Edinburgh Research Explorer

Toward Location-Enabled IoT (LE-IoT): IoT Positioning Techniques, Error Sources, and Error Mitigation

Citation for published version:

Digital Object Identifier (DOI):
10.1109/JIOT.2020.3019199

Link:
Link to publication record in Edinburgh Research Explorer

Document Version:
Peer reviewed version

Published In:
IEEE Internet of Things Journal

General rights
Copyright for the publications made accessible via the Edinburgh Research Explorer is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy
The University of Edinburgh has made every reasonable effort to ensure that Edinburgh Research Explorer content complies with UK legislation. If you believe that the public display of this file breaches copyright please contact openaccess@ed.ac.uk providing details, and we will remove access to the work immediately and investigate your claim.
Abstract—Localization techniques are becoming key to add location context to Internet of Things (IoT) data without human perception and intervention. Meanwhile, the newly-emerged Low-Power Wide-Area Network (LPWAN) and 5G technologies have become strong candidates for mass-market localization applications. However, various error sources have limited localization performance by using such IoT signals. This paper reviews the IoT localization system through the following sequence: IoT localization system review - localization data sources - localization algorithms - localization error sources and mitigation - localization performance evaluation. Compared to the related surveys, this paper has a more comprehensive and state-of-the-art review on IoT localization methods, an original review on IoT localization error sources and mitigation, an original review on IoT localization performance evaluation, and a more comprehensive review of IoT localization applications, opportunities, and challenges. Thus, this survey provides comprehensive guidance for peers who are interested in enabling localization ability in the existing IoT systems, using IoT systems for localization, or integrating IoT signals with the existing localization sensors.

Index Terms—Positioning system; indoor navigation; Survey; Review; State of the art; error sources; LoRaWAN; NB-IoT; Sigfox; LTE-M; 5G; machine learning; artificial intelligence; multi-sensor integration.

I. INTRODUCTION

The Internet of Things (IoT) is shaping the future of many industrial and mass-market applications. As a core technology to acquire spatial IoT data, localization techniques are both an important application scenario and a distinguishing feature for the next-generation IoT [1]. In particular, Location-Enabled IoT (LE-IoT) is becoming key to add location context to IoT data without human perception and intervention.

This section answers three questions: (1) why is it necessary to review localization techniques for IoT systems; (2) what are the advantages and challenges for IoT-signal-based localization; and (3) what are the contributions of this survey.

A. Localization Technologies and Applications

As a military-to-civilian application, localization has been intensively researched and successfully commercialized in many professional applications, such as underground construction and machine industry. In contrast, mass-market IoT localization is more challenging due to the following factors:

- Many low-cost IoT nodes (including user end-devices) cannot afford Global Navigation Satellite Systems (GNSS) receivers.
- The complexity of environments. For example, the occurrence of Non-Line-of-Sight (NLoS) [2], multipath [3], wide-area [4] and multi-floor [5] effects, and interference by moving objects and human bodies [6].
- The necessity of using low-cost sensors that have significant sensor errors. The sensor errors also change over time and are susceptible to environmental factors (e.g., temperature [7]).
- The variety of node motions, such as changes of speed and orientation [8]. Also, it may be difficult to constrain node motion with predefined paths.

There are various types of localization technologies. Their advantages and disadvantages are

- GNSS. GNSS can provide global weather-independent positioning solutions. However, its performance can be degraded by signal outages, degradations, and multipath in indoor and urban environments [9].
- Wireless localization. Localization using wireless signals can provide long-term location accuracy. However, its performance is highly dependent on signal availability and geometry [10]. Meanwhile, its accuracy can be degraded by signal fluctuations and interferences due to NLoS conditions [2], multipath [3], and the outage [11] and time variance [8] of radio maps.
Environmental signals (e.g., magnetic, air pressure, light, and sound intensity). Database matching (DB-M) is one main technique for localization using environmental signals. Environmental signals can be sensed by user devices without specific infrastructure. The challenges include the dependency on features in environmental signals, the low signal dimension [8], and the time variance and outage [12] of environmental signal feature maps.

- Dead-reckoning (DR). Motion sensor (e.g., inertial sensor, magnetometer, and odometer) based DR can provide autonomous outdoor/indoor localization solutions [13]. Nevertheless, it is challenging to obtain long-term accurate DR solutions with low-cost sensors because of the existence of sensor errors [7], the misalignment angles between the vehicle (e.g., human body and land vehicles) and device [14], and the requirement for position and heading initialization.

- Visual localization. Vision sensors (e.g., cameras and Light Detection and Ranging (LiDAR)) can provide high location accuracy when loop closures have been correctly detected [15]. However, the performance of visual localization systems is highly dependent on whether the measured features are distinct in space and stable over time [16].

In general, the existing technologies have advantages and limitations [8]. Thus, it is difficult to generate low-cost but high-performance localization solutions by using a standalone technology. Multi-sensor integration has become a trend to achieve reliable, continuous, and accurate outdoor/indoor seamless localization.

The Low-Power Wide-Area Network (LPWAN) and 5th Generation cellular network (5G) technologies have been used for communication in pilot sites. However, their localization capability has not been fully developed. Many of the existing IoT systems still rely on location solutions from the existing localization technologies such as GNSS [17] and Wireless Fidelity (WiFi) [18]. There are two reasons for this phenomenon: (1) the deployment density (i.e., the number within a given area) of current IoT Base Stations (BS) is not high enough for accurate localization. (2) There are several error sources and challenges in IoT localization uses. For the second factor, this survey provides detailed investigation and guidance.

B. Advantages and Challenges of IoT Signals for Localization

The latest communication infrastructure is beginning to support the research on IoT signal based localization because: (1) IoT signals have been supported by mainstream IoT devices and are expected to be supported by more intelligent consumer devices. (2) IoT systems can already provide various localization-signal measurements such as Received Signal Strength (RSS), Time Difference of Arrival (TDoA), and Channel State Information (CSI). (3) The popularization of IoT/5G small BSs and the possibility to enable the communication capability of smart home appliances (e.g., lamps, routers, speakers, and outlets) are increasing the density of localization BSs. This survey focuses on LPWAN signals but also covers other IoT technologies such as cellular networks (e.g., 5G) and local wireless networks (e.g., WiFi, Bluetooth Low Energy (BLE), Zigbee, and Radio-Frequency IDentification (RFID)).

Fig. 1. Coverage ranges and power consumption of IoT signals (modified on [23] [57] [76]).

The research paper [22] has reviewed the physical structures, techniques, and parameters for LPWAN, as well as the specific techniques for meeting the requirements such as long-range, low power consumption, and low cost.

On the other hand, the emergence of LPWAN has brought new challenges for localization techniques. Such challenges include the existence of wide-area scenarios, the high requirement for power consumption, the necessity of using low-data-rate and low-cost nodes, the high density of nodes, and the existence of complex node motions. Section IV reviews the IoT localization error sources and their mitigation in detail.

C. Related Surveys and Tutorials

The survey paper [19] describes several key IoT parameters, such as quality of service, scalability, latency, network coverage, battery life, and cost. Meanwhile, the review paper [22] investigates the LPWAN physical techniques for the requirement of long-range, low-power-consumption, low-cost, and...
D. Main Contributions and Structure

Compared to the related surveys, this paper has

- A more comprehensive and state-of-the-art review on IoT localization methods.
- The first review on IoT localization error sources.
- The first review on IoT localization error mitigation.
- The first review on IoT localization performance analysis and evaluation.
- A more comprehensive review of IoT localization applications, opportunities, and challenges.

This survey is organized as follows. Section II overviews the existing IoT technologies, followed by IoT localization applications, system architecture, and signal measurements. Section III demonstrates the state-of-the-art IoT localization methods. Afterwards, Section IV systematically reviews IoT localization error sources and mitigation. Then, Section V illustrates the localization performance evaluation methods. Finally, Section VI shows the new localization opportunities.

II. Overview

This section first compares the existing LPWAN systems from the perspective of localization-system users, followed by the application scenarios of IoT localization. Afterwards, the IoT localization system architecture and the types of localization signal measurements are described. This section answers the following questions: (1) how to choose an LPWAN system for localization purposes; (2) what are the potential application scenarios for LE-IoT systems; and (3) what are the possible measurements that can be used for localization.

A. LPWAN Technologies

1) LPWAN: The early-state IoT systems mainly used local communication technologies (e.g., RFID, WiFi, BLE, and Zigbee) and GNSS localization until LPWAN systems became available. Among current LPWAN technologies, LoRaWAN and Sigfox, which use license-free bands, and NB-IoT and LTE-M, which use licensed bands, are most widely used. Meanwhile, there are other LPWAN technologies, such as WiFi HaLow, Weightless, and Ingenu RPMA. Figure 2 illustrates the comparative aspects of several IoT technologies. The papers [19] [22] [24] [27] have detailed descriptions on these LPWAN technologies.
2) Summary and Insight on LPWAN Technologies: Table II compares the features of the main LPWAN techniques. Meanwhile, Figure 3 illustrates the advantages and challenges of NB-IoT, LTE-M, LoRaWAN, and Sigfox regarding communication. It is notable that some advantages, such as less BS density, may be disadvantages for localization. Although they have attracted most interests, there are various other LPWAN systems. Therefore, it is expected that the users can select LPWAN technologies according to their requirements and integrate multiple LPWAN signals for better communication and localization performances.

3) Relation Between LPWAN and 5G: In addition to LPWAN, the development of 5G (i.e., the fifth-generation cellular network) technology has brought opportunities for both communication and localization. 5G has been well-known for its high speed, massive connection, high reliability, and low latency in communication. Compared to 4G, 5G has innovatively designed various standards and solutions for different application scenarios, including Ultra-Reliable and Low Latency Communication (URLLC), Enhanced Mobile BroadBand (eMBB), and massive Machine-Type Communication (mMTC) [32]. LPWAN provides an important application
direction for 5G mMTC applications. Meanwhile, besides NB-IoT and LTE-M, which are within the 5G standardization group, there are other LPWAN networks and systems. They can provide complementary supports to 5G applications. On the other hand, because the coverage range for 5G BSs may be shrunk from kilometers to hundreds of meters or even under 100 m [33], such small BSs can help LPWAN communication and localization. Also, 5G features (e.g., mmWave MIMO, large-scale antenna, beamforming, and D2D communication) may enhance LPWAN performance and experience. Moreover, 5G has a stronger connection with the new-generation information technologies (e.g., big data, cloud/edge computing, and Artificial Intelligence (AI)) and thus can extend the application space of LPWAN.

B. IoT Localization Applications

According to the above analysis, IoT systems are especially suitable for applications that have a massive connection, low data rate, low power consumption, and are not sensitive to latency. Many such applications have a strong requirement for localization. Examples of these applications include:

- Emergency service: determining people’s location is a feature of increasing importance for emergency systems such as the Enhanced 911 (E-911) in North America and the E-112 in Europe. The current E-911 system uses cellular signals and has a typical localization accuracy of 80% for an error of 50 m [34]. IoT signals have the potential to enhance location accuracy in urban and indoor areas. Meanwhile, LE-IoT can localize patients and medical devices and then provide services such as remote monitoring, fall detection, and motion analysis.

- Smart community and home: through the deployed IoT BSs and nodes around a community, it is possible to localize the residents and their surrounding facilities (e.g., security alarms, fire alarms, lamps, air conditioners, and surveillance cameras) and home appliances (e.g., smart speakers, lamps, and outlets). Location is essential to many personalized services.

- Intelligent transportation and logistics: LE-IoT nodes or chips in vehicles (e.g., cars or bikes) can be used for positioning and information tracking. The vehicle and related infrastructure (e.g., charging piles and parking spaces) locations can be used for traffic monitoring and parking guidance.

- Smart animal husbandry and animal tracking: LE-IoT can be used to track livestock locations and motions and thus provide services such as diet monitoring and meat traceability.

- Environmental monitoring and smart agriculture: LE-IoT can be used for localizing environmental hazards such as debris flows, sewer abnormities, and hazardous wastes. Also, LE-IoT nodes around farms can be used to localize fertilization devices and monitoring environmental factors (e.g., temperature and humidity).

