Localised information fusion techniques for location discovery in wireless sensor networks

Adnan M. Abu-Mahfouz*

Meraka Institute,
Council for Scientific and Industrial Research (CSIR),
Pretoria, 0841, South Africa
Email: a.abumahfouz@ieee.org
*Corresponding author

Gerhard P. Hancke

Department of Computer Science,
City University of Hong Kong,
China

and

Department of Electrical, Electronic and Computer Engineering,
University of Pretoria,
Pretoria, 0002, South Africa
Email: ghancke@ieee.org

Abstract: Many wireless sensor network (WSN) applications and techniques require sensor location to work correctly. Information fusion in WSNs is crucial for location discovery and many information fusion techniques have been used in several location-discovery algorithms. In this paper, we provide an overview of information fusion techniques used by localisation algorithms to enhance and simplify the position computation process. Furthermore, we analyse several approaches that can be used by localised information fusion algorithms and discuss how these approaches enhance the performance of information fusion and make information fusion play a leading role in localisation algorithms. This paper provides an introduction to different localised information fusion methods for node location discovery in WSNs and could be used as a starting point for implementing existing proposals or designing new techniques.

Keywords: localisation system; information fusion; localised algorithms; WSNs; wireless sensor networks; sensor’s position.


Biographical notes: Adnan M. Abu-Mahfouz received his ME and PhD in Computer Engineering from the University of Pretoria. He is currently a Principal Research Engineer at the Council for Scientific and Industrial Research (CSIR). He is also an Adjunct Research and Innovation Associate in the Faculty of Engineering and Built Environment at the Department of Electrical Engineering/F’SATI (French South African Institute of Technology) at Tshwane University of Technology. He is the Chair of Tshwane Water Resource Management Network. His research interests include wireless sensor networks, software-defined wireless sensor networks, network management, network security, localisation systems and low-power wide area networks.

Gerhard P. Hancke is currently an Assistant Professor with the Department of Computer Science at City University of Hong Kong. He received his Bachelor’s and Master’s Degrees in Computer Engineering from the University of Pretoria (South Africa) in 2002 and 2003, respectively, and PhD in Computer Science for the Security group from the University of Cambridge’s Computer Laboratory in 2008. He worked for the Smart Card Centre and Information Security Group at Royal Holloway, University of London. His main research interests include sensing applications and security of embedded systems.
1 Introduction

The proliferation of wireless communication technologies has enabled the development of wireless sensor networks (WSNs), which consist of a large number of small and cheap sensors with limited resources, such as computing, communication, storage, and energy (Yick et al., 2008). These sensor nodes are able to sense, measure and collect raw data from the environment, perform simple computations and then transmit only the required and partially processed data to the node responsible for fusion (Akyildiz et al., 2002).

WSNs have been deployed extensively in areas such as military operations, health monitoring, natural disaster management, smart grid, smart water systems and hazardous environments (Hussain et al., 2009; Toh et al., 2009; Wirawan et al., 2008; Werner-Allen et al., 2006; Abu-Mahfouz et al., 2015, 2016). Most of these applications require that the position of the nodes must be determined. In some scenarios, node location information plays a critical role, such as data-centric storage application (Ramasamy et al., 2002). Several WSN techniques require highly accurate knowledge of the location, such as the geographical routing technique, network security, energy efficient management and identifying boundary nodes (Karp and Kung, 2000; Li et al., 2000; Al-Otaibi and Hamdy, 2010; Hu et al., 2003; Xu et al., 2001; Fayed and Hussein, 2009). The main advantages of node location information are enhancing the efficiency of the WSNs, identifying the location of an event of interest, facilitating numerous application services and assisting in various system functionalities (Srinivasan and Wu, 2008).

Recently, several ‘location-discovery’ algorithms for WSNs have been proposed (Ding et al., 2013; Xiao et al., 2014; Coluccia and Ricciato, 2013; Mao et al., 2013; Kong et al., 2014; Supriya et al., 2013). One approach that has been followed by these algorithms is to use special nodes called ‘beacons’, which know their location (e.g., through a global positioning system (GPS) receiver or manual configuration). The other nodes that do not know their location, sometimes referred to as ‘unknowns’, use different techniques to compute their own position based on the location information of the beacons and the measured distance to these beacons. The term ‘reference nodes’ or simply ‘references’ will be used in this study to refer to the sensor nodes that are willing to help other nodes to estimate their position. Therefore, the reference set includes beacons and knowns (i.e., unknowns that have obtained their position), which are willing to act as a reference for other unknowns.

Several information fusion techniques have been applied to the location-discovery problem in WSNs (Abu-Mahfouz and Hancke, 2013; Liu et al., 2008; Hongyang et al., 2005; Sichitiu and Ramadurai, 2004; Hu and Evans, 2004; Blumenthal et al., 2006; Li et al., 2005). The basic idea of information fusion is the combining of disparate data (either raw data or processed estimates) to improve accuracy and achieve more specific estimates than individual estimates could deliver.

In the literature, there are contributions covering localisation algorithms. For instance, Mao et al. (2007) and Srinivasan and Wu (2007) review some of the existing localisation techniques and list their strengths and weaknesses. Sayed et al. (2005) give an overview of the challenges faced in developing techniques for acquiring accurate wireless location information, while the authors of Arisar and Kemp (2011), Boukerche et al. (2008) discuss the security issues of several localisation techniques.

This paper will discuss information fusion techniques used by localisation algorithms to enhance and simplify the position computation process, and will analyse several approaches that can be used by localised information fusion algorithms. It will point out the impact of these approaches to enhance the performance of information fusion and make information fusion play a leading role in localisation algorithms. To the best of our knowledge, no similar kind of analysis has been published.

The rest of this paper is organised as follows: a brief overview of information fusion concepts, objectives and types is provided in Section 2. Section 3 discusses the components of a localisation system and a review of the information fusion techniques used by localisation algorithms is presented in Section 4. In Section 5, we discuss the localised information fusion algorithms and analyse a number of approaches used by them. In Section 6, we provide a comparative summary of these approaches. Finally, Section 7 concludes the paper.

2 Information fusion

Sensor nodes may be deployed to gather relevant data, either to provide a better understanding of the behaviour of the monitored entity, or to detect the occurrence of possible events. Information fusion is concerned with the way in which the data, once gathered, can be processed to enhance its relevance.

Dasarathy (2001) defines information fusion as

“encompassing the theory, techniques and tools conceived and employed for exploiting the synergy in the information acquired from multiple sources (sensor, databases, information gathered by human[s], etc.) such that the resulting decision or action is in some sense better (qualitatively or quantitatively, in terms of accuracy, robustness, etc.) than would be possible, if these sources were used individually without such synergy exploitation.”

This definition is depicted in Figure 1.
There are many objectives of information fusion. It can be used to compose a comprehensive view from the partial views provided by different nodes; fuse overlapping measurements to get more accurate information and combine complementary data to allow inferences (e.g., a node can fuse the location of, and the estimated distance to, the neighbour references to obtain its position). Information fusion can reduce the overall communication load in the network and thus conserve energy and prolong the lifetime of the entire network. The minimum requirements for WSNs would be the improvement of accuracy and energy saving (Nakamura et al., 2007).

Two approaches to location fusion can be adopted: centralised fusion and decentralised (distributed) fusion. Theoretically, a centralised fusion system should outperform a distributed one, because the central unit has global knowledge in the sense that all measured data is available. However, this system also requires that all the raw data (or processed estimate) be transmitted from the nodes to the central unit. Such a high volume of communications might not be practical and might consume too many system resources.

In a decentralised or distributed fusion system, each node has its own processing facility, which cuts out the requirement for any central fusion or central communication facility. In this approach, each node estimates its position and the fusion occurs locally at each node, based on local observation and the information received from neighbouring nodes. The main advantages of a decentralised fusion system are that it reduces the communications overheads and thus overcomes the problem of limited communication bandwidth. It avoids the effect of centralised computational bottlenecks, which makes this approach scalable. It is also adaptable to dynamic changes in the network structure and to the addition or loss of sensing nodes. In view of the nature of WSNs, with their limited resources and bandwidth, these advantages make the distributed algorithms preferable to centralised algorithms.

Localised algorithms are a special type of distributed algorithm in which only a subset of nodes in the WSN is invoked for a specific task (e.g., sensing, tracking, reasoning, communication and computation) (Meguerdichian et al., 2001). Localised algorithms dramatically reduce redundant processing and communication, and thus save power and prolong the lifetime of the network, which could make them the best solution for WSNs. However, choosing the proper subset of nodes to participate efficiently in a specific task is not a minor problem. Their scalability, robustness and energy-effectiveness have attracted several researchers to use localised algorithms to develop various protocols for WSNs, such as directed diffusion and sensor protocols for information via negotiation and collaborative signal and information processing (Intanagonwiwat et al., 2000; Kulik et al., 2002; Kumar et al., 2002). Meguerdichian et al. (2001) developed a generic localised algorithm for location discovery in WSNs.

3 Localisation systems

Localisation systems consist of three major components: distance/angle estimation, position computation and a localisation algorithm (Boukerche et al., 2008). To explain the general role of information fusion within localisation, two supplementary components were added: a localised algorithm and an information fusion technique as shown in Figure 2.

