Attack-Resistant Received Signal Strength based Compressive Sensing Wireless Localization

JunYan[1], Kegen Yu[2], Yangqin Cao[3] and Liang Chen[4,5]
[1] College of Telecommunications and Information Engineering
Nanjing University of Posts and Telecommunications, Nanjing 210003, China
[e-mail: yanj@njupt.edu.cn]
[2] School of Geodesy & Geomatics and Collaborative Innovation Center for Geospatial Technology
Wuhan University, Wuhan 430079, China
[e-mail: kegen.yu@ieee.org]
[3] College of Telecommunications and Information Engineering
Nanjing University of Posts and Telecommunications, Nanjing 210003, China
[e-mail: caoyangqin0607@163.com]
[4] State Key Laboratory of Information Engineering in Surveying, Mapping and Remote Sensing
Wuhan University, Wuhan 430079, China
[5]Collaborative Innovation Center of Geospatial Technology (INNOGST)
Wuhan 430079, China
[e-mail:l.chen@whu.edu.cn]
*Corresponding author: Liang Chen

Received November 28, 2016; revised April 20, 2017; accepted May 25, 2017; published September 30, 2017

Abstract

In this paper a three-phase secure compressive sensing (CS) and received signal strength (RSS) based target localization approach is proposed to mitigate the effect of malicious node attack. RSS measurements are first arranged into a group of subsets where the same measurement can be included in multiple subsets. Intermediate target position estimates are then produced using individual subsets of RSS measurements and the CS technique. From the intermediate position estimates, the residual error vector and residual error square vector are formed. The least median of residual error square is utilized to define a verifier parameter. The selected residual error vector is utilized along with a threshold to determine whether a node or measurement is under attack. The final target positions are estimated by using only the attack-free measurements and the CS technique. Further, theoretical analysis is performed for parameter selection and computational complexity evaluation. Extensive simulation studies are carried out to demonstrate the advantage of the proposed CS-based secure localization approach over the existing algorithms.

Keywords: compressive sensing; secure wireless localization; received signal strength; attacked measurement; wireless network;

The work was supported by the National Natural Science Foundation of China (No. 61302103, 61372122, 61372123) and the Research Center of Optical Communications Engineering & Technology Jiangsu Province (ZSF0101).

https://doi.org/10.3837/tiis.2017.09.013 **ISSN : 1976-7277**

1. Introduction

Location based service (LBS) has recently drawn considerable attention due to the government regulations and commercial applications [1-6]. Location data can be utilized in a wide range of services such as navigation, tracking, health care monitoring, intelligent transport system (ITS) and access control [7]. In the literature, a variety of ranging methods and techniques has been proposed for wireless localization, including these based on received signal strength (RSS) [8], time of arrival (TOA) [9,10], time difference of arrival (TDOA) [11] and angle of arrival (AOA) [12]. The RSS based scheme is a cost-effective solution for wireless localization systems, since RSS measurement is normally available in a radio system or transceiver. On the other hand, approaches based on TOA, TDOA, and AOA typically require a more advanced radio receiver and processing capability. The AOA based approach is more costly because it requires multiple antennas or directional antennas, while TOA/TDOA based methods face some technical challenges in high-speed analog-to-digital conversion, time synchronization and coherent demodulation [13].

The localization method with RSS measurements can be classified into three broad categories, namely, range-based [14], fingerprint-based [15] and compressive sensing (CS) based methods [16] [17]. The range-based method uses a theoretical propagation model to estimate the distance between the emitter and the receiver and employs an iterative or non-iterative technique for location determination, while the fingerprint-based approach constructs the RSS-signature database and uses signature matching for location determination. In CS-based methods, the problem of wireless localization is formulated as the sparse signal recovery problem. The target positions can be estimated in the discrete spatial domain by solving an under-determined linear system. However, these existing techniques may fail in hostile environments where some of the nodes may be compromised by adversaries and/or used to transmit misleading information, aiming to prevent accurate localization of the sensors of interest. The adversary can attack the signal strength by attenuating or amplifying the signal strength at the receiver or at the transmitter. For example, an adversary could introduce an absorbing barrier between the transmitter and the target. When the signal propagates through the barrier, it is attenuated, and hence the target would observe much lower RSS. Thus, it is desirable to study the impact of these attacks on RSS based localization algorithms and explore methods to detect and further to eliminate the effect of these attacks [18]. Although efficient cryptography techniques are used to provide a layer of security for the localization system [19] [20] [21], there has been little study on the robustness of the RSS-based localization algorithms in the presence of malicious attacks.

Recently, with the development of the cooperative localization technique which enable the ranging and position estimation exchange between neighboring nodes [22], TOA based localization algorithms in attacked environments become more and more important. A performance limit of TOA based algorithm with ranging outliers was studied in [23]. A close-form Cramer-Rao lower bound (CRLB) approximation is proposed to describe the relationship among the localization accuracy, anchor's malfunctioning probability, ranging error of well functioning anchors, and the maximum communication range. There are two main secure mechanisms for RSS based localization in wireless networks: attack detection method and robust approach [24]. The basic idea of the attack detection method is to detect the attack-corrupted measurements and then exclude them from the position estimation process [25], [26]. However, the latter method exploits all RSS measurements for location

determination using a robust strategy such as assigning weights to the measurements [27], [28]. The attack detection methods make use of the RSS measurement inconsistency caused by malicious node attack to detect the inaccurate measurements. The drawback is that it can not guarantee a good detection performance. If undetected attack-corrupted measurements are used for target position estimation, the localization performance will degrade dramatically. On the other hand, the basic idea of robust method is to minimize the effect of the attacked measurements on localization accuracy. The main shortcoming of this method is its high computational complexity.

The RSS measurement vector formed at a group of access points (AP) changes as a target moves from one position to another. Also, the target population is usually much smaller than the number of discrete grids defined over the localization area. Thus, the localization problem can be formulated as a sparse signal recovery problem. The target positions can be estimated in the discrete spatial domain by some recovery methods with a limited number of RSS measurements. This is the basic idea of the CS-based localization approach which in general has a better accuracy and robustness when the sparse information is available [29]. Although the CS-based localization has been studied by a number of researchers, the security issue has not yet been investigated.

In this paper, we investigate the CS-based localization in presence of malicious node attack with an objective to propose an effective secure localization approach. The major contributions of the paper are two-fold. The first main contribution is to propose an effective attack detection based secure strategy for CS-based localization algorithm. Since a robust technique is utilized to produce more reliable verifier parameter estimation, the attack detection performance as well as the localization performance can be improved. The second main contribution is the theoretical analysis of the proposed approach. A guideline is first provided for the selection of the measurement subset size used for the intermediate position determination. It is then shown that the intermediate target position can be estimated accurately by utilizing the CS technique. Furthermore, for some given parameters, a relationship between the number of subsets and the probability that at least one subset does not contain any attacked measurement is derived which enables accurate verifier parameter estimation for malicious node detection so as to achieve improved attack detection performance. Also, the computational complexity of the proposed method is analyzed, supporting practical application in real-time environments.

The remainder of this paper is organized as follows. Section 2 provides an overview of the related work. The CS-based localization model with RSS measurements and the RSS attack model are described in Section 3. The proposed secure localization approach is presented in detail in Section 4. Simulation results are reported in Section 5. Finally, Section 6 draws our conclusion.

