International Journal of Future Generation Communication and Networking
Vol. 8, No. 6 (2015), pp. 99-114
http://dx.doi.org/10.14257/ijfgcn.2015.8.6.10

Outlier Detection Techniques for Localization in Wireless Sensor Networks: A Survey

Hala Abukhalaf, Jianxin Wang and Shigeng Zhang

School of Information Science and Engineering, Central South University, Changsha, 410083, China
h_hf2005@yahoo.com, jxwang@mail.csu.edu.cn, sgzhang@csu.edu.cn

Abstract

In wireless sensor networks (WSNs), localization is one of the most important topics because the location information is typically useful for many applications. The primary data used in a localization process include the locations of anchor nodes and the distances between neighboring nodes. However, these data may contain outliers that deviate from their true values. The existence of the outliers might make the estimated positions not accurate. Thus, it is important to detect and handle outliers in order to achieve high localization accuracy. In this paper, we survey the existing outlier detection techniques for localization in wireless sensor networks. We provide taxonomy for classifying outlier detection techniques in WSNs localization based on different features. In addition, we present comparisons of these techniques. Finally, we discuss the future research directions in this area.

Keywords: *Comparison, detection, localization, outlier, taxonomy, technique, WSNs, wireless sensor networks*

1. Introduction

Recent advancements in MEMS and wireless communication technology have made it possible to develop wireless sensor networks (WSNs), which usually consist of a large number of distributed autonomous inexpensive and low-power sensors. WSNs are usually used to monitor collect physical or environmental data, such as temperature, vibration, sound, and humidity. Wireless sensor nodes cooperatively pass the sensed data to a base station for further processing through multiple hop routing. WSNs have been used in many applications, including intelligent transportation, target tracking, disaster rescue, medical care, environmental monitoring, battlefields, industrial monitoring, and bush fire detection [1, 2].

In many WSN applications, all the collected information requires to know the accurate locations of sensor nodes. Sensing data without knowing the sensors locations are meaningless. In addition, many network functions require the positions of sensor nodes, , geographical routing, geographic key distribution, and location-based authentication. The procedure through which the nodes obtain their positions or locations is called localization. Localization can be considered as one of the most basic and important technologies in WSNs. From the localization point of view, nodes in a sensor network can be classified into two categories: anchors and unknown nodes. Anchors, which are also called landmarks or beacon nodes in some literature, know their own locations via manual placement or GPS. Unknown nodes or sensor nodes, on the other hand, do not know their position information and need to be localized [3-5]. In this paper, we call nodes that have their own location information as anchors, and call nodes that do not know their location information as unknown nodes.

ISSN: 2233-7857 IJFGCN

Most existing localization algorithms of WSNs could be classified as either range-based localization or range-free localization [3-5]. Range-based localization algorithms use absolute point-to-point range measurements (, distance or angle) to estimate unknown nodes' locations. The measurement techniques include TOA [6], TDOA [7], RSS [8], and AOA [9]. Range-free localization algorithms use network connections for the estimation of unknown nodes' locations, DV-Hop [10] and APIT [11]. Range-based localization algorithms can output more accurate position estimates than range-free localization algorithms. However, they require complex equipments to obtain range measurements [5, 12].

In localization algorithms, the calculation of unknown node's positions heavily relies on primary or raw data, the primary data are distances between neighboring nodes and the position knowledge of anchors. In practice, the primary data may contain outliers, including both the outlier distances and the outlier anchors. The outliers generally come from the following sources: adversary attacks, environmental factors, and hardware malfunction. In spite of the reasons of outliers, the presence of them can reduce localization accuracy. So using outlier detection scheme that can find the outliers and handle them is important in order to achieve high localization accuracy [13-15].

This paper gives a survey on the outlier detection techniques used in wireless sensor networks localization. We summarize the key characteristics of current outlier detection techniques in WSNs localization. Previous works mainly focus on general-purpose outlier detection techniques for wireless sensor networks or for secure localization in wireless sensor networks. To the best of our knowledge, this paper gives the first survey to provide a comprehensive overview of outlier detection techniques for wireless sensor networks localization.

The rest of this paper is organized as follows. Section 2 introduces the background. Section 3 states the motivation. In Section 4, we present our taxonomy of outlier detection techniques for localization in WSNs. The outlier detection techniques are introduced and compared in Section 5 and Section 6, respectively. Shortcomings and future research directions are given in Section 7. Finally, Section 8 provides the conclusions of this survey.

2. Background

In this section, we first present the definition of outliers in wireless sensor networks as general, then we describe the concept of localization for WSNs, and finally we state outlier in localization for WSNs.

2.1. Outliers in WSNs

Outlier, also known as anomaly, originally theme from the field of statistics. In this paper, we will use the term outlier, because it is more popular. In wireless sensor networks field, outliers are those measurements that significantly deviate from the normal pattern of sensed data. The outliers may be caused by events, malicious attacks on the network, noise and errors. Conventional outlier detection techniques are not directly applicable to wireless sensor networks due to the nature of sensor data, limitations and specific requirements of the WSNs. Recently, the topic of outlier detection in the field of WSNs has attracted much attention, and several outlier detection techniques specifically developed for the wireless sensor networks have emerged. According to the possible sources of outliers as mentioned earlier, the identification of outliers provides event reporting; secure functioning of the network and data reliability [16].

2.2. Localization for WSNs

The process of finding the spatial locations or positions of unknown nodes in WSNs has been called localization, self-organizing, positioning and geolocation in the literature [17]. The term localization is the most familiar and so it will be used in this paper.