3.1 Distance/angle estimation

This component is responsible for determining the physical relationship between two nodes, which can later be used to compute a node’s location. Different approaches can be used for this purpose, such as directional antennas, radio frequency (RF) fingerprinting (communication neighbour authentication), connectivity (in range) and distance bounding (Hu and Evans, 2004; Rasmussen and Capkun, 2007; Maheshwari et al., 2007; Meadows et al., 2006). Practically, these approaches use several techniques, including received signal strength (RSS), time of arrival (ToA), time difference of arrival (TDoA), angle of arrival
A.M. Abu-Mahfouz and G.P. Hancke

The next two components rely on accurate estimates so if the system is to be secured the designer must consider the possible vulnerabilities of these techniques that could be exploited by an attacker. Advanced concepts of secure location systems are not within the scope of this paper but are covered in the literature (Arisar and Kemp, 2011; Boukerche et al., 2008; Clulow et al., 2006).

Figure 2 Components of localisation systems (see online version for colours)

Recent work in Abu-Mahfouz and Hancke (2012) investigates the security issues of these different approaches and techniques and shows that distance bounding approaches that are based on RTT and follow the four principles proposed by Clulow et al. (2006) could achieve secure distance estimation between two nodes.

3.2 Position computation

This component is responsible for computing the position of a node based on available information about the distance estimated from the previous component and position of references. Recognised techniques used in this component include triangulation, trilateration and multilateration (Cheng et al., 2005; Tian et al., 2007; Liu et al., 2006). In the triangulation technique, an unknown measure AoA of at least three beacons then uses the simple geometric relationships to estimate its position. One potential problem of the AoA approach is the expense of equipment to obtain precise angle estimates (Priyantha et al., 2003). Trilateration also uses the geometry of a triangle to estimate nodes’ position. However, instead of using AoA, it uses the location of and the distance to at least three beacons. The multilateration technique estimates location by solving the mathematical intersection of multiple hyperbolas (Srinivasan and Wu, 2008); it is also based on the location of and the distance to three or more beacons.

3.3 Localisation algorithm

This is the main component of a localisation system. It determines how the available information will be manipulated to enable most or all of the nodes of the WSN to estimate their position. These algorithms can be centralised (global) or distributed. The centralised algorithms (Nguyen et al., 2005; Shang et al., 2004; Biswas and Ye, 2004) are powerful and estimate the nodes’ position with high accuracy. However, they have a high communication and computational requirement, which is usually not available in WSNs. To reduce the communication overhead, various distributed localisation algorithms have been proposed, which decompose the global estimation system into sub-systems and then iterate over these sub-systems. Several iterative techniques have been followed. For instance, Niculescu and Nath (2001) uses references’ location information and local computation to localise unknown nodes iteratively. In this technique, the node estimates its position locally after getting some information from its neighbour nodes that know their location, which is sometimes called references. After the node knows its position, it could act as a reference for other nodes that do not know their position yet. Lim and Hou (2005) uses shortest-path approximation to the reference node to approximate Euclidean distances. The third technique uses local refinement (Costa et al., 2006), where each device starts locally estimating its location from measured ranges to its neighbours, then the devices successively refine their location estimates effectively finding a solution to a global optimisation problem that uses all ranges measured between neighbours. The disadvantage of iterative techniques is the effect of error propagation and accumulation, which is less prominent in centralised algorithms.
3.4 Information fusion technique

Several information fusion techniques have been used by various localisation algorithms to attain certain objectives, such as enhancing the accuracy of position estimation, reducing the required communication and computational requirements and saving energy. Several information fusion techniques will be explained in Section 4.

3.5 Localised algorithm

A localised algorithm is a special type of distributed algorithm in which only a subset of nodes in the WSNs is invoked for the position estimation task. Localised algorithms are explained in more details in Section 5.

4 Information fusion techniques for location discovery

Certain information fusion techniques (such as Bayesian inference, maximum likelihood, least squares (LS), moving average filter, Kalman filter, particle filter and occupancy grid) have been used by localisation algorithms to enhance the performance of the position discovery process. This section will explain these techniques briefly and mention some localisation algorithms using them. More details about these techniques can be found in Nakamura et al. (2007).

- **Bayesian inference**: Bayesian inference offers a formal way to combine evidence according to the rules of probability theory. The uncertainty of systems can be represented in terms of conditional probability, which estimates the degree of belief in the [0,1] interval, where 1 represents absolute belief while 0 represents absolute disbelief. Sichitiu and Ramadurai (2004) propose a localisation algorithm using Bayesian inference to process information from one mobile beacon. The unknown node uses the beacon’s position and the RSS measurement to construct a constraint on its position estimate, and then it applies Bayesian inference to compute its new position estimate.