2. Related Work

In this section, we review some earlier research efforts toward CS based localization and RSS based secure localization. In [31], the authors proposed a novel localization protocol using the CS theory to reformulate the localization problem in wireless networks, leading to a theoretical CS-based localization framework. In [32], a rigorous proof for the necessity of Restricted Isometry Property (RIP) is provided and a comprehensive analysis for the choice of the grid size is conducted. In [16], a two-step CS-based indoor localization algorithm is proposed, which consists of a coarse localization by cluster matching and a fine CS-based localization. In the coarse localization step, the orthogonalization preprocessing procedure is

used to induce incoherence needed in the CS theory. In the fine localization step, the AP selection techniques are utilized to decrease the computational complexity and increase the accuracy. In [17], the application scenario is focused on mobile targets. Specifically, the localization problem is solved first by applying a proximity constraint to limit the distance between a coarse estimate of the current position and a previous estimate. Then, a CS-based scheme is applied to obtain a refined position estimate by a map-adaptive Kalman filter. In [33], a CS technique is applied to perform sparsity-based indoor localization, resulting in an energy-constrained algorithm which reduces the amount of information transmitted from a wireless device with limited power, storage and processing capabilities to a central server. In [34], a data processing technique is proposed for indoor localization using CS and fingerprinting. To mitigate the influence of large measurement noise, a sparse transformation model based on Gaussian kernel function is proposed to transform the location vector into a strictly sparse one. Besides, in order to reduce the high computational complexity, several fingerprinting space filtering algorithms are also exploited to remove some useless fingerprints in the radio map. From the above discussions, it is known that most of the existing works on CS-based localization focus on theoretical modeling and analysis [31], [32], satisfied condition for CS theory [16], [17], computational complexity [16], [34] and localization performance [16], [17], [34]. In practice, one of the significant challenges is that RSS measurements can be easily corrupted by malicious attacks. Despite the fact that most existing algorithms are based on the assumption of accurate RSS measurements, a number of methods have already been proposed in the literature to cope with malicious attacks.

In [35], two attack-resistant localization estimation techniques, attack-resistant minimum mean square estimation method and voting based location estimation method, are proposed to handle the malicious attacks on range-based localization in wireless sensor networks. In [28], an attack-resistant fingerprinting localization algorithm based on a probabilistic inclusive disjunction model is proposed to achieve more robust location estimation in malicious attack environments. A principle is proposed in [36] to design localization algorithms which are robust to signal strength attacks. The Ratio based Signal Strength Metric (RSSM) is proposed to perform robust wireless localization under the all-around signal strength attack. In [24], the necessary and sufficient conditions are established for secure distance-based localization in the presence of cheating beacon nodes and a class of algorithms that can always guarantee a bounded localization error is outlined. In [27], the adaptive least squares and least median squares (LMS) algorithms are utilized to maintain robustness in malicious node attack conditions for triangulation based localization. Meanwhile, a median-based nearest neighbor scheme that employs a median-based distance metric is proposed for fingerprinting-based location estimation. From the above discussions, we can conclude it can be seen that all the existing RSS based secure wireless localization strategies have been developed for the ranged-based [24] [27] [35][36] and fingerprint-based methods [27] [28] [36]. There has been no study on CS-based and RSS- based secure localization.

3. Problem Formulation

For the sake of simplicity and clarity, 2D localization is assumed and without loss of generality, the localization area is a rectangle that is divided into N equal grids which are the potential sites for targets. N is known in advance. When the target moves in a grid, the center of the grid is represented as the target position. There is only one possible target within a grid at any time. Furthermore, we assume the target population K is known and much smaller than N. There are M APs in the localization area and M is a known number and also much smaller

than *N*.

At every sampling instant, the *M* APs measure the power/strength of the signals transmitted from all *K* targets and the *M* RSS measurements, one from each AP, are forwarded to the data fusion center to determine the position of the *K* targets. The Euclidean distance between the *m*th AP and the target on the *n*th grid is given by

$$d_{m,n} = \sqrt{(x_m - x_n)^2 + (y_m - y_n)^2},\ 1 \le m \le M,\ 1 \le n \le N \tag{1}$$

where (x_m, y_m) and (x_n, y_n) are the coordinates of the *m*th AP and the center of the *n*th grid, respectively.

Under the realistic condition, the RSS measurements at each AP are often affected by obstructions, multipath propagation, and other environmental factors. According to the signal-fading model described in [37], when the signal is transmitted from the target at the *n*th grid, the RSS measurement at the *m*th AP is described as:

$$P_{m,n} = P_0 - 10n_p \lg(d_{m.n} / d_0) \tag{2}$$

where P_0 is the average receiving power at distance d_0, $d_{m,n}$ is the distance between the transmitter and receiver as defined in (1), and n_p is the attenuation exponent which depends on the propagation environment.

The total RSS measured at the *m*th AP is then given by

$$u_m = \sum_{n=1}^{N} P_{m,n}\vartheta_n + \varepsilon_m \tag{3}$$

where ε_m is the measurement noise including modeling error, ϑ_n is equal to one if a target is located in the *n*th grid; otherwise, it is zero. Equation (3) can be rewritten in a compact form as

$$\mathbf{u} = \mathbf{P}\boldsymbol{\theta} + \boldsymbol{\varepsilon} \tag{4}$$

where

$$\mathbf{u} = \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_{\mathrm{M}} \end{bmatrix},\ \mathbf{P} = \begin{bmatrix} P_{1,1} & P_{1,2} & \cdots & P_{1,N} \\ P_{2,1} & P_{2,2} & \cdots & P_{2,N} \\ \vdots & \vdots & \vdots & \vdots \\ P_{M,1} & P_{M.2} & \cdots & P_{M.N} \end{bmatrix},\ \boldsymbol{\theta} = \begin{bmatrix} \vartheta_1 \\ \vartheta_2 \\ \vdots \\ \vartheta_N \end{bmatrix},\ \boldsymbol{\varepsilon} = \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_M \end{bmatrix} \tag{5}$$

Here $\boldsymbol{\theta}$ is an unknown vector which has K ones and $N-K$ zeroes. In general, $N-K$ is much greater than K, i.e., $\boldsymbol{\theta}$ is a *K*-sparse vector. Therefore, the localization problem can be considered as a *K*-sparse signal recovery problem and can be handled using the CS theory. Note that $\mathbf{u}$ is the actual measurement vector, whereas $\mathbf{P}$ is the theoretical measurement matrix.

In [38], a linear relationship between the unattacked signal strength and the attacked signal strength for various propagation media is established. The linear relationship implies that there is an easy way for an adversary to perform and control the effect of an attack on the observed signal strength by appropriately adding different materials in the transmission channels. Due to the observed linear relationship, we refer to this as the linear attack model which can be described as

$$\tilde{u}_i = \alpha u_i + \beta \tag{6}$$

where $\tilde{u}_i$ and u_i are respectively the attacked and unattacked measurements observed at the *i*th AP node, α and β are the two attack parameters. In the absence of any attack, $\alpha = 1$ and $\beta = 0$. When $\alpha > 1$ and $\beta = 0$, the received signal power at the specific AP node is

amplified; otherwise, when $\alpha < 1$ and $\beta = 0$ the received signal power is attenuated. Clearly, the more α deviates from one, the more severe the attack is. Although there exist different signal strength attack models, the linear model in (6) has the advantage of simplicity in theoretical analysis.

4. Proposed Algorithm

The block diagram of the proposed CS-based secure localization approach is illustrated in **Fig. 1**. The approach contains three main phases: (1) verifier parameter estimation; (2) malicious node detection and elimination; and (3) final target position estimation. More details of the proposed approach are provided below.