Many of the current LE-IoT nodes use existing localization sensors (e.g., GNSS, inertial sensors, WiFi, and RFID) and an extra communication module (e.g., LTE), which lead to the high cost and power consumption. With the advent of LPWAN, the cost of communication modules can be significantly reduced. For example, the configuration of LPWAN (for communication) plus GNSS (for positioning) has been used for bus tracking [35], highway tracking [36], and patient monitoring [37]. Furthermore, if the LPWAN localization capability can be explored, it will be possible to remove all or parts of other localization sensors and thus further reduce node cost and power consumption. If this is the case, the localization accuracy may degrade as well.

C. IoT Localization System Architecture

An LE-IoT system is comprised of four components: nodes (including end-devices), BSs (including gateways or anchors), network servers, and application servers [22]. Figure 4 shows an LE-IoT system architecture. The LE-IoT system has an extra localization engine compared to an ordinary IoT system. The localization module may be located at either nodes or network servers, depending on user requirements on node computational load, communication load, and data security. The main functionalities of its components are as follows.

1) Nodes: A node contains a transponder, which transmits signals, and optionally a micro-controller with on-board memory. Meanwhile, the node optionally has application sensors such as a GNSS receiver for precise positioning, inertial sensors for motion tracking, and environmental sensors for monitoring temperature, humidity, smoke, gas, light, magnetic, and sound. The sensors may be connected to or integrated within the transponder chip. Also, the nodes may be fixed for static monitoring, mounted on dynamic objects as tags, or put on the human body as user devices. In some IoT applications, the nodes only broadcast signals frequently, instead of processing data, to save energy. There are also applications in which motion-tracking or localization data processing is implemented on nodes to reduce communication load. Meanwhile, some applications may use ToA localization, which requires precise timing on nodes. This requirement can only be met in relatively high-end IoT applications.

2) BSs: The main communication function of BSs is to route the data between nodes and network servers. The BSs may connect to network servers and transmit the data from

![Fig. 4. LE-IoT system architecture](Image)
node sensors to network servers and vice versa. BSs usually have fixed and known locations as well as globally unique IDentities (IDs, e.g., MAC addresses). For LE-IoT, BSs also need to measure localization signals, such as node ID, BS ID, the data reception time, channel, RSS, payload, and Signal-to-Noise Ratio (SNR). Furthermore, BS time synchronization may be required for TDoA or ToA based localization [38]. Meanwhile, multi-array antennae and phase detection may be needed for Angle of Arrival (AoA) localization [39].

3) Network servers: A network server is responsible for decoding data from BSs, recording data into databases, optionally implementing localization computation, and transmitting processed data to application servers. Network servers can be used for both sensor-to-application and application-to-sensor communication. For TDoA based localization, it is important that the packets from different BSs arrive at a network server. Furthermore, for localization applications, there are extra localization-signal databases on network servers. Meanwhile, motion-tracking and localization data-processing engines are located at network servers for many LPWAN applications. Network servers may be either cloud or edge servers.

4) Application servers: Their main functions are to obtain data from network servers, parse it, and process it for further applications.

D. IoT Localization Signal Measurements

Compared to IoT, LE-IoT systems measure localization signals and process them to estimate motion states such as location, velocity, attitude, and motion modes. This subsection illustrates the commonly used localization signals.

1) RSS: RSS is measured when a BS or node receives the data packet from the other side. The advantages of using RSS include: (1) RSS can be straightforwardly collected without extra hardware on either nodes or BSs. (2) RSS can be flexibly used for various localization algorithms, such as proximity, region-determination, multilateration, and DB-M. On the other hand, the challenges for using RSS include: (1) it is difficult to determine the Path-Loss Model Parameter (PLM-P) accurately in wide-area [40], urban [41], and indoor [42] scenarios, where many IoT applications take place. (3) The RSS-ranging resolution degrades over node-BS distance. An RSS change of one dBm may lead to differences of meters in small areas but hundreds of meters in wide areas [43]. (4) RSS variations and interference due to environmental factors are issues inherent to wireless signals.

2) ToA: ToA is obtained by measuring the time interval between signal transmission and reception. The advantages of using ToA include: (1) theoretically, ToA measurements can be linearly converted to node-BS distances without any known PLM-P. (2) ToA ranging can achieve high accuracy (e.g., decimeter- or even centimeter-level) in line-of-sight (LoS) environments [44]. (3) ToA localization has a well-researched theoretical-derivation and accuracy-assessment mechanism [45]. The challenges for ToA localization include: (1) ToA measurements require precise timing on both nodes and BSs or precise time synchronization between them. Ten-nanosecond-level timing accuracy is required to achieve meter-level ranging. Such timing accuracy is not affordable for many IoT nodes. (2) High accuracy is commonly expected when ToA is used. In this case, the degradations from environmental factors (e.g., NLoS and multipath) are relatively more significant.

3) TDoA: TDoA is measured by computing the signal arrival time differences among multiple BSs. TDoA localization has the following advantages: (1) it does not need precise timing on nodes or precise time synchronization between nodes and BSs. Instead, it only requires precise BS time synchronization, which is affordable for many IoT systems such as LoRaWAN, Sigfox, and NB-IoT [38]. (2) The impact of node diversity can be mitigated through the use of differential measurements between BSs. (3) TDoA localization methods, such as hyperbolic localization, have a well-researched theoretical-derivation and accuracy-assessment mechanism [46]. On the other hand, the challenges for TDoA localization include: (1) the requirement of precise time synchronization increases the BS cost. (2) The use of differential measurements enhances the impact of noise in localization signals.

4) AoA: AoA systems provide the node position by measuring BS-node angles [47]. The advantages of AoA positioning include: (1) typical AoA localization systems can provide high-accuracy (e.g., decimeter- or centimeter-level) locations [48]. (2) AoA requires less BSs than ToA and TDoA. It is feasible to use two BS-node angle measurements, or one BS-node angle and one BS-node distance, for two-dimensional (2D) localization. By fixing the AoA BS on the ceiling with known height, it is even possible to provide accurate localization with one BS [48]. (3) AoA localization approaches, such as multi-antenna, have a well-researched theoretical-derivation and accuracy-assessment mechanism [49]. The challenges for AoA localization include: (1) AoA systems need specific hardware such as multi-array antennae and phase-detection [39]. The high node cost has limited the use of AoA in low-cost IoT applications. (2) Although there are low-cost RSS-based AoA systems [50] [51], the accuracy of both angular-measuring and positioning degrade significantly when the BS-node distance increases. Thus, a high-density BS network is still needed for wide-area applications.

5) Round-Trip Time (RTT): RTT can be collected by measuring the round-trip signal propagation time to estimate the distance between nodes and BSs [52]. The use of RTT has advantages such as: (1) compared to ToA, RTT needs less accurate clock synchronization between BSs and nodes [1]. (2) RTT can be collected from the MAC layer, instead of the PHY [53]. (3) It is straightforward to use ToA-localization methods for RTT localization. The challenges for RTT localization include: (1) Modification on nodes is not affordable for many low-cost IoT applications. (2) The response delay between signal reception and transmission, which is difficult to eliminate, directly leads to ranging errors [1]. Thus, it is important to estimate the delay and compensate for it. (3) RTT-estimation accuracy is degraded by the same error sources as ToA.

6) CSI: It is becoming possible to collect CSI between IoT nodes and BSs [54]. The advantages of using CSI include: (1) CSI localization can achieve a high accuracy (e.g., decimeter
level or higher) \cite{55}. (2) CSI measurements have more features than RSS. (3) CSI is more robust to multipath and indoor noise \cite{1}. (4) Many existing localization approaches, such as DB-M and multilateration, can be used for CSI localization. The challenges for CSI localization include: (1) CSI may not be available on off-the-shelf Network Interface Controllers (NICs). (2) The CSI measurements may suffer from deviations because of factors such as limitations in channel parameter estimation \cite{56}. (3) It is challenging to assure the CSI-based ToA measurement accuracy due to the limited IoT signal bandwidth \cite{56}. To mitigate this issue, techniques such as frequency hopping may be needed.

7) Phase-of-Arrival (PoA): PoA is obtained by measuring the phase or phase difference of carrier signals between nodes and BSs. PoA measurements can be converted to BS-node distances \cite{57}. The advantages for PoA localization include: (1) PoA measurements can achieve high (e.g., centimeter-level or higher) ranging accuracy \cite{58}. (2) The existing ToA and TDoA algorithms can be directly used for PoA localization. The challenges for PoA localization include: (1) Extra node and BS hardware are needed to measure PoA \cite{57}. Meanwhile, accurate PoA-ranging requires a relatively high data rate, which is not suitable for many IoT applications. (2) High accuracy is commonly expected when PoA is used. In this case, the degradations from environmental factors (e.g., NLoS and multipath) are relatively more significant \cite{57}. (3) PoA measurements may suffer from the integer-ambiguity issue \cite{59} and cycle slips.

8) Summary and Insight on IoT Localization-Signal Measurements: Similar to other engineering problems, the selection of IoT localization signals is a tradeoff between performance and cost. Some measurements (e.g., ToA, AoA, RTT, and PoA) can be used to achieve high localization accuracy but require extra hardware or modifications on nodes, which are not affordable for many low-cost IoT applications. In contrast, measurements such as RSS can be collected without any change in hardware; however, their localization accuracy is lower, especially for wide-area applications. Another type of measurements (e.g., TDoA and CSI) can be realized by adding extra hardware or modifications on the BS side, which is affordable for some IoT applications. Besides performance and cost, other factors should be considered when selecting localization signals. Examples of these factors include environment size, outdoors or indoors, node motion modes, and the number of nodes.

Because each type of localization measurement has advantages and limitations, it is common to combine various types of measurements (e.g., TDoA/RSS \cite{60} and AoA/TDoA \cite{61}) for a higher localization performance. Furthermore, data from other sensors (e.g., inertial sensors, magnetometers, barometers, and maps) can be introduced to enhance localization solutions by mitigating the impact of error sources that are inherent to wireless signals.

Moreover, all the localization signals suffer from both deterministic and stochastic measurement errors. The impact of deterministic errors (e.g., sensor biases, scale factor errors, and thermal drifts) may be mitigated through calibration or on-line estimation. In contrast, the stochastic measurement errors can be modeled as stochastic processes. The statistical parameters for stochastic models may be estimated through methods such as correlation, power spectral density analysis, and Allan variance \cite{62}.

III. IOT LOCALIZATION METHODS

This section will answer the following questions: (1) what are the state-of-the-art localization approaches; and (2) what are the advantages and challenges for each type of localization method. Over one decade ago, the majority of localization methods were geometrical ones, which are realized on geometric measurements such as distances and angles. By contrast, DB-M methods, which are data-driven, have been well developed during this decade because of the development of Machine Learning (ML) techniques and the diversification of modern localization scenarios. Meanwhile, there are DR methods that use sensors such as inertial, odometry, and visual ones. Figure 5 demonstrates part of the main localization algorithms.

A. Database-Matching Localization Methods

Although different localization technologies have various physical measurements and principles, they can generally be used for localization through DB-M. The basic principle for DB-M localization is to compute the difference between the measured fingerprints and the reference fingerprints in the database and find the closest match. The DB-M process consists of three steps:

- Localization feature (LF) extraction. Valuable LFs are extracted from raw localization signals. A valuable LF should be stable over time and distinct over space. Examples of the LFs include RSS/CSI for wireless localization, magnetic intensity for magnetic matching, and visual features for vision localization. The extracted LFs are recorded and used for training and prediction.
- Database training. [LF, location] fingerprints at multiple Reference Points (RPs) are used to generate or update a database, which can also be regarded as a type of map. The database may be a data structure that stores the LFs at
multiple RPs or be the coefficients of parametric models. Figure 6 demonstrates the principle of the training and prediction steps.

- Prediction. The real-time measured LFs are compared with the database to locate the node. The likelihood (or weight) for each RP can be computed through deterministic, stochastic, and ML methods. These methods are described separately in the following subsections.

1) Deterministic DB-M: In these methods, only deterministic values (e.g., the mean value) of each LF at each RP are stored in the database. Thus, the LF values at each RP construct a vector, while the reference LF values at multiple RPs build a matrix. Each column vector in the matrix is the reference LF vector at one RP. At the prediction step, the similarity between the measured LF vector and each reference LF vector in the matrix is calculated by computing the vector distance. The RPs that have the highest similarity values are the nearest neighbors. To compute the similarity, the Euclidean distance is widely used. Meanwhile, there are other types of vector distances, such as Manhattan distance, Minkowski distance, Spearman distance, and information entropy.