- **Maximum likelihood**: When the state being estimated is not the outcome of a random variable, then the maximum-likelihood estimation (MLE) technique can be used. The likelihood function can be defined as the probability density function of the observation sequence given the true value of a certain state. Several localisation algorithms (Abu-Mahfouz and Hancke, 2013; Liu et al., 2008; Capkun and Hubaux, 2006; Savvides, et al., 2001; Fang et al., 2005; Lieckfeldt et al., 2008; Patwari et al., 2003) use MLE to estimate the position of unknown nodes by minimising the difference between the measured distances and the estimated distances, assuming that the unknown node has an adequate number of beacons. The minimum mean square estimate (MMSE) from a set of distance measurements can be used to find the MLE of the unknown node’s position.

- **Least squares**: The LS technique is a mathematical optimisation technique that searches for a function that best fits a set of input measurements. This is achieved by minimising the sum of the square error between points generated by the function and the input measurements (Nakamura et al., 2007). Li et al. (2005) propose the use of least median squares (LMS) as an improvement over LS for achieving robustness to attacks. However, LS outperforms LMS in the absence of attacks. In view of the measurement noise and error propagation that are introduced by the iterative techniques, Liu et al. (2006) propose a robust least-square (RLS) technique for localisation, which considers the error at each iteration. The authors show that RLS is more stable than LS in the presence of measurement noise.

- **Kalman filter**: The Kalman filter was proposed by Kalman (1960). Since then, it has been the subject of extensive research and applications. The Kalman filter is a set of mathematical equations that provides an efficient computational solution of the least-squares method. The filter is a popular information fusion method, because it supports estimations of desired states even when the precise nature of the modelled system is unknown. The Kalman filter technique has been applied in several distance estimation and location-discovery algorithms (Hongyang et al., 2005; Olson et al., 2006; Savvides et al., 2003; Gasparri et al., 2009).

- **Particle filter**: The particle filter is recursive implementations of statistical signal processing known as sequential Monte Carlo methods (Doucet et al., 2001). The Kalman filter provides an effective solution to the linear Gaussian filtering problem. However, for a non-linear model, or non-Gaussian noise, the particle filter should be used. Nordlund et al. (2002) show that particle filters allow great flexibility when addressing the problem of positioning, and they can be used in non-linear and non-Gaussian applications. Several algorithms (Hu and Evans, 2004; Gustafsson and Gunnarsson, 2003; Gustafsson et al., 2002; Miguez and Artes-Rodriguez, 2006; Ding et al., 2013) use the particle filter for refinement of node position estimates or for obtaining node location.

- **Moving average filter**: The moving average filter (Smith, 1999) is a very simple filter to use and to understand, making it the most commonly used in digital signal processing. This filter is also optimal for reducing random noise while retaining a sharp step response. Blumenthal et al. (2006) propose a new distance estimation method using an exponentially weighted moving average filter to flatten the resulting sequence of distance estimates and to filter outliers. The filter uses multiplying factors to give different weights, which change exponentially, to different beacon locations based on the estimated distances.
• Occupancy grid: The occupancy grid is a multidimensional random field that maintains stochastic estimates of the occupancy state of the cells in a spatial lattice (Elfes, 1989). The basic idea of the occupancy grid is to represent a multidimensional map of the environment as evenly square or cubic cells, which have random variables indicating the presence of an obstacle at that cell. The cell’s probability of being occupied can be computed using Bayesian theory or fuzzy set theory (Ribo and Pinz, 2001), based on information provided by several sources. Wongngamnit and Angluin (2001) propose a new robot localisation algorithm using the occupancy grid concept. Readers can find more information about using the occupancy grid for positioning estimation in Schiele and Crowley (1993).

Information fusion can play two roles in the localisation algorithms: a supporting role and a leading role. In the supporting role, information fusion acts as a tool to assist the localisation algorithms, by using one of the information fusion techniques, explained early, to assist in the location discovery. In the leading role, the localisation algorithms are designed to support an information fusion application. One or multiple information fusion techniques may be executed to accomplish the application’s objectives. Rather, these techniques are responsible for guiding the location-discovery process and the fusion process simultaneously. This means the localisation algorithm should be designed with two objectives: location discovery and achieving information fusion. The following section shows the impact of using localised algorithms to enhance the performance of information fusion and make information fusion play a leading role in localisation algorithms.

5 Localised information fusion algorithms

The basic idea of using localised algorithms is not only to request and process information to estimate position locally, but also to use data only from nodes that are more likely to contribute to rapid and accurate formation of the final position estimate (Meguerdichian et al., 2001). In other words, the localised algorithms used for location discovery should consider the following three conditions:

• First, request and process information with regard to the localisation algorithm only locally; i.e., report the location of the unknown but do not send the raw data to a centralised entity for processing.