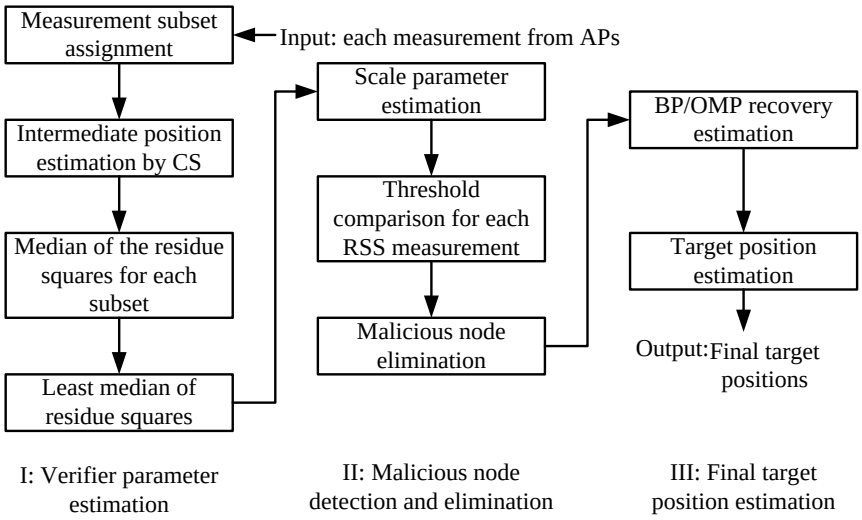

Fig. 1. Block diagram of the proposed secure CS localization scheme

4.1. Verifier Parameter Definition and Determination

Define the actual attack-corrupted received signal strength measurement vector as $\tilde{\mathbf{u}} = [\tilde{u}_1, \tilde{u}_2, \cdots, \tilde{u}_M]^T$. The measurements are arranged into Q measurement subsets, each of which consists of M_s measurements that are selected from M received measurements. Note that the number of ways of choosing M_s out of M measurements is equal to the binominal coefficient $C_M^{M_s}$. For example, when M=64 and M_s=35, the binominal coefficient $C_M^{M_s}$ is equal to 1.39×10^{18}. Therefore, for a small number of measurement subsets Q, it can hardly obtain two identical measurement subsets. Let us define the ith subset as

$$\mathbf{v_i} = [v_{i,1}, v_{i,2}, \cdots, v_{i,M_s}]^T, \quad i = 1, 2, \cdots, Q \tag{7}$$

where $v_{i,j} \in \{\tilde{u}_1, \tilde{u}_2, \cdots, \tilde{u}_M\}$, $j = 1, 2, \cdots, M_s$, $M_s \le M$. As mentioned earlier, each subset is exploited to produce an intermediate target position estimate and both the ℓ_0-norm minimization and ℓ_1-norm minimization can be utilized to perform the localization. Since the ℓ_0-norm minimization is a NP-hard problem with a much higher computational complexity,

in this paper, only the ℓ_1-norm minimization is used to obtain an estimate of the sparse vector $\boldsymbol{\theta}_i$, namely,

$$\hat{\boldsymbol{\theta}}_i = [\hat{\vartheta}_{i,1}, \hat{\vartheta}_{i,2}, \cdots \hat{\vartheta}_{i,N}]^T = \arg\min_{\boldsymbol{\theta}\in R^N} \|\boldsymbol{\theta}_i\|_1 \quad \text{s.t. } \mathbf{v}_i = \tilde{\mathbf{P}}_i \boldsymbol{\theta}_i + \tilde{\boldsymbol{\varepsilon}} \tag{8}$$

where $\tilde{\mathbf{P}}_i$ is the theoretical measurement matrix associated with the ith measurement subset, consisting of the corresponding column vectors of measurement matrix $\mathbf{P}$, and $\tilde{\boldsymbol{\varepsilon}}$ is the noise vector. Once the unknown vector $\boldsymbol{\theta}_i$ is determined, it can be utilized to determine whether the intermediate target position is at the corresponding grid or not. Basically, there are two different approaches to solve the ℓ_1-norm minimization problem to obtain the signal recovery vector estimate $\hat{\boldsymbol{\theta}}_i$, which are the convex relaxation approach and the greedy approach [39]. Since the greedy approach has a lower computational complexity, it may be preferable for scenarios where complexity is a of a primary concern. Two greedy algorithms, the orthogonal matching pursuit (OMP) algorithm [40] and the basis pursuit (BP) algorithm [41], will be employed in this paper as discussed in the simulation section.

Using the estimated K-sparse vector $\hat{\boldsymbol{\theta}}_i$, the residual error vector with respect to the position estimated with the measurements of the ith subset is defined as

$$\mathbf{r}_i = [r_{i,1} \quad r_{i,2} \quad \cdots \quad r_{i,M}]^T = \tilde{\mathbf{u}} - \mathbf{P}\hat{\boldsymbol{\theta}}_i \tag{9a}$$

Also, define the residual error square vector as

$$\mathbf{r}_i^2 = [r_{i,1}^2 \quad r_{i,2}^2 \quad \cdots \quad r_{i,M}^2]^T \tag{9b}$$

Then, the index of the error vector or error square vector of interest is determined using the minimum median technique as

$$\lambda = \arg\min_{\lambda=1,2,\cdots,Q} \text{med}\{\mathbf{r}_\lambda^2\} = \arg\min_{\lambda=1,2,\cdots,Q} \text{med}\{r_{\lambda,1}^2, r_{\lambda,2}^2, \cdots, r_{\lambda,M}^2\} \tag{10}$$

where med $\{.\}$ is the operation of taking the median of the vector components. The verifier parameter is simply defined as

$$\tilde{S} = \sqrt{\text{med}(\mathbf{r}_\lambda^2)} \tag{11}$$

The selected residual error vector $\mathbf{r}_\lambda$ and the above verifier parameter will be exploited for malicious node detection as discussed in the following subsection.

4.2. Attacked Measurement Detection

To make the subset-based position estimation accurate and reliable, the subset size should be sufficiently large. In view of the computational complexity, on the other hand, the number of measurement subsets should be kept small. To compensate for the effect of the insufficient number of subsets, the scale parameter proposed in [27] is employed, producing the revised verifier parameter as

$$S = 1.4826\left(1 + \frac{5}{M - \kappa}\right)\tilde{S} \tag{12}$$

where κ is the dimension of the estimated parameter vector. Specifically, κ is two or three for two-dimensional or three-dimensional localization.

Since the selected λth subset of measurements produces a position estimate with the minimum median error, the position estimate could be treated as the best among all the

intermediate position estimates. The quality of a measurement or whether a node is under attack could be judged based on the corresponding component magnitude of the residual error vector $\mathbf{r}_\lambda$. A simple threshold comparison scheme is proposed to generate a weight vector $\mathbf{w} = [w_1, w_2, \cdots, w_M]$ by

$$w_i = \begin{cases} 1, & |r_{\lambda,i}/S| \le \gamma \\ 0, & otherwise \end{cases} \tag{13}$$

where γ is a predefined threshold. Then, whether a node is being attacked or not is determined according to the following rules:

(1) If $w_i = 0, i = 1,2,\cdots,M$, the corresponding APs are assumed to have been attacked so that their measurements are excluded from further processing.

(2) If $w_i = 1, i = 1,2,\cdots,M$, then the *i*th AP is assumed to have not been attacked and the corresponding measurement is used to form a new subset of measurements, denoted as $\mathbf{u_f}$, for final target position estimation.

4.3. Final target position estimation

Rewrite the measurement matrix defined in (5) as:

$$\mathbf{P} = [\mathbf{p}_1, \mathbf{p}_2, \cdots, \mathbf{p}_M]^T \tag{14}$$

where $\mathbf{p}_i = [p_{i,1}, p_{i,2}, \cdots, p_{i,N}]^T$ ($i = 1,2,\cdots,M$). Based on the weight vector formed in (12), the row vectors of $\mathbf{P}$ associated with a nonzero weight are picked up to form a new measurement matrix denoted by $\mathbf{P}_\mathrm{f}$. The new signal recovery problem is then formulated as

$$\hat{\boldsymbol{\theta}}_\mathbf{f} = [\hat{\vartheta}_{f,1}, \hat{\vartheta}_{f,2}, \cdots \hat{\vartheta}_{f,N}]^\mathrm{T} = \arg\min_{\boldsymbol{\theta}_\mathrm{f} \in \mathrm{R}^N} \|\boldsymbol{\theta}_\mathbf{f}\|_1 \quad \text{s.t. } \mathbf{u_f} = \mathbf{P_f}\boldsymbol{\theta}_\mathbf{f} + \boldsymbol{\varepsilon}_\mathbf{f} \tag{15}$$

where $\boldsymbol{\theta}_\mathbf{f}$ and $\boldsymbol{\varepsilon}_\mathbf{f}$ represent the recovery vector and the measurement noise vector, respectively. Both the BP and the OMP minimization algorithms can be employed to obtain the recovery vector estimate $\hat{\boldsymbol{\theta}}_\mathbf{f}$. At last, the grids with the *K* largest recovery coefficients in $\hat{\boldsymbol{\theta}}_\mathbf{f}$ are chosen and the centers of the grids are the final position estimates of the *K* targets. The performance analysis and comparison between the two approaches are provided in the following two sections.