The deterministic DB-M methods have advantages of small database size and computational load. On the other hand, the main limitation is that the stochastic LFs, which may be caused by various factors in real-world practices, have not been involved.

2) Stochastic DB-M: Compared to the deterministic methods, stochastic DB-M approaches introduce stochastic LFs at each RP through various methods, such as the Gaussian-distribution and histogram methods. Meanwhile, likelihood between measured LFs and the reference ones in the database is used to replace vector distance as the weight of similarity.

It is notable that the majority of related works assume that various LFs (e.g., RSS values with different BSs) are independent with one another, so as to simplify the likelihood computation to the product (or summation) of the likelihood values for all LF components. Then, the likelihood-

computation problem becomes how to estimate the likelihood value for each LF component. If the Gaussian-distribution method is used, the likelihood value for an LF can be estimated by applying the Gaussian-distribution model.

Although the Gaussian-distribution model is one of the most widely used models for stochastic errors in localization applications, there may be systematic LF measurement errors due to environmental factors. The existence of systematic errors theoretically breaks the Gaussian-distribution assumption. However, the Gaussian distribution is still widely used in engineering practices because real-world environmental factors are difficult to predict and model. One approach for reducing the degradation from systematic errors is to set relative larger variance values in localization filters to absorb the impact of such errors.

Moreover, WiFi and LoRaWAN RSS values may have not only symmetric histograms but also asymmetric ones, such as left-skewed, bimodal, and other irregular histograms. Such asymmetric measurement distributions may be caused by environmental factors. An approach for mitigating the impact of asymmetric measurement distributions is to use histograms or advanced stochastic models.

The histogram method can obtain likelihood values without assumption on signal distributions. The histograms for all LF components at all RPs are calculated in the training step. In the prediction step, the measured fingerprint is compared with the corresponding histogram to find the likelihood. However, since it determines likelihood values through histogram matching, instead of using a parametric model, it may suffer from a large database size and overfitting.

3) ML-based DB-M: In recent years, ML has started to bring empowerment to numerous applications because of the increased data volume, the increased computing power, and the enhanced ML algorithms. This subsection reviews the typical ML algorithms that have been utilized for localization.

- Artificial Neural Network (ANN): ANN is a type of framework for using ML methods to process complex (e.g., nonlinear and non-Gaussian) data. ANN consists of one input layer, at least one hidden layer, and one output layer. Each layer contains at least one neuron. The neurons are connected via weights and biases, which are trained and stored. There are numerous publications (e.g.,) on the principle of ANN. Also, there are various types of ANN, such as the Long Short-Term Memory (LSTM) networks and Convolutional Neural Networks (CNN).

ANN techniques have been used for localization over a decade ago; however, it has not been widely adopted until recent years. RSS (e.g., RSS from WiFi, BLE, ZigBee, RFID, cellular, and photodiode), RSS features (e.g., 2D RSS map, differential RSS, and RSS statistics), CSI, and AoA have been used. The majority of these works directly output node locations, while the others also generate the identification of floors, rooms, and regions, NLoS, the similarity of fingerprints, localization success rate, and localization accuracy prediction. The paper provides a detailed review of the works that use ANN in wireless localization.

ANN has several advantages: (1) the algorithm has been well-developed and successfully in various fields (e.g., speech...
recognition and image processing). (2) The current ANN platforms and toolboxes are open and straightforward to use. On the other hand, the shortcomings of ANN include: (1) an ANN model is similar to a black box for most users. It is difficult to determine an explicit model representation of how the ANN works. (2) It is difficult to understand and adjust the internal algorithms. For example, although the majority of localization works above use one to three hidden layers, they set the numbers of hidden layers and neurons through brute-force data processing, instead of following a theoretical guide. Such specifically-tuned ANN parameters maybe not suitable for varying localization environments.

- **Gaussian Processes (GP):** GP is a supervised ML method for regression and probabilistic classification [72]. A GP is a set of random variables that have joint Gaussian distributions. Therefore, for localization applications, GP can involve the correlation among all RPs. This characteristic makes GP different from many other localization methods that treat each RP separately. Another characteristic for GP is that it can be uniquely determined by a mean function and a kernel function (i.e., covariance function). In the localization area, the mean function may be set by using geometrical LF models [66]; meanwhile, a zero mean function is used in some scenarios. In contrast, there are various types of covariance functions, such as the constant, linear, squared exponential, Matérn, and periodic ones [72]. The geometrical LF models have not been involved in covariance functions in the existing works. The research in [74] has presented a hyperparameter estimation model for learning the GP model parameters.

The use of GP for RSS localization has been proposed in [73] for cellular networks. The paper [74] extends the work by introducing a Bayesian filter that builds on a graphed space representation. Afterwards, GP has been used in processing data from various localization sensors, such as magnetometers [12] and RFID [75]. Furthermore, GP has become one of the main techniques for localization-database prediction (or interpolation), that is, to predict LFs at unvisited or out-of-date RPs based on training data at other RPs. Reference [76] compares the performance of DB prediction by using GP and geometrical interpolation methods.

GP has advantages such as: (1) it has a physical meaning and an explicit model representation, compared to many other ML methods. (2) GP captures both the predicted solution and its uncertainty. The latter is not provided in ANN. (3) GP has a small number of parameters; thus, its engineering implementation is straightforward. The challenges for GP include: (1) it is based on a Gaussian-process assumption, which may be degraded in challenging localization scenarios. Integrating GP with geometrical localization models may be a possible method to mitigate this degradation [66]. (2) GP has a small number of parameters. Thus, in localization scenarios that have complex environments and massive data, GP may not be able to exploit the potential of complex databases as well as other ML methods (e.g., ANN).

- **Random forests:** The random forests algorithm is an ensemble classifier that uses a set of decision trees (i.e., classification and regression trees) for supervised classification [77]. The paper [78] has a detailed description of its principle and theoretical formulae. Random forests can be implemented through three steps: subsampling, decision-tree training, and prediction. In the subsampling step, the algorithm randomly selects a subsample that contains a fixed number of randomly-selected features from the original dataset. The subsample is trained with a decision in the decision-tree training step. The training process creates the if-then rules of the tree. One typical method for this process is Gini impurity, which is a measure of how often a randomly selected element will be incorrectly labeled if the element is randomly labeled according to the label distribution in the subset.

When splitting a branch in the tree, all possible conditions are considered and the condition with the lowest Gini impurity is chosen as the new node of the decision tree. If the split is perfect, the Gini impurity of that branch would be zero. There are other criteria (e.g., information gain). For prediction, each tree in the ensemble gives a prediction result. Based on the votes from all trees, a probabilistic result can be generated.

In the localization area, the random-forest approach has been applied for fingerprinting [79] and NLoS condition identification [2]. The advantages of random forests include [80]: (1) compared to other decision-tree based methods, it is less sensitive to outliers in training data. (2) The random-forest parameters can be set easily. (3) Random forests can generate variable importance and accuracy together with prediction solutions. The shortcomings of random forests include: (1) it is not efficient in computational load. A large number of trees are needed for an accurate vote. This phenomenon leads to large databases and computational loads. (2) It may be over-fitted. Also, it is sensitive to noise [81].

- **Deep Reinforcement Learning (DRL):** DRL, which is the core algorithm for AlphaGo, has attracted intensive attention. It combines deep learning and reinforcement learning. The former provides a learning mechanism, while the latter provides goals for learning [82]. In general, DRL allows the agent to observe states and act to collect long-term rewards. The states are mapped to an action through a policy.

The DRL algorithm has experienced stages such as Deep Q-Networks (DQN), Asynchronous Advantage Actor-Critic (A3C), and UNSupervised REinforcement and Auxiliary Learning (UNREAL). Specifically, DQN [83] introduces value networks to represent the critic module, and constructs value networks according to specific applications by using ANNs. Then, A3C applies the actor-critic framework and asynchronous learning. The basic idea of actor-critic is to evaluate the output action and tune the possibility of actions based on evaluation results. Compared to A3C, the UNREAL algorithm is closer to the human-learning mode. Specifically, UNREAL enhances the actor-critic mechanism through multiple auxiliary tasks. The research in [84] has pointed out three components for a DRL solution: basis/core (e.g., state definition, action definition, and reward definition), basic units (e.g., Q-network, action selection, replay memory, and target network), and state reformulation (i.e., the method for state-awareness data processing).

DRL has been used for navigation in Atari games, mazes, and the real world. Meanwhile, DRL has been applied for navigation using data from a monocular camera, 360-degree
camera, LiDAR, magnetic sensor, wireless sensor, and Google street view. The paper [85] has a review of the sensors and approaches that use DRL in navigation and localization. Many of the recent DRL research works focus on navigation without a map, in new environments, and with varying targets. However, most of these methods are designed for navigation, instead of localization. Navigation and localization use the same sensor data as the input but have different purposes. Navigation is the issue of finding the optimal motion path between the node and a target location; in contrast, a localization module outputs the node location. It is straightforward to model a navigation process as a Markov decision process; thus, it can be processed by DRL. By contrast, localization is closer to a deep-learning problem.

The advantages of DRL include: (1) it can obtain not only the optimal solution at the current moment but also the long-term reward. (2) DRL can reduce the computational complexity caused by re-optimization due to factors such as environment changes [84]. The challenges of DRL include: (1) the set of reward definition is key to the DRL performance. However, it is challenging to determine a theoretical model for reward definition. (2) The DRL algorithm itself has met several challenges [86], such as hyperparameter sensitivity, sample efficiency, off-policy learning, and imitation learning. (3) The future DRL algorithms may need supports from ML chips due to their large computational loads.

There are also other ML methods for enhancing localization. For example, the Hidden Markov Model (HMM) approach has been used for RSS fingerprinting, trajectory modeling, and room recognition. Also, the Support Vector Machine (SVM) method has been applied for wide-area localization [87]. Meanwhile, the fuzzy-logic method has been used for smartphone localization [88]. Various ML methods have different advantages and thus are suitable for different localization use cases.

B. Geometrical Localization Methods

Geometrical localization methods have been researched for decades. To use such methods, BS locations are commonly known or can be estimated. The main measurements are BS-node distances and angles. The main localization methods include multilateration, hyperbolic localization, multiangulation, multiangulation, and other simplified methods such as min-max, centroid, and proximity. Previous survey papers [1] and [28] already have detailed descriptions of these methods. Thus, this survey only summarizes their main characteristics. Figure 7 illustrates the principle of several geometrical localization methods.

1) Multilateration: Multilateration can be used to estimate node location by using locations of at least three BSs and their distances to the node. Its basic principle is to estimate the intersection between spheres (for 3D localization) and circles (for 2D localization). This method has been widely used in wireless BS location estimation [59]. The common estimation techniques are the least-squares and Kalman filter (KF). The multilateration performance may be enhanced by improving ranging accuracy by mitigating the impact of environment-related and receiver errors [90]. There are also well-developed blunder-detection and accuracy-evaluation mechanisms [91] as well as geometry indicators such as the Dilution of Precision (DOP) [10].

2) Hyperbolic Localization: Hyperbolic localization, which was developed for Loran navigation, is the main method for TDoA localization. It is based on the distance differences between the node and various BSs. Because hyperbolic localization has eliminated the requirement for precise timing on nodes, it has strong potential for LPWAN localization. There is also an accuracy analysis [46] for this method.

3) Multiangulation: Multiangulation can be implemented by measuring the angles between the node and at least two BSs [47]. Theoretically, two BS-node angles can determine a 2D point. When considering angle-measurement errors, a quadrangle can be determined. The multiangulation method has been used in wireless Ad Hoc networks to reduce the requirement on BS (or anchor) locations [94]. The performance analysis of multiangulation has also been provided [49].

4) Multiangulation: Multiangulation localizes a node by at least one BS-node angle and one BS-node distance. This method has been widely used in traverse networks in engineering surveying [95]. For indoor localization, it is feasible to use one AoA (e.g., VLP [28] and BLE [48]) BS on the ceiling with known height for localization.

5) Min-Max: Min-max is a variant of multilateration. Its geometrical principle is to calculate the intersection between cubes (for 3D localization) and squares (2D localization), instead of spheres and circles. The benefit of using cubes and squares is that their intersections can be directly computed through deterministic equations [96], which have significantly lighter computational loads than least squares. The limitation of min-max is that it has not considered the stochastic measurements. Meanwhile, compared to spheres, cubes deviate from spatial distribution with a certain BS-node distance. The min-max method can be used to generate coarse localization solutions, which provide initial positions for fine-localization approaches such as multilateration.