• Second, only a subset of nodes takes part in the position estimation process.

• Third, only the references that are most likely to contribute to accurate position computation of an unknown are selected.

Developing practical localised information fusion algorithms for WSNs is very important and a challenging task for a number of reasons (Meguerdichian et al., 2001):

The communication delay in WSNs is significantly greater than that of other traditional networks, since communication consumes more energy than sensing and computation; dynamic changes occur in WSNs owing to nodes dying or extra deployment; and nodes are not always able to participate in every task (e.g., owing to lack of energy, obstacles, etc.). A WSN is resource constrained in terms of resources such as processing, communication and energy. Finally, security issues might require that only a subset of nodes take part in a task, thereby simplifying the task of ensuring that all nodes participating in the process are authentic.

Localised location discovery requires only the position estimates from a subset of the unknown’s neighbours, and need not involve all other nodes in the network. This makes the system-wide location-discovery task a good candidate for distributed algorithms. However, not all distributed localisation algorithms proposed in the literature can be considered as localised algorithms. In fact, most of these algorithms satisfy only the first two of the three conditions mentioned earlier. To satisfy the third condition, localised information fusion algorithms should select only references that are most likely to contribute to accurate position estimates (i.e., use only a subset of available references to compute the position instead of using all of them). In this type of algorithm, information fusion plays a leading role, since it not only assists in the position estimation process, but also accomplishes fusion’s objectives (such as energy saving and accuracy improvement).

Several localisation algorithms rely on using all or most of the available references to enhance their performance. They are based on the assumption that using more references could enhance the accuracy of estimation. However, to implement an efficient localisation algorithm for WSNs, this assumption should be reconsidered for the following reasons:

• The complexity of computation of a localisation algorithm increases in proportion to the number of references used (Lieckfeldt et al., 2008), so more references require more computation, more memory space and more energy consumption. A resource-constrained network such as WSN, however, needs to reduce the number of actively participating references as far as possible.

• The validity of the assumption could be compromised in a ‘hostile’ environment. One or more malicious nodes could deliberately provide incorrect location information to mislead other nodes. Preventing such types of malicious node from taking part in the localisation process will lead to more accurate estimation of position than when all the available references are used. Furthermore, from a security and privacy perspective, only a subset of nodes should take part in a task.

• The availability of a high number of references is a critical issue that cannot be guaranteed in WSNs for two main reasons: The first is the dynamic changes in
Localised information fusion techniques for location discovery in wireless sensor networks

WSNs owing to nodes dying or nodes moving. The second is that in WSNs it is not realistic to expect that all nodes will always be able to participate in every task (owing, for example, to lack of energy or the existence of obstacles).

- Estimates of location are based on one type of information fusion that combines complementary data to draw inferences. In other words, a node can fuse the location of and the measured distance to the neighbouring references to obtain its position. However, when the amount of additional incorrect data outweighs the amount of additional correct data, this can reduce the overall performance of the fusion process (Dasarathy, 2000).

- Distance-measurement techniques are all subject to errors. In a noisy environment, the position estimation will be more accurate if the node excludes those references that could bias the estimate towards an inaccurate location.

For these reasons, to enable unknown nodes to estimate their own location, it is desirable to select those references (i.e., subset of references) that could contribute more to accuracy, rather than using all the available references. Different approaches have been used to select a subset of references. The purpose is to show the impact of using these approaches to enhance the performance of localisation systems. Therefore, we will analyse a number of existing approaches and highlight their merits and weaknesses.

5.1 Nearest references

This is a very simple approach, which is based on choosing the nearest references as a subset to estimate a node’s position, assuming that the estimation error would be higher for distant references than for near ones. This approach could improve the accuracy of position estimation in WSNs. Assuming that the node will use four reference nodes to estimate its position as shown in Figure 3, node \( \{0\} \) will select only the nearest four reference nodes \( \{1, 3, 6, 7\} \) instead of selecting all the references \( \{1, 2, 3, 4, 5, 6, 7, 8\} \).

![Figure 3](image)

Cheng et al. (2005) propose a localisation algorithm called APS (Near-3), which is a modification of the original ad-hoc positioning system (APS) (Niculescu and Nath, 2001), which considers all the available beacons during the position estimation. The new, improved APS algorithm simply chooses the nearest three beacons to the unknown node inside the original APS computation (i.e., the triangulation mechanism and least-square method) to estimate the unknown node position. The simple heuristic used to select the best beacons requires much fewer communication overheads than to the original APS approach. Blumenthal et al. (2005, 2006) assign a different weight to each reference, depending on its estimated distance from the unknown node, with a higher weight to the near references. However, these algorithms can be modified to select a subset of weighted references by assigning a weight equal to zero for distant references. This approach assumes that the estimation error would be higher for distant references than for near ones and that the estimation error comes only from the distance measurement and ignores neighbour location error (because it only uses beacons that have no, or low, location error). Logically, if near references with location estimation errors are to be used, this assumption will not be valid and distant references could make a better contribution to position estimates than near ones.