5. Algorithm Analysis

5.1. Parameter Selection

When implementing the proposed approach, two issues need to be addressed. The first one is about how to choose the appropriate size of the measurement subset to ensure the desirable performance of the final position estimation. The second one is how many measurement subsets should be chosen to generate the intermediate target position estimates. The answer depends on the performance requirement and the computational complexity constrained by the system. In what follows, we provide some useful information about the choice of these parameters.

Clearly, the number of the selected APs or the subset size should be greater than the target population. Then, what is the relationship between the two parameters to ensure a reliable

recovery of the signal vector $\boldsymbol{\theta}_i$? As reported in [43], there is a four-to-one practical rule which says that for exact reconstruction, one needs about four incoherent measurements per unknown nonzero term in recovery vector, i.e.

$$M \geq 4K \tag{16}$$

Accordingly, the size of measurement subset M_s should be $4K$ at least.

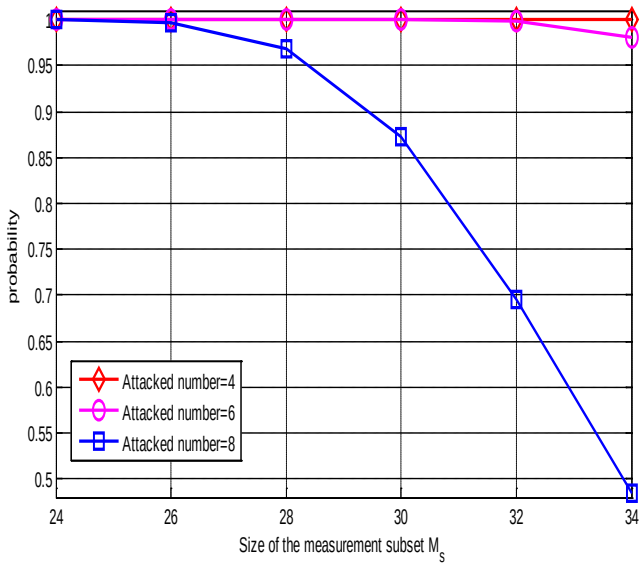

Fig. 2. Probability p_d versus subset size for the cases of three different numbers of attacked measurements.

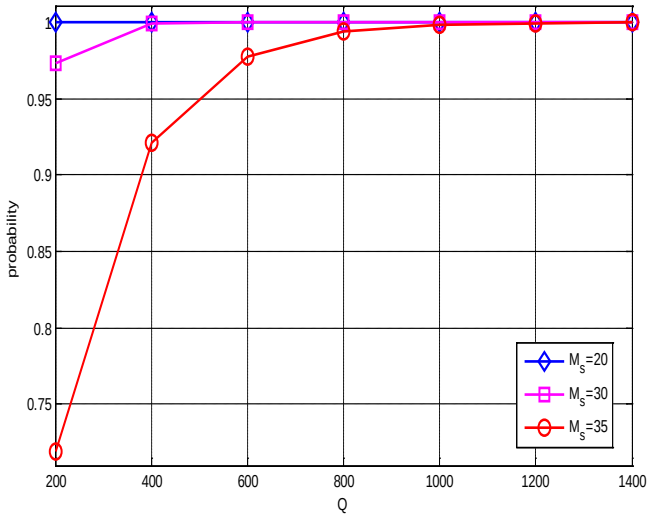

Fig. 3. Probability p_d under different number of measurement subset

The second question may be answered through calculating the probability, denoted by p_d, that at least one measurement subset does not contain any attacked measurement. Let us

assume that the total number of attacked measurements is η, the subset size is M_s as mentioned earlier, and $M\text{-}\eta > M_s$. When choosing M_s measurements from M measurements, the number of combinations is equal to the binominal coefficient $C_M^{M_s}$. Similarly, the number of subsets which do not contain any attacked measurements equals $C_{M-\eta}^{M_s}$. Then, in this case, when randomly selecting a subset of measurements, the probability of getting an attack-free subset is $C_{M-\eta}^{M_s} / C_M^{M_s}$. Therefore, the probability of the measurement subset which contains at least one attacked measurement is given by $1\text{-}C_{M-\eta}^{M_s} / C_M^{M_s}$. When selecting Q subsets, the probability that each of the subsets contains at least one attacked measurement is given by $(1\text{-}C_{M-\eta}^{M_s} / C_M^{M_s})^{Q}$. Therefore, the probability p_d is equal to $1\text{-}(1\text{-}C_{M-\eta}^{M_s} / C_M^{M_s})^{Q}$. **Fig. 2** shows the probability p_d with respect to the subset size when $M = 64$ and $Q = 500$. The probability p_d decreases as the subset size increases, which is due to the fact that it is more likely to include an attacked measurement in a larger subset. However, the probability is rather insensitive to the subset size unless the subset size reaches a certain value, where p_d decreases sharply. Therefore, it is important to choose a subset size smaller than 26 for η=8. **Fig. 3** shows the probability p_d versus the number of selected subsets when η=6, $M = 64$, and M_s =20, 30 and 35. As expected, p_d increases as Q increases since increasing the number of subsets will result in more free-attack subsets. The results when varying M_s are in consistant with those shown in **Fig. 2**. **Fig. 4** illustrates how the number of attacked measurements affects p_d when $M = 64$, $M_s = 30$, and Q =500 and 1000. One main observation is that p_d is insensitive to the number of attacked measurements when it is less than 8. However, p_d drops greatly when the number of attacked measurements is larger than 8.

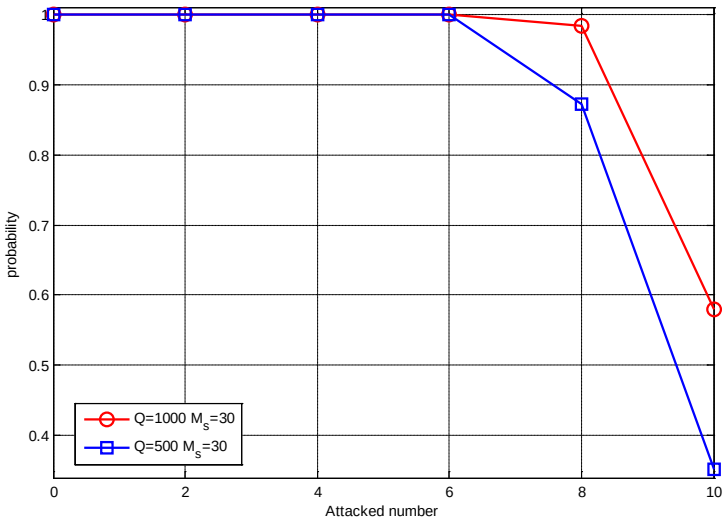

Fig. 4. Probability p_d versus number of attacked measurements.

5.2. Computational Complexity

Table 1 shows the computational complexity comparison between the proposed algorithm and existing methods. Most of the computations of the proposed algorithm are involved in the verifier parameter estimation (or the intermediate position estimation) and the final target position estimation. It can be seen that the complexity of the BP and OMP recovery methods are $O(M^2N^{3/2})$ and $O(KMN)$ [40], respectively, where K, M and N are the target population, the number of measurements and the number of grids, respectively. Therefore, in the verifier parameter estimation phase, the computational complexity of the OMP-based and the BP-based method is $Q\times O(KM_sN)$ and $Q\times O(M^2N^{3/2})$, respectively. In the final target position estimation, the computational complexity of the OMP-based and BP-based methods is $O(K(M-\eta)N)$ and $O((M-\eta)^2N^{3/2})$, respectively. Considering that typically $M_sQ > M-\eta$, the computational complexity of the proposed OMP-based algorithm is $O(KM_sQN)$. On the other hand, the computational complexity of the proposed BP-based algorithm is either $O(KM_sQN)$ or $O((M-\eta)^2N^{3/2})$, depending on which one is significantly larger. Therefore, in the case where the OMP algorithm is employed for final target position estimation, the computational complexity of the proposed algorithm is (QM_s/M) times that of the traditional OMP algorithm.