Wireless sensor networks localization is an important area that attracted significant research interest. This interest is expected to grow further with the proliferation of WSNs applications. In environmental monitoring applications such as water quality monitoring, bush fire surveillance, and precision agriculture, the measurement data are meaningless without knowing the location from where the data are obtained [18]. In addition, location estimation plays an important role in some network functionalities such as geographic routing and data centric storage [19].

A straightforward solution for localization is to equip all sensors with global positioning systems (GPSs) [20]. However, this solution is impractical since GPS are expensive and high energy consuming [3, 18, 19]. In recent years, a number of localization algorithms have been proposed to reduce or remove the dependence on GPS in WSNs [9-11, 21-25]. The main idea in most localization algorithms is that a few nodes called anchors are aware of their locations (by GPS receivers or manual configuration) transmit beacons with their coordinates to help the rest nodes called unknown nodes discover their locations [3, 4].

The localization process is to estimate the locations of the unknown nodes. This process can be divided into the following two stages:

(1) Primary data acquisition: The primary data for localization is collected, which may include distances between neighboring nodes and the position knowledge of anchor nodes. To measure distances between neighboring nodes, the distance ranging techniques such as Radio Signal Strength (RSS), Time of Arrival (TOA) and Time Difference of Arrival (TDoA) are adopted. According to position knowledge of anchors, it can be obtained by GPS or manual placement.

(2) Location computation: Once the primary data is collected, location or position is computed. Several methods can be used to compute the location of an unknown node such as triangulation [9], multilateration [21], and trilateration [26]. The choice of which method to use depending on ranging technique used in stage one [19].

2.3. Outlier in Localization for WSNs

As we mentioned before, the primary data used by localization process are the distances between neighboring nodes and the position information of anchors. However, these primary data may contain outliers that strongly deviate from their true values, which include both the outlier distances and the outlier anchors [13].

2.3.1. Distance Outlier

In WSNs localization research field, the mainstream studies assume that the distances between neighboring nodes are accurately measured which are then used to derive node locations accordingly. Typical distance-measuring techniques or ranging techniques include TOA, TDoA and RSS. However, among these distance measurements, there inevitably exist outlier distance whose distance measurement error (the difference between the true distance and the measured distance) is abnormally large [13].

Generally, the probable sources of outlier distances are as following:

(1) Environmental factors: TOA may generate outlier distances with strongly enlarged estimates due to non-line-of-sight propagation. RSS is sensitive to channel noise, reflection, and interference, all of which have significant impacts on signal amplitude. The irregularity of signal attenuation remarkably increases, especially in complex indoor environments.

(2) Hardware malfunction: When encountering ranging hardware malfunction distance measurements will be meaningless. In addition, incorrect hardware calibration and configuration also worsen ranging accuracy. For example, the inaccuracy of clock synchronization results in ranging errors for TDoA, and RSS suffers from transmitter, receiver, and antenna variability.
(3) Malicious attacks: When a WSN is deployed in hostile environments, the localization process is becoming the target of adversary attacks. By reporting fake location or ranging results, an attacker, for example a compromised (malicious) node, can completely distort the coordinate system [13-15].

2.3.2. Anchor Outlier

The anchors are the sensor nodes with a priori knowledge of their positions. The position or locations of these nodes are defined in a global coordinate frame (GCF), for example, GPS coordinate frame. Their purpose is to guarantee that the locations of normal nodes are also defined in the GCF, which can be understood by the system users. However, it is inevitable to have outlier anchor nodes declaring erroneous locations that deviate from their true locations in GCF. Therefore, an outlier anchor can be defined as an anchor node that declares its location in an erroneous coordinate frame that is different from the global coordinate frame (GCF). The potential reasons of outlier anchors can be misconfigurations when deploying the anchor nodes or malignant attacks, for example, Sybil attack and replay attack [13].

3. Motivation

In practice, the measurements used for localization often contain outliers [13-15, 27, 28]. The outlier distance measurements can strictly degrade the accuracy of network localization algorithms [13-15].

Figure 1 shows an example of how outliers destroy localization accuracy. In Figure 1 nodes A1, A2, A3, and A4 (the black circle) are anchors at known positions (four vertices of a square of length $\sqrt{2}$) and node U (the white box in the center of the square) is unknown node which the location of it needs to be determined. Suppose the accurate distance measurements are |A1A2| = | A2A3| = | A3A4| = |A1 A4| = and |A1U| = | A2U | = | A3U | = | A4U | = 1. Apparently, the calculated location of U is just the same as its real location when distance measurements are correct. Now suppose an outlier ranging result occurs: the distance between U and A2 is wrongly measured as 2 (|A2U | = 2). In this case, if all ranging results are indiscriminately used to locate U, the estimated location U' (the black box in Figure 1 (b)) calculated by multilateration [21] is away from the real location U. However, if we layout the outlier ranging, a better estimated location that is the same as the real location of U can be achieved [14].

Besides outlier distance, outlier anchor is another key research issue that must be focus on it. Most of localization schemes assume all anchors are supposed to provide correct reference information. However, when the sensor networks are deployed in a hostile environment, where anchors can be compromised [3], such an assumption does not hold in many related work. Thus, most outlier detection localization techniques focus on outlier distances, few of them focus on the threat of outlier anchor nodes such as in [3] and [13].