5.2 Low-error references

As shown in Figure 4, localisation error results mainly from two sources: location error \( (\varepsilon_j^l) \), which is the error in neighbouring nodes’ position, and distance error \( (\varepsilon_{ij}^d) \), which is the error in the distance measurement. The readers can be referred to Zheng et al. (2012) and He et al. (2013) for detail error analysis and modelling of range-based localisation algorithms in WSN. Iterative techniques that may be used by localisation algorithms propagate this error, and so references that have large errors contaminate their neighbours’ location estimate. Using a reliable subset of references that consists of references with a low-error rate will prevent this type of contamination.

![Figure 4](image)

This technique has been used by Liu et al. (2006), where an unknown node computes the total error of its neighbour references, which is the sum of the location error and distance error. Then, it ranks references in an ascending order based on their error. Finally, it selects references with an error below a certain threshold and discards the others.
Sinha and Chowdhury (2007) propose that a localisation algorithm should choose a subset of three references in such a way that the error in the estimated location is within a certain limit. However, this algorithm requires high computational complexity. Selecting references in Kaplan (2006) is also based on this approach.

5.3 Malicious node removal

As shown in Figure 5, an attacker may provide an incorrect location reference to unknown nodes, which will then estimate their locations incorrectly. The malicious node removal approach aims to keep as many benign location references as possible, while the malicious ones are removed, resulting in a more accurate position estimation.

Liu et al. (2008) investigated two types of attack-resistant techniques to target malicious attacks against range-based location discovery in WSNs. In the first technique, the unknown nodes defeat malicious attacks by checking the consistency of references and then removing the inconsistent malicious references. This technique starts by using the entire set of references and then it gradually removes the most suspicious references till it reaches a certain level of consistency, which depends on the measurement error of an estimated location. The authors developed an incremental MMSE approach to reduce the computation cost, but it increases the size of the required memory.

The second technique is called voting-based location estimation (Figure 6), which quantises the deployment field into a grid of cells, and then the unknown node determines how likely it is to be correct in each cell, based on each reference. After the unknown node has processed all references, it chooses the cell(s) with the highest vote, and uses its (their) geometric centroid as the estimated location of the sensor node. However, specifying the voting by each reference at each cell of the grid requires a high computation cost. Liu et al. (2005), Abu-Mahfouz and Hancke (2013), Srinivasan et al. (2011) and Niculescu and Nath (2008) also follow malicious node removal approach to develop their secure localisation algorithms.

5.4 Consistency of references

This approach selects a subset of references based on their consistency with each other and excludes the inconsistent ones to increase the robustness and accuracy of the location estimate. One of the techniques to find the degree of consistency of each reference is to find the reference location error with respect to other references. This is the sum of the squared differences between the calculated distance and the estimated distance from one reference to the rest of them.

Albowicz et al. (2001) propose a localisation algorithm for choosing a reliable subset of references based on a reference consistency approach. The algorithm starts when the unknown node gathers information from neighbour references, which includes their degree of consistency (in Albowicz et al. (2001) termed ‘residual value’), and then the unknown node chooses only those references with the highest degree of consistency to estimate its location. While most of the unknown nodes should manage to get their position estimate, only the most accurate should extend system coverage and become references, to prevent incorrect convergence and divergence. Liu et al. (2008) also use this approach to identifying the malicious references.

5.5 Impact of geometry

This approach excludes insignificant references from participating in the localisation estimate, based on the geometry of references. Geometry could have a greater impact on accuracy of localisation than distance between references and unknowns. The Cramer-Rao-Lower-Bound (CRLB), which was defined by Patwari et al. (2003), can be used to specify the impact of geometry to quantify and compare the contribution of each reference to the accuracy of localisation and then to be able to choose a subset of references that contribute most to the accuracy. Figure 7 shows the impact of geometry in the accuracy in terms of CRLB. The figure shows the CRLB for two beacons.
Localised information fusion techniques for location discovery in wireless sensor networks

movement along a linear vector and a circular vector, respectively. After a certain distance, movement along the circular vector increases the CRLB dramatically when compared with the movement along a linear vector.

