Swarm Intelligence based Detection of Malicious Beacon Node for Secure Localization in Wireless Sensor Networks

S. Qureshi, A. Asar, A. Rehman, and A. Baseer

University of Engineering and Technology Peshawar, Pakistan.

Corresponding Author: S. Qureshi

Abstract
Wireless sensor networks (WSNs), made of collection of small, cheap, low power sensors, have diverse applications in several domains of life. Most of their applications demand for spatial position of sensors. Beacon nodes which already know their location are mostly used to help other sensor nodes to locate their position. However localization of sensors in a hostile environment is a crucial issue. The compromised beacon nodes can severely affect the process of localization. The aim of this research is to propose an algorithm to detect malicious beacon nodes based on swarm of intelligent water drops to provide secure localization in wireless sensor networks. The paper also critically reviews existing secure localization schemes.

Keywords: security, localization, swarm intelligence, malicious node.

INTRODUCTION
A wireless sensor network (WSN), composed of large number of small, low cost, independent, intelligent sensors, is able to monitor its surroundings. Some of its applications are in the field of health (Stanislaw Rost and Hari Balakrishnan, 2006), military operations (G. Simon et al., 2004), natural disaster assuagement (M. Castillo-Effen et al., 2004), traffic management (M. Yousaf et al., 2010), monitoring volcanic eruptions (Geoffrey–Allen et al., 2005). These sensor nodes have limited memory, computational capacity and battery. Owning to their applications and operational challenges, WSNs suffer from security threats and are vulnerable to be attacked by adversaries.

Considering these challenges, security protocols for WSN should be designed carefully. Conventional security mechanisms cannot be directly applied to resource constrained sensor networks (Q, Zhang et al., 2008). Nowadays security issues in wireless sensors network have gathered attention of many researchers, secure localization being a key issue. As, many applications (e.g. target tracking) depend on correct location of nodes, researchers have proposed several localization techniques which can minimize localization anomalies. However this area of research is still under investigation and researchers are trying to develop models for elimination of malicious nodes to ensure correct localization.

The rest of the paper is organized as follows. Section II provides basic configuration of Wireless Sensor Networks. In section III several localization schemes have been briefly elaborated. Security threats to the sensors localization have been illustrated in section IV. Critical analysis and review of existing secure localization methods is available in section V. Section VI introduces the swarm intelligence and some of its applications in wireless sensor networks. In section VII we propose IWD based detection of malicious beacon nodes in wireless sensor networks and future work. Section VIII provides concluding remarks.

General Architecture of WSN
WSN is a decentralized and collaborative network which accomplishes its task by communicating locally with neighbor nodes. In WSN each individual node comprises of sensing, processing and communication units (transceiver), memory and power supply (L. Yong-Min et al., 2009). Due to restricted size and battery the processing, data storage and communication capabilities are limited.

Figure: 1: Sensor Illustrated
Sensors are generally deployed in hundreds and thousands of numbers to collect desired data. Mostly these nodes communicate with each other via wireless medium using radio antennas, although
infrared, laser communication can also be used. Usually only a few sink nodes or gateways are responsible for collecting data from other nodes and communicating with user via internet. Sensor nodes can communicate with direct nodes, if a node has to communicate beyond its coverage, it must be via multi hop route transmission through intermediate nodes. Unlike traditional networks, WSNs are application dependant and should be designed considering their environment and objectives (L. Yong-Min et al., 2009).

**Localization in WSN**
The process of finding spatial position either relatively or absolutely is known as localization. Localization of sensors is essential in most of the WSN applications for instance target tracking, rescue operations, geographical routing protocols (B. Karp and H. Kung, 2000), (N. Chaki, 2010), location based authentication (N. Sastry et al., 2003) and to setup paths of sensors to track the sensed data to the sink node or base station. Also the base station needs to know the source of received data. In wireless sensor networks localization is a key issue due to resource-constrained nature of the sensors. GPS is a costly solution and also communication signals are weak for indoor applications (Z. Liang et al., 2010). Different approaches have been presented for localization; the motivation is to achieve better accuracy. These approaches can be broadly classified as:

- **GPS-based and GPS free**
  GPS free schemes are independent of global positioning system. They employ the distances between nodes to calculate the positions relative to local network. The alternative is to use GPS receiver on every node, with result of absolute spatial position and large accuracy. However, to install GPS receiver on every node is a costly solution so the researchers suggested to use only few nodes with GPS called anchors/beacons which would help other nodes in network to estimate their positions.