(a) Ground truth (b) Estimated location

Figure 1. Multilateration with an Outlier Measurement, Where A1, A2, A3 and A4 are at Corners of a Square of √2, and U is at the Center of that Square

Ignoring the existence of anchor or distance outliers is not good choice. For example, if there are outliers in primary data used by localization schema, the estimated locations of unknown nodes may go far away from their actual locations, this can cause severe consequence if WSNs are used for battle fields surveillance. When unknown nodes report that their regions are safe, this wrong information can cause significant damage. Therefore, it is important to detect and handle outlier in order to achieve high localization accuracy.

4. Classification

Outlier detection techniques for localization in WSNs can be classified according to several criteria or features. Those different features form a reasonable taxonomy for characterizing outlier detection techniques. Figure 2 shows our proposed taxonomy. The rest of this section describes the criteria of the taxonomy in more details.

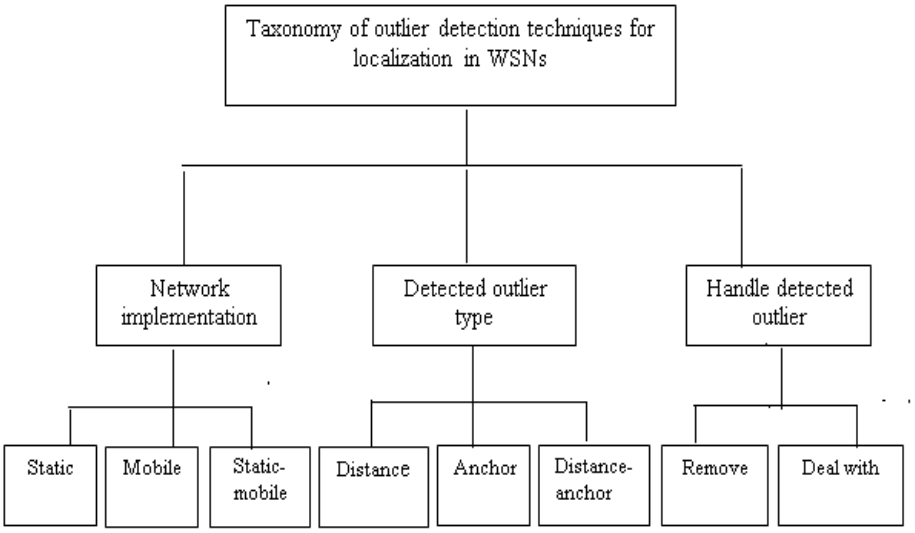

Figure 2. Taxonomy of Outlier Detection Techniques for WSNS Localization

4.1. Network Implementation

This feature means the kind of wireless sensor network that the outlier detection technique is implemented on it. Based on the network implementation, we classify outlier detection techniques into following three types: *static, mobile* and *static-mobile*. In *static*, they are implemented on static WSNs. In *mobile*, they are implemented on mobile WSNs. However, in *static-mobile*, they are implemented on static WSNs also they are implemented on mobile WSNs.

4.2. Detected Outlier Type

As we mention before, the primary data used in localization process are the anchors locations and the distances between neighboring nodes. Therefore, outliers may be distances and /or anchors. So we use feature called detected outlier type to classify outlier detection techniques, this feature means what the outlier type that technique can detects distance or anchor or both. According to the detected outlier type, the outlier detection techniques can be classified into three categories: *distance*, *anchor*, and *distance-anchor*. In *distance*, they detect distance measurement outliers. In *anchor*, they detect anchor outliers. Whereas in *distance-anchor*, they can detect distance outliers and they can detect anchor outliers.

4.3. Handle Detected Outlier

After the outlier has been detected how the technique handle it, based on this feature we classify outlier detection techniques into *remove* and *deal with*. In the former, they remove or eliminate the detected outliers from localization process. In the latter, they consider that the detected outliers may contain important information, so they deal with the outliers instead of removing them.

5. Outlier Detection Techniques for Wireless Sensor Networks Localization

In this section, we present overview of the existing outlier detection techniques for wireless sensor networks localization.

5.1. LAD Localization Anomaly Detection

A scheme named Localization Anomaly Detection (LAD) is proposed by Du [3] to detect anchor node outliers in the localization process for wireless sensor networks. The scheme attempts to identify the outliers and perform compromise resistant localization without remove the malicious anchors from the network.

LAD takes advantage of the deployment knowledge that is available in many sensor networks applications. When sensor nodes are deployed in groups, each node follows two-dimensional Gaussian distribution, which is centered at the deployment point of the node's group. It uses the known deployment information and the group membership of neighbor sensor nodes to check whether the computed position of the unknown nodes is consistent with the known deployment knowledge. If it is inconsistent, LAD will report an outlier. Three metrics are proposed to measure the degree of inconsistency between an unknown nodes derived location and its observation; the metrics are Difference metric, Add-all metric and Probability metric. For each metric, they obtain a threshold through training. If the level of inconsistency exceeds such threshold, they claim that the localization results are inconsistent with the observation, thus an alarm will be raised.

They have evaluated LAD technique, including its tolerance to malicious attacks, false positive rates, detection rates, etc. The simulation results show that even if the outlier detection thresholds are not optimally selected, the technique still has a high detection rate and low false alarm rate for large localization errors. This makes it an ideal candidate for localization outlier detection.