Figure 7 The impact of movement along a linear and a circular vector on the CRLB

The local-CRLB algorithm, which is proposed in Lieckfeldt et al. (2008), considers the impact of geometry. Local-CRLB starts when an unknown broadcasts a request for localisation. The neighbour references receiving the request estimate their distance to the unknown, which can be used in addition to the CRLB to assign beacons a probability of response. Responses, which include the originator’s address, location and distance estimate to the unknown, are broadcast. Subsequent beacons can use the additional information provided by the former responses. Local-CRLB constitutes a significant improvement over the algorithms selecting the nearest beacons as a subset. However, this algorithm assumes ideal estimation of distances, which is a strong assumption that would never be available in real-life applications. Lieckfeldt et al. (2008) investigate the Local-CRLB algorithm by considering energy consumption and impact on accuracy of localisation, using a maximum-likelihood estimator (MLE).

5.6 Noisy distance estimate

In the realistic case, the distance estimate is corrupted by noise and so localisation algorithms using only a distance estimate (e.g., those based on the nearest-references approach) to select neighbouring references could tend to select references whose estimate distance is shorter than the true distance. This approach considers a noisy distance estimate to remove bias from location estimates even in high-noise environments. Costa et al. (2006) propose a localisation algorithm called distributed weighted multidimensional scaling (dwMDS). dwMDS selects a subset of references based on a noisy RSS distance estimate and small neighbourhoods to avoid the biasing effect of a noisy environment. The proposed algorithm consists of two steps. In the first step, it finds the estimated node location based only on a distance estimate. In the second step, it excludes neighbours with a high biasing effect to construct a subset of references that require fewer iterations to converge to an accurate position estimate. Han et al. (2010) modify the dwMDS algorithm by simplifying the computation and reducing the processing time. Moore et al. (2004) also used this approach.

6 Comparison of the analysed approaches

Each of these approaches has advantages and disadvantages and it is not possible to consider one of them as the best approach for every application, scenario or network. The selection of one of these approaches to be implemented in WSNs is a little more complicated task because of resource limitations. When deciding which approach will be used, several issues should be considered, such as available resources, security level, computational cost, time of convergence and accuracy level. For example, if the designer would like to use minimal resources and is concerned about the execution time and computational cost, then the nearest-references approach is a possible choice. If the localisation algorithm is to be used for WSNs in a hostile environment, then the security level is an important issue and so malicious node removal and references consistency can be considered. The noisy distance estimate approach can be selected for WSNs with high noise to avoid the biasing effect of a noisy environment. A designer who would like to estimate position with high accuracy could choose one of the following approaches: the low-error references or the noisy distance estimate approach. However, one who is also looking for lower time of convergence could select the low-error references approach. On the other hand, the designer should also consider the limitations of each approach. For instance, the nearest references approach is very simple but cannot achieve a high level of accuracy compared with other approaches. The malicious node removal and references consistency approaches require higher computational cost, and the noisy distance estimate approach requires higher time of convergence.

A comparative summary is provided in Table 1. This table highlights some of the advantages and disadvantages of the analysed approaches. The last two fields of this table (targets and limitations) could be used as a guideline to help the designer to select an applicable approach that would be more suitable for his specific system requirements. Targets represent the issues that can be achieved using the corresponding approach, while limitations indicate the issues that cannot be achieved (or not completely fulfilled).
Table 1  Comparison of the analysed approaches

<table>
<thead>
<tr>
<th>Approach</th>
<th>Advantages</th>
<th>Disadvantages</th>
<th>Targets</th>
<th>Limitations</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nearest references</td>
<td>• Very simple</td>
<td>• Does not consider references’ location error</td>
<td>• Resources usage</td>
<td>• Security level</td>
</tr>
<tr>
<td></td>
<td>• Low computation</td>
<td></td>
<td>• Computation cost</td>
<td>• Accuracy level</td>
</tr>
<tr>
<td></td>
<td>• Few references</td>
<td></td>
<td>• Convergence time</td>
<td>• Noise level</td>
</tr>
<tr>
<td>Low-error references</td>
<td>• Accuracy</td>
<td>• Computationally intensive</td>
<td>• Accuracy level</td>
<td>• Computation cost</td>
</tr>
<tr>
<td></td>
<td>• Few references</td>
<td></td>
<td>• Convergence time</td>
<td>• Security level</td>
</tr>
<tr>
<td>Malicious node removal</td>
<td>• Works in hostile environment</td>
<td>• Computationally intensive</td>
<td>• Security level</td>
<td>• Resources usage</td>
</tr>
<tr>
<td></td>
<td>• Accuracy</td>
<td>• Large memory</td>
<td>• Accuracy level</td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Few references</td>
<td></td>
<td>• Convergence time</td>
<td></td>
</tr>
<tr>
<td>References consistency</td>
<td>• Works in hostile environment</td>
<td>• Computes the consistency of each reference</td>
<td>• Security level</td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Accuracy</td>
<td></td>
<td>• Accuracy level</td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Few references</td>
<td></td>
<td>• Convergence time</td>
<td></td>
</tr>
<tr>
<td>Impact of geometry</td>
<td>• Accuracy</td>
<td>• Assumes ideal estimation of distances</td>
<td>• Accuracy level</td>
<td>• Security level</td>
</tr>
<tr>
<td></td>
<td>• Few references</td>
<td></td>
<td>• Convergence time</td>
<td>• Noise level</td>
</tr>
<tr>
<td>Noisy distance estimate</td>
<td>• Works in noisy environment</td>
<td>• Elimination criteria</td>
<td>• Accuracy level</td>
<td>• Security level</td>
</tr>
<tr>
<td></td>
<td>• Accuracy</td>
<td></td>
<td>• Convergence time</td>
<td></td>
</tr>
</tbody>
</table>