6) Centroid: Centroid is a simplification of multilateration. It estimates the node location by a weighted average of the locations of various BSs. The weights for BS locations are commonly determined by BS-node distances [97]. Similar to min-max, the centroid has a low computational load but...
has not considered stochastic measurements. Furthermore, the centroid result will be limited within the region that has a boundary formed by BS locations. Similar to min-max, the centroid is commonly used for coarse localization.

7) Proximity: Proximity can be regarded as a further simplification of the centroid approach. The possible location region is determined by using the location of one BS as the circle center and using the BS-node distance as the radius. Proximity is commonly used in RFID [28] and cell-identification [99] localization. The method has the lowest computational load, the lowest number of BS, but the largest location uncertainty. Thus, it is commonly used for near-field localization, coarse localization, or for bridging the outages when the other localization methods do not have sufficient BSs.

C. Summary and Insight on Database-Matching and Geometrical Localization Methods

DB-M and geometrical localization methods have several similarities, such as

- Both methods have the training and prediction steps. In DB-M methods, a [LF, location] database is generated through training; in contrast, in geometrical methods, the coefficients for parametric models are estimated and stored at the training step.
- They have some common error sources, such as device diversity, orientation diversity, and human-body effects.

On the other hand, these methods have differences (or complementary characteristics), such as

- Geometrical methods are more suitable for scenarios (e.g., outdoor and indoor open environments) that can be explicitly modeled and parameterized. In contrast, DB-M approaches are more suitable for complex scenarios (e.g., wide-area urban and indoor areas) that are difficult to be parameterized.
- For geometrical methods, environmental factors such as NLoS and multipath conditions are error sources that need to be modeled and mitigated. By contrast, DB-M methods may use the measurements of these factors as fingerprints to enhance localization.
- Geometrical methods are based on parametric models; thus, the databases contain only model parameter values and thus are relatively small. On the other hand, it is difficult to describe complex scenarios by using parametric models with limited numbers of parameters. In contrast, DB-M methods directly describe localization scenarios by using data at all points in the space; thus, both database and computational loads are large, especially for wide-area applications. However, DB-M approaches have more potential to provide higher resolution and more details on complex localization scenarios.
- Progresses in ML algorithms have brought the great potential to localization methods, especially for the applications that have complex scenarios that are difficult to model, parameters that are difficult to determine and tune, and have nonlinear, non-parametrical, correlated measurements that are caused by environmental and motion factors. Although ML methods are classified into the DB-M group in this survey according to the existing works, ML can also be used to enhance geometrical localization, such as training of parametric models and their parameters.

Due to their complementary characteristics, geometrical and DB-M methods can be integrated. For example, geometrical methods can be used to reduce the computational load of DB-M [100], to aid database prediction [66], and to provide localization uncertainty prediction [69]. It is also feasible to directly integrate geometrical and DB-M methods for more robust localization solutions [101].

IV. IoT Localization Error Sources and Mitigation

This section describes the main IoT-localization error sources and their mitigation. These two topics are key components in a LE-IoT system. This section answers two questions: (1) what are the error sources for IoT localization; and (2) how to mitigate or eliminate the impact of these error sources. The localization error sources are classified into four groups as

- **End-device-related errors**: device diversity, motion/attitude diversity, data loss and latency, and channel diversity.
- **Environment-related errors**: multipath, NLoS, wide-area effects, multi-floor effects, human-body effects, weather effects, and signal variations.
- **Base-station-related errors**: the number of BSs, BS geometry, BS location uncertainty, BS PLM-P uncertainty, and BS time synchronization errors.
- **Data-related errors**: database timeliness/training cost, RP location uncertainty, database outage, intensive data, data and computational loads, and localization integrity.

Figure 8 shows the main IoT-localization error sources. The details for the listed error sources are provided in the following subsections.

A. End-Device-Related Errors

This subsection introduces the error sources on the node side.

1) Node (Device) Diversity: A large number of low-cost sensors are used in IoT nodes. Such sensors may have diversity when measuring the same physical variable. For example, the RSS diversity for the same-brand BLE nodes may reach 20 dBm [102], which directly leads to localization errors.

To alleviate the impact of node diversity, there are calibration-based methods that fit data from multiple nodes by using histograms [103] and ANNs [104].

Meanwhile, there are calibration-free differential-measurement-based approaches, for example, the differential approaches that set the datum by using the data from a selected BS [105], the average data from all BSs [107], selected BSs [108], or an advanced datum-selection method [109]. Also, differences between data from all BS couples [106] have been calculated. There are two challenges for using differential signals: (1) it is challenging to switch
Fig. 8. IoT localization error sources

the datum for differential computation in a wide-area IoT application. (2) The differential computation increases the noise levels in measurements.

2) Motion/Attitude Diversity: IoT nodes may experience various motion modes, such as being held horizontally, dangling in the hand, mounted on chest, and stored in pockets or bags. This phenomenon leads to changes in node attitude (or orientation). Meanwhile, movements of nodes lead to changes in the relative attitude angles between node- and BS-antenna directions. Such attitude changes lead to localization errors [102].

To reduce the impact of motion/attitude diversity, the paper [111] extends the method to dynamic localization scenarios. Meanwhile, the paper [110] presents a heading-aided method that adds the node attitude into fingerprints and decision trees, respectively. Moreover, reference [112] uses histogram equalization for orientation-effect compensation, while the research in [102] presents an orientation-compensation model to compensate for orientation diversity.

The orientation angles of the BS antennae are commonly constant, while the node attitude angles are dynamic and can be estimated in real-time through the Attitude and Heading Reference System (AHRS) algorithm [113]. In such an algorithm, gyro measurements are used to construct the system model, while accelerometer and magnetometer data provide measurement updates. Meanwhile, autonomous gyro calibration is important to enhance attitude estimation [114].

3) Data Loss and Latency: The response rate is a practical factor in wireless localization. Data from some BSs may be wrongly missed in the scanning duration [115]. Also, the response rate may vary across BSs and are related to the level of signals [115]. BSs with lower RSS values tend to have lower response rates. The response rate can also be used as fingerprint information [116].

Meanwhile, to meet the requirement of low power consumption, IoT nodes commonly have low sampling rates. Furthermore, the existence of massive numbers of nodes may lead to data collision, loss [111], and latency [117]. Both data loss and latency may directly lead to localization errors. In particular, the loss or delay of data from an important BS (e.g., a BS that is geographically close to the node) [118] may lead to significant degradation on localization performance.

A possible method to mitigate this issue is to predict localization signals by using approaches such as time-series analysis [119] and ML [120]. Furthermore, data from other localization sensors or vehicle-motion constraints may be used to construct node-motion models, which can be used for localization-signal prediction.

4) Channel Diversity: In IoT applications, especially those with a high node density (e.g., in animal-husbandry applications), multi-channel mechanisms may be utilized to reduce data collision [120]. The difference between training and prediction data channels may lead to localization errors. To alleviate such errors, one method is to calibrate channel diversity through parametric models [121] or ML [120]. Meanwhile, data from multiple channels may be combined [122] or treated as data from various BSs [123] to enhance localization.

B. Environment-Related Errors

This subsection describes the environment-related localization error sources, including multipath, NLoS, wide-area effects, multi-floor effects, human-body effects, weather effects, and signal variations.

1) Multipath: Multipath is a common issue when using wireless signals in indoor and urban areas. Some papers have evaluated the impact of multipath in localization. For example, the paper [124] compares the impact of multipath on AoA and PoA methods with a photodiode and found that AoA localization was affected by an order of magnitude lower than PoA. Meanwhile, the research [125] has investigated the impact of multipath on RSS and CSI localization and found that CSI localization suffered from less degradation. In the localization area, the research on multipath is mainly on two aspects: mitigating its degradation on localization solutions and using it to enhance localization.

There are methods for mitigating the degradation from multipath. Some research reduce multipath errors by using specific node design, such as beamforming [126] and multi-channel signals [122]. Meanwhile, there are approaches for extracting
<table>
<thead>
<tr>
<th>Error Type</th>
<th>Details</th>
</tr>
</thead>
</table>
| **End-device-related errors** | - Calibration-based: histograms \[103\] and ANN \[104\].  
- Calibration-free: differential methods such as pairwise difference \[106\], mean difference \[107\]; advanced datum BS selection \[109\].  
- Motion/Attitude Diversity  
  - DB-M based: attitude-appended fingerprints \[110\].  
  - Combination of DB-M and parametric models: compensation using histogram equalization \[112\].  
  - Parametric model based: orientation-compensation model \[102\].  
- Data Loss and Latency  
  - Signal prediction: time-series analysis \[119\], multi-channel mechanisms \[120\], and ML \[120\].  
  - Control impact of data loss \[41\], lagency \[117\], low response \[115\] \[116\]; use important BSs \[118\].  
- Channel Diversity  
  - Calibration using parametric models \[121\] and ML \[120\].  
  - Combine data from multiple channels \[122\] or treated them as data from various BSs \[123\].  
| **Environment-related errors** | - Multipath  
  - Detection: multipath-delay extraction \[127\] and 3D city model assisting \[128\].  
  - Modeling: real-time multipath-parameter estimation by using SLAM \[129\] and EKF \[130\].  
  - Mitigation: beamforming \[126\] and multi-channel signals \[122\].  
  - Assisted localization: principle and methodology \[131\], evaluation models \[3\], CRLB \[132\], and uses in outdoor \[133\] and indoor \[134\] signal-BS systems; treat multipath components as signals emitted from virtual BSs \[135\]; use multipath-sensitive signal features \[136\].  
- NLoS  
  - Identification: ANN \[140\], random forests \[2\], and SVM \[141\]; use multi-channel data \[142\]; relation between NLoS and location errors \[138\]; NLoS-based CRLB \[139\].  
  - Modeling: Path-Loss Models (PLMs) that involve building walls and floors \[42\], thickness and intersection angles of obstacles \[143\], and losses of walls and interactions \[90\].  
  - Mitigation: estimation techniques such as PF \[144\] and UKF \[145\]; advanced models such as radial extreme value distribution \[146\]; integration with vision \[147\] and inertial \[148\] sensors.  
- Wide-Area Effects  
  - Evaluate wide-area PLM changes with environment type \[150\], terrain category \[4\] \[151\], and BS antenna height \[4\].  
  - Use advanced wide-area PLMs: multi-slope PLM \[152\], higher-order PLM \[40\], and height-dependent PLM \[153\]; ML methods for determining PLM-P values \[154\].  
  - Improved wide-area localization methods such as BS identity \[149\] and minimum-mean-square-error \[155\]; use wide-area and local IoT signals for coarse and fine localization, respectively.  
- Multi-Floor Effects  
  - Floor detection using data from wireless sensors \[5\], a barometer \[156\], inertial sensors \[158\], and a floor plan \[157\]; use estimation techniques such as KF \[156\], PF \[157\], and ANN \[5\].  
  - Evaluate the impact of multi-floor effects on PLMs \[42\].  
| **Human-Body Effects** | - Evaluation: human-body impact on localization \[159\]; evaluate the human-body effect when the node is carried on the human body and when a human is close to a BS \[159\]; evaluate the influence when the node is located at various places and with different orientations on human body \[6\], and short-term and long-term human-body effects \[136\].  
  - Modeling: model it on ToA ranging and localization \[161\]; model its relation with node orientation toward fixed BSs \[6\]; use other sensors (e.g., vision) to detect human bodies and aid the modeling of their effects.  
  - Mitigation: treat it as a NLoS signal and use NLoS-mitigation methods \[160\]; combine data from multiple nodes on different human body locations \[6\].  
  - Use human-body effects for device-free localization by using signals such as RSS \[162\] and CSI \[163\].  
| **Weather effects** | - Evaluate the relation between ranging/localization performance and weather factors such as temperature, relative humidity, air pressure, \[136\], rain condition and wind speed \[155\], and water \[151\].  
  - Integrate IoT signals with self-contained localization technologies (e.g., inertial navigation) to reduce weather effects.  
| **Signal Variations** | - Model stochastic signals by using methods such as Allan variance \[62\].  
  - Mitigation through denoising methods such as averaging \[166\], wavelet \[167\], and ANN \[168\].  

