012
116. Rai A, Chintalapudi K K, Padmanabhan V N, Sen R. Zee: zero-effort crowdsourcing for indoor localization. In: Proceedings of ACM Annual International Conference on Mobile Computing and Networking. 2012, 293–304
117. Chang K, Han D. Crowdsourcing-based radio map update automation for wi-fi positioning systems. In: Proceedings of ACM International Conference on Advances in Geographic Information System. 2014, 24–31
118. Wilk P, Karciarz J, Swiatek J. Indoor radio map maintenance by automatic annotation of crowdsourced wi-fi fingerprints. In: Proceedings of IEEE International Conference on Indoor Positioning and Indoor Navigation. 2015, 1–8
119. Elhamshary M, Youssef M, Uchiyama A, Yamaguchi H, Higashino T. Transitlabel: A crowd-sensing system for automatic labeling of transit

stations semantics. In: Proceedings of ACM International Conference on Mobile Systems, Applications, and Services. 2016, 193–206
120. Alzantot M, Youssef M. Crowdinside: automatic construction of indoor floorplans. In: Proceedings of ACM International Conference on Advances in Geographic Information Systems. 2012, 99–108
121. Philipp D, Baier P, Dibak C, Dürr F, Rothermel K, Becker S, Peter M, Fritsch D. Mapgenie: Grammar-enhanced indoor map construction from crowd-sourced data. In: Proceedings of IEEE International Conference on Pervasive Computing. 2014, 139–147
122. Chen S, Li M Y, Ren K, Qiao C M. Crowd map: Accurate reconstruction of indoor floor plans from crowdsourced sensor-rich videos. In: Proceedings ofIEEE International Conference on Distributed Systems. 2015, 1–10
123. Jiang Y F, Yun X, Pan X, Li K, Lv Q, Dick R P, Shang L, Hannigan M. Hallway based automatic indoor floorplan construction using room fingerprints. In: Proceedings of ACM International Joint Conference on Pervasive and Ubiquitous Computing. 2013, 315–324
124. Dong J, Xiao Y, Noreikis M, Ou Z H, Ylä-Jääski A. imoon: Using smartphones for image-based indoor navigation. In: Proceedings of ACM Conference on Embedded Networked Sensor Systems. 2015, 85–97
125. Zhou X L, Guo D K, Teng X Q. Magspider: a localization-free approach for constructing global indoor map for navigation purpose. In: Proceedings of the ACM Turing 50th Celebration Conference - China. 2017, 44:1–44:10
126. Zhao Z, Chen X. Building 3D semantic maps for mobile robots using RGB-D camera. Intelligent Service Robotics, 2016, 1–13
127. Gunther M, Wiemann T, Albrecht S, Hertzberg J. Building semantic object maps from sparse and noisy 3D data. In: Proceedings of IEEE/RSJ International Conference on Intelligent Robots and Systems. 2013, 2228–2233
128. Elhamshary M, Moustafa Y. Semsense: Automatic construction of semantic indoor floorplans. In: Proceedings of IEEE International Conference on Indoor Positioning and Indoor Navigation. 2015, 1–11
129. Elhamshary M, Youssef M, Uchiyama A, Yamaguchi H, Higashino T. Transitlabel: A crowd-sensing system for automatic labeling of transit stations semantics. In: Proceedings of ACM International Conference on Mobile Systems, Applications, and Services. 2016, 193–206
130. Yang D J, Xue G L, Fang X, Tang J. Crowdsourcing to smartphones: incentive mechanism design for mobile phone sensing. In: Proceedings of ACM Annual International Conference on Mobile Computing and Networking. 2012, 173–184
131. Zhang X, Xue G L, Yu R Z, Yang D J, Tang J. Truthful incentive mechanisms for crowdsourcing. In: Proceedings of IEEE International Conference on Computer Communications. 2015, 2830–2838
132. Zhao D, Li X Y, Ma H D. How to crowdsource tasks truthfully without sacrificing utility: Online incentive mechanisms with budget constraint. In: Proceedings of IEEE International Conference on Computer Communications. 2014, 1213–1221
133. Wen Y T, Shi J Y, Zhang Q, Tian X H, Huang Z Y, Yu H, Cheng Y, Shen X M. Quality-driven auction-based incentive mechanism for mobile crowd sensing. IEEE Transactions on Vehicular Technology, 2015, 64(9): 4203–4214
134. Luo T, Kanhere S S, Tan H, Wu F, Wu H Y. Crowdsourcing with tullock contests: A new perspective. In: Proceedings of IEEE International Conference on Computer Communications. 2015, 2515–2523
135. Dai W, Wang Y, Jin Q, Ma J. An integrated incentive framework for mobile crowdsourced sensing. Tsinghua Science and Technology, 2016, 21(2): 146–156
136. Zhang X L, Yang Z, Sun W, Liu Y H, Tang S H, Xing K, Mao X F. Incentives for mobile crowd sensing: A survey. IEEE Communications Surveys and Tutorials, 2016, 18(1): 54–67
137. Wu C S, Yang Z, Zhou Z M, Qian K, Liu Y H, Liu M Y. Phaseu: Real-time LOS identification with wifi. In: Proceedings of IEEE International Conference on Computer Communications. 2015, 2038–2046

Xiaolei Zhou received the B.A. degree from Nanjing University, China, in 2009, and received M.S. and Ph.D. degrees in military operational research from National University of Defense Technology, Changsha, China, in 2011 and 2016, receptively. He is an Assistant Professor with the Nanjing Telecommunication Technology Research Institute, National University of Defense Technology. His research interests include mobile and pervasive computing, mobile crowdsensing, especially wireless indoor positioning and location based services.

Tao Chen received the B.S. degree in military science, the M.S. and Ph.D. degrees in military operational research from the National University of Defense Technology, Changsha, China, in 2004, 2006, and 2011, respectively. He is an Assistant Professor with the College of Information System and Management, National University of Defense Technology, Changsha, P.R. China. His research interests include wireless sensor networks, peer-to-peer computing, and data center networking.

Deke Guo received the B.S. degree in industry engineering from Beijing University of Aeronautic and Astronautic, Beijing, China, in 2001, and the Ph.D. degree in management science and engineering from National University of Defense Technology, Changsha, China, in 2008. He is a Professor with the College of Information System and Management, National University of Defense Technology. His research interests include distributed systems, software-defined networking, data center networking. He is a member of the ACM and the IEEE. **He was rewarded with the Award of the NSFC Excellent Young Scholars Program in 2014.**



Xiaoqiang Teng received the B.S. degree from the School of Mechanical Engineering, Shenyang University of Technology, China, in 2013. Currently, he is a Ph.D. student in the College of Information System and Management, National University of Defense Technology, Changsha, China. His research interests include mobile computing, pervasive computing, and computer vision.

Bo Yuan received the B.S. degree from the School of Electronic Information and Electrical Engineering, Shanghai Jiao Tong University, China, in 2015. Currently, he is a master student in the College of Information System and Management, National University of Defense Technology, Changsha, China. His research interests include mobile computing, pervasive computing, and computer vision.