Figure 8  Information fusion levels (see online version for colours)

In general, the localised algorithms, which use these approaches, could outperform the other distributed localisation algorithms because they aim to achieve the following objectives: First, reduce the computational cost, required resources, communication load and convergence time. Second, improve the position estimation accuracy. Third, achieve more energy saving and so prolong the network lifetime. These objectives are also targeted by information fusion techniques; in other words, using these localised algorithms could enhance the performance of information fusion techniques and cause information fusion to play a leading role in the localisation algorithms by guiding the location-discovery process simultaneously with the fusion process.

As mentioned earlier in Section 4, information fusion can play two roles in the localisation algorithms: a supporting role and a leading role. To distinguish between these two roles in this discussion, the information fusion used by localisation algorithms will be classified into three levels, based on the objectives that can be achieved as shown in Figure 8. In level 1, localisation algorithms use one of the information fusion techniques, as explained in Section 4, in the position computation. The objective of information fusion on this level is to combine complementary data to allow inferences: the node fuses the location of, and the measured distance to, at least three references to determine its position. On this level, information fusion plays only a supporting role to assist in the location discovery. In level 2, information fusion starts to play a leading role by achieving some of the information fusion objectives: accuracy, robustness, fewer computations and security. However, reducing the communication
overhead is not targeted by this level of information fusion. In level 3, information fusion also plays a leading role by achieving several objectives. The main objective of information fusion on this level is to reduce the communication overhead to enhance the energy efficiency of the localisation algorithm. Developing an efficient localisation algorithm that employs the three levels of information fusion will enhance the performance of this algorithm and assist in achieving several design objectives.

7 Conclusion

In this paper, we analyse localisation systems from the information fusion perspective. Information fusion plays a crucial role in location discovery of nodes in WSNs, and because of the importance of node location for many WSN applications and techniques, several localisation algorithms have been proposed to address this challenging task. In this paper, we review the information fusion techniques used by a number of localisation algorithms to assist them in the location-discovery process (i.e., information fusion is used in a supporting role). To distinguish between the distributed and localised algorithms, we delineate three conditions that should be satisfied by localised algorithms and then we analyse a number of approaches used by localised information fusion algorithms, highlighting some of their strength and weaknesses, briefly compare them and show how designers could decide which approach should be followed to implement their localised algorithms. We also show that using these localised algorithms could make information fusion play a leading role and enhance the performance of the entire localisation algorithm. Further evaluation of these approaches using either simulation tools, such as network simulator (ns-2) or WSN testbed (Abu-Mahfouz et al., 2012; Dludla et al., 2013), is left for future work.

This paper raises a number of important issues that are worth keeping in mind: First, since WSNs are considered as resource-constrained networks, we emphasise that localised information fusion algorithms should be used to provide an accurate position estimate at reduced cost. Second, localised algorithms should satisfy three conditions:

- request and process information with regard to the localisation algorithm only locally
- only a subset of nodes takes part in the position estimation process
- only the references that are most likely to contribute to accurate position computation of an unknown are selected.

Third, in these localised algorithms, we focus on the impact and significance of selecting only references that are more likely to contribute to an accurate position estimate. Fourth, a number of design parameters and certain requirements play a vital role in selecting the proper approach, such as resources usage, computation cost, convergence time, accuracy level, security level and noise level. Fifth, selecting a certain approach could not always give the best position estimate (for instance, because of its simplicity) and so selecting another approach with a more sophisticated technique for selecting a subset of references is required. Finally, despite significant research into the development of localisation systems, developing a localised information fusion algorithm for node location discovery in WSNs based on carefully selecting a sufficient number of the best references to enhance the accuracy of position estimate at reduced cost is still a challenge and an open area for future investigation.

References


Localised information fusion techniques for location discovery in wireless sensor networks