---

**TABLE III**

Localisation error sources and mitigation, Part #1: End-device- and environment-related errors
individual multipath propagation delays \[127\]. Using 3D city models to assist multipath detection, which is a combination of localization and mobile mapping, has also been researched \[128\]. Furthermore, some papers focus on estimating multipath parameters in real-time by using Simultaneous Localization And Mapping (SLAM) \[129\] and Extended KF (EKF) \[130\] algorithms.

Moreover, due to the new features in 5G, using multipath signals to enhance localization is attracting research interest. The research in \[131\] describes the principle and methodology for multipath-assisted localization. Meanwhile, reference \[3\] has derived the statistical performance bounds and evaluation models, while the paper \[132\] derives the Cramér-Rao lower bound (CRLB) for multipath-assisted localization. Multipath-assisted localization has been utilized in outdoor \[133\] and indoor \[134\] single-BS systems. Moreover, the research in \[135\] treats multipath components as signals emitted from virtual transmitters and uses them to assist localization within a SLAM algorithm, while the paper \[136\] improves the accuracy of RSS fingerprinting by introducing multipath-sensitive signal features.

2) NLoS: Similar to multipath, the NLoS effect is an important error source for wireless localization. NLoS is one factor that will cause multipath propagation. This subsection focuses on other effects of NLoS. It has been revealed that NLoS environments may significantly reduce the coverage range of IoT signals \[137\] and change their PLM-P values \[41\]. The paper \[138\] investigates the effect of NLoS signals on RSS localization and attempts to quantify the relation between localization errors and NLoS, while the research in \[139\] derives the NLoS-based CRLB. In general, the research on NLoS can be classified into three groups: identification, modeling, and mitigation.

To identify NLoS signals, algorithms such as ANN \[140\], random forests \[2\], and SVM \[141\] have been applied. There is also a method that detects NLoS by comparing the difference between signals in multiple frequency bands \[142\].

Examples of NLoS-modeling approaches include the use of advanced PLMs involving walls and building floors \[42\]. Furthermore, there is research that has considered the thickness of obstacles and intersection angles between obstacles and the direct path \[143\] as well as the wall and interaction loss factors \[90\].

To mitigate the NLoS effect, there are methods that use various estimation techniques, such as Particle Filter (PF) \[144\] and Unscented KF (UKF) \[145\]. Advanced models (e.g., the radial extreme value distribution model \[146\]) have also been used. Furthermore, it is feasible to realize NLoS mitigation by introducing external localization sensors, such as vision \[147\] and inertial \[148\] sensors.

3) Wide-Area Effects: Wide-area localization is more challenging than local localization due to factors such as lower RSS, SNR, and response rate values \[43\], and stronger multipath and fading effects \[149\]. Moreover, it is challenging to obtain wide-area PLM-P values because they change significantly with factors such as environment type (e.g., highway, rural, and urban) \[150\], terrain category (e.g., hilly, flat, with light/moderate/heavy tree densities, and on water/ground) \[151\], and even BS-node distances and BS antenna heights \[4\].

Advanced PLMs have been used to enhance wide-area localization. Examples of these PLMs include the multi-slope PLM for BS-node distances from meters to hundreds of meters \[152\], the higher-order PLM \[40\], and the height-dependent PLM \[153\]. ML methods (e.g., ANN) have been used to determine the PLM-P values \[154\]. Meanwhile, there are other improved wide-area localization algorithms, such as the BS-identity \[149\] and minimum-mean-square-error \[155\] ones. Considering the popularization of small BSs, it may become a trend to use wide-area IoT signals for coarse positioning and use local IoT signals for fine localization.

4) Multi-Floor Effects: Most of the existing works are focusing on 2D localization. However, in the IoT era, nodes may be used in 3D scenarios. A practical 3D-localization approach is to find the floor at which the node is located and then localize the node on the floor. Thus, robust floor detection is required. For this purpose, researchers have used data from various sensors, such as wireless sensors \[5\], a barometer \[156\], inertial sensors \[158\], and a floor plan \[157\]. There are also various estimation techniques, such as KF \[156\], PF \[157\], and ANNs \[5\]. The impact of multi-floor effects on PLM-P values has also been researched \[42\].

5) Human-Body Effects: It is necessary to consider the human-body effect in some LE-IoT applications, especially when the nodes are placed on the human body. There are four types of research on the human-body effect: evaluation, modeling, mitigation, and utilization.

Some papers evaluate the impact of the human body on localization \[159\]. It has been found that the human body may cause degradation on localization: also, the degradation is more significant when the node is carried on the human body, compared to the use case when a human blocks the BS-node LoS \[159\]. Meanwhile, the research \[6\] has revealed that the influence of the human-body effect varies with the node position and orientation. Also, the tests in \[136\] show that the human body leads to short-term RSS fluctuations, instead of long-term signal shifts.

There are approaches for modeling the human-body effect. For example, the literature \[161\] models the impacts on both Toa ranging and localization. Also, the paper \[6\] presents a compensation model for the human-body effect by introducing the user orientation toward fixed infrastructure. For modern localization applications, it may be feasible to use other sensors, such as vision, to detect the human body and compensate for its effect.

Moreover, to mitigate the human-body effect, the research in \[160\] treats it as an NLoS signal and mitigates its effect by using NLoS-mitigation methods. Also, the paper \[6\] combines the data from multiple nodes at different places on a human body to control the human-body effect.

Furthermore, human-body effects have been utilized for device-free localization. Signals such as RSS \[162\] and CSI \[163\] are used.

6) Weather Effects: Researchers have studied the weather effect on wireless signal propagation. Theoretically, large-wavelength wireless signals are not susceptible to external fac-
tors such as precipitation and vegetation [164]. Meanwhile, the research in [136] indicates that FM RSS has a weak correlation with temperature, relative humidity, air pressure, wind speed, and aviation-specific runway visibility. The research has not considered weather-dependent environment changes such as movements of trees and power-line wires as well as changes in ground conductivity and the multipath effects.

However, on-site experimental results in [165] show that variations in rain conditions and wind speeds significantly degraded accuracy in distance estimation using Global System for Mobile communication (GSM) RSS. Meanwhile, reference [151] shows that LoRaWAN RSS measurements suffer from significant noise and drifts when operating near lakes. Thus, whether the weather change has a significant impact on a certain type of IoT signal need specific evaluation. Meanwhile, integrating IoT signals with self-contained localization technologies, such as inertial navigation, may reduce the weather impact on positioning solutions.

7) **Signal Variations:** The fluctuation is an issue inherent to wireless signals, especially for those in indoor and urban areas. The LoRaWAN RSS variation can reach several dBm and over 10 dBm under static and dynamic motions, respectively [43]. Although signal variations cannot be eliminated, their effects can be mitigated through denoising methods such as averaging [166], wavelet [167], and ANN [168].

Many denoising methods assume the noise follows a Gaussian distribution. However, both LoRaWAN [43] and WiFi [68] RSS variations can follow either symmetric or asymmetric distributions. To model an irregular distribution, stochastic-signal analysis methods such as Allan variance [62] can be used.

C. **Base-Station-Related Errors**

Examples of localization error sources on the BS side include an insufficient number of BSs, poor geometry, BS location uncertainty, BS PLM-P uncertainty, and BS time synchronization errors.

1) **Number of BSs:** Public telecommunication and IoT BSs are mainly deployed for communication, instead of localization. Communication uses require signals from at least one BS, while localization needs signals from multiple BSs. Thus, the dependence on the number of BSs is a challenge for telecommunication- and IoT-signal based localization. To alleviate this issue, there are several approaches, such as using single-BS localization techniques, adding motion constraints, choosing localization algorithms that need fewer BSs, and integration with other sensors.

Single-BS localization techniques can be conducted by using various measurements, such as ToA, AoA, and RSS. Similar to multi-BS localization, ToA [169] and AoA [48] based localization may provide high-accuracy locations but require professional nodes for distance and angle measurements, respectively. RSS-based single-BS localization can also be implemented through either DB-M [170] or parametric-model-based methods [51]. Compared to multi-BS localization methods, single-BS localization has great potential to be used in existing telecommunication systems but faces new challenges, such as the difficulty to detect outliers in its results [51].

To mitigate this issue, other approaches are needed. First, adding motion constraints can reduce the requirement for localization-signal measurements. For example, adding a constant-height constraint can reduce one state to be estimated [91]. Meanwhile, coarse-localization algorithms (e.g., min-max, centroid, and proximity) require localization signals from fewer BSs. Furthermore, it is feasible to fuse wireless signal measurements with data from other sensors (e.g., inertial and vision sensors) for tightly-coupled localization [92].

2) **BS Geometry:** The wireless-localization performance is directly correlated with BS geometry. The research in [171] and [90] investigate BS location optimization from the indoor communication and multi-floor signal coverage perspective, respectively. In the localization field, poor BS geometry may lead to the problem of location ambiguity [115]. The relation between BS geometry and localization performance has been investigated [172]. Furthermore, there are indicators, such as DOP [10], to quantify the geometry. DOP can be used to predict the multilateration accuracy given BS locations. A small DOP value is a necessary condition for accurate multilateration.

On the other hand, a small DOP value is not a sufficient condition for accurate localization because DOP can reflect only the geometrical BS-node relation, instead of many other error sources, such as NLoS, multipath, and stochastic errors. Therefore, other approaches for BS location optimization have been presented. For example, the research in [173] combines DOP and a floor plan to address the problem. Meanwhile, there are methods for BS-geometry evaluation and optimization through CRLB analysis [174] and the Genetic algorithm [175]. The research in [176] analyzes the influence of BS geometry on indoor localization and has involved the impact of obstacles. According to the existing works, it is worthwhile to evaluate the BS geometry in advance and optimize it based on the actual localization environment.

In the future, more BSs will become available for localization due to the popularization of small IoT BSs. The existence of more BSs is generally beneficial for localization; however, too many BSs may increase the complexity of location estimation. For example, if several BSs are located nearby, a location ambiguity problem may occur around the region in which these BSs are located. Thus, it is also necessary to use BSs selectively based on their importance. The research in [118] quantifies the importance of each BS by the signal discrimination between distinct locations. Meanwhile, the paper [177] examines several BS-selection criteria, such as the Kullback-Leibler divergence, dissimilarity, maximum value, and entropy. The literature [178] adds more criteria, including the number of RPs, the average, and variance values. Moreover, there are BS-selection approaches through the use of theoretical-error analysis [179], and strong-signal detection [180]. Meanwhile, the research in [181] investigates optimal BS selection through a tradeoff between localization accuracy and energy consumption.

3) **BS Location Uncertainty:** Geometrical localization methods commonly need to know BS locations, a require-
ment that cannot be met in many IoT applications. To solve this issue, BS-localization approaches have been researched. The literature [89] reflects the idea of estimating the locations of BSs by using their distances to multiple RPs that have known locations. The research in [182] improves the accuracy of BS localization by using the RSS gradient. Meanwhile, there are improved BS-localization methods by using Fresnel-zone identification [183] and recursive partition [184]. The BS-localization methods have been applied in 5G mmWave systems [185], Ad-Hoc networks [94], and cooperative-localization systems [186]. Meanwhile, the literature [187] estimates BS PLM-P values before BS-location determination.

Another challenge is that the locations of IoT BSs may change, which causes localization errors. Thus, SLAM [188] and crowdsourcing-based [115] methods have been applied to estimate node and BS locations simultaneously. Meanwhile, there is research that involves the changes in BS locations in node localization [189].

4) BS Path-Loss Model Parameter Uncertainty: To achieve accurate RSS-based ranging and geometrical localization, BS PLM-P values are commonly estimated. The SLAM [188] and crowdsourcing-based [115] methods estimate PLM-P values simultaneously with BS locations. Meanwhile, there are other PLM-P estimation methods, such as those based on geometrical probability [190] and the finite-difference time-domain technique [191]. Also, there are improved estimation techniques, such as the closed-form weighted total least squares [192] and PF [193]. The papers [194] and [195] have derived the CRLB and hybrid CRLB models for PLM-P estimation errors, respectively.

Meanwhile, to enhance the ranging and localization performance, researchers have applied advanced PLMs, such as the third-order polynomial long-distance PLM [196], its combination with Gaussian models [197], and the models that involve signal degradation by walls, floors, and topological features [198]. Furthermore, specific localization signals (e.g., the Doppler effect [199] and direction detection Liu-W–2018) are introduced to improve PLM-P estimation.

5) BS Time-Synchronization Errors: The BS timing-synchronization error is a practical and important error source for time (e.g., ToA, TDoA, and RTT) based localization methods. The research in [188] investigates multiple wired and wireless time-synchronization approaches and the relation between BS time-synchronization errors and localization errors, while the research in [200] evaluates several wireless time-synchronization approaches for indoor devices. Also, the literature [201] examines the degradation in localization accuracy caused by time-synchronization errors through the CRLB analysis under Gaussian noise, while the paper [202] derives the CRLB for TDoA localization when multiple BSs subject to the same time-synchronization error.

D. Data-Related Errors

IoT localization systems also meet challenges at the data level. Examples of such challenges include those for database timeliness and training cost, RP location uncertainty, database outage, intensive data, data and computational loads, and the degradation of localization integrity in challenging environments.

1) Database Timeliness/Training Cost: To maintain localization accuracy in public areas, a periodical database update is required. The most widely used database-training method is to collect data at every RP. Such a method can improve database reliability by averaging LFs at each RP [203]. However, the static data-collection process becomes extremely time-consuming and labor-intensive when dense RPs are needed to cover a large area. To reduce time and manpower costs, dynamic-survey methods have been applied through the use of landmarks (e.g., corners and intersections with known positions and corridors with known orientations) on floor plans and a constant-speed assumption [204] or DR solutions [208]. To further reduce user intervention, there are other types of database-update methods based on crowdsourcing [205] or SLAM [206]. A challenge for crowdsourcing is to obtain robust RP positions, which is discussed in the next subsection.

2) RP Location Uncertainty: The uncertainty in RP locations will lead to drifts in localization databases. In particular, for dynamic-survey or crowdsourcing-based localization database training, it is difficult to assure the reliability of RP locations. DR can provide autonomous localization solutions [13]; however, it is challenging to obtain long-term accurate DR solutions with low-cost sensors because of the existence of sensor errors [114], the requirement for position and heading initialization, and the misalignment angles between the vehicle (e.g., the human body) and the node [14]. Thus, constraints are needed to correct for DR errors. Vehicle-motion constraints, such as Zero velocity UPdAte (ZUPT), Zero Angular Rate Updates (ZARU), and Non-Holonomic Constraints (NHC) [8], are typically adopted. However, these motion constraints are relative constraints, which can only mitigate the drifts of DR errors, instead of eliminating them. Absolute constraints, such as GNSS position [69] and user activity constraint [205], updates, can be used to ensure the quality of DR solutions in the long term. However, such position updates are not always available in indoor environments.

From the perspective of geospatial big data, only a small part of crowdsourced data is robust enough for updating databases. Thus, the challenge becomes how to select the crowdsourced data that have the most reliable DR solutions. Reference [69] presents a general framework for assessing sensor-data quality.

3) Database Outage: To obtain a reliable database, sufficient training data are required for all accessible areas. However, such a requirement is difficult to meet in wide-area applications [207]. If there is an outage of databases in certain areas, it will be impossible to locate the user correctly with traditional DB-M methods. To mitigate the database-outage issue, geometrical interpolation (e.g., Kriging interpolation [11]) and ML methods (e.g., GP [74]) have been presented for DB prediction, that is, predicting LFs at arbitrary locations based on training data at other locations. The research in [76] compares the performance of DB prediction by using GP and geometrical interpolation methods. Furthermore, as demonstrated in [66], the combination of geometrical and ML-
while, a method uses wireless signals for coarse localization
(WiFi, BLE, and RFID) signals for fine localization. Mean-
signals for coarse localization and use local wireless (e.g.,
proximity may be used to limit the search region for fine
localization strategy can be used. The coarse-localization
rate. Therefore, for wide-area IoT localization, a coarse-to-
node is localized to a place that is far from its actual position)
be large in wide-area IoT applications. This
higher accuracy and resolution.

4) Data and Computational Loads: The localization
database becomes large in wide-area IoT applications. This
phenomenon brings three challenges: large data load, large
computational load, and potentially large mismatch (i.e., the
node is localized to a place that is far from its actual position)
rate. Therefore, for wide-area IoT localization, a coarse-to-
fine localization strategy can be used. The coarse-localization
solutions from algorithms such as min-max, centroid, and proximity
may be used to limit the search region for fine
localization. A practical strategy is to use LPWAN or cellular
signals for coarse localization and use local wireless (e.g.,
WiFi, BLE, and RFID) signals for fine localization. Meanwhile,
a method uses wireless signals for coarse localization
and use magnetic measurements for fine localization.

Furthermore, region-based algorithms can be applied to
reduce the search space within the database. Examples of
these algorithms include region division and timing advanced-based algorithms. Region-based algorithms
can effectively reduce data and computational loads. On the
other hand, correct region detection and handover between
adjacent regions are key for region-based methods.

5) Localization Integrity: The integrity of localization
solutions, which describes the consistency between actual localization
errors and the estimated localization uncertainty, is even
more important than the localization accuracy. To be specific,
it may be difficult to provide accurate localization solutions
due to the limitation of the physical environment. At this time,
the localization system should be able to output an accurate
indicator of the location uncertainty, which makes it possible to reduce the weight of an inaccurate localization result.

However, most of the widely-used DB-M approaches do not have an indicator of the uncertainty of location outputs. Also, DB-M methods such as random forests and GP can only provide an indicator for relative probability (i.e., the probability of the selected RPs compared to other RPs), instead of the absolute location uncertainty. In contrast, geometrical methods have theoretical location-accuracy prediction indicators such as DOP [10], CRLB [93], and observability [212]. However, these indicators have not involved many real-world error sources, such as the environment-related and motion-related errors. This phenomenon leads to poor localization integrity. For example, smartphones may provide over-estimated GNSS location accuracy when a user stands at indoor window areas. The reason for this phenomenon is that the device can receive enough number of wireless signal measurements but has not detected the presence of environment-related errors such as multipath.

Thus, assuring location integrity for IoT systems is challenging but important. In particular, future IoT systems will have a higher requirement for scalable localization, which is less dependent on human perception and intervention. Thus, the existing localization solutions, which do not have an accuracy metric, will limit the promotion of localization techniques in IoT.

To predict the uncertainty of localization solutions, data from other localization sensors (e.g., inertial sensors) is used to detect unreliable wireless signal measurements through adaptive KFs [213]. Furthermore, to enhance localization-uncertainty prediction for wireless localization, mathematical model [69] and ML [65] based methods can be used. However, it is still a challenge to assure localization integrity in indoor and urban areas due to the complexity and unpredictability of localization environments.

E. Summary and Insight on IoT Error Sources and Mitigation

Different IoT applications may suffer from various degrees of localization errors. For example, professional IoT use cases commonly have limited application areas, which make it more straightforward to model and mitigate environment-related errors; also, the use of high-end sensors and in-the-lab sensor calibration can effectively reduce many end-device-related errors; meanwhile, the availability of specifically designed and deployed BSs can control BS-related errors; finally, the availability of powerful communication and computation hardware can alleviate the data-related errors. Therefore, it is more straightforward to promote localization techniques in professional IoT applications. By contrast, it is commonly not affordable for mass-market IoT applications to add a specific node or BS hardware or implement in-the-lab sensor calibration; meanwhile, the mass-market localization environment varies significantly; finally, only low-cost localization, communication, and processing sensors can be used. Thus, it is necessary to involve more error sources when designing localization algorithms for mass-market IoT applications.

In real-world localization scenarios, especially those for dynamic applications, the actual localization error is a combination of multiple error sources. Each error source may change in real-time. The complexity and diversification of actual IoT application scenarios have greatly increased the challenge to mitigate some errors (e.g., many environment-related errors). Thus, although the reviewed approaches can effectively reduce or eliminate the influence of some errors in some scenarios, it is challenging to mitigate other errors due to the physical environment limitations. To mitigate this issue, integrating IoT signals with data from other sensors such as inertial sensors is a feasible approach. IoT signals may provide long-term and wide-area absolute location solutions, while DR solutions from inertial sensors can provide short-term reliable and smooth relative location solutions. Also, DR solutions can bridge short-term outages and resist outliers in IoT signals.

Some error sources occur in both DB-M and geometrical localization methods. These error sources include the majority of end-device-related errors (e.g., device diversity, motion/attitude diversity, data loss/latency, and channel diversity), the minority of environment-related errors (e.g., wide-area effects, weather effects, and signal variations), the minority of BS-related errors (e.g., the number of BSs and geometry), and the minority of data-related errors (e.g., RP location uncertainty and localization integrity issues). By contrast, there are error sources that mainly exist in geometrical methods. These error sources include several main environment-related errors (e.g., multipath, NLoS, multi-floor effects, and human-body effects) and the majority of BS-related errors (e.g., BS location uncertainty, BS PLM-P uncertainty, and BS time-synchronization errors). The error sources that mainly exist in DB-M approaches include the majority of data-related errors (e.g., database timeliness/training cost, database outage, intensive data, and data and computational loads). Thus, in-depth knowledge of localization error sources is vital for selecting localization sensors and algorithms.

Although some factors (e.g., multipath, NLoS, and human-body effects) have been listed as localization error sources, they can also be used as valuable measurements to enhance localization. Furthermore, these factors can even be used as the key component for new localization approaches, such as device-free localization. This phenomenon can also partly explain why DB-M methods are commonly more suitable for complex environments (e.g., indoor environments). The main reason is that the existence of these factors is beneficial to DB-M.

V. IoT Localization-Performance Evaluation

This section describes the existing localization-performance evaluation approaches, including theoretical analysis, simulation analysis, in-the-lab testing, semi-field testing, and field testing. These approaches reflect the tradeoff between performance and cost. In general, this section answers the following questions: (1) what are the existing localization-performance evaluation methods; and (2) what are the advantages and limitations of these methods.

A. Theoretical analysis

There are various theoretical-analysis methods, including DOP, CRLB, and observability analysis. The existing research
on DOP has been described in Subsection IV-C2. The CRLB analysis methods have been applied on various localization applications, such as those for multilateration [45], hyperbolic positioning [46], multiangulation [49], DB-M [214], generic 3D localization [211], and the combination of multiple localization methods such as multilateration and multiangulation [211]. The influence of other factors, such as SNR [215], has also been investigated.

Besides DOP and CRLB, observability describes the ability to correctly estimating the states in a system [216]. With better observability, it is more straightforward to estimating a state in the localization filter. By contrast, an unobservable state cannot be estimated correctly even when the system does not have measurement noise [217]. Observability-analysis methods are commonly applied in inertial sensor-based localization [216] and have also been used in wireless positioning [212].

Theoretical-analysis methods have advantages such as: (1) they have a rigorous theoretical derivation. (2) They can be used to analyze not only a certain error source but also the relation between multiple error sources. (3) They have a low hardware complexity and algorithm computational load. (4) They can be directly transposed to other localization methods. On the other hand, the limitations of theoretical-analysis methods include: (1) it is difficult to involve actual localization error sources in theoretical analysis. Thus, theoretical-analysis methods can only provide the necessary conditions and a lower bound on performance. (2) Real-world localization scenarios such as the environmental and motion factors are difficult to model and derive.

B. Simulation Analysis

Simulation analysis is also a widely used localization-performance evaluation approach, especially for frontier applications (e.g., LPWAN and 5G) in which field-test data is difficult to obtain. Simulation analysis can be implemented in advance and its outputs can guide the design of subsequent tests such as in-the-lab and field testing. The simulation may be conducted based on localization software [218] or specific simulation platforms [219].

The simulation analysis methods have advantages such as: (1) they have a rigorous theoretical derivation. (2) They can be used to analyze not only a certain error source but also the relation between multiple error sources [218]. (3) They have a low hardware complexity and algorithm computational load. (4) They can be directly transposed to other localization methods. On the other hand, the limitations of simulation analysis methods include: (1) most of the existing simulation-analysis models are relying on simplified models and have not reflected complex environmental and motion factors. (2) It is difficult to reflect the actual localization situation through simulation.

C. In-The-Lab Testing

In-the-lab testing is a localization-performance evaluation method between simulation analysis and field testing. Compared to simulation, in-the-lab testing uses real hardware and sensors for data collection. Compared to real-field testing, in-the-lab testing is implemented in a controlled environment and is commonly affordable to use specific calibration and testing equipment such as a turntable [221] or shake table [222].

The advantages of in-the-lab testing include: (1) the localization sensor data are real. (2) It is feasible to test a certain type of error source (e.g., antenna radiation pattern [221]) in lab environment. However, in-the-lab testing methods have limitations such as (1) it may be difficult to reflect some error sources in lab environments due to the limitation of available equipment. Thus, it is still challenging to reflect real localization scenarios by in-the-lab testing. (2) In-the-lab testing may require specific equipment, which is not affordable for many low-cost IoT applications.

D. Semi-Field Testing

Signal-grafting, which is a typical semi-field testing approach, is a tradeoff between field testing and in-the-lab testing [226]. The idea of signal grafting is to add data from low-cost sensors to the data obtained from higher-end sensors to mimic low-cost sensor data. Specifically, real localization data are collected by taking high-end sensors and moving under real scenarios. Thus, both the localization environmental and motion factors are real; also, the data from aiding sensors (e.g., GNSS, magnetometers, and odometers) are real. Meanwhile, the data of low-cost sensors are collected in the lab and grafted into the field-tested high-end sensor data.

The advantages of signal grafting include: (1) it is significantly more cost-effective than field testing. In particular, it is not necessary to conduct a specific field test for each low-cost node. Only typical high-end sensor data in real localization scenarios and low-cost sensor data in the lab are needed. (2) It has a real localization environment, movement, and aiding information. Thus, it is closer to real localization situation than in-the-lab testing. On the other hand, its challenges include: (1) it is difficult to reflect the sensor errors that are correlated with real localization scenarios. (2) This method has not yet been applied in published IoT-signal-based localization studies.

E. Field Testing

The field-testing method has been widely used in IoT positioning applications, such as Sigfox and LoRaWAN DB-M [223], TDoA [225], and RSS localization [43]. In a field testing, the whole process from sensor selection to data processing is real. However, a challenge for field testing is the high cost, especially for wide-area applications. Meanwhile, for low-cost wide-area IoT applications, it is not affordable to carry out a field test that may cost much more than the actual IoT nodes.

There are use cases that apply LoRaWAN [223]. Sigfox [223], and NB-IoT [224] for city-level positioning. The research in [224] deployed 83 NB-IoT BSs in an area of approximately 40 km² in the city of Antwerp, Belgium. Thus, the BS density was around 500,000 m²/BS. In other words, there was one BS for a circle area that had a radius of 400 m. Such a BS density can partly represent the recent deployment density for city-level LPWAN or 5G.

The red triangles in Figure 9 (a) represent the location of BSs, while the blue dots are the location of mobile nodes.
The RSS data were used for positioning with three algorithms, including proximity, the range-based geometrical method, and fingerprinting. These methods respectively achieved the position accuracy of 340 m, 320 m, and 204 m in the mean error, and 294 m, 259 m, and 132 m in the median error. Figure 9 (b) shows the cumulative Distribution Function (CDF) of position errors.

In general, city-level IoT positioning is challenging due to the existence of almost all the error sources mentioned in this article. When the majority of other research is limited in theoretical analysis and simulation, the use cases in [225] and [224] provide valuable real-world practice and experience for this topic.

For local-area applications, the use case in [43] carried out RSS characterization and positioning using a private LoRaWAN network. As shown in Figure 10, 10 GWs were deployed in an area of 50,000 m². Thus, the BS density was 5,000 m²/BS. LoRaWAN positioning had been implemented using fingerprinting and compared with Android GPS locations on outdoor and indoor points. Figure 11 shows the location solutions on selected points. Meanwhile, Figure 12 demonstrates the CDF of position errors.

The Android GPS positions were smooth and reliable in outdoor tests (root mean square (RMS) 7.0 m, 90 % in 10.8 m) but degraded significantly indoors (RMS 40.6 m, 90 % in 52.9 m). Furthermore, half of the tested indoor points did not have a GPS solution. In comparison, LoRaWAN has similar position errors for the outdoor (RMS 27.8 m, 90 % in 35.3 m) and indoor (RMS 26.8 m, 90 % in 34.1 m) tests. The outcomes indicate the potential for using LoRaWAN RSS for continuous local-area localization in GNSS-challenging areas.

F. Summary and Insight on Localization-Performance Evaluation Methods

The evaluation of localization performance is a tradeoff between performance and cost. For example, from the perspective of reflecting real localization scenarios, the order of the methods from the best to the worst is field testing, signal grafting, in-the-lab testing, simulation analysis, and theoretical analysis. In contrast, from the cost-effectivity perspective, the opposite order prevails. In LE-IoT applications, especially the low-cost ones, it is preferred to implement localization-performance evaluation through the methods in the order of theoretical analysis, simulation analysis, in-the-lab testing, signal grafting, and finally field testing.

An important factor for in-the-lab, signal-grafting, and field-testing methods is the acquirement of location references. When higher-accuracy external position-fixing techniques (e.g., GNSS Real-time kinematic, optical tracking, vision localization, ultra-wide band, ultrasonic, and RFID) are available, their positioning results can be used as location references. Otherwise, manually selected landmark points may be used for evaluating the localization solutions when the node has passed these points. Meanwhile, the constant-speed assumption [204] or short-term DR solutions [208] may be used to bridge the gap between landmark points. In this case, it is important to assure the accuracy of landmark positions and motion-assumption/DR solutions.
VI. LOCALIZATION OPPORTUNITIES FROM LPWAN AND 5G

The newly-emerged LPWAN and 5G signals are bringing changes into the localization field. For example, they are expected to bring new technologies such as smaller BSs, smarter devices, mmWave, MIMO, the support for D2D communications, and device-centric architectures. These new technologies may bring opportunities and changes to IoT localization.

A. Cooperative Localization

The research of cooperative localization mainly covers two topics. First, there are cooperative-localization methods that use the connection between multiple nodes and localization sensors on each node. Meanwhile, there may be master nodes, which are equipped with higher-end sensors, and slave nodes, which have lower-end sensors. The theoretical model and accuracy bounds of non-cooperative and cooperative localization systems are needed.

Meanwhile, there are cooperative-localization approaches that use multiple sets of nodes (e.g., smartphones, smartwatches, and smart glasses) at various places on the human body. Algorithm constraints are required for robust localization using multiple devices on the same human body.

In the coming years, the characteristics of dense BSs and D2D communication capability may make it possible to provide accurate cooperative localization.

B. Machine Learning / Artificial Intelligence

Subsection III-A3 has demonstrated the use of ML methods in IoT localization. Furthermore, the paper [227] has illustrated some challenges for using ML in sensor networks and location-based services. Examples of these challenges include how to improve ML effectiveness under localization scenario changes and how to collect, transfer, and store massive localization data. It is expected that ML will be more widely and deeply used in IoT localization applications due to factors such as the popularization of IoT BSs and nodes, the emergence of geospatial big data, and the further development of ML platforms and algorithms. How to combine the existing geometrical and DB-M localization methods with the state-of-the-art ML techniques will be a significant direction for IoT localization.

C. Multi-Sensor Integration

Multi-sensor integration is becoming a mainstream technique for enhanced IoT localization. Technologies such as LPWAN, 5G, GNSS, WiFi, and BLE can provide long-term location updates. Meanwhile, inertial sensors and magnetometers can be used to provide attitude updates, which can be used for compensating for orientation diversity in IoT signals. Moreover, it is possible to obtain a DR solution by fusing data from inertial sensors, odometers, air-flow sensors, and vision sensors. Furthermore, to enhance localization performance, barometers and Digital Terrain Models can provide height or floor constraints, while road networks and floor plans can provide map constraints. Observability analysis can be applied to indicate the unobservable or weakly-observable states in the localization system; then, it becomes possible to determine which types of sensors can be added to enhance localization.

D. Motion Constraints

For many low-cost IoT applications, it is not affordable to enhance localization through adding extra sensor hardware. Therefore, the use of motion constraints within the algorithm has a great potential. The typical constraints including vehicle kinematic constraints (e.g., NHC, ZUPT, and ZARU), vehicle dynamic models (e.g., steering constraints, accelerating/braking constraints, multi-device constraints, aerodynamic forces, air-flow constraints, and path information), and control inputs. It is also possible to evaluate the contribution of algorithm constraints through methods such as observability and CRLB analyses.

E. Airborne-Land Integrated Localization

With the development of small satellite and Low Earth Orbits (LEO) communication technologies, it has become possible to extend IoT-signal coverage by using LEOs [232]. The paper [233] characterizes the performance of localization signals from LEOs.

Besides LEOs, enhancing localization signals from airborne platforms may also become a trend. It may be possible to use UAVs as BSs for IoT localization. There is great academic and industrial potential to integrate airborne- and land-based signals for enhanced IoT localization.

F. Multipath-Assisted Localization

Due to new features such as MIMO, dense miniaturized BS, and mmWave systems in 5G, using multipath signals to enhance localization, instead of reducing the multipath effect is attracting research interest. The research in [131] describes the principle and methodology for multipath-assisted localization, which has shown its potential to provide high-accuracy localization solutions.

G. Fog/Edge Computing

Fog and edge computing are being intensively researched in the IoT field. The paper [228] has described in detail the relation between IoT and fog/edge computing. However, the influence of fog/edge computing on localization has not been investigated. The research in [229] points out that there is a trend to use fog computing technology to achieve low-latency localization and location awareness solutions. On the other hand, an accurate IoT localization solution may contribute to the use of fog/edge computing. Therefore, the integration of IoT localization and fog/edge computing needs further investigation.
H. Blockchain

Blockchain is another newly-emerged technology that has gained wide attention in numerous fields. The paper [23] has reviewed the relation between blockchain and IoT but has not involved IoT localization. For localization, the research in [23] presents a blockchain-based geofencing and localization strategy. In general, the combination of blockchain and IoT localization is still at an early stage and thus requires further investigation.

VII. CONCLUSION

This paper reviews the IoT localization system through the sequence of IoT localization system review, localization data sources, localization algorithms, localization error sources and mitigation, localization performance evaluation, and new localization opportunities. Specifically, Section II overviews the IoT technologies, IoT localization applications, system architecture, and signal measurements. Section III demonstrates the state-of-the-art IoT localization methods. Afterwards, Section IV systematically reviews IoT localization error sources and mitigation. The localization errors are divided into four parts: (1) (e.g., device diversity, motion/attitude diversity, data loss/latency, and channel diversity), (2) (e.g., multipath, NLoS, wide-area effects, multi-floor effects, human-body effects, weather effects, and signal variations), (3) base-station-based errors (e.g., number of BSs, BS geometry, BS location uncertainty, BS PLM-P uncertainty, and BS time synchronization errors), and (4) data-based errors (e.g., database timeliness/training cost, RP location uncertainty, database outage, data and computational loads, and localization integrity). Then, Section V illustrates IoT localization performance evaluation methods. Finally, Section VI shows the possible localization opportunities.

In general, the emergence of LPWAN and 5G technologies have brought not only great advantages but also new challenges to localization applications. These technologies have attractive features such as long-range, low-power, and low-cost IoT signals, massive node connections, small and high-density BSs, the communication capacity. Therefore, it is worthwhile to conduct further research on exploring localization functionality for future LE-IoT systems.

REFERENCES

You Li (M’16) received Ph.D. degrees with the University of Calgary and Wuhan University in 2016. He is a Senior Researcher at the University of Calgary, and has been the R&D Lead at Appropolis Inc., Lead Scientist of EZRoad Ltd., and Algorithm Designer at Trusted Positioning Inc. (acquired by InvenSense Inc.). He has co-authored 70 academic papers and over 20 patents on location, motion, and related problems. He has won the EvAAL indoor localization competition and received four best paper awards from IEEE, ISPRS, or ION conferences. He serves as an Associate Editor of the IEEE Sensors Journal, the TPC of several international conferences such as MMT and SENSORETNETS, and a committee member of the ISPRS mobile mapping and IAG autonomous navigation for unmanned systems sessions.

Yuan Zhuang (M’16) is a professor at State Key Laboratory of Information Engineering in Surveying, Mapping and Remote Sensing, Wuhan University, China. He received the bachelor degree in information engineering from Southeast University, Nanjing, China, in 2008, the master degree in microelectronics and solid-state electronics from Southeast University, Nanjing, China, in 2011, and the Ph.D. degree in geomatics engineering from the University of Calgary, Canada, in 2015. His current research interests include multi-sensors integration, real-time location system, personal navigation system, wireless positioning, Internet of Things (IoT), and machine learning for navigation applications. To date, he has co-authored over 60 academic papers and 11 patents and has received over 10 academic awards. He is an associate editor of IEEE Access, the guest editor of the IEEE Internet of Things Journal and IEEE Access, and a reviewer of over 10 IEEE journals.

Xin Hu (M’16-SM’19) received the Ph.D. degree from the Institute of Electrics, Chinese Academy of Sciences, in 2012. He is currently an Associate Professor with the Information and Electronics Technology Lab, Beijing University of Posts and Telecommunications. He has been a visiting scholar at the Department of Electrical and Computer Engineering, University of Calgary. His research interests include smart signal processing, space and ground information integration, deep reinforcement learning (DRL), and communication/navigation integration.

Ling Pei (M’14-SM’19) received the Ph.D. degree in communication engineering and the Ph.D. degree in signal and information processing from Wuhan University, Wuhan, China, in 2007 and 2013, respectively. From 2010 to 2012, he was a co-trained Ph.D. Student with the National University of Singapore. He is currently an Associate Professor with the School of Data and Computer Science, Sun Yat-sen University, Guangzhou, China. His areas of interest include autonomous driving, robotics, artificial intelligence, where he has contributed more than 70 publications. He received the IEEE Vehicular Technology Society 2018 Best Land Transportation Paper Award, the IEEE Intelligent Vehicle Symposium 2018 Best Student Paper Award, and the Best Workshop Paper Award. He serves as an Associate Editor for the IEEE Transactions on Intelligent Transportation Systems, IEEE Transactions on Intelligent Vehicles and the IEEE Technical Committee on Cyber-Physical Systems newsletter. He serves as the Guest Editor for the IEEE Internet of Things Journal.

Jia Hu is a Senior Lecturer in Computer Science at the University of Exeter. He received his PhD in Computer Science from the University of Bradford, UK, in 2010, and M.Eng. and B.Eng degrees in Electronic Engineering from Huazhong University of Science and Technology, Wuhan, China, in 2006 and 2004, respectively. His research interests include edge-cloud computing, resource optimization, applied machine learning, and network security. He has published over 60 research papers within these areas in prestigious international journals and reputable international conferences. He serves on the editorial board of Elsevier Computers & Electrical Engineering and has guest-edited many special issues on major international journals (e.g., IEEE IoT Journal, Computer Networks, Ad Hoc Networks). He has served as General Chair/Co-Chair of IEEE CIT’15, ICCC’15, etc., and Program Chair/Co-Chair of IEEE ISPA’20, ScalCom’19, SmartCity’18, CYBCONF’17, EAI SmartGIFT 2016, etc. He has been the recipient of the Best Paper Awards at IEEE SOSE’16 and ICCC’14.

Zhouzheng Gao received his bachelor degree and master degree from China University of Geosciences Beijing, China in 2008 and 2012, and he received the PhD degree in school of Geodesy and Geomatics at Wuhan University in 2016. During 2014 and 2017, He worked in German Research Center for Geosciences (GFZ) in Potsdam, Germany, as a visiting scholar. Currently, he is an associate professor and doctoral supervisor in School of Land Science and Technology at China University of Geosciences Beijing. To date, he has six authorized software copyrights and publishes 30 papers. Currently, his research interest focusses on GNSS precise positioning algorithms (PPP and RTK), multi-sensor integration, and the corresponding applications.

Long Chen (M’14-SM’19) received the B.Sc. degree in communication engineering and the Ph.D. degree in signal and information processing from Wuhan University, Wuhan, China, in 2007 and 2013, respectively. From 2010 to 2012, he was a co-trained Ph.D. Student with the National University of Singapore. He is currently an Associate Professor with the School of Data and Computer Science, Sun Yat-sen University, Guangzhou, China. His areas of interest include autonomous driving, robotics, artificial intelligence, where he has contributed more than 70 publications. He received the IEEE Vehicular Technology Society 2018 Best Land Transportation Paper Award, the IEEE Intelligent Vehicle Symposium 2018 Best Student Paper Award, and the Best Workshop Paper Award. He serves as an Associate Editor for the IEEE Transactions on Intelligent Transportation Systems, IEEE Transactions on Intelligent Vehicles and the IEEE Technical Committee on Cyber-Physical Systems newsletter. He serves as the Guest Editor for the IEEE Internet of Things Journal.

You Li (M’16) received Ph.D. degrees with the University of Calgary and Wuhan University in 2016. He is a Senior Researcher at the University of Calgary, and has been the R&D Lead at Appropolis Inc., Lead Scientist of EZRoad Ltd., and Algorithm Designer at Trusted Positioning Inc. (acquired by InvenSense Inc.). He has co-authored 70 academic papers and over 20 patents on location, motion, and related problems. He has won the EvAAL indoor localization competition and received four best paper awards from IEEE, ISPRS, or ION conferences. He serves as an Associate Editor of the IEEE Sensors Journal, the TPC of several international conferences such as MMT and SENSORETNETS, and a committee member of the ISPRS mobile mapping and IAG autonomous navigation for unmanned systems sessions.

Yuan Zhuang (M’16) is a professor at State Key Laboratory of Information Engineering in Surveying, Mapping and Remote Sensing, Wuhan University, China. He received the bachelor degree in information engineering from Southeast University, Nanjing, China, in 2008, the master degree in microelectronics and solid-state electronics from Southeast University, Nanjing, China, in 2011, and the Ph.D. degree in geomatics engineering from the University of Calgary, Canada, in 2015. His current research interests include multi-sensors integration, real-time location system, personal navigation system, wireless positioning, Internet of Things (IoT), and machine learning for navigation applications. To date, he has co-authored over 60 academic papers and 11 patents and has received over 10 academic awards. He is an associate editor of IEEE Access, the guest editor of the IEEE Internet of Things Journal and IEEE Access, and a reviewer of over 10 IEEE journals.

Xin Hu (M’16-SM’19) received the Ph.D. degree from the Institute of Electrics, Chinese Academy of Sciences, in 2012. He is currently an Associate Professor with the Information and Electronics Technology Lab, Beijing University of Posts and Telecommunications. He has been a visiting scholar at the Department of Electrical and Computer Engineering, University of Calgary. His research interests include smart signal processing, space and ground information integration, deep reinforcement learning (DRL), and communication/navigation integration.

Ling Pei (M’14-SM’19) received the Ph.D. degree from Southeast University, Nanjing, China, in 2007. From 2007 to 2013, he was a Specialist Research Scientist with the Finnish Geospatial Research Institute. He is currently an Associate Professor with the School of Electronic Information and Electrical Engineering, Shanghai Jiao Tong University. He has authored or co-authored over 100 scientific papers. He is also an inventor of 24 patents and pending patents. His main research is in the areas of indoor/outdoor seamless positioning, ubiquitous computing, wireless positioning, Bio-inspired navigation, context-aware applications, location-based services, and navigation of unmanned systems. Dr. Pei was a recipient of the Shanghai Pujiang Talent in 2014.

Zhouzheng Gao received his bachelor degree and master degree from China University of Geosciences Beijing, China in 2008 and 2012, and he received the PhD degree in school of Geodesy and Geomatics at Wuhan University in 2016. During 2014 and 2017, He worked in German Research Center for Geosciences (GFZ) in Potsdam, Germany, as a visiting scholar. Currently, he is an associate professor and doctoral supervisor in School of Land Science and Technology at China University of Geosciences Beijing. To date, he has six authorized software copyrights and publishes 30 papers. Currently, his research interest focusses on GNSS precise positioning algorithms (PPP and RTK), multi-sensor integration, and the corresponding applications.
Kejie Chen is now an assistant professor of the Department of Earth and Space Sciences, Southern University of Science and Technology, China, and the focus of his current mainly lies in precise GNSS data processing and its geophysical application. Dr. Chen received his PhD degree from University of Potsdam, Germany in 2016. From May 2016 to May 2018, he worked as a postdoc at National Aeronautics and Space Administration’s Jet Propulsion Laboratory, where he had employed real-time GNSS data to build an operational tsunami early warning system. From May 2018 to September 2019, he was an assistant research scientist in Seismological Laboratory, California Institute of Technology, where he extended his knowledge from geodetic to geophysical field.

Maosong Wang received the B.S. degree from Harbin Engineering University in 2012, the M.S. degree from the National University of Defense Technology in 2014, and the Ph.D. degree from the National University of Defense Technology in 2018. From September 2016 to March 2018, he was a Visiting Student Researcher at the University of Calgary, Canada. Currently, he is a lecturer at the National University of Defense Technology. His research interests include inertial navigation algorithm, and multi-sensor integrated navigation theory and application.

Xiaoji Niu is currently a Professor with the GNSS Research Center, Wuhan University, China. He received the Bachelor’s and Ph.D. degrees from the Department of Precision Instruments, Tsinghua University, in 1997 and 2002, respectively. He did his Postdoctoral Research with the University of Calgary. He was a Senior Scientist with SiRF Technology Inc. He has published more than 100 academic papers and own 28 patents. He leads the Multi-Sensor Navigation Group, which focuses on GNSS/INS integrations, low-cost navigation sensor fusion, and its new applications.

Ruizhi Chen is the Director of the State Key Laboratory of Information Engineering in Surveying, Mapping and Remote Sensing, Wuhan University. He has been an Endowed Chair and Professor in Texas A & M University Corpus Christ, U.S. and Head & Professor of the Department of Navigation and Positioning at the Finnish Geodetic Institute, Finland. He has co-authored two books, 5 book chapters, and over 200 scientific papers. Dr. Chen is the general chair of the IEEE conferences “Ubiquitous Positioning, Indoor Navigation and Location-based Services”, Editor-in-Chief the J Global Positioning Systems and associate editor of the J Navigation. He was the President of the International Association of Chinese Professionals in Global Positioning Systems (2008) and board member of the Nordic Institute of Navigation (2009-2012). His research interests include smartphone positioning indoors/outdoors, context awareness and satellite navigation.

Fadhel M. Ghannouchi (S’84-M’88-SM’93-F’07) is a Professor, a Fellow of several institutions, including the Institution of Electrical and Electronics Engineers (IEEE), the Royal Society of Canada (RSC), the Engineering Institute of Canada (EIC), the Canadian Academy of Engineering (CAE), and the Institution of Engineering and Technology (IET), a Canada Research Chair of Green Radio Systems, and the Director of the iRadio Laboratory with the Department of Electrical and Computer Engineering, University of Calgary, Canada. He has experience in wireless communication, positioning, and navigation for over thirty years and have authored or co-authored over 700 refereed publications and 25 U.S. patents (5 pending), 6 books, and 3 spun-off companies in these fields.

Naser El-Sheimy is a Professor at the Department of Geomatics Engineering, the University of Calgary. He is a Fellow of the Canadian Academy of Engineering and the US Institute of Navigation, and a Tier-I Canada Research Chair in Geomatics Multi-sensor Systems. His research expertise includes Geomatics multi-sensor systems, GPS/INS integration, and mobile mapping systems. He is also the founder and CEO of Profound Positioning Inc. He published two books, 6 book chapters and over 450 papers in academic journals, conference and workshop proceedings, in which he has received over 30 paper awards. He supervised and graduated over 60 Masters and PhD students. He is the recipient of many national and international awards including the ASTech “Leadership in Alberta Technology” Award, and the Association of Professional Engineers, Geologists, and Geophysicists of Alberta (APEGA) Educational Excellence Award.

John Thompson (M’94-SM’13-F’16) is currently the Personal Chair in Signal Processing and Communications with the University of Edinburgh, U.K. His main research interests are in wireless communications, sensor signal processing and energy efficient communications networks, and smart grids. He has published around 300 papers in these topics. He was a recipient of the Highly Cited Researcher Award from Thomson Reuters in 2015 and 2016. He also currently leads the European Marie Curie Training Network ADVANTAGE, which trains 13 Ph.D. students in smart grid and grid topics for ComSoc in 2014 and 2015. He is also currently an Editor of the Green Series of IEEE Communications Magazine and an Associate Editor for the IEEE TRANSACTIONS ON GREEN COMMUNICATIONS ANDNETWORKS.