In addition to being effective in detection attacks against the localization, LAD must also resist against attacks on the outlier detection scheme itself. As much as adversaries like to attack localization schemes, they will attack the detection scheme if they know such a scheme is deployed; there are a number of attacks the adversaries can launch. Therefore, the authors have developed a mathematical framework to model those attacks, and this model is used in their simulation-based evaluation to generate attacks. The results

show that the proposed detection scheme is highly resilient against attacks that can cause large damage.

5.2. RobustLoc Outlier Detection

Xiao [13] focus on the problem of localizing a wireless sensor network robustly against outlier links or outlier distances and outlier anchors, so they proposed RobustLoc algorithm. RobustLoc iteratively invokes patch merging operation that can effectively reject outlier distances in sparse networks and meanwhile achieve high localization percentage. Even with robust patch merging, there still exists a challenge towards robust network localization. There is a precondition for the robust patch merging to reject outlier distances, that is, the connectivity between two patches should be sufficiently redundant. However, this condition may not be satisfied in sparse subregions of a network. Thus, in such a sparse subregion, an outlier distance may not be rejectable and will skew the location estimates of nearby sensor nodes. They address this challenge by proposing a robust patch merging operation that can detect the non-rejectable outliers and explicitly report their existence. When receiving such a report, a network localization algorithm can isolate the non-rejectable outliers.

In addition, the authors propose an enhancement to the proposed algorithm to tolerate multiple outlier anchors that can collude due to malicious attacks. RobustLoc algorithm is from few outlier detection algorithms that reject both outlier distances and outlier anchors. In contrast, of most outlier detection localization algorithms those only reject outlier distances.

The experiments show that RobustLoc algorithm has good localization accuracy and can achieve a high percentage of localizable nodes in both dense and sparse networks. In addition, simulations show that RobustLoc can reject colluding outlier anchors reliably in both convex and concave networks.

5.3. Beyond Triangle Inequality

Yang [14] show that noisy and outlier distance measurements can greatly degrade the location accuracy of existing localization approaches. Hence, outlier detection serves as an essential and prior component for all localization approaches. They analyze the limitations of pervious methods that are based on triangle inequality, such as triangle inequality just indicates the existence of outliers, but cannot identify them. These limitations motivate the authors to design approach to identify outlier distance measurements beyond triangle inequality. They formally define the outlier detection problem for wireless ad hoc sensor network localization and build a theoretical foundation to identify outliers based on graph embeddability and rigidity theory. Their analysis shows that the redundancy of distance measurements plays an important role.

An algorithm based on bilateration and generic cycles is accordingly designed to eliminate outliers during the localization process. Different from several pervious methods that require dense networks, the proposed outlier detection algorithm pays more attention to exploring and utilizing the topological structure, and thus works properly in networks that have moderate connectivity. The results of a wireless network prototype and extensive simulations show that the proposed algorithm largely improves the location accuracy by modestly and wisely rejecting outliers during the localization process.

5.4. Detecting Outlier Measurements Based on Graph Rigidity

Yang [15] describe localization is to figure out the locations of unknown nodes based on inter-node distance measurements and global locations of the anchors, the authors notice that a wireless sensor network can uniquely be located without using all inter-node distances, they use redundant inter-node distance information to identify outliers. By applying rigidity theory, they define edge verifiability and derive the conditions for an

edge being verifiable. On this base, the authors design localization approach with outlier detection, which explicitly eliminates the outliers before location computation thus increasing location accuracy. Considering the entire network, they define verifiable graphs in which all edges are verifiable. If a wireless sensor network meets the requirements of verifiable graph, it is not only localizable but outlier resistant as well.

The results show edge verification can be used to sift outlier-ranging information and improve the accuracy of localization. They use multilateration [21] as the basic localization method, and compare the localization performance of the two strategies of outlier sifting and no sifting. The comparison results show remarkable improvement in location accuracy by sifting outliers.

For simplicity in this paper, we will call this technique as outlier graph rigidity.

5.5. Range-Based Localization Using Density- Based Outlier Detection

Almuzaini and Gulliver [17] proposed range-based algorithm called LDBOD. It is based on the density based outlier detection (DBOD) approach from data mining, which used the distance to the K- nearest neighbors (KNN) to select the best candidate points, and these are averaged to get the estimated location of the unknown node. LDBOD differs from conventional solutions to the localization problem in wireless networks. Typically, the locations of the anchors within range and the estimated distances between the unknown node and these anchors are used to directly estimate its location. Instead, a LDBOD used multi-step process, the first obtain distance estimates from unknown node and the anchor nodes that are within range. Second, these estimates provide the radii for circles around the nodes; the intersection of these circles for an unknown node forms a set of points to be used in the next step. The third step is to calculate the density of each intersection point; the points with a density higher than the mean density are selected as candidates. In last step the density-based outlier detection (DBOD) is used, which is commonly used in anomaly detection. The outlier score is just the inverse of the density score of a point. The density is the inverse of the mean distance to the K-nearest neighbors of the point.

The proposed LDBOD was compared with two previous approaches, the results shown that LDBOD performs better than these approaches even when the anchor geometry about an unknown node is poor.