Table 1. Computational complexity comparison

algorithm	computational complexity
BP algorithm	$O(M^2N^{3/2})$
OMP algorithm	O(KMN)
Proposed BP based algorithm	$\max(O(KM_sQN), \quad O((M-\eta)^2N^{3/2}))$
Proposed OMP based algorithm	$O(KM_sQN)$

6. Simulation Results

6.1 Simulation Setup

The assumed localization area is a $50\text{m}\times 50\text{m}$ square region which is divided into $N=14\times 14=196$ grids. The number of APs is assumed to be 64 with their positions uniformly distributed in the localization area. The predefined threshold for attacked measurement detection is set to be $\gamma=10$. Subset size is set to be $M_s=30$. And according to **Fig. 4**, under η =8 (it is the largest attacked number of measurements) and M=64, if we want the probability p_d nearly to 1, the parameter Q must choose larger than 1000, in this paper, the number of subsets used for intermediate position estimation is set to be $Q=1500$. The parameters of the pathloss model in a fading channel are set as follows: $p_0=-40\text{dB}$, $d_0=1$, and $n_p=2$ [37]. The attacked measurements are modeled by setting $\alpha=0$ and $\beta=-17$ in (6), simulating a scenario where the signal is obstructed by a metal object so that heavy signal attenuation occurs [38].

The performance metrics are the root mean square error (RMSE) and the error cumulative distribution function (CDF) of the position error. To realize a fair performance analysis, occasional large position errors are excluded from calculating the RMSE. In particular, 5% of estimation results with the largest errors are not used to calculate the RMSE. For algorithm comparison, both the BP and the OMP algorithms are employed to obtain the position estimation straightly. Meanwhile, the performance of these two methods in the idealized condition is used as the performance reference. The idealized condition means that whether or not the measurement is under attack is known in advance and only the unattacked measurements are used for position estimation.

6.2 Effect of Parameter Selection

Fig. 5 shows the CDF of the position estimation error in the presence of three and six targets respectively, when the number of attacked measurements is 4 and the SNR is 35 dB. As expected, when increasing the target population, the performance degrades. The performance of BP algorithm is better than that of OMP algorithm, because BP algorithm can obtain the optimal global solution by the convex optimization but the complexity is significantly higher as discussed in Section 5.2. **Fig. 6** shows the CDF of the position estimation error when there are three targets and the SNR is 25dB and 35dB, respectively. When increasing the SNR from 25dB to 35dB, the BP algorithm clearly shows a performance gain over different ranges of errors, while the gain of the OMP algorithm is not always consistent.

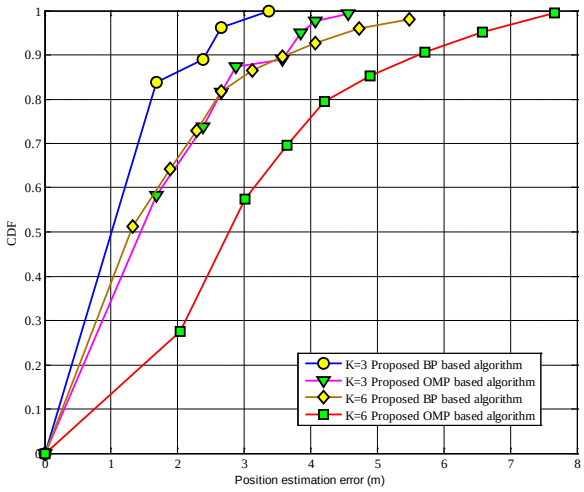

Fig. 5. CDF of the position error under two target populations (3 and 6)

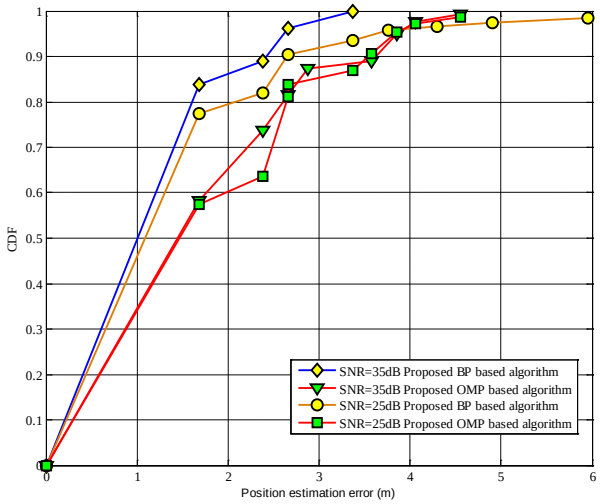

Fig. 6. CDF of the position error under two different SNR values (25dB and 35dB).

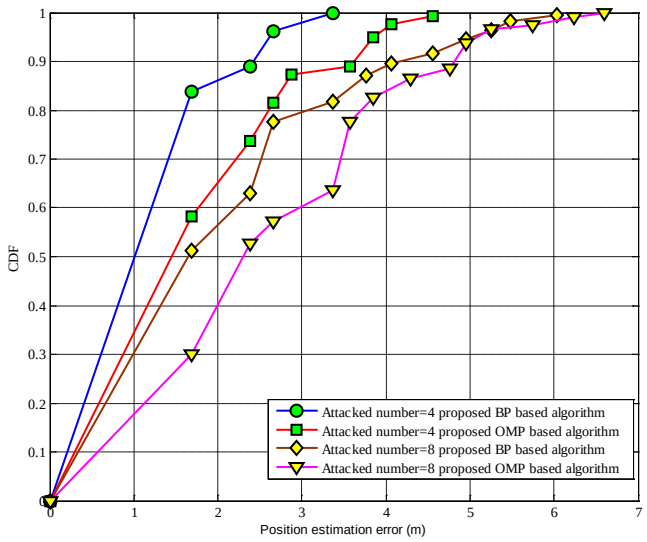

Fig. 7. CDF of the position error under different numbers of attacked measurements (4 and 8).

Fig. 7 shows the CDF of the position estimation when the target population is three and the number of attacked measurements is four and eight, respectively. It can be seen that increasing the number of attacked measurements from 4 to 8 would significantly degrade the performance. For instance, the CDF of an error less than 3m is decreased by about 18% and 34% for BP and OMP algorithms, respectively. **Fig. 8** illustrates the CDF of the position estimation error for two different thresholds when the target population is three and the number of attacked measurements is four. It can be seen that the performance is quite sensitive to the selection of the threshold especially for OMP method. However, it is rather insensitive for the BP based approach. The reason may be explained as follows. When increasing the

threshold, a larger number of measurements will be chosen for the signal vector estimation. Thus, the recovery performance gets better with a larger threshold. However, in this case, it can also decrease the detection probability of the attacked measurements. In addition, under the same condition, because the BP algorithm requires a smaller number of measurements for signal recovery than OMP approach, the performance variation of BP algorithm is not as significant as the OMP method. Therefore, the threshold should be chosen more carefully to obtain a better performance especially for OMP method.

6.3 Performance Comparison

In this subsection, performance comparison is made between the proposed two methods and the existing BP and OMP algorithms under different scenarios. The upper portion of **Fig. 9** displays the RMSE with respect to target population for the four algorithms when SNR is 35 dB and the number of attacked measurements is four. As expected, when the target population increases, the RMSE of the proposed BP and OMP based algorithms increases. The reason can be attributed to the CS recovery theory. When the sparsity degree increases, the signal recovery performance will degrade. Thus, when the target population becomes larger, the localization performance will be lower. Meanwhile, the deviation between the proposed algorithm and the idealized bound is also widened. It is worth developing new techniques to reduce the gap in the future. The lower portion of **Fig. 9** shows the performance of the conventional BP and OMP algorithms, indicating that the two algorithms perform much worse in presence of attacked measurements. As the target population increases, on the contrary, the performances of the two methods become better. The possible reason is that increasing the target population will likely increase the number of attack-free measurements and such a performance gain is greater than the performance loss caused by increased target population.