- **Centralized and distributed Schemes**
  In localization techniques with centralized approach, there is one main center point (usually sink node or Base station) which receives information from the entire group of sensors and then after calculating their positions send back results to the sensors. In this way the energy of sensors is saved since all the computation for position estimation is performed by the base station. While in distributed schemes sensors themselves have to estimate their positions cooperating with beacon nodes resulting in reduced communication (saving energy) with center node. However the results are not that accurate and the error produced in calculating a position will propagate and will effect successive location estimations.

- **Range-based and Range-free**
  Range free computations involve locations of anchors and hop-distances between nodes and based on these, location of a sensor is estimated. Thus connectivity of nodes is required. Although range free techniques hold less accuracy, they are useful in terms of less hardware and energy demand and thus cost-efficient. In Range-based localization, nodes positions are calculated relative to their neighbor nodes which mostly know their location through GPS. The absolute distance between the two nodes is measured (thus called range-based) by using any characteristics of received communication signal. These characteristics are:
  - Direction/Angle of arrival (AOA): Range measurement between the two anchor nodes is first obtained then by using directional antenna, angle to the anchor nodes are measured. Using this information position of unknown node is determined (R. Feng and M. L. Siciliano, 2006).
  - Time of arrival (TOA): The difference between transmitting and receiving time of a communication signal is determined and using light velocity distance is calculated, assuming perfect time synchronization between sender and receiver (Y. Zhang et al., 2005).
  - Time difference of arrival (TDOA): The time synchronization problem in TOA is resolved by TDOA. The time difference between the signal (transmitted by unknown), received by minimum three receivers is noted and then used for calculating nodes position (F. Gustafsson and F. Gunnarsson, 2003).
  - Received Signal Strength: While knowing the transmitted and received signal strength, range between two nodes can be measured, however due to environmental hindrances, results are affected.

Range measurement is the first step in range-based schemes. The second step is to use any localization algorithm to calculate position e.g. **triangulation, trilateration or multilateration**.

- **Fine grained and coarse grained**
  Methods that do not find and utilize absolute distance using any characteristic features of received signal result in coarse-grained localization. It can be inferred that range-free techniques provide proximity coarse grained results. On the other hand methods employing features of received signal derive fine-grained results.

The issue of localization in WSN is being widely searched by researchers these days. Some authors have presented methods with addition of new concepts like (Y. Li et al., 2009), (M. Shaifur Rahman et al., 2009) while some are enhancements and modifications of previous ones to improve localization accuracy. Table I represents summary of the some localization schemes.
Table I. Localization Schemes

<table>
<thead>
<tr>
<th>Localization Schemes</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>Centralized</td>
<td>(Q. Zang et al., 2008), (Z. Liang and X. Liu, 2010)</td>
</tr>
<tr>
<td>Distributed</td>
<td>(J.P. Montillet et al., citeseer), (Y. Ding et al., 2007), (C. Xue et al., 2008)</td>
</tr>
<tr>
<td>GPS-free</td>
<td>(N. Bulusu et al., 2000), (S. Capkun et al., 2001)</td>
</tr>
<tr>
<td>Range-based</td>
<td>(F. Reichenbach et al., 2006), (J. Blumenthal et al., 2005), (Y. Forghani, 2008)</td>
</tr>
<tr>
<td>Range-free</td>
<td>(L. Nian qiang and L. Ping, 2008), (T. He et al., 2003)</td>
</tr>
<tr>
<td>Coarse-grained</td>
<td>(N. Bulusu et al., 2000)</td>
</tr>
<tr>
<td>Fine-grained</td>
<td>(A. Savvides et al., 2001)</td>
</tr>
<tr>
<td>Modified algorithm</td>
<td>(Q. Yufeng et al., 2010), (X. Bao et al., 2010)</td>
</tr>
<tr>
<td>Multiple/hybrid localization algorithm</td>
<td>(W. Huang et al., 2009), (S. Ganesh, 2008)</td>
</tr>
</tbody>
</table>

Security Threats to Localization

Wireless sensor networks due to their unique operational challenges are prone to localization attacks. These attacks can be launched either by external or internal adversaries. External adversary is an outside node, which does not have any authentication to be included in network. In order to secure network against external attacks, cryptography (encryption/decryption) (G. Yang et al., 2007) and key management methods are employed. External attacker can be a false node, which can present itself to be legitimate if it strikes valid signature. Also external attacker can corrupt or modify transmitted data (location) sent by the beacon nodes, thus transmitted data needs to be encrypted to ensure confidentiality prospect. However, beacon nodes can be captured and compromised by the intruder and cryptographic information can be stolen thus creating internal adversary. Due to misbehavior of compromised nodes wrong sensor locations will be estimated. These compromised beacon nodes can launch a number of attacks e.g. Sybil, Replay, Wormhole and link spoofing.

Sybil: A beacