RESEARCH ARTICLE - ELECTRICAL ENGINEERING

Performance Analysis of Device Discovery Algorithms for D2D Communication

Omar Hayat[1,2] · Razali Ngah[2] · Siti Zaiton Mohd Hashim[3]

Abstract

The device-to-device (D2D) communication is a candidate technology to implement 5G standards commercially. To initiate D2D, device discovery is a primary issue and very few algorithms have been proposed for device discovery. A discovery algorithm has many parameters to discover the accurate position of the devices in walking and velocity scenarios. Due to rapid changes in the environment, LOS and NLOS algorithms become complex and accurate discovery ventures. Therefore, it is needed to evaluate the performance of the discovery algorithms. In this paper, a methodological approach is introduced for the performance evaluation of discovery algorithms. The performance evaluation for discovery estimation errors and complexity is evaluated using metrics and parameters, and analysis is made for range-based RSS technique using performance metrics. Discussion of performance evaluation metrics and criteria is analyzed followed by numerical/experimental, simulation models, and the parameters which affect performance and assessment. The metrics and criteria are defined in terms of a discovery signal success ratio, average residual energy, accuracy, and root-mean-square error (RMSE). Two differentiating discovery studies, Hamming and Cosine, are given and contrasted with reference RMSE for evaluation. This paper concludes with a discovery algorithm improvement cycle overview from simulation to implementation. It decreases discovery error $\in$ and enhances RMSE accuracy by an average of 21%. It also reduces the complexity of 12 pairs by Euclidean distance by 29%.

Keywords D2D communication · Device discovery · Performance analysis · Discovery algorithms · Root-mean-square errors (RMSE)

1 Introduction

In recent years, device-to-device (D2D) communication has received special attention as a candidate technology of 5G wireless communication. D2D communication empowers direct connection of discovery-based services and applications of proximal devices. It will improve spectrum usage, system throughput, and energy effectiveness. There are two potential possibilities generally for the communication between two distance-based proximal devices as presented in Fig. 1. First, if the devices are near to each other, they can communicate in two ways, directly or via the base station. Second, if the devices are extremely far apart, cellular infrastructure is used. Therefore, to initiate D2D communication, device discovery is a fundamental problem. There are two types of device discovery initiation procedure: autonomous discovery and network-assisted discovery. In an autonomous device discovery procedure, one device transmits a known reference signal (beacon) without coordination of network and uses a randomized procedure. The communication and discovery without coordination are usually time- and energy-consuming. So, the assistance of the existing system is used for device discovery procedure by coordinating frequency and time and their distribution for transmitting and receiving discovery signals [1]. The outcomes are energy efficiency, efficient resource management, interference control and mode selection using link qualities. So, network-assisted device discovery is considered in this research because it improves the device discovery process and has been proved in this work [2].

✉ Omar Hayat
  ohayat@numl.edu.pk

1 National University of Modern Languages (NUML), H-9, Islamabad, Pakistan

2 Wireless Communication Centre (WCC), School of Electrical Engineering, University Technology Malaysia, 81310 Johor Bahru, Malaysia

3 Big Data Centre, School of Computing, University Technology Malaysia, 81310 Johor Bahru, Malaysia

 Springer

Fig. 1 **a** D2D communication and cellular communication, **b** network-assisted device discovery for D2D communication
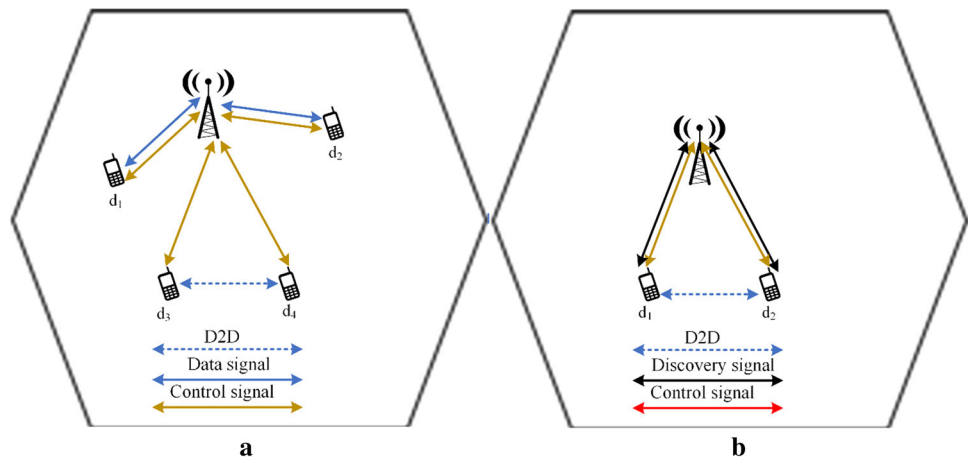
Many technologies for positioning and localization (discovery) have been implemented for out-band D2D communication, but the fine time estimation is quite difficult due to unlicensed spectrum [2–8]. Presently, in view of technological innovation, the device discovery for emergency devices can be committed as RF based, inertial measurement units (IMU) based and hybrid [6, 9]. The primary advantage of the RF-based discovery system is that it travels through obstacles, and therefore, network performance is not disturbed by device motion by walking and velocity. The device discovery performance can be enhanced by relaying, and the discovery process can be rehashed after relaying. However, it requires more than three base station (triangulation) estimations to discover the devices and suffers from NLOS, weather conditions and the unavailability of relay devices. IMU-based device discovery is another research area that depends on inertia and movement of sensors (3D magnetometer and barometer, 3D accelerometer and gyroscope) that create IMU. The upside of such discovery procedure is low cost, no additional infrastructure, discovery continuity, and the capability for indoor condition, while the disadvantages are error exponential, effective by velocity and need for an emergency incident commander. The hybrid may be a combination of RF and IMU based, but due to heterogeneity decisions and area conditions, accurate discovery is difficult. Therefore, by taking advantage of in-band RF based, discovery algorithm performance is evaluated.

Performance evaluation is essential for researchers, either for validation of new discovery algorithm against the preceding algorithm or while selecting the current algorithm which best fits the prerequisites of a given D2D application. However, there is an absence of unification in the D2D field due to new technology in terms of discovery algorithm assessment and correlation. Also, no standard methodology/technique exists to evaluate algorithm via simulation and emulation, modeling and real deployment [10]. Thus, it can be difficult to evaluate precisely how and under what con-

ditions one algorithm is superior to another. In addition, it can be difficult to choose what performance measures of discovery algorithms are to be looked at or assessed against. It is significant for the accomplishment of the subsequent implementation given that diverse applications will have divergent necessities. This is because the discovery algorithm is required to be utilized as a part of real applications and it is not definite to confirm their simulation performance. This research here contends that algorithms ought to be imitated and consequently implemented in equipment, in D2D-enabled environment, as part of the entire trial of their performance.

In 5G cellular, the proposed accuracy of discovery is in centimeters (cm) [11] and latency less than 10 ms. Therefore, the required accuracy and latency for in-band device discovery can be achievable because of control propagation characteristics, robust signaling, and high-resolution range. Both industry and academia are focusing on high accuracy of device discovery for different applications and scenarios. In this research, range-based (RF-based) device discovery is assessed and correlation matrix introduced to improve the accuracy. In all cases, performance metrics are evaluated for complexity, precision, and accuracy. In this paper, performance evaluation metrics are assessed together with three primary criteria, which are discovery signal success ratio (DSR), average E2E latency, and RMSE. Given that devices are normally constrained in terms of lifetime and per-device computational resources, tending to these requirements prompts the trade-off in the performance of the discovery algorithm. For instance, if boosting the discovery accuracy is the premier need, particular equipment must be added to every RF-based discovery system, expanding device size, price, and weight. On the other hand, if the device accessibility is already decided, then the application requires that performance criteria be adjusted. In this research, analysis of the range-based RSS error is evaluated using Euclidean distance, Hamming distance, and Cosine distance. In addition,

accuracy, precision, complexity, and cost of the proposed algorithm are evaluated.

The rest of the paper is organized as follows: Sect. 2 explains the related work for device discovery algorithms. The performance evaluation methodology is discussed in Sect. 3 with metrics and parameters, and evaluation of discovery error. Section 4 explains the analysis of the range-based RSS technique with performance metric using different scenarios. Discovery error estimation (DEE) using triangulation is derived in Sect. 5, and results and discussion are explained in Sect. 6. Section 7 concludes the paper.

2 Related Work

A D2D communication is broadly utilized as a part of numerous conditions to perform different monitoring tasks in IoTs. To carry out these tasks, different algorithms have been proposed to afford better precision regardless of whether anchor device density is low [12], while in high-density areas, the probability of anchor device is high. The anchor devices discovery benefits use today to enhance discovery ratio and the assignment of the anchor devices is to help other obscure devices to discover their location. To discover the unknown devices, range-based techniques are applied and angle of arrival (AOA) or distance/direction information (DOA) is required to determine the direction of obscure devices. However, while these techniques give high precision, additional equipment is required to find device coordinates. It also permits to route the discovery signal through relaying [13, 14]. So, in a D2D network, devices gather the information about surroundings and also share their observed data to the D2D-enabled base station. The base station holds the data in the database unit, where devices can access the data without searching the surroundings as presented by the model in Fig. 2. The discovery algorithms are applied to minimize the discovery errors and to maximize the precision.

The device discovery is an important problem for D2D communication and its applications [3, 15, 16]; however, not all applications require fine time determination of discovery. For such services and applications, even though discovery error is introduced by a discovery algorithm, this may not be injurious. But in 5G cellular network standardization, accuracy and time are the main parameters of discovery. Therefore, there are very few discovery algorithms proposed for device discovery based on the integration technologies proposed in 5G [17, 11]. These discovery algorithms are classified into the following groups: range based and range free, region based and connectivity based, and many more [5–8, 18]. It has been explained in the literature that different classes of algorithms introduce errors with different characteristics. Therefore, different types of errors are added up during the development of an algorithm [19, 20]. The

algorithm performance is compared with traditional techniques using device density and discovery error as explained in Table 1.

There are also many types of researches that have explored the different types of errors for D2D discovery [21, 22]. Most research considers the Euclidian distance between definite discovery and estimates discovery as an error metric to judge the accuracy discovery algorithms. The direct functions for Euclidian distance are also used to normalize the error values based on communication range [12, 23]. However, there is no significant research that contemplates the other metrics such as Hamming distance, Cosine distance, precision, complexity, and cost that are argued in this research, and it is sure that this paper is the first attempt to incorporate various other metrics to judge the accuracy of device discovery algorithm in D2D communication. We have proposed more than two algorithms [1, 2, 24] that will be also evaluated to measure the performances. Some metrics are popular and well known in other domains such as WSN, but not recognized in D2D communication. Besides these popular metrics, novel metrics that can fit D2D applications and will be tolerated according to applications are suggested, and therefore, this research can be considered as an alternative metric in discovery domain. Hence, this study is considered here as a unique and innovative contribution for performance evaluation of device discovery algorithms.

Among the extensive literature survey on device discovery, two performance indicators are very important, namely energy efficiency and discovery latency [25, 26]. Device discovery in a single cell and multicell and in dense areas is equally important. Energy efficiency strongly depends on the discovery latency, and discovery latency has statistical properties which may change in the different application context [27]. Very few researchers have done work on discovery latency. However, in some applications, initial discovery latency is the only concern while final discovery latency is the main concern in others. While it is attractive to have a high energy efficiency and a low discovery latency [28], it is not hard to perceive a trade-off between energy efficiency and discovery latency. A higher energy efficiency usually prompts a lower discovery latency. Therefore, how to adjust these two conflicting measurements becomes key. To shorten the discussion, a composite metric is proposed in [29], known as power-latency product, resulting in average energy consumption on worst discovery latency. Another important parameter is the discovery accuracy [30] to assess the performance of discovery procedure and algorithm. It demonstrates the difference between true value and the estimated value of discovery. Precision is also another parameter to calculate the reproducibility of progressive discovery measures [31]. This value can be utilized to evaluate the robustness of the discovery algorithm as it uncovers the variety of discovery
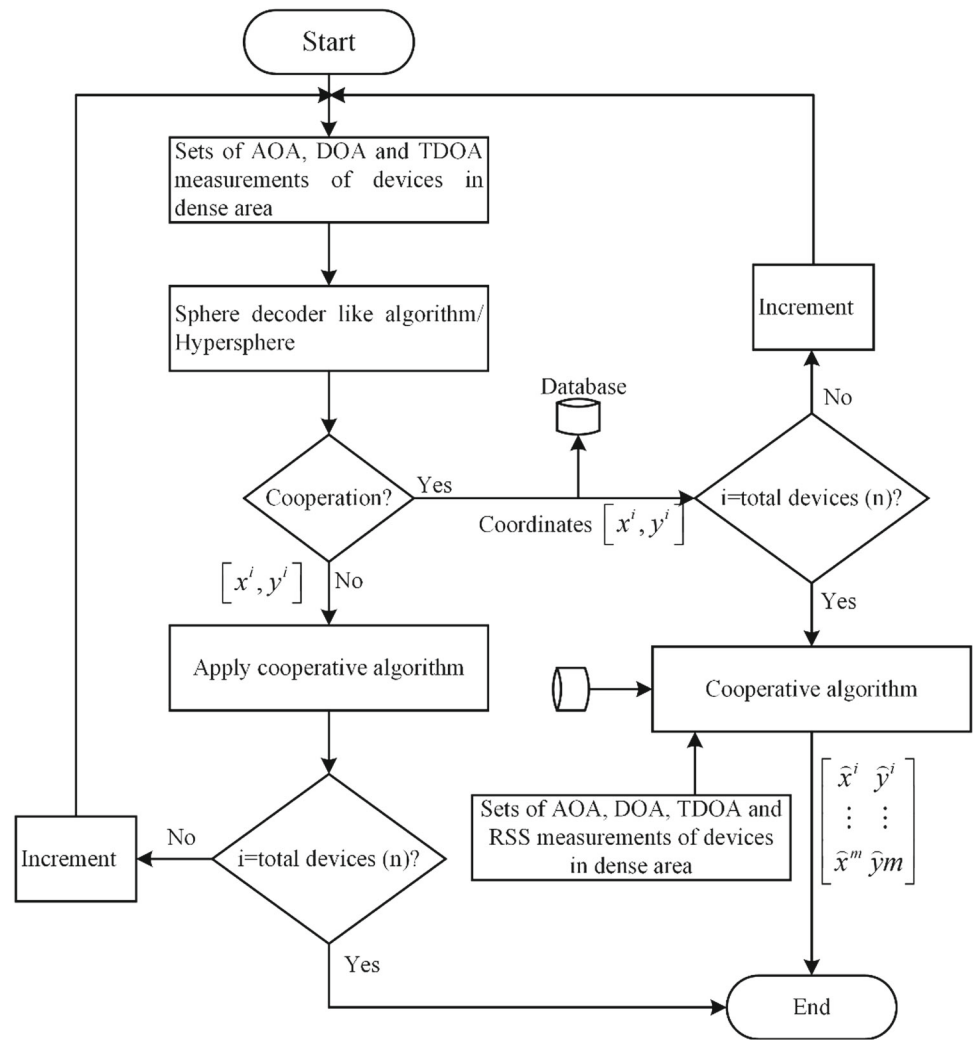
Fig. 2 Discovery algorithm model

Table 1 Comparison of the proposed algorithm

Metrics	Standard discovery algorithm based on			
	Centroid	Weighted centroid	DV hop	Improved DV hop
Equations	$X_{\text{est}} = \frac{\sum_{i=1}^{n} X_i}{n}$ $Y_{\text{est}} = \frac{\sum_{i=1}^{n} Y_i}{n}$	$X_{\text{est}} = \frac{\sum_{i=1}^{n} W_i X_i}{W_i}$ $Y_{\text{est}} = \frac{\sum_{i=1}^{n} W_i Y_i}{W_i}$	$W_i = \frac{1}{\text{Hop}_i}$	$W_i = \left(\frac{1}{\text{Hop}_i}\right)^{\frac{1}{n}}$
Complexity	Simple	Simple	Complex	Complex
Range based	Yes	Yes	No	No
Error rate	High	High	Low	Low
Accuracy	High	High	High	Low
Distribution	LOS	Uniform	Non-uniform	Uniform
Anchor devices density	High	High	Moderate	low

appraisals over several iterations. To calculate the precision, it is needed to initially find the median discovery of random devices in the first iteration. After that, Euclidean error of each estimated device using median position is computed [21].

3 Performance Evaluation Methodology

In this section, discovery model for multiuser and orthogonal frequency-division multiple-access (OFDMA) system and discovery using sphere decoder-like (SDL) algorithm [1, 2] in in-band 5G networks performance is evaluated for both moveable and static devices. The performance metrics
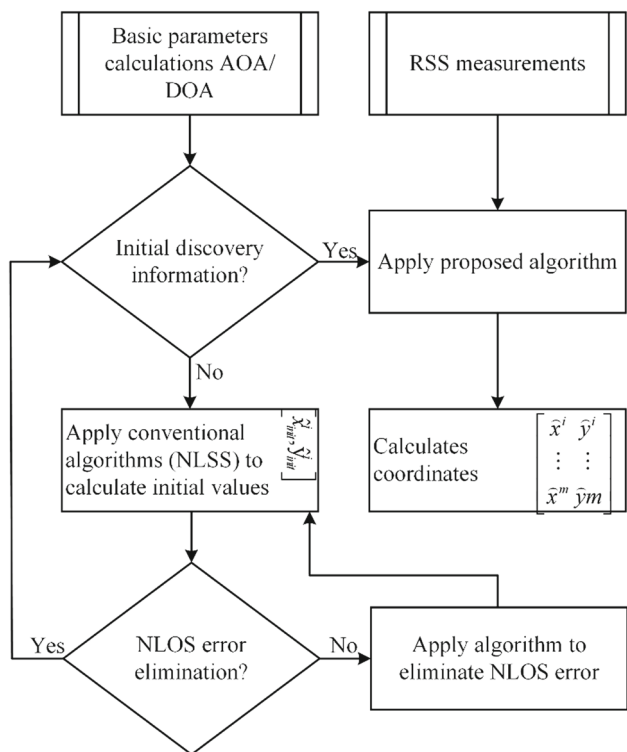
Fig. 3 Methodology for device discovery using the proposed algorithm for performance evaluation

Table 2 Simulation parameters

Parameters	Rate	
Coverage area	$100 \times 100 \text{ m}^2$	
Devices	10, 15, 20 and 25 pairs	
Mobility model	Haphazard walk, velocity	
Velocity	Haphazard walk	0.25, 1 and 1.5 m/s
	Velocity	10, 50 and 100 m/s

3.1 Metrics and Parameters

The performance evaluation algorithm is developed in MAT-LAB and evaluated in terms of discovery signal success ratio (DSR), end-to-end (E2E) latency, energy consumption, and signaling overhead by changing the parameters [33]. The important changing parameters are the number of devices and mobility (haphazard walk and velocity) as presented in Table 2. The main metrics are:

3.1.1 DSR (%)

It is the ratio of discovery signal delivered to discovery devices to discovery signal sent by all discoverer devices and calculated as:

$$DSR\% = \frac{\sum \text{all discovery signal received by all destination devices}}{\sum \text{all discovery signal sent by all discoverer devices}}$$
(1)

3.1.2 Average E2E Latency

It is a measurement of the taken time for delivery of discovery signal to the destination devices. It is measured from the first discovery signal to be recognized by the discovery device. It embraces propagation and hop delay, NLOS, and scattering. It is computed as

$$\text{E2E Latancy} = \frac{\sum_1^N \text{Time}_r - \text{Time}_s}{N}$$
(2)

where Time_r and Time_s are reception time and the sending time, respectively, and N is the number of successful discovery signals.

3.2 Evaluation of Discovery Errors

The D2D applications that insist on discovery (position information) of devices, which can be estimated by various discovery algorithms, unavoidably may present different kinds of error in their estimation. How an application is influenced by errors and a discovery error metrics reaction to errors may rely upon the error characteristics [30] as
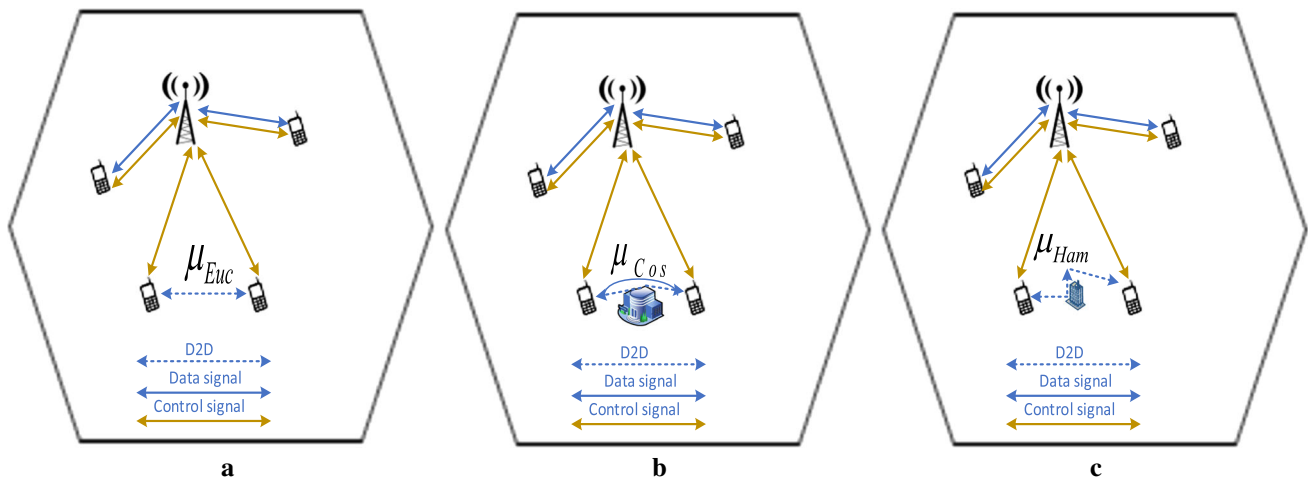
evaluation, simulation parameters, and simulation results are discussed. The RSS and DOA are applied to discover devices in areas of interest as presented in Fig. 3. The discovery performance can be assessed through least square, analytical modeling, Taylor series and (extended) Kalman filter [30]. The algorithms were assessed under many situations with various propagation conditions for static and moving devices and under the human body's impact. A NLOS error mitigation and identification algorithm was likewise used to enhance the ranging measurements. Performance metrics are defined which help in comparing device discovery algorithms. In this way, the performance of the proposed discovery algorithm under the diverse situations is compared in view of the accompanying measurements: precision, accuracy, and RMSE. The accuracy is expressed by the average distance error, and precision is characterized as the success probability of estimated discovery regarding accuracy using traditional methods. In any case, this methodology needs to give helpful information for an indoor discovery precision, since the precision is constantly connected with the accuracy and these measurements are autonomous. So, precision and accuracy are presented as a cumulative distribution function (CDF) and expressed in numerical value [32].

Fig. 4 **a** Euclidean distance, **b** Cosine distance, **c** Hamming distance

3.2.1 Accuracy

Accuracy is the most important parameter to assess the performance of discovery algorithm. It shows the difference between true and estimated discovery. This parameter is measured as the Euclidean error and defined as

$$D_{\text{Accuracy}}^2 = (x_{\text{Est}} - x_{\text{Actual}})^2 + (y_{\text{Est}} - y_{\text{Actual}})^2 \tag{3}$$

where x_{Est}, y_{Est} are estimated coordinates by the discovery algorithm and x_{Actual}, y_{Actual} are the true coordinates.

3.2.2 Precision

Precision calculates the reproducibility of progressive discovery measures. This value can be utilized to evaluate the robustness of the discovery algorithm as it uncovers the variety of discovery appraisals over several iterations. To calculate the precision, it is necessary to initially find the median discovery of random 50 (25 pairs) devices in first iteration. After that, Euclidean error is computed for each estimated device using median position as:

$$D_{\text{Precision}}^2 = (x_{\text{Est}} - x_{\text{median}})^2 + (y_{\text{Est}} - y_{\text{median}})^2 \tag{4}$$

3.2.3 RMSE

Unlike accuracy, using RMSE gives error for X and Y coordinates. The RMSE on coordinates can be calculated as

$$\text{RMSE}_i^2 = \frac{\sum (\text{Est}_i - \text{Actual}_i)^2}{\text{Number of Estimates}} \tag{5}$$

where i is the axis. The joint values of RMSE_X and RMSE_Y give net RMSE of the discovery algorithm, and the net RMSE can be calculated as:

$$(\text{Net RMSE})^2 = \text{RMSE}_X^2 + \text{RMSE}_Y^2 \tag{6}$$

Therefore, it is important to compute the correct error metric to assess the error performance of substitute discovery procedure that is conceivable to use for an application. To date, unfortunately, only shortsighted error metric relies upon the Euclidean distance between a base station or anchor device position and discovery device. To evaluate the performance of the proposed algorithm, two common metrics, Euclidian error and Hamming error, are used, and results are compared with Cosine error.

3.2.4 Euclidean Distance

It gives the shortest distance metric between true value and estimated value as shown in Fig. 4a and is calculated as:

$$d_{\text{Euc}} = \sqrt{\left(x_i - \hat{x}_i\right)^2 + \left(y_i - \hat{y}_i\right)^2} \tag{7}$$

$$\mu_{\text{Euc}} = \frac{1}{N} \sum_{i=1}^{N} d_{\text{Euc}} \tag{8}$$

where x_i and y_i are true value coordinates and $\hat{x}_i$ and $\hat{y}_i$ are estimated values of the device. N is the number of devices in search space and μ_{Euc} discovery error calculated by the discovery algorithm for all devices. In this metric, every device discovery and its estimates are thought of in isolation from other device discoveries and their estimates. Since this metric does not have the direction information of device in the network, it does not do well in NLOS conditions. This is why it may not be a decent metric in applications for which

evaluating relative discovery is more critical than evaluating absolute discovery.

3.2.5 Hamming Distance

Hamming error is another two-dimensional and prevalent metric. It is the measured distance of coordinates along axis at right angles as presented in Fig. 4c. It has finite field with x and y elements. The Hamming error between any two vectors x and y is $\mu_{\text{Ham}} = d(0112, 0122) = 1$ and usually satisfies conditions as

$$\mu_{\text{Ham}} = 0 \text{ only when } x = y \tag{9}$$

$$\mu_{\text{Ham}}(x, y) = \mu_{\text{Ham}}(y, x) \tag{10}$$

$$\mu_{\text{Ham}}(x, z) \leq \mu_{\text{Ham}}(y, x) + \mu_{\text{Ham}}(y, z) \text{ when } x, y, z \mathcal{E} F \tag{11}$$

Therefore,

$$d_{\text{Ham}}(x, y) = \left(\left| \hat{x}_i - x_i \right| + \left| \hat{y}_i - y_i \right| \right) \tag{12}$$

$$\mu_{\text{Ham}} = \frac{1}{N} \sum_{i=1}^{N} d_{\text{Ham}} \tag{13}$$

It does well even in dense areas, but due to binary constraints, its accuracy and complexity are not good.

3.2.6 Cosine Distance

It is a similarity measure metric that incorporates multiple values at one time instead of single value. It is used in data recovery domain. Device discovery domain considers two devices having two vectors $\vec{X}_{ij}$ and $\vec{X}_{ij}'$, actual and estimated, respectively. The Cosine values depend on θ between the vectors as presented in Fig. 4b. It is opposite to Euclidian matric and good for where directivity is needed [21]. It has range between -1 and +1. From Fig. 4b, the Cosine distance between two devices is calculated as $\frac{1-\cos\theta}{2}$. For multiple devices in dense area, the topology distance is computed as

$$\cos\theta = \frac{\vec{X}_{ij} \cdot \vec{X}_{ij}'}{\left| \vec{X}_{ij} \right| \left| \vec{X}_{ij}' \right|} \tag{14}$$

$$d_{\text{Cos}}(x, y) = \frac{1 - \cos\theta}{2} \tag{15}$$

$$\mu_{\text{Cos}} = \frac{2}{N(N-1)} \sum_{i=1}^{N} \sum_{j=i+1}^{N} d_{\text{Cos}} \tag{16}$$

Table 3 Overview of proposed error metrics

Error metrics	Proposed error metrics based on		
	Euclidean error	Hamming error	Cosine error
Equations	(8)	(13)	(16)
Complexity	Low	High	Low
Operating environment	Range based/free	Range based	Range based
Error rate	High	High	Low
Accuracy	Low	Medium	High
Energy consumption	Low	High	Low
Convergence time	Low	High	Low
Precision	High	Moderate	High
RMSE	Low	High	Low

It does well even in NLOS conditions due to its characteristics (14). Therefore, this error metric has been applied for performance evaluation in this research. A shortsighted overview of proposed error metrics is explained in Table 3.

4 Analysis of the Range-Based RSS Technique

The RSS presents the relationship between transmitted and received power of discovery signal with respect to the distance of discoverer:

$$P_r(\text{dB}) = \mathcal{A} - 10a \log \Delta \tag{17}$$

where P_R is the received power, Δ is the distance between discoverer and discovery and a is the propagation constant that is dependent on the environment. From (17), the values of $\mathcal{A}$ describe the association between the RSS and the distance of a discovery signal transmission. There are two RSS propagation models: log-normal shadow fading and free-space models. Log-normal shadow fading model is appropriate for both indoor and outdoor [34]. It is best suited for discovery signal due to its flexibility to different environmental conditions. On the other hand, free-space models have some advantages such as longer transmission distance than antenna size and carrier wavelength. In addition, free-space models are not affected as much by obstacles. The received power at distance Δ device is

$$P_R(\Delta) = \frac{P_T \mathcal{G}_T \mathcal{G}_R \lambda^2}{(4\pi\Delta)^2 \mathcal{L}} \tag{18}$$

$$\mathbb{P}_{\mathbb{L}}(\text{dB}) = 10 \log \frac{P_T}{P_R} = -20 \log \left[\frac{\lambda}{4\pi\Delta} \right] \tag{19}$$

where $\mathcal{G}_T$ and $\mathcal{G}_R$ are the antenna gains, $\mathcal{L}$ is the system loss, P_T is the transmitted power and (19) is the attenuation factor. For log-normal shadowing fading, it gives many parameters for different environments as

$$\mathbb{P}_{\mathbb{L}}(\Delta)(\text{dB}) = \overline{\mathbb{P}_{\mathbb{L}}}(\Delta) + \mathcal{X}_\sigma = \overline{\mathbb{P}_{\mathbb{L}}}(\Delta_0) + 10 \log a\left(\frac{\Delta}{\Delta_0}\right) + \mathcal{X}_\sigma \tag{20}$$

where Δ_0 is reference distance and depends on empirical values, a is path loss exponent and depends on propagation characteristics and $\mathcal{X}_\sigma$ is Gaussian random variable with zero mean. It also depends upon the frequency and power so,

$$P(\Delta) = P(\Delta_0) - 10a \log \frac{\Delta}{\Delta_0} + \mathcal{X}_\sigma \tag{21}$$

$P(\Delta)$ is received power for exact located devices at Δ and $\mathcal{X}_\sigma \sim N(0, \sigma^2)$. $P(\Delta_0)$ is the free-space path loss and relies upon frequency used by the discovery signal. Therefore, it is considered as frequency-dependent parameter [35]. Furthermore, path loss exponent depends on transmission frequency $a \propto f_{(\text{MHz})}$. In the RSS-based device discovery, signal propagation parameters are computed online or off-line. Online RSS measurements consume more radio resources than off-line, but when devices are moving, RSS updating is not possible. If σ and a are found precisely in any environment, then RSS discovery will be quite perfect. Moreover, the RSS pattern has been gathered from the experimental setup and RSS $(r_{m,n})$ database values are

$$\Phi = \begin{bmatrix} r_{1,1} & r_{1,2} & \dots & r_{1,n} \\ r_{2,1} & r_{2,2} & \dots & r_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ r_{m,1} & r_{m,2} & \dots & r_{m,n} \end{bmatrix} \tag{22}$$

where Φ is used as off-line and online training data. The PDF of gathered RSS data using path loss model can be composed as

$$f_{x,y}(\Phi) = \prod_{i=1}^{k} \frac{10}{\log 10 \sqrt{2\pi \sigma_i^2} \Phi} \exp\left[-\frac{\Phi}{4}\left(\log \frac{d_i}{\hat{d}}\right)^2\right] \tag{23}$$

where $\Phi = [r_{m,1} \dots r_{m,n}]$, $\hat{d} = d_0 \left(\frac{\text{RSS}(d_0)}{P_i}\right)^{\frac{1}{a_i}}$

Different measurements are used to quantify the performance of discovery techniques. The accuracy is not the only parameter to explain the performance of the algorithms. Alluding to the literature, most discussed performance evaluation parameters are complexity, robustness, accuracy, scalability, coverage, and cost. They are basically associated with thrifty or technical requirements, for example, equipment cost, stumpy battery power, and minimum computational complexity. The discovery accuracy (DA) is a very

important parameter for requirement of discovery system, and mean error is considered as performance metric. It is explained as follows:

$$\text{DA\%} = \frac{1}{k} \sum_{i=1}^{K} \frac{\hat{x}_i - x_i^2}{r_i^2} \tag{24}$$

where K is the number of devices to be discovered, d_i and $\hat{d}_i$ are true value and estimated value, respectively, and r_i is the RSS range of the device inside the network coverage. It is presented in percentage and normalized with coverage range. Cooperative and centralized discovery system gives more accurate discovery than distributed ones [1]. Clearly, the greater the accuracy, the better the discovery framework. However, it is trade-off between discovery estimation accuracy and other features as coverage, complexity, and many more. Therefore, a bargain between the required accuracy and other features is required. These features are coverage, complexity, scalability, robustness, and cost [36].

The device discovery depends on the network coverage and device transmission range. If the devices are out of coverage, then accurate device discovery is much difficult. Cooperative coverage will help to improve the discovery range. Complexity is attributed as software, hardware, and operation factors. Range-based discovery is much complex due to hardware involvement than range free. Discovery algorithm has also computing complexity. So, a centralized discovery process is preferred due to low complexity. Scalability is a measure in which right discovery is ensured when network coverage area is expanding or changing. A discovery system should scale as per network size, density, and dimensional space. Robustness is a measure of discovery stability even discovery signal is noisy or unavailable. In some cases, especially in indoor discovery, the discovery signal is blocked due to obstructions and NLOS condition. Therefore, some devices in the network could be uncertain. The cost of discovery system depends on software, hardware and weight, energy, and time. But the RSS-based device discovery does not need any extra hardware. To get better resolution, additional hardware is generally required that altogether increases the cost of every device and besides enhances the weight of the devices.

The development and evaluation of discovery algorithm cycle implies modeling, simulation and validation. Every cycle is characterized and validates precise feature of the algorithm. The algorithm is modeled based on the prescribed parameters of discovery such as RSS, DOA, and AOA. After modeling, simulation validates the algorithm under specific and simulated conditions, and this verifies the function of the algorithm. After simulated verification, the algorithm is applied to real applications. To evaluate the performance of the proposed algorithm, worst-case scenario is consid-

Fig. 5 Discovery time divisions
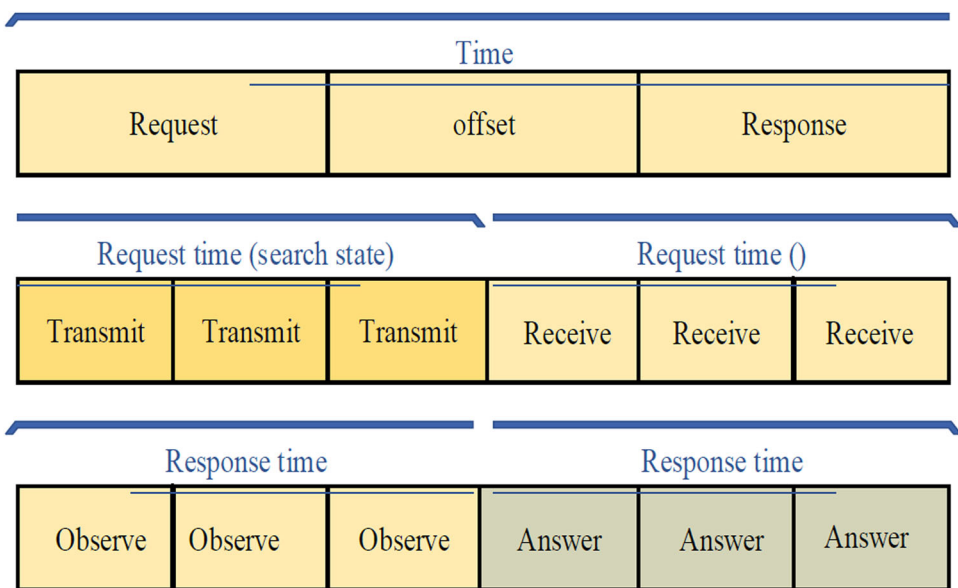
ered, when no device is discovered during discovery interval. It is divided into three probable groups: (1). A = no other devices receive when every device is transmitting its discovery request in transmission state, (2) B = one device receives when all other devices' discovery signals are transmitting, and (3) C = when all the devices answer, and N is the total number of devices in dense area. The discovery is apportioned into three states: request state, offset state, and response state, as presented in Fig. 5. The request has further two states, transmit and receive, while response time is converted to observe and answer. It will work as presented in flowchart shown in Fig. 6. When group A occurs, transmitter and receiver patterns overlap. Therefore,

$$P(A) = \frac{1}{2}\left(\frac{1}{2}\left(\frac{1}{(d(t_{\max}) - d(t_{\min}) + 1)}\right)\right)^{N-1} \quad (25)$$

When group B occurs, some devices choose the same discovery signal on whole discovery channel,

$$P(B) = \left[\sum_{M=2}^{N-1}\left(\left\langle\begin{matrix}x\\y\end{matrix}\right\rangle\left(\frac{1}{2}\right)^N\right) \times \left(\frac{1}{j}\right)^{M-1}\right]^X \quad (26)$$

where M is devices at the same frequency, N is total devices, and X is maximum discovery time and minimum discoverable time is 1. C case is the combination of A and B [32].

5 Discovery Error Estimation (DEE) Using Triangulation

RSS from more than two base stations creates a triangle as presented in Fig. 7. It enhances the discovery ratio and quality

with minimum error. It is relatively meek, since it associates the base station with strongest RSS as device discovery. One fundamental function in mobile devices is to discover the base station with strongest RSS. Accordingly, this technique can be accomplished without hardware enhancement of either device or base station. Therefore, to enhance the accuracy, algorithm for estimation and maximization of associated parameters is needed. The algorithm calculates the device discovery with location error. Regarding each base station, it often has different error values due to shadowing fading. With the RSS of the base station, it is conceivable to evaluate the distance between the device and the base station and devices can get their positions from several adjacent cells or base stations. This will transform the problem into the outstanding triangulation situating issue [37], which is depicted in Fig. 7. This issue can be portrayed as below:

$$x_m^2 + y_m^2 = d^2 \quad (27)$$

$$(x_m - x_{b1})^2 + (y_m - y_{b1})^2 = d_1^2 \quad (28)$$

$$(x_m - x_{b2})^2 + (y_m - y_{b2})^2 = d_2^2 \quad (29)$$

$$(x_m - x_{b3})^2 + (y_m - y_{b3})^2 = d_3^2 \quad (30)$$

Here, (x_{bi}, y_{bi}), $i \in [1, 2, 3 \ldots]$, is the location of base stations, $d_i^2 \in [1, 2, 3 \ldots]$ is the separation from the device to base stations, and (x_m, y_m) is the location of the device. Equation (27) has two obscure factors (x_m and y_m) and three different equations for three cells. If d_i can be gained precisely, the solution of equations will also be precise. The issue is that the distance d_i is considered with *RSS*. In the transmission, the signal would endure interference and shadow
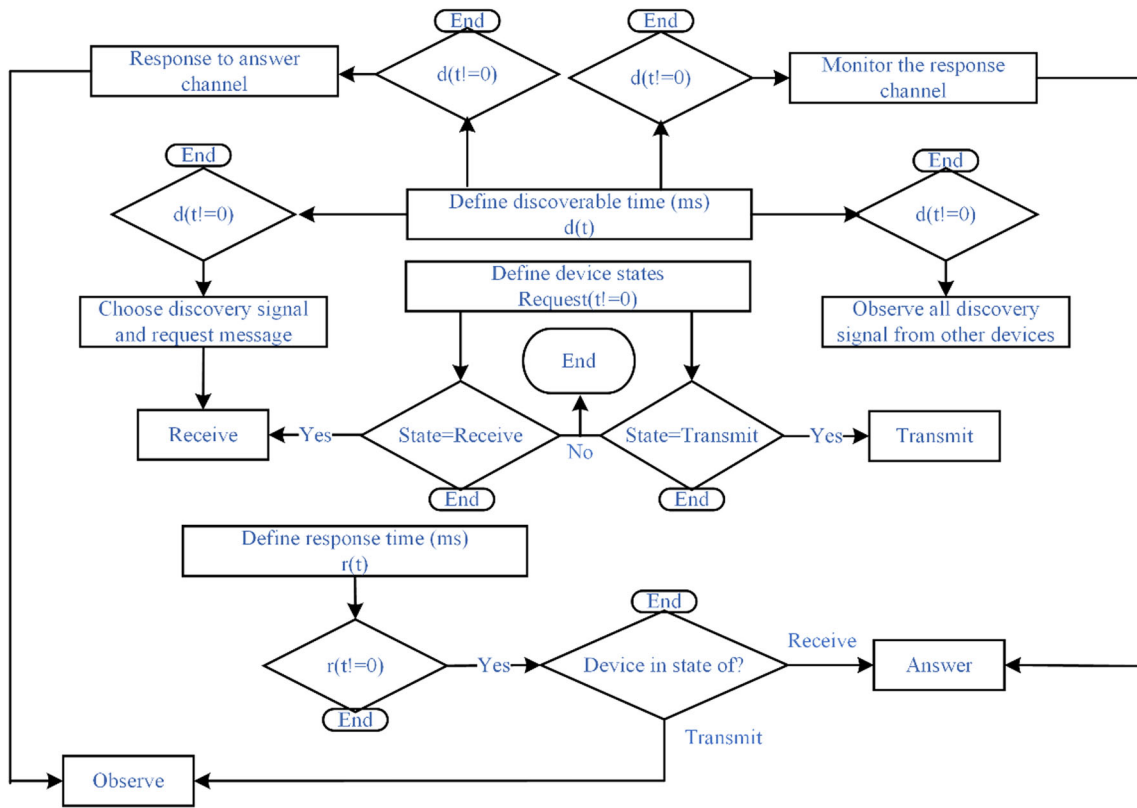
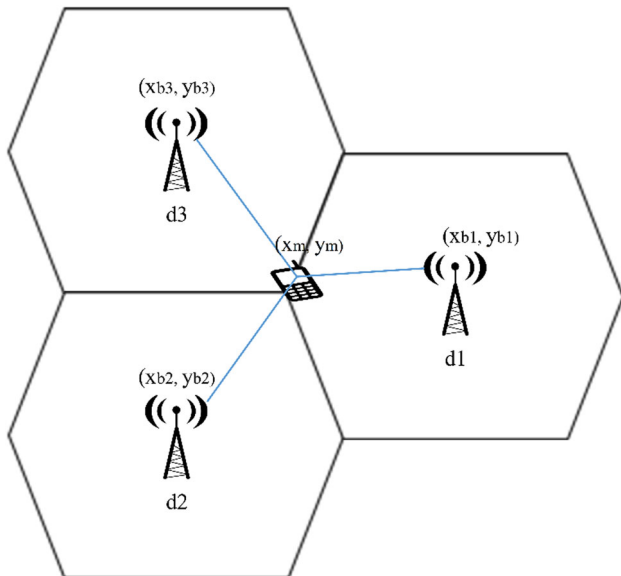**Fig. 6** Discovery signal transmission and reception with device states

Fig. 7 Triangulation positioning

fading, and d_i would not be correct distance between base station and device [38]. In this case, (28)–(30) will advance into

$$(x_m - x_{b1})^2 + (y_m - y_{b1})^2 = (d_1 + e_1(t))^2 \qquad (31)$$

$$(x_m - x_{b2})^2 + (y_m - y_{b2})^2 = (d_2 + e_2(t))^2 \qquad (32)$$

$$(x_m - x_{b3})^2 + (y_m - y_{b3})^2 = (d_3 + e_3(t))^2 \qquad (33)$$

where $e_i(t), i \in \{1, 2, 3 \ldots\}$, is distance error brought about by the shadow fading and interference which is a function of time. Equations (31)–(33) may not have an analytical solution. However, the numerical solution may solve it. The solution equations give the shortest distance from the base station and provide the device discovery solution. By combining from (27)–(30) and subtracting (27) from (28), (29) and (30) yield

$$x_{b1}^2 + y_{b1}^2 - 2x_m x_{b1} - 2y_m y_{b1} = d_1^2 - d^2 \qquad (34)$$

$$x_{b2}^2 + y_{b2}^2 - 2x_m x_{b2} - 2y_m y_{b2} = d_2^2 - d^2 \qquad (35)$$

$$x_{b3}^2 + y_{b3}^2 - 2x_m x_{b3} - 2y_m y_{b3} = d_3^2 - d^2 \qquad (36)$$

To solve the above equations, convert them into matrix form for LLME solution as

$$\underbrace{\begin{bmatrix} x_{b1} \ y_{b1} \\ x_{b2} \ y_{b2} \\ x_{b3} \ y_{b3} \end{bmatrix}}_{\Psi} \underbrace{\begin{bmatrix} x_m \\ y_m \end{bmatrix}}_{\alpha_i} = \frac{1}{2} \underbrace{\begin{bmatrix} d_1^2 - d^2 - k_1^2 \\ d_2^2 - d^2 - k_2^2 \\ d_3^2 - d^2 - k_3^2 \end{bmatrix}}_{k_i} \tag{37}$$

$$\begin{bmatrix} x_m \\ y_m \end{bmatrix} = \frac{1}{2} \begin{bmatrix} x_{b1} \ y_{b1} \\ x_{b2} \ y_{b2} \\ x_{b3} \ y_{b3} \end{bmatrix}^{-1} \begin{bmatrix} d_1^2 - d^2 - k_1^2 \\ d_2^2 - d^2 - k_2^2 \\ d_3^2 - d^2 - k_3^2 \end{bmatrix} \tag{38}$$

where $k_i^2 = x_{bi}^2 + y_{bi}^2$ in which x_{bi} and y_{bi} are the true positions of ith base station and x_m and y_m are the estimated device positions. If k_i represents the measured range and d_i represents the true range, d_i can be written in terms of k_i as $d_i = \alpha_i k_i$, where α_i is *NLOS* propagation and has values $0 < \alpha_i \leq 1$. The values α_i are limited in such a way that the *NLOS* error is a vast positive bias that makes the measured ranges to be more prominent than the exact ranges. From the RSS (r_i) measured values,

$$r_i = \alpha_i - 10\beta \log(d_i) + \eta_i \tag{39}$$

where β is path loss exponents and $\eta \sim \text{Norm}(0, \sigma)$. The *RSS* from ith base station and the distance can be calculated by

$$d_i = \log^{-1}\left(\frac{\alpha_i + \eta_i - r_i}{10\beta}\right) \tag{40}$$

$$\theta = \tan^{-1}\left(\frac{y_b - y_m}{x_b - x_m}\right) + \eta \tag{41}$$

User position is changing simultaneously due to the mobility, so it is necessary that every base station tracks the behavior of the device using environmental information. Therefore, device discovery procedure can be performed by every base station independently, and many intriguing practices arise when numerous base stations running a device discovery procedure coordinate to distinguish an entering user. The accessibility of more base stations reduces the discovery time because of the parallelism search. From (39), the α_i could be calculated using the pseudoinverse of Ψ as follows:

$$\hat{\alpha}_i = \frac{1}{2}\left(\Psi\Psi^{\mathbf{T}}\right)^{-1}\Psi k_i \tag{42}$$

6 Results and Discussion

There are many measures to evaluate the device discovery algorithms. These metrics are accuracy, precision, root-mean-square error, complexity and robustness against attacks. Other parameters that may help to evaluate the performance of discovery procedure are discovery signal

success ratio, average end-to-end latency, average residual energy and signaling overhead. From Eqs. (31) to (33) and (42), the e_i can be defined as

$$\text{e}_i = \|\alpha_i - \hat{\alpha}_i\|^2 \tag{43}$$

where e_i is the estimation error of ith devices, $\alpha_i = [x_m, y_m]^T$, which has ith device coordinates and is excessively complex. It is appealing to calculate and investigate discovery estimation errors (DEE) as

$$\bar{\text{e}}_i(m) = \mathbb{E}\{\text{e}_i\} \tag{44}$$

where $\bar{\text{e}}_i(m) = \text{E}\{\text{g}_i(m)\}$ and $\text{g}_i(m) = \frac{1}{N}\sum_1^m \text{e}_i$. $\text{g}_i(m)$ is the global discovery estimation error with the help of the base station. Suppose, in a dense area like stadium and shopping mall where there are a large number of devices, by applying the *law of large numbers* (*LLN*)

$$\text{g}_i(m) \xrightarrow{p \to 1} \bar{\text{e}}_i(m) \tag{45}$$

where $\longrightarrow p \to 1$ is the convergence probability toward one. $\bar{\text{e}}_i(m)$ is not only a statistical average, but it converges to any realization of $((x_m, y_m), m = 1, \ldots, N)$. To calculate average discovery estimation from (44),

$$\text{e}_i = \frac{1}{4}\|(\Psi\Psi^{\mathbf{T}})^{-1}\Psi \delta_i\|^2 \tag{46}$$

where Δ_i is the small changes and can be calculated as $\varsigma_1 - \varsigma_N, \ldots, \varsigma_{N-1} - \varsigma_N$ with $\varsigma_N = \hat{d}_{i-j}^2 - d_{i-j}^2$. Applying the trace on e_i gives

$$\text{e}_i = Tr((\Psi\Psi^{\mathbf{T}})^{-1}\Psi \delta_i \delta_i^T \Psi(\Psi\Psi^{\mathbf{T}})^{-1}) \tag{47}$$

$$\text{e}_i = Tr\left(\underbrace{\Psi(\Psi\Psi^{\mathbf{T}})^{-2}\Psi}_{\mathcal{W}} \delta_i \delta_i^T\right) \tag{48}$$

$$= \left(\sum_{i=1}^{N-1} \mathcal{W}_{jj}(\Delta_{ij})^2 + \sum_{i=1}^{N-1}\sum_{j=1, j\neq k}^{N-1} \mathcal{W}_{jk}(\Delta_{ij})(\Delta_{ik})\right) \tag{49}$$

$$E\{(\Delta_{ij})(\Delta_{ik})\} = \begin{cases} \sigma^2 & if \quad j=k \\ 0 & if \quad j\neq k \end{cases} \tag{50}$$

$$\text{e}_i = \sigma^2\left(2Tr(\mathcal{W}) + \sum_{i=1}^{N-1}\sum_{j=1, j\neq k}^{N-1}\mathcal{W}_{jk}\right) \tag{51}$$

Standard concepts of computational complexity of any algorithm in time and space can be utilized as correlation measurements for the relative cost of discovery algorithms. For instance, in dense areas, the discovery algorithm with $\mathcal{O}(n^3)$ complexity takes extensive time to converge as compared to $\mathcal{O}(n^2)$ and the same is valid for space complexity [39]. This is because, as the number of devices increases,
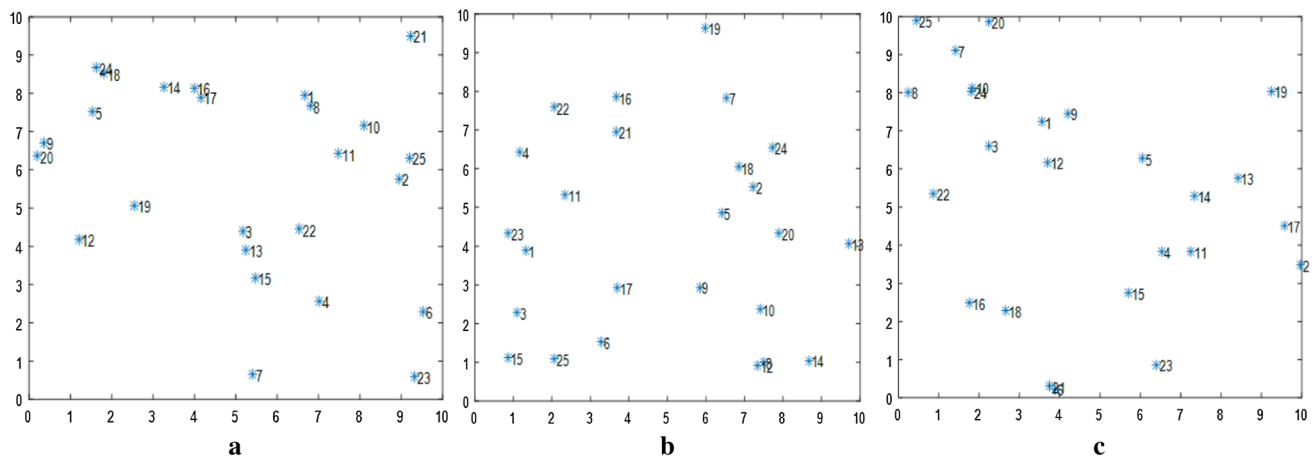
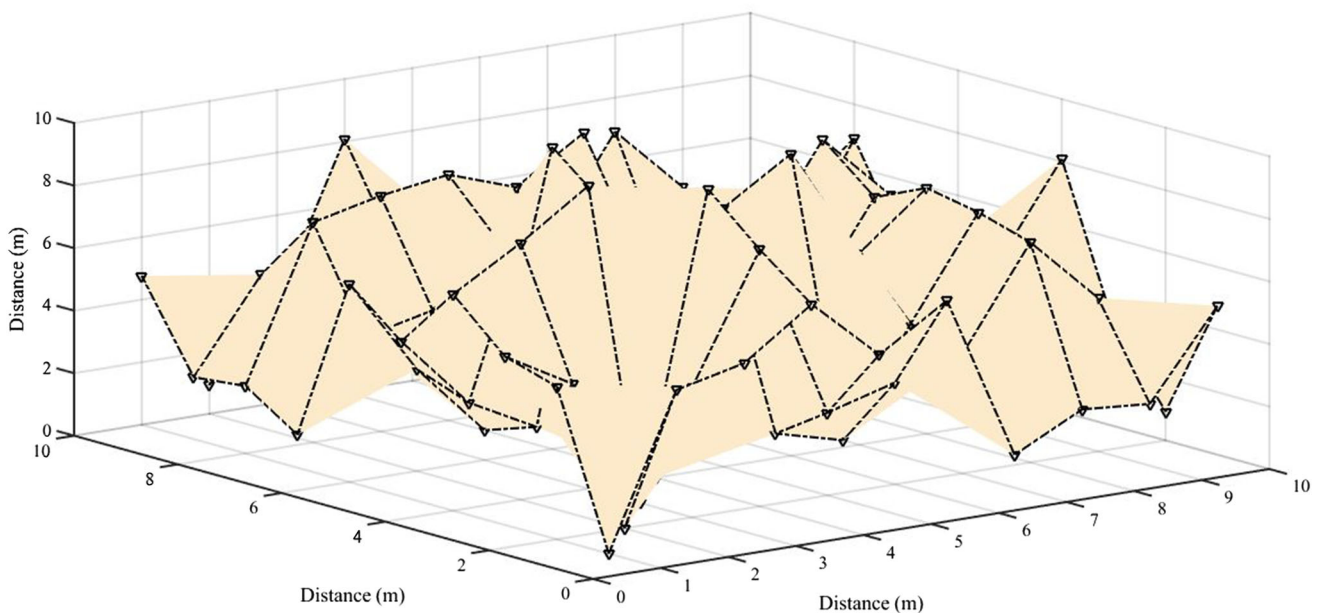**Fig. 8** **a** Cosine distance, **b** Hamming distance, **c** Euclidean distance

Fig. 9 3D view of Cosine distance of devices in dense areas

volume of memory (RAM) is required and will increment at a specific rate. The algorithm that requires less memory at a given scale might be ideal. This may help stimulate a trade-off between distributed and centralized algorithms; for example, centralized approach is better for dense areas.

To evaluate the discovery algorithms, the primary parameter is an evaluation of the discovery error, which is explained in Sect. 3. Further, it has three basic parameters, which are accuracy, precision, and mean square error. Accuracy defines the difference between estimated values and true values, while precision revolves around the median and gives a more robust accuracy of discovery algorithm. The mean square error gives the difference in X and Y coordinates and depends on the number of estimations. If estimations are large like in dense areas or in multicell, it gives small values. To calcu-

late these parameters, Cosine error metrics is applied and compared with Euclidean error metrics and Hamming error metrics. To compare these metrics, simulation parameters mentioned in Table 1 are used and results are presented in Fig. 8. The Cosine error metrics give more accurate discovery than Hamming and Euclidean error metrics. A 3D view of Cosine error metrics of discovered devices in dense area is presented in Fig. 9.

The discovery accuracy depends upon variance and the number of devices. Variance shows the difference of observation with each other. There are two states of variance: probability of discovery (p) and probability of no discovery ($1 - p$). By applying the binomial distribution to measure the variance, $n * p(1 - p)$, where n is the number of iterations dependent on device density. So, the proposed scheme
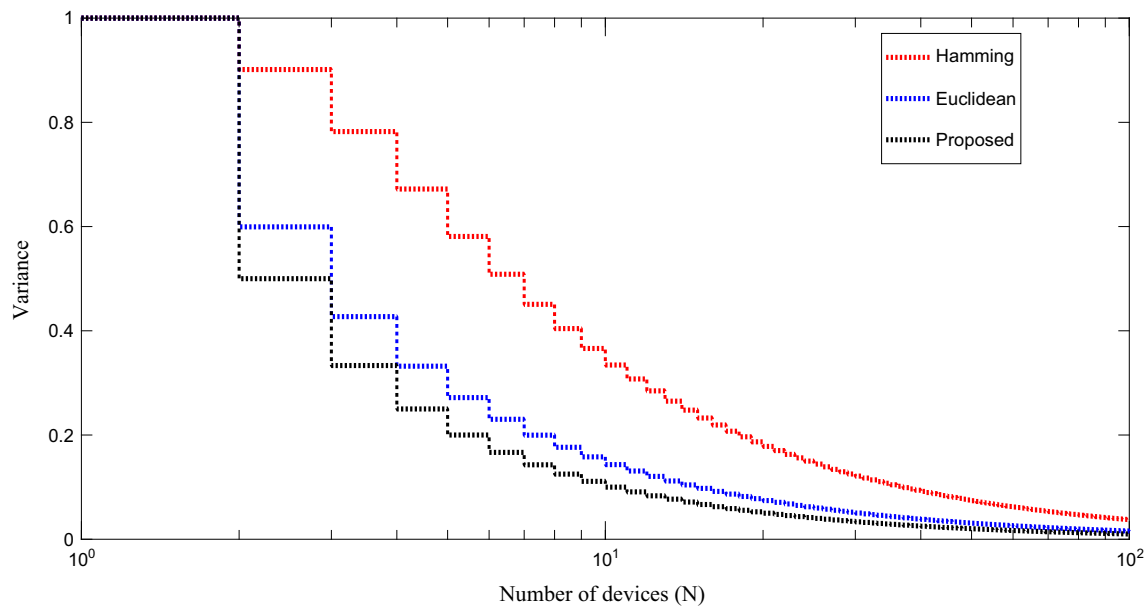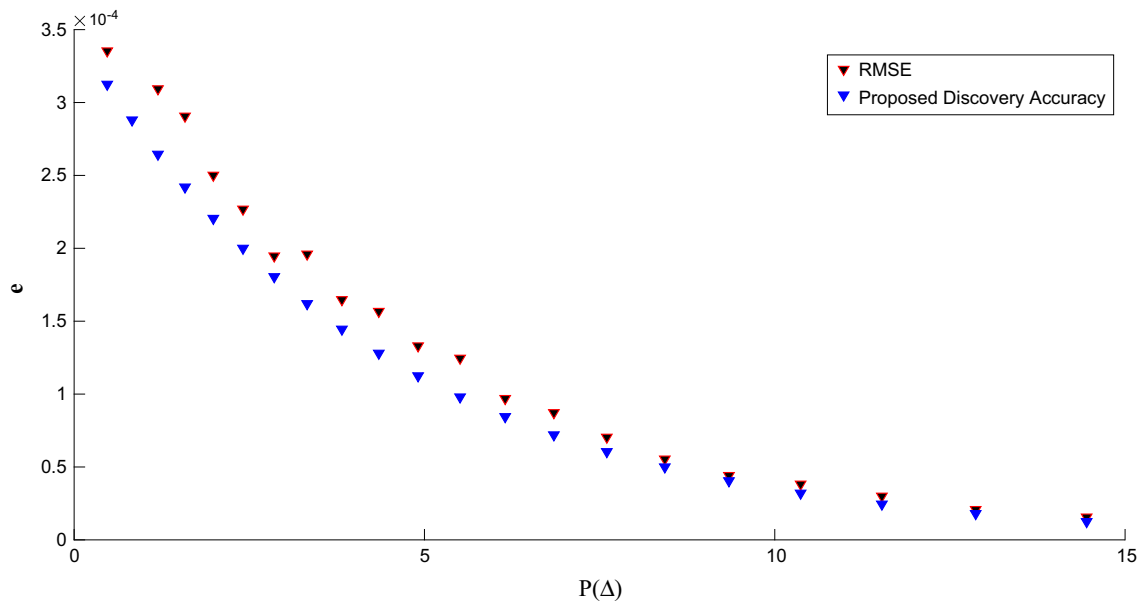
Fig. 10 Variance of proposed model

Fig. 11 Error performance of the proposed discovery algorithm

is also verified by the variance as presented in Fig. 10. The main parameter of discovery algorithm is discovery accuracy, and it represents the algorithm precision and accuracy. The accuracy error (in meters) depends on the transmission power (dBm). As power increases, e decreases toward zero and proposed method reduces accuracy of RMSE by 21% as proved in Fig. 11. In the last, the algorithm performance is assessed using algorithm complexity. Initially, it is supposed that the algorithm is complex and when Euclidean distance is applied, the complexity is reduced at 12 pairs of devices

to 11%, while at the same conditions, that of the proposed algorithm decreases to 29% as verified in Fig. 12.

7 Conclusion

In conclusion, the discovery algorithms' performance evaluation is not undervalued by the researchers. To entirely evaluate a discovery algorithm and its performance, it must be verified by simulation, imitation and realistic atmospheres. Although simulation is cheap and the most applied tool for
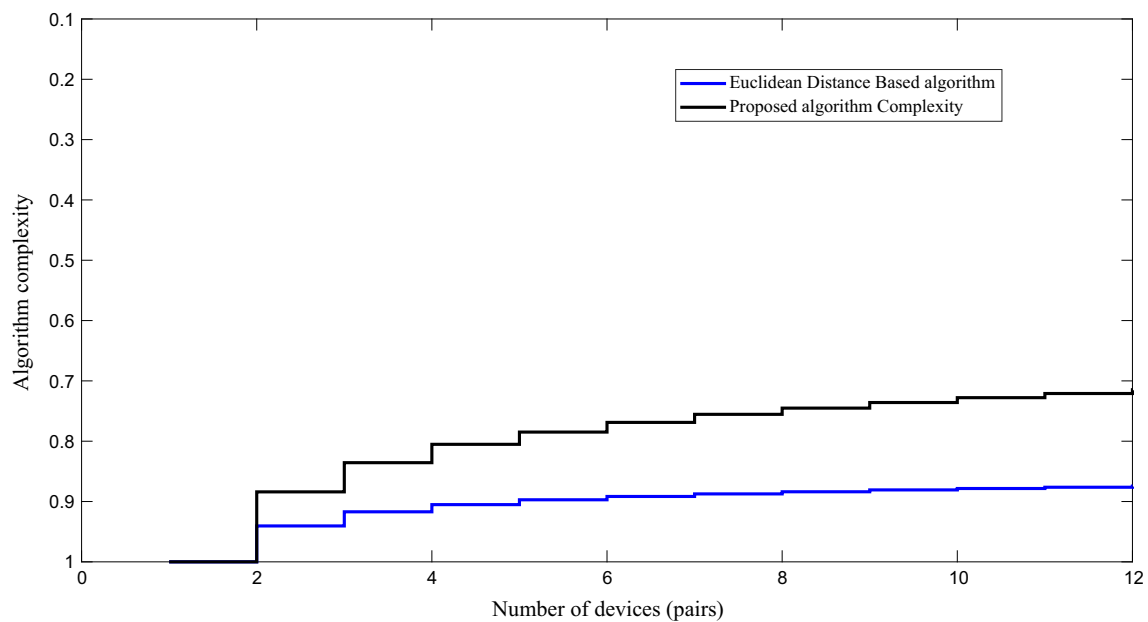
Fig. 12 Proposed algorithm complexity

algorithm evaluation, awareness of some constraints as RF communication and mobilization (walk and velocity model) is required. To fulfill the best fit for discovery algorithm, the design and advancement for new discovery algorithms require that attention be paid to trade-offs of accuracy, complexity, and scalability the discovery system is required to achieve. The parameter metrics are used to describe the discovery quality, and it is important for overall evaluation criteria, but most significant for accuracy evaluation. The Hamming and Euclidean error metrics are the weakest, but do not always measure if the discovery solution fits for ground truth. Therefore, equivalent metric, Cosine error metrics, is applied to measure the inter-devices distance estimates. It gives the significant performance for performance evaluation of discovery algorithms. This work can be extended for IoTs application, where small sensors are moving randomly.

References

1. Hayat, O.; Ngah, R.; Zahedi, Y.: Cooperative device-to-device discovery model for multiuser and OFDMA network base neighbour discovery in in-band 5G cellular networks. Wirel. Personal Commun. **97**(3), 4681–4695 (2017). **(in English)**
2. Hayat, O.; Ngah, R.; Zahedi, Y.: Device discovery for D2D communication in in-band cellular networks using sphere decoder like (SDL) algorithm. Eurasip J. Wirel. Commun. Netw. **2018**(1), 74 (2018). **(in English)**
3. Zhang, P.; Lu, J.; Wang, Y.; Wang, Q.: Cooperative localization in 5G networks: a survey. Ict Express **3**(1), 27–32 (2017). **(in English)**
4. Zhang, Z.F.; Wang, L.S.; Liu, D.; Zhang, Y.: Peer discovery for D2D communications based on social attribute and service attribute. J. Netw. Comput. Appl. **86**, 82–91 (2017). **(in English)**
5. Liu, Y.; Shen, Y.; Guo, D.N.; Win, M.Z.: Network localization and synchronization using full-duplex radios. IEEE Trans. Signal Process. **66**(3), 714–728 (2018). **(in English)**
6. Sun, Q.; Tian, Y;; Diao, M.: Cooperative localization algorithm based on hybrid topology architecture for multiple mobile robot system. IEEE Internet Things J. 1 (2018)
7. Trigui, I.; Affes, S.: Generalized Analysis and Optimization of D2D Communications in Cellular Networks (2018). arXiv preprint arXiv:1802.01618
8. Wu, D.; Zhu, D.L.; Liu, Y.L.; Zhao, D.M.: Location verification assisted by a moving obstacle for wireless sensor networks. IEEE Internet Things J. **5**(1), 322–335 (2018). **(in English)**
9. Frattasi, S.; Monti, M.: Ad-coop positioning system (ACPS): positioning for cooperative users in hybrid cellular ad-hoc networks. Trans. Emerg. Telecommun. Technol. **19**(8), 923–934 (2008)
10. Mao, G.: Localization Algorithms and Strategies for Wireless Sensor Networks: Monitoring and Surveillance Techniques for Target Tracking: Monitoring and Surveillance Techniques for Target Tracking. IGI Global (2009)
11. Orange, J. S.-B.; Armada, A.G.; Evans, B.; Galis, A.; Karl, H.: White Paper for Research Beyond 5G (Final Edit) (2016)
12. Chawla, G.; Khurana, M.: Performance and accuracy analysis of range-free localization algorithms in wireless sensor network. Int. J. Comput. Sci. Eng. Technol. (IJCSET) (Semantic Scholar) **7**(5), (2016)
13. Hayat, O.; Ngah, R.; Hashim, S.Z.M.: Swarm optimization based radio resource allocation for dense devices D2D communication. Int. J. Adv. Comput. Sci. Appl. **9**(6), 252–256 (2018)
14. Hayat, O.; Ngah, R.; Zahedi, Y.: In-band device to device (D2D) communication and device discovery: a survey. Wirel. Personal Commun. **106**(2), 451–472 (2019)
15. Poulakis, M.I.; Gotsis, A.G.; Alexiou, A.: Multicell device-to-device communication a spectrum-sharing and densification study. IEEE Veh. Technol. Mag. **13**(1), 85–96 (2018). **(in English)**
16. Gandotra, P.; Jha, R.K.: Device-to-device communication in cellular networks: a survey. J. Netw. Comput. Appl. **71**, 99–117 (2016). **(in English)**

17. Hayat, O.; Ngah, R.; Zahedi, Y.: Device discovery signal design for proximal devices in D2D communication. Wirel. Personal Commun. (2019). https://doi.org/10.1007/s11277-019-06435-6

18. Jaffry, S.; Hasan, S.F.; Gui, X.: Neighbourhood-aware out-of-network D2D discovery. Electron. Lett. **54**(8), 507–509 (2018)

19. Liu, J.Q.; Wan, J.F.; Wang, Q.R.; Deng, P.; Zhou, K.L.; Qiao, Y.P.: A survey on position-based routing for vehicular ad hoc networks. Telecommun. Syst. **62**(1), 15–30 (2016). **(in English)**

20. Zaidi, S.; El Assaf, A.; Affes, S.; Kandil, N.: Accurate range-free localization in multi-hop wireless sensor networks. IEEE Trans. Commun. **64**(9), 3886–3900 (2016). **(in English)**

21. Aksu, H.; Aksoy, D.; Korpeoglu, I.: A study of localization metrics: Evaluation of position errors in wireless sensor networks. Comput. Netw. **55**(15), 3562–3577 (2011). **(in English)**

22. Zhang, J.Y.; Deng, L.K.; Li, X.; Zhou, Y.C.; Liang, Y.N.; Liu, Y.: Novel device-to-device discovery scheme based on random back-off in LTE-advanced networks. IEEE Trans. Veh. Technol. **66**(12), 11404–11408 (2017). **(in English)**

23. Frattasi, S.; Della Rosa, F.: Mobile Positioning and Tracking: From Conventional to Cooperative Techniques. Wiley, Chichester (2017)

24. Hayat, O.; Ngah, R.; Zia, J.A.: Glowworm swarm based cooperative mobile device discovery for D2D communication in cellular networks. Sci. Int. (Lahore) **29**(6), 1207–1211 (2017)

25. Li, Y.; Sun, K.; Cai, L.: Cooperative device-to-device communication with network coding for machine type communication devices. IEEE Trans. Wirel. Commun. **17**(1), 296–309 (2018). **(in English)**

26. Shen, Y.Y.; Jiang, C.X.; Quek, T.Q.S.; Ren, Y.: Device-to-device-assisted communications in cellular networks: an energy efficient approach in downlink video sharing scenario. IEEE Trans. Wirel. Commun. **15**(2), 1575–1587 (2016). **(in English)**

27. Xenakis, D.; Kountouris, M.; Merakos, L.; Passas, N.; Verikoukis, C.: Performance analysis of network-assisted D2D discovery in random spatial networks. IEEE Trans. Wirel. Commun. **15**(8), 5695–5707 (2016). **(in English)**

28. Chour, H.; Nasser, Y.; Artail, H.; Kachouh, A.; Al-Dubai, A.: VANET aided D2D discovery: delay analysis and performance. IEEE Trans. Veh. Technol. **66**(9), 8059–8071 (2017). **(in English)**

29. Kandhalu, A.; Lakshmanan, K.; Rajkumar, R.R.; U-connect: a low-latency energy-efficient asynchronous neighbor discovery protocol. In: Proceedings of the 9th ACM/IEEE International Conference on Information Processing in Sensor Networks, pp. 350–361. ACM (2010)

30. Ferreira, A.G.; Fernandes, D.; Catarino, A.P.; Monteiro, J.L.: Performance analysis of ToA-based positioning algorithms for static and dynamic targets with low ranging measurements. Sensors (Basel) **17**(8), 1915 (2017)

31. Vargas-Rosales, C.; Mass-Sanchez, J.; Ruiz-Ibarra, E.; Torres-Roman, D.; Espinoza-Ruiz, A.: Performance evaluation of localization algorithms for WSNs. Int. J. Distrib. Sens. Netw. **11**(3), 493930 (2015). **(in English)**

32. Madhusudhan, S.; Jatadhar, P.; Reddy, P.D.K.: Performance evaluation of network-assisted device discovery for lte-based device to device communication system. J. Netw. Commun. Emerg. Technol. (JNCET), **6**(8). www.jncet.org (2016)

33. Hadir, A.; Zine-Dine, K.; Bakhouya, M.; El Kafi, J.; El Ouadghiri, D.: Performance evaluation of DV-hop localization algorithm for geographical routing in wireless sensor networks. Procedia Comput. Sci. **113**, 261–266 (2017)

34. Hussein, A.A.; Rahman, T.A.; Leow, C.Y.: Performance evaluation of localization accuracy for a log-normal shadow fading wireless sensor network under physical barrier attacks. Sensors **15**(12), 30545–30570 (2015). **(in English)**

35. Donlan, B.M.; McKinstry, D.R.; Buehrer, R.M.: The UWB indoor channel: large and small scale modeling. IEEE Trans. Wirel. Commun. **5**(10), 2863–2873 (2006). **(in English)**

36. Niewiadomska-Szynkiewicz, E.: Localization in wireless sensor networks: classification and evaluation of techniques. Int. J. Appl. Math. Comput. Sci. **22**(2), 281–297 (2012). **(in English)**

37. Liu, H.; Zhang, Y.; Su, X.; Li, X.; Xu, N.: Mobile localization based on received signal strength and Pearson's correlation coefficient. Int. J. Distrib. Sensor Netw. (2015)

38. Xing, B.; Gao, W.-J.: Innovative Computational Intelligence: A Rough Guide to 134 Clever Algorithms. Springer, New York (2014)

39. Zhu, F.; Wei, J.F.: Localization algorithm for large scale wireless sensor networks based on fast-SVM. Wirel. Pers. Commun. **95**(3), 1859–1875 (2017). **(in English)**