5.6. Outlier-Detection-Based Indoor Localization System for WSN

Chen and Juang [27] develop an outlier detection scheme to deal with outliers for the localization in wireless sensor networks based indoor environment. They present a general framework for estimating the data distribution in view of adjustable window operation. In the RSS-based localization problems, the values of RSS are integers and the values of RSS from a fixed node are usually not a constant as time varies. Thus, the dataset of RSS values is coarsely quantized and the distributed density of outliers in a small size window may be relatively high in comparison with normally distributed data. In the indoor environments, the characteristics of radio signal and the obstacles often cause some data to deviate to become outliers. These outliers, however, may provide the information about the walls or obstacles in the indoor environments. Hence, the proposed outlier detection scheme assigns each data a confidence indicator, which indicates the degree of the reliability of the corresponding data instead of just identifying or removing it. The confidence indicator one can obtain by combining the Hampel filter and probability data computed from kernel density estimator (KDE).

In the localization procedure, the proposed outlier detection scheme can be applied in two ways: censoring sensor reading sequences and RSS database, respectively. The proposed outlier detection scheme can censor raw RSS sequences and give each reading a confidence indicator; then the fingerprinting methods can use the confidence indicators as

weightings in the position determination process. The proposed scheme is shown to be robust and effective in dealing with data that are subject to outlier.

For simplicity in this paper, we will call this technique as outlier-detection- indoor.

5.7. Outlier Compensation in SN Self-Localization via the EM

Self-localization of individual sensor positions is an essential prerequisite for the utility of most sensor networks (SN). Localization estimates are usually obtained by processing of time-of-arrival (TOA), angle-of-arrival (AOA), or received signal strength (RSS) measurements between nodes. In practice, measurements used for self-localization often contain outliers. In [28] Ash and Moses propose a robust self-localization algorithm that effectively mitigates the effects of outlier measurements, the proposed algorithm only considers TOA measurements, but it is easily extended to AOA and RSS observations.

They employ the EM algorithm [29] to iteratively detect outlier measurements, the unobserved measurements of the complete data set in the EM algorithm were taken as indicator variables denoting whether a particular measurement was an outlier or not. Through successive refinements on the posterior distribution of indicator variable, the EM algorithm proved successful in compensating for outlier measurements. The results show that using the proposed EM-based algorithm reduced the localization error.

For simplicity in this paper, we will denote this technique as Outlier- EM.

5.8. Secure RSS-Based Localization in Sensor Networks

Capkun [30] develop a secure localization algorithm for sensor networks based on robust received signal strength (RSS) ranging. It makes use of robust localization and time synchronization primitives, which appropriately combined, enable the detection of attacks on localization, within a realistic attacker model.

The proposed algorithm relies on end-to-end propagation delay measurements to detect distance enlargement attacks and on inverse verifiable multilateration to detect distance reduction attacks by external attackers. The authors show that this algorithm requires no additional hardware support from what already exists on a typical sensor node.

To address internal attackers and outlier they proposed an outlier detection scheme, termed as Neighborhood Consistency Check (NCC), which can tolerate a high fraction of compromised nodes. Compromised nodes can report non-existing links, false locations and can modify ranges measured to them by their neighboring nodes. The objective of NCC is to determine the set of correctly estimated locations and the set of incorrectly estimated locations. NCC achieves this by using information about the set of neighbors and the system.

For simplicity in this paper, we will denote this technique as Secure RSS.

5.9. Robust Distributed Node Localization with Error Management

Liu [31] introduce an iterative least-squares (ILS) method for node localization, in which location information progressively propagates from anchor nodes to other nodes. Compared to the pervious iterative approaches the proposed approach is introduces an error control and a robust formulation of the localization problem so that the proposed approach is less sensitive to noise and computes the location information incrementally.

To prevent error from propagating and accumulating during the iteration, the error control mechanism of the proposed algorithm uses an error registry to select nodes that participate in the localization procedure, based on their relative contribution to the localization accuracy. This mechanism filters out outliers (”the bad seeds") that may contaminate the entire network and preventing them from propagating further. In each iteration of the localization process, a location estimate using a robust least-squares (RLS) formulation is computed. Compared to the traditional least-squares (LS), RLS is more stable against measurement noise. The authors develop an incremental algorithm for the

LS/RLS method that incorporates a new sensor measurement into the location estimation without the need to recomputed from scratch with all the previous measurements. This greatly improves the computational efficiency and allows any-time implementation, which can terminate at any time and still provide usable information.

The incremental, or any-time, aspect is particularly desirable in resource-constrained, decentralized networks where nodes are limited in on-board energy and computation. Furthermore, the incremental computation is efficient and lightweight.

The proposed algorithm has been tested using both simulations and real experiments on a small network of Mica2 motes with ultrasound TOA ranging devices. Results have shown that the error control mechanism is effective in mitigating the effect of error propagation. The error control significantly improves the localization quality and speeds up the convergence in iteration.

For simplicity in this paper, we will denote this technique as localization error management.

5.10. Robust Statistical Methods for Securing Wireless Localization

A list of attacks that are unique to localization algorithms and their countermeasures are presented by Li [32]. The authors propose the idea of tolerating attacks, instead of eliminating them, by exploiting redundancies at various levels within wireless networks; they develop robust statistical methods to make localization attack-tolerant. The authors focus on applying robust mechanisms to two broad classes of localization schemes: triangulation and the RF–based fingerprinting.

For triangulation-based localization, they propose the use of least median squares (LMS) as an improvement over least squares (LS) for achieving robustness to attacks. The vulnerability of the least squares algorithm to attacks is essentially due to its non-robustness to “outliers”. Due to lack of false information filtering ability LS, this scheme will cause large location error in hostile environments. In order to get rid of outlier samples to improve the estimation accuracy, least median squares (LMS) are introduced to minimize the median of error squares rather than the sum of error squares in LS. LMS achieves higher localization accuracy, however, it requires intensive computation, and in order to reduce the computational complexity of LMS a linearization of the least squares location estimator is formulated.

For fingerprinting-based location estimation, they show that the use of traditional Euclidean distance metrics is not robust to intentional attacks launched against the base stations involved in localization. They propose a median-based nearest neighbor scheme that employs a median-based distance metric that is robust to location attacks.

For simplicity in this paper, we will denote this technique as robust statistical methods.

5.11. A Robust Localization Algorithm in WSNs

In [33] Li present a distributed robust localization algorithm called Bilateration that employs a unified way to deal with all kinds of location attacks as well as other kinds of information distortion caused by node malfunction or abnormal environmental noise. Bilateration tries to find a maximum set of close-by positions from all candidate positions and use the average of these close-by positions as the estimated location. Taking close-by positions as reasonable candidate positions is based on the observation that candidate positions calculated from correct reference positions and distance measurements tend to be close to each other, and the use of maximum (optimal) set of close-by positions is to optimize the localization accuracy as well as defeat collaboration location attack launched by compromised nodes. Bilateration algorithm is robust in the sense that it can locate the unknown node with acceptable accuracy even in the presence of some false information. Bilateration can detect and filter outliers; it has stronger outliers filtering ability compared to multilateration with LS, LMS, and LLMS.

Simulation results show that Bilateration achieves the best trade-off between localization accuracy and computational complexity.

5.12. Robust Interval-Based Localization for Mobile Sensor Networks

Mourad [34] present original adaptive approach for mobile sensor localization in the presence of outliers. The interval analysis [35] is employed; the estimates are sets of nonoverlapping boxes containing the possible locations of the sensors.

The authors propose set inversion via interval analysis SIVIA-based method (SBL) and the combinatorial-based one (CBL), these two algorithms are robust to outliers. SBL algorithm bisects the search region leading to many boxes describing efficiently the solution set; CBL algorithm leads to larger boxes including the solution as well. In terms of computing time, CBL is more efficient than SBL for a small number of outliers, whereas the complexity of SBL is almost constant whatever the number of tolerated outliers. Choosing one algorithm or the other depends on the anchor density and on the proportion of outliers.

Another contribution of [34] is that it proposes a technique for evaluating the maximal number of outliers to be robust to them. Moreover, it uses a connectivity-based observation model.

For simplicity in this paper we will denote this technique as robust interval-based.

6. Comparison of Outlier Detection Techniques for Localization in WSNS

In this section, we present comparative tables to compare the existing outlier detection techniques for localization in WSNs using the taxonomy proposed in Section 4.

The characteristics of outlier detection techniques are shown in Table1; it presents the key idea and the network implementation of these techniques. As shown in Table 1 most techniques are implemented on static WSNs. Only [34] is implemented on mobile WSNs, also [27] is implemented on both static and mobile WSNs.

Table 1. Characteristics of Outlier Detection Techniques for Localization in WSNS

Technique	Key idea	Network Implementation
LAD [3]	consistency with the known deployment knowledge	static
RobustLoc [13]	patch merging operation	static
Beyond triangle inequality [14]	graph embeddability and rigidity theory	static
Outlier graph rigidity [15]	graph rigidity	static
LDBOD [17]	density based outlier detection (DBOD)	static
Outlier-detection- indoor [27]	confidence indicator	static- mobile
Outlier-EM [28]	EM algorithm	static
Secure RSS [30]	neighborhood consistency check(NCC)	static
Localization error management [31]	error registry	static
Robust statistical methods [32]	statistical methods	static
Bilateration[33]	reasonable sample	static
Robust interval-based [34]	interval-based	mobile

In Table 2 we compare between outlier detection techniques in respect of the detected outlier type and the way to handle the detected outlier. From Table 2 we note that most techniques detect distance outliers, only [3, 32] detect anchor outliers, also [13] can detect both distance and anchor outliers.

Table 2. Comparison of Outlier Detection Techniques for Localization in WSNS

Technique	Outlier type	Handle outlier
LAD [3]	anchor	deal with
RobustLoc [13]	distance-anchor	remove
Beyond triangle inequality [14]	distance	remove
Outlier graph rigidity [15]	distance	remove
LDBOD [17]	distance	remove
Outlier-EM [28]	distance	deal with
Secure RSS [30]	distance	deal with
Localization error management [31]	distance	deal with
Robust statistical methods [32]	anchor	deal with

7. Shortcomings of Outlier Detection Techniques for WSNS Localization and Future Research Directions

In this section, we first present shortcomings of the existing outlier detection techniques for wireless sensor networks localization, and then we specify the future research directions of this subject.

7.1. Shortcomings of Outlier Detection Techniques for WSNS Localization

From our studying of the existing outlier detection localization techniques and from the comparative tables that presented in pervious section, we realize that the current outlier detection techniques for WSNs localization have the following shortcomings:

(1) As general, there are little techniques to detect outliers localization for wireless sensor networks.
(2) Majority of existing techniques do not take into account detecting anchor outliers, many of techniques only consider distance outliers. Little work such as [3, 13, 32] has been done on detecting anchor outliers.
(3) Most of the existing techniques assume sensor nodes are static and do not consider nodes mobility, a few techniques consider mobile sensor networks such as [27] and [34]. Applying outlier detection techniques localization for mobile sensor networks would be challenging.

7.2. Future Research Directions

The future research directions of outlier detection techniques for wireless sensor networks localization possibly are as following:

(1) It is necessary to design new outlier detection algorithms for localization in WSNs as general.
(2) Detecting anchor outliers should be taken into account. More work must be done on distinguishing between anchor outliers and distance outliers. Also new techniques to detect anchor outliers must be developed.
(3) Researches should propose new algorithms to detect outliers in localization algorithms for mobile sensor networks.
(4) We can investigate the applicability of Artificial Intelligence (AI) techniques for outlier detection localization in WSNs.

(5) Hybrid two or three existing outlier detection techniques then produce new enhanced technique that have the advantages of the combined techniques and deal with the shortcomings of them.

8. Conclusions

Outlier detection is necessary for localization in wireless sensor networks because the existence of outliers can degrade the localization accuracy. In this paper, we discussed the problem of outliers in wireless sensor network localization and the importance of outlier detection. Also we have presented an overview of the current outlier detection methods for localization in WSNs and we made comparisons between them. This paper provided taxonomy of the different criteria that can be used to classify outlier detection techniques for WSNs localization. In addition, we introduced shortcomings of these techniques.

There are several literatures about outlier detection techniques for wireless sensor networks as general and for secure localization in wireless sensor networks, but to the best of our knowledge, the literatures about outlier detection for wireless sensor networks localization are few, thus there is need for further research in this area. Therefore, we listed future research directions of outlier detection localization in this survey.

Acknowledgments

This paper is partially supported by the Nature Science Foundation of China under Grant Nos. 61173169, 61402056, and 61420106009.

References

[1] U. Singh and M. Jha, "Performance evaluation of localization technique in wireless sensor network", International Journal of Engineering and Computer Science, vol. 2, no. 4, **(2013)**, pp.1381-1384.
[2] X. Wang, D. Bi, L. Ding and S. Wang, "Agent collaborative target localization and classification in wireless sensor networks", Sensors, vol. 7, no. 8, **(2007)**, pp. 1359-1386.
[3] W. Du, L. Fang and P. Ning, "LAD: Localization anomaly detection for wireless sensor networks", Journal of Parallel and Distributed Computing, vol. 66, no.7, **(2006)**, pp. 874-886.
[4] J. Jiang, G. Han, C. Zhu, Y. Dong and N. Zhang, "Secure localization in wireless sensor networks: A survey", Journal of Communications, vol. 6, no. 6, **(2011)**.
[5] G. Han, H. Xu, T. Q. Duong, J. Jiang and T. Hara, "Localization algorithms of wireless sensor networks: A survey", Telecommunication Systems, vol. 52, no. 4, **(2013)**, pp.2419- 2436.
[6] I. Guvenc and Z. Sahinoglu, "Threshold-based TOA estimation for impulse radio UWB systems", Proceedings of IEEE International Conference on Ultra-Wideband, Zurich, Switzerland, **(2005)** September, pp.420-425.
[7] X. Wei, L. Wang and J. Wan, "A new localization technique based on network TDOA information", Proceedings of IEEE International Conference on ITS Telecommunications, **(2006)** June, pp. 127-130.
[8] A. Hatami, K. Pahlavan, M. Heidari and F. Akgul, "On RSS and TOA based indoor Geolocation - A comparative performance evaluation", Proceedings of IEEE Wireless Communications and Networking Conference, Las Vegas, NV, **(2006)** April, pp. 2267- 2272.
[9] D. Niculescu and B. Nath, "Ad doc positioning system (APS) using AOA", Proceedings of IEEE INFOCOM, San Francisco, CA, **(2003)** March-April, pp. 1734-1743.
[10] D. Niculescu and B. Nath, "DV based positioning in ad hoc networks", Telecommunication Systems, vol. 22, no. 1, **(2003)**, pp. 267-280.
[11] T. He, C. Huang, B. M. Blum, J. A. Stankovic and T. Abdelzaher, "Range-free localization schemes for large scale sensor networks", Proceedings of ACM MobiCom, **(2003)** September, pp. 81-95.
[12] K. Kim and W. Lee, "MBAL: A mobile beacon-assisted localization scheme for wireless sensor networks", Proceedings of the 16th International Conference on Computer Communications and Networks (ICCCN), **(2007)**, pp. 57-62.
[13] Q. Xiao, K. Bu, Z. Wang and B. Xiao, "Robust localization against outliers in wireless sensor networks", ACM Transactions on Sensor Networks, vol. 9, no. 2, **(2013)**.
[14] Z. Yang, L. Jian, C. Wu and Y. Liu**,** "Beyond triangle inequality: Sifting noisy and outlier distance measurements for localization", ACM Transactions on Sensor Networks (TOSN), vol. 9, no. 2, **(2013)**.
[15] Z. Yang, C. Wu, T. Chen, Y. Zhao, W. Gong, and Y. Liu, "Detecting outlier measurements based on graph rigidity for wireless sensor network localization", Vehicular Technology, vol. 62, no. 1, **(2013)**, pp. 374-383.