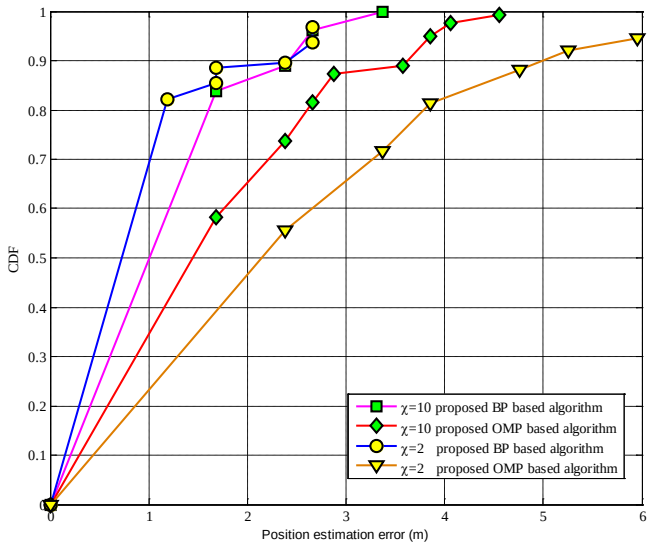

Fig. 8. CDF versus threshold

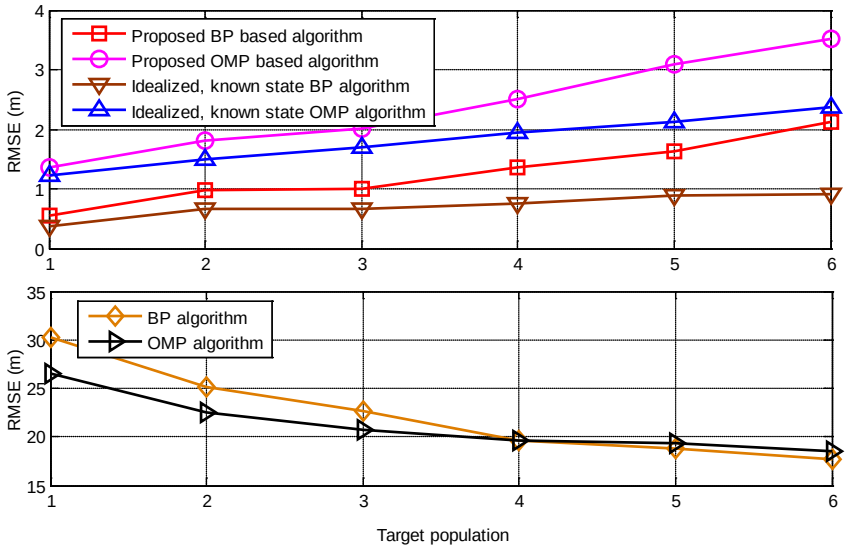

Fig. 9. RMSE versus target population

Fig. 10 illustrates the performance comparison in terms of RMSE versus number of attacked measurements for the six methods when target population is three and SNR is 35dB. It can also be observed that the performance of the proposed BP and OMP based algorithms is much better than the traditional BP and OMP algorithms in the malicious node attack scenarios. The proposed BP based algorithms perform the best among the different approaches. As the number of attacked measurements becomes larger, the number of measurements used for final target position estimation will be smaller. Accordingly, the localization performance of the proposed algorithms will degrade. The performance of the traditional BP and OMP algorithms is rather insensitive to the number of attacked measurements. Their performance in absence of malicious attack is similar to that of the proposed algorithms.

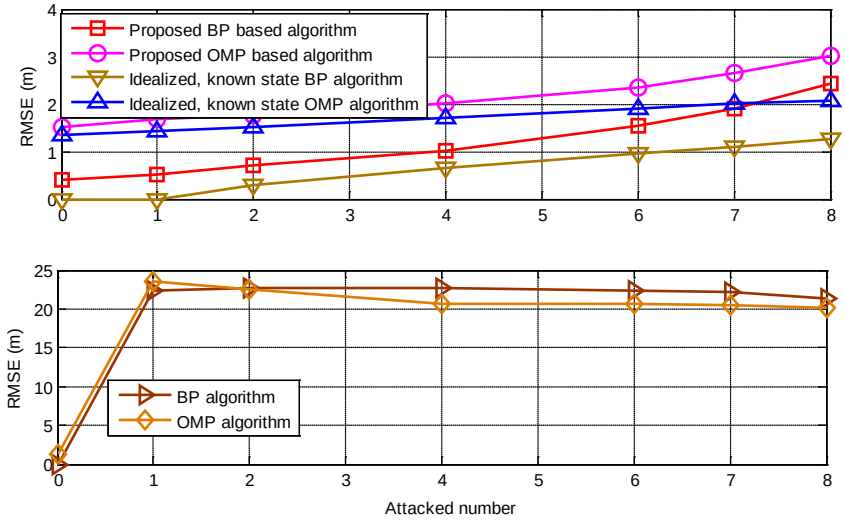

Fig. 10. RMSE versus number of attacked measurements.

Next, we compare different algorithms under different attack parameters when there are three targets and four attacked measurements. **Fig. 11** shows the RMSE of the algorithms versus the attack parameter α when setting $\beta = 0$. The parameter α represents different degree of attack on the measurements. A larger value of $|\alpha - 1|$ corresponds to a more powerful attack. It can be observed that because only the attack-free measurements are utilized for target position estimation, the idealized performance bounds of the BP and OMP algorithms virtually remain the same. The performance of the proposed BP and OMP based approaches is also insensitive to the attack parameter when α is less than 0.6. However, the performance degrades considerably when $\alpha = 0.8$, while the performance improves significantly when $\alpha = 1$. The reason may be that the proposed algorithm can almost correctly find all the attacked measurements with $\alpha \le 0.6$. However, when $\alpha = 0.8$, the average number of the attacked measurements identified by the proposed two algorithms is about 3.84. Thus, a few attacked measurements are selected for final position determination, causing a performance degradation. Another abnormality occurs under the attack-free condition (i.e. $\alpha = 1$), which is that the performance of the proposed methods is better than the performance bounds. The reason can be explained as follows. When $\alpha = 1$ and $\beta = 0$, the effect of the malicious attack is zero. In this case, all measurements employed for position determination using the proposed algorithms are attack free. However, these four assumed attacked measurements are excluded for position determination in the idealized condition, resulting in some performance degradation.

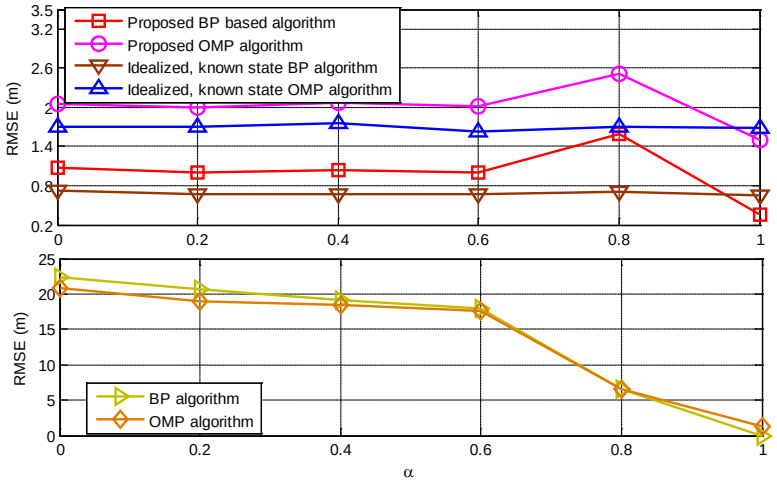

Fig. 11. RMSE versus attack parameter alpha.

Finally, the performance of the methods is evaluated for a scenario where the position of each target is randomly generated within the grid, while the above simulation results are obtained by assuming that each target is located at the center of the relevant grid. **Fig. 12** shows the CDF of the position error for the six algorithms, when the target population is three, the SNR is 35 dB, and the number of attacked measurements is four. Similarly, the proposed BP and OMP based methods significantly outperform the conventional BP and OMP algorithms. On average, the proposed OMP algorithm outperforms the proposed BP based algorithm, which is contrary to the case where the targets are located in the centers of the grids. The reason may be that the BP method requires more accurate measurements for signal recovery.

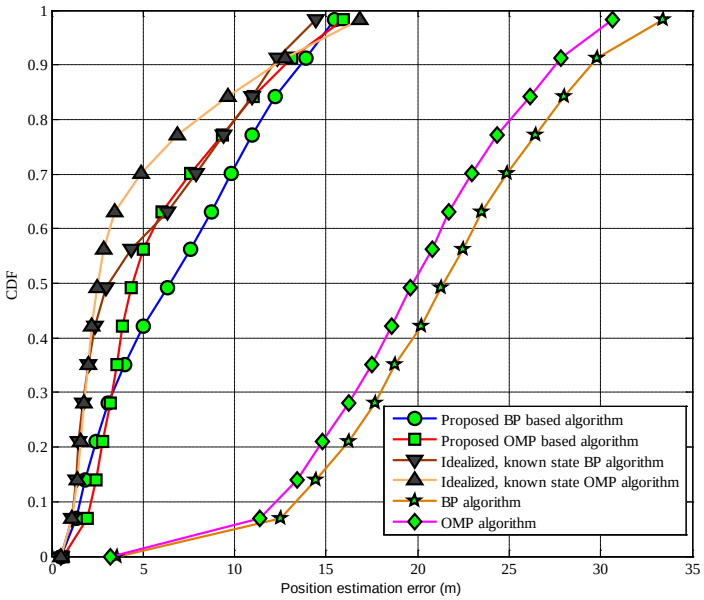

Fig. 12. CDF of algorithms when position of each target is randomly selected within a grid.

7. Conclusions

In this article, a three-phase secure CS based algorithm with RSS measurements has been proposed for target localization in malicious node attack scenarios. First, each subset of measurements is employed to generate intermediate target positions using the CS technique. The least median of the squared residual errors is employed to form the verifier parameter. The attacked measurements are identified through the threshold comparison between the residual errors and the verifier parameter and then excluded from target position determination. The final target position estimates are obtained using the assumed attack-free measurements through the CS approach. The proposed algorithm can achieve much better localization performance than other existing algorithms in terms of the RMSE of the position estimation as evidenced by extensive simulation results. Theoretical analysis has also been performed to provide a guideline in the selection of parameters for the proposed approach

References

[1] T.V. Nguyen, Y. Jeong, H. Shin, and M.Z. Win, "Least square cooperative localization," *IEEE Trans. Vehicular Technology* vol. 64, no. 4, pp.1318-1330, 2015. Article (CrossRef Link)

[2] S.H. Li, M. Hedley, and I.B. Collings, "New efficient indoor cooperative localization algorithm with empirical ranging error model," *IEEE Journal on Selected Areas in Communications*, vol. 33, no. 7, pp. 1407-1417, 2015. Article (CrossRef Link)

[3] I. Sabek, M. Youssef, and A.V. Vasilakos, "ACE: an accurate and efficient multi-entity device-free WLAN localization system," *IEEE Trans. Mobile Computing*, vol. 14, no. 2, pp. 261-273, 2015. Article (CrossRef Link)

[4] I. Sharp and K. Yu, "Indoor TOA error measurement, modeling and analysis for indoor positioning," *IEEE Trans. Instrumentation and measurement*, vol. 63, no. 9, pp. 2129-2144, 2014. Article (CrossRef Link)

[5] I. Sharp and K. Yu, "Enhanced least squares algorithms for indoor positioning," *IEEE Trans. Mobile Computing*, vol. 12, no. 8, pp. 1640-1650, 2013. Article (CrossRef Link)
[6] L. Chen, L. Pei, H. Kuusniemi, Y. Chen, T. Kröger and R. Chen. "Bayesian Fusion for Indoor Positioning Using Bluetooth Fingerprints," *Wireless Personal Communications*, vol. 70, no. 4, pp. 1735-1745, 2013. Article (CrossRef Link)
[7] A.H. Sayed, A. Tarighat, N. Khajehnouri, "Network-based wireless location: challenges faced in developing techniques for accurate wireless location information," *IEEE Signal Processing Magazines*, vol. 22, no. 4, pp. 24-40, 2005. Article (CrossRef Link)
[8] F. Bandiera, A. Coluccia, and G. Ricci, "A cognitive algorithm for received signal strength based localization," *IEEE Trans. Signal Processing*, vol. 63, no. 7, pp. 1726-1736, 2015. Article (CrossRef Link)
[9] L. Chen, O. Julien, P. Thevenon, D. Serant and H. Kuusniemi, "TOA Estimation for Positioning with DVB-T Signals in Outdoor Static Tests," *IEEE Transactions on Broadcasting*, vol. 61, no. 4, pp. 33-38, 2015. Article (CrossRef Link)
[10] Y.T. Chan, W.Y. Tsui, H.C. So, and P.C. Ching, "Time-of-arrival based localization under NLOS conditions," *IEEE Trans. Veh. Technol*. Vol. 55, no. 1, pp. 17-27, 2006. Article (CrossRef Link)
[11] K. Yu, I. Sharp, and Y.J. Guo, "Ground-Based Wireless Positioning," *Wiley-IEEE Press*, 2009. Article (CrossRef Link)
[12] S. Gezici, "A survey on wireless position estimation," *Wireless Personal Commun.*, vol. 44, no. 3, pp. 263-282, 2008. Article (CrossRef Link)
[13] M.K. Oh and J.Y. Kim, "Ranging implementation for IEEE 802.15.4a IR-UWB systems," *in Proc of IEEE Veh. Technol. Conf. (VTC)*, Singapore May 11-14, pp. 1077-1081, 2008. Article (CrossRef Link)
[14] T. Rappaport, "Wireless Communications: Principles and Practice," *IEEE Press Piscataway*, NJ, USA.1996. Article (CrossRef Link)
[15] D. Milioris, L. Kriara, A. Papakonstantinou, G. Tzagkarakis, P. Tsakalides, and M. Papadopouli, "Empirical evaluation of signal-strength fingerprint positioning in wireless LANs," in *Proc.of 13th ACM International Conference on Modeling, Analysis and Simulation of Wireless and Mobile Systems*, Bodrum, Turkey, Oct. 2010. Article (CrossRef Link)
[16] C. Feng, W.S.A. Au, S. Valaee, and Z. Tan, "Received-signal-strength-based indoor positioning using compressive sensing," *IEEE Transactions on Mobile Computing*, vol. 11, no. 12, pp. 1983-1993, 2012. Article (CrossRef Link)
[17] A.W.S. Au, C. Feng, S. Valaee, S. Reyes, S. Sorour, S.N. Markowitz, D. Gold, K. Gordon, and M. Eizenman, "Indoor tracking and navigation using received signal strength and compressive sensing on a mobile device," *IEEE Trans. Mobile Computing*, vol. 12, no. 10, pp. 2050-2062, 2013. Article (CrossRef Link)
[18] A. Boukerche, H.A.B. Oliveira, E.F. Nakamura, and A.A.F. Loureiro, "Secure localization algorithms for wireless sensor networks," *IEEE Communications Magazine*, pp. 96-101, 2008. Article (CrossRef Link)
[19] L. Lazos and R. Poovendran, "Hirloc: high-resolution robust localization for wireless sensor networks," *IEEE Journal on Selected Areas in Communications*, vol. 24, pp. 233-246, Feb. 2006. Article (CrossRef Link)
[20] L. Lazos and R. Poovendran, "Serloc: secure range-independent localization for wireless sensor networks," in *Proc.of ACM Workshop on Wireless Security (WiSe)*, pp. 21-30, 2004. Article (CrossRef Link)
[21] L. Lazos, R. Poovendran, S. Capkun, "Rope: robust position estimation in wireless sensor networks," in *Proc. of Fourth International Symposium on Information Processing in Sensor Networks (IPSN)*, pp. 324-331, Apr. 2005. Article (CrossRef Link)
[22] W. Yuan, N. Wu, B. Etzlinger, H. Wang, J. Kuang, "Cooperative joint localization and clock synchronization based on gaussian message passing in asynchronous wireless networks," *IEEE Trans. Vehicular Technology*, vol. 65, no. 9, pp. 7258-7273, 2016. Article (CrossRef Link)

[23] N. Wu, Y. Xiong, H. Wang, J. Kuang, "A performance limit of TOA-based location-aware wireless networks with ranging outliers," *IEEE Communications. Letters*, vol. 19, no. 8, pp. 1414-1417, 2015. Article (CrossRef Link)
[24] M. Jadliwala, S. Zhong, S. Upadhyaya, C. Qiao, and J.P. Hubaux, "Secure distance-based localization in the presence of cheating beacon nodes," *IEEE Trans. Mobile Computing*, vol. 9, no. 6, pp. 810-823, 2010. Article (CrossRef Link)
[25] R. Garg, A. Varna, and M. Wu, "An efficient gradient descent approach to secure localization in resource constrained wireless sensor networks," *IEEE Trans. Information Forensics and Security*, Vol. 7, No. 2, pp. 717-730, 2012. Article (CrossRef Link)
[26] S. Capkun and J.P. Hubaux, "Secure positioning in wireless networks," *IEEE Journal on Selected Areas in Communications*, vol. 24, no. 2, pp. 221-232, 2006. Article (CrossRef Link)
[27] Z. Li, W. Trappe, Y. Zhang, B. Nath, "Robust statistical methods for securing wireless localization in sensor networks," in *Proc. of Fourth International Symposium on Information Processing in Sensor Networks (IPSN)*, pp. 91-98, 2005. Article (CrossRef Link)
[28] S.H. Fang, C.C. Chuang, and C. Wang, "Attack-resistant wireless localization using an inclusive disjunction model," *IEEE Trans. Communications*, vol. 60, no. 5, pp. 1209-1214, 2012. Article (CrossRef Link)
[29] V. Cevher, P. Boufounos, R.G. Baraniuk, A.C. Gilbert, and M.J. Strauss, " Near-optimal Bayesian localization via incoherence and sparsity," in *Proc. of 2009 International Conference on Information Processing in Sensor Networks (IPSN'2009)*, pp. 205-216, 2009. Article (CrossRef Link)
[30] J.C. Emmanuel and B.W. Michael, "An introduction to compressive sampling," *IEEE Signal Processing Magazine*, vol. 25, no. 2, pp. 21-30, 2008. Article (CrossRef Link)
[31] S. Nikitaki and P. Tsakalides, "Localization in wireless networks via spatial sparsity," in *Proc. of the Forty Fourth Asilomar Conference on Signals, Systems, and Computers (ASILOMAR)*, pp. 236-239, Nov. 2010. Article (CrossRef Link)
[32] J. Wang, D. Fang, X. Chen, Z. Yang, T. Xing, and L. Cai, "LCS: Compressive sensing based device-free localization for multiple targets in sensor networks," in *Proc. of IEEE INFOCOM*, pp 145-149, 2013. Article (CrossRef Link)
[33] D. Milioris, G. Tzagkarakis, A. Papakonstantinou, M. Papadopouli, and P. Tsakalides, "Low-dimensional signal-strength fingerprint-based positioning in wireless LANs," *Ad Hoc Networks*, vol. 12, pp. 100-114, 2014. Article (CrossRef Link)
[34] J. Deng, Q. Cui, and X. Zhang, "Data pre-processing in compressive sensing based indoor fingerprinting positioning," *International Journal of Wireless Information Networks*, vol. 20, pp. 256-267, 2013. Article (CrossRef Link)
[35] D. Liu, P. Ning, and W.K. Du, "Attack-resistant location estimation in sensor networks," in *Proc. of Fourth International Symposium on Information Processing in Sensor Networks (IPSN)*, pp. 99-106, 2005. Article (CrossRef Link)
[36] X.Y. Li, Y.Y. Chen, J. Yang, and X.Y. Zheng, "Designing localization algorithms robust to signal strength attacks," in *Proc. of IEEE International Conference on Computer Communications (Infocom)*, pp. 341-345, 2011. Article (CrossRef Link)
[37] N. Patwari, J.N. Ash, S. Kyperountas, A.O. Hero, R.L. Moses, and N.S. Correal, "Locating the nodes: cooperative localization in wireless sensor networks," *IEEE Signal Processing Magazine*, vol. 22, no. 4, pp. 54-69, 2005. Article (CrossRef Link)
[38] Y.Y. Chen, K. Kleisouris, X.Y. Li, W. Trappe, and R.P. Martin, "A security and robustness performance analysis of localization algorithms to signal strength attacks," *ACM Trans. Sensor Networks*, vol. 5, no. 1, pp. 1-36, Feb. 2009. Article (CrossRef Link)
[39] Y. Eldar and G. Kutyniok, "Compressed sensing: Theory and Applications," *1st Edition, Cambridge University Press*, 2012. Article (CrossRef Link)
[40] J.A. Tropp and A.C. Gilbert, "Signal recovery from random measurements via orthogonal matching pursuit," *IEEE Trans. Information Theory*, vol. 53, no. 12, pp. 4655-4666, Dec. 2007. Article (CrossRef Link)

[41] S.S. Chen, D.L. Donoho, and M.A. Saunders, "Atomic decomposition by basis pursuit," *SIAM Journal on Scientific Computing*, vol. 20, no. 1, pp. 33-61, 1998. Article (CrossRef Link)
[42] J.K. Pant, "Compressive sensing using lp optimization, Dissertation," *University of Victoria*, Canada, 2012. Article (CrossRef Link)

Jun Yan received his Ph.D degree in electrical engineering from Southeast University, Nanjing, China in 2012. He is currently a associate professor at College of Telecommunications and Information Engineering, Nanjing University of Posts and Telecommunications. His research interest is statistical signal processing for wireless location.

Kegen Yu (SM'12) received his Ph.D. degree in electrical engineering from the University of Sydney, Sydney, Australia in 2003. Dr. Yu initially worked as associate engineer at Jiangxi Geology and Mineral Bureau, and as associate lecturer, and later, lecturer in the Department of Industrial Automation at Nanchang University. Subsequently, he was post-doc research fellow at the Centre for Wireless Communications, University of Oulu; research scientist at the CSIRO ICT Centre; research fellow in the Department of Electronic Engineering at Macquarie University; Senior research fellow at the Australian Centre for Space Engineering Research and the School of Surveying and Geospatial Engineering, University of New South Wales. Currently he is the professor at School of Geodesy and Geomatics Wuhan University. Dr. Yu coauthored the book Ground-Based Wireless Positioning, (Wiley-IEEE Press, a Chinese version of the book is also available), and book chapters in three books published by Wiley. He also authored or coauthored 30 refereed journal papers and over 30 refereed conference papers. His current research interests include ground-based and GNSS-based positioning and GNSS remote sensing. He is currently on the editorial board of the EURASIP Journal on Advances in Signal Processing, IEEE Transactions on Aerospace and Electronic Systems, and IEEE Transactions on Vehicular Technology. He is lead guest editor for the special issue of Physical Communication on indoor navigation and tracking, and adjunct professor at Macquarie University.

Yangqin Cao received the BS degree from Nanjing University of Post and Telecommunication in 2014. She is currently working towards the MS degree in the Nanjing University of Post and Telecommnication. Her research interests is wireless location.

Liang Chen is a Senior Research Scientist in the Department of Navigation and Positioning at the Finnish Geospatial Research Institute (FGI), Finland. Before he joined in FGI, he worked in the Department of Mathematics at Tampere University of Technology, Finland from 2009 to 2011. He received his PhD in Signal and Information Processing from Southeast University, China, in 2009. His research interests include statistical signal processing for positioning, wireless positioning using signals of opportunity and sensor fusion algorithm for indoor positioning.