[16] Y. Zhang, N. Meratnia, P. Havinga, "Outlier detection techniques for wireless sensor networks: A survey", Journal of IEEE Communications Survey & Tutorials, vol. 12, no. 2, **(2010),** pp. 159-170.
[17] K. Almuzaini and T.A. Gulliver, "Range-based localization in wireless networks using density-based outlier detection", Wireless Sensor Network, vol. 2 no. 11, **(2010)**, pp. 807- 814.
[18] G. Mao, B. Fidan and B. D. O. Anderson, "Wireless sensor network localization techniques", Computer Networks, vol. 51, no. 10, **(2007)**, pp. 2529-2553.
[19] X. Li, N. Mitton, I. Simplot-Ryl, and D. Simplot-Ryl, "Dynamic beacon mobility scheduling for sensor localization", IEEE Transactions on Parallel and Distributed Systems, vol. 23, no. 8, **(2012)**, pp. 1439-1452.
[20] B. Hofmann-Wellenhof, H. Lichtenegger and J. Collins, "Global Positioning System Theory and Practice", 4th edn, Springer, New York, **(1997)**.
[21] A. Savvides, C. Han, and M. B. Strivastava, "Dynamic fine-grained localization in Ad-hoc networks of sensors", Proceedings of ACM MobiCom, **(2001)**.
[22] N. Bulusu, J. Heidemann and D. Estrin, "GPS-less low cost outdoor localization for very small devices", IEEE Personal Communications Magazine, **(2000),** pp.28–34.
[23] L. Doherty, K.S. Pister and L.E. Ghaoui, "Convex optimization methods for sensor node position estimation", Proceedings of INFOCOM'01, **(2001)**.
[24] R. Nagpal, H. Shrobe and J. Bachrach, "Organizing a global coordinate system from local information on an ad hoc sensor network", IPSN'03, **(2003)**.
[25] A. Nasipuri, K. Li, "A directionality based location discovery scheme for wireless sensor networks", Proceedings of ACM WSNA'02, **(2002)**.
[26] D. Niculescu, B. Nath, "Ad hoc positioning system (APS)", Proceedings of the 2001 IEEE Global Telecommunications Conference of the IEEE Communications Society, vol. 5, **(2001)**, pp. 2926–2931**.**
[27] Y. Chen and J. Juang, "Outlier-detection-based indoor localization system for wireless sensor networks", International Journal of Navigation and Observation, **(2012)**.
[28] J.N. Ash and R.L. Moses, "Outlier compensation in sensor network self_localization via the EM algorithm", IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP'05), **(2005)** March, pp. 749–752.
[29] A. P. Dempster, N. M. Laird and D. B. Rubin, "Maximum likelihood from incomplete data via the EM algorithm", Journal of the Royal Statistical Society, vol. 39, no. 1, **(1977)**, pp. 1- 38.
[30] S. Capkun, S. Ganeriwal, F. Anjum and M. Srivastava, "Secure RSS-based localization in sensor networks", Technical Reports 529, ETH Zürich, **(2006)**.
[31] J. Liu, Y. Zhang and F. Zhao, "Robust distributed node localization with error management", Proceedings of the 7th ACM International Symposium on Mobile Ad Hoc Networking and Computing (MobiHoc 2006), **(2006)** May, pp. 250-261.
[32] Z. Li, W. Trappe, Y. Zhang and B. Nath, "Robust statistical methods for securing wireless localization in sensor networks", Proceedings of the 4th International Symposium on Information Processing in Sensor Networks, **(2005)**.
[33] X. Li, B. Hua, Y. Shang and Y. Xiong, "A robust localization algorithm in wireless Sensor networks", Journal of Frontiers of Computer Science in China, vol. 2, no. 4, **(2008**), pp. 438-450.
[34] F. Mourad, H. Snoussi, M. Kieffer and C. Richard, "Robust interval-based localization algorithms for mobile sensor networks", International Journal of Distributed Sensor Networks, (**2012).**
[35] L. Jaulin, M. Kieffer, O. Didrit and E. Walter, "Applied Interval Analysis", Springer, **(2001)**.
[36] Y. Zeng, J. Cao, J. Hong, S. Zhang, and L. Xie, "Secure localization and location verification in wireless sensor networks: A survey", Journal of Supercomputing, vol. 64, no. 3, **(2013)**, pp. 685–701.

Authors

Hala Abukhalaf, she received her BSc in Computer Systems Engineering from Palestine Polytechnic University. She received her MSc in Computer Science from Al-Quds University, Jerusalem, in 2008. She worked as a part time lecturer in Al-Quds Open University from 2010 to 2011. She is currently studying towards her PhD degree at the School of Information Science and Engineering, Central South University, Changsha, China. Her research interests include wireless sensor networks and mobile ad hoc networks.

Jianxin Wang, he received his BSc and MSc degrees in Computer Science from Central South University of Technology, China, and his PhD degree in Computer Science from Central South University. Currently, he is the vice dean and a professor in the School of Information Science and Engineering, Central South University, Changsha,

Hunan, China. He has served as a program committee member for many international conferences. He was a program committee co-chair for the International Symposium on Bioinformatics Research and Applications (ISBRA) in 2011, 2012 and 2014, and a program committee co-chair for the 8th International Frontiers of Algorithmics Workshop (FAW2014). His current research interests include algorithm analysis and optimization, parameterized algorithm, bioinformatics and computer networks. He has published more than 200 papers in various international journals and refereed conferences. He is a senior member of the IEEE.

Shigeng Zhang, he received the BSc, MSc, and DEng degrees, all in Computer Science, from Nanjing University, China, in 2004, 2007, and 2010, respectively. He is currently an assistant professor in School of Information Science and Engineering at Central South University, China. His research interests include cloud computing, Internet of Things, wireless sensor networks, and RFID systems. He has published more than 40 technique papers in international journals and conferences. He is a member of IEEE, a member of ACM, and a member of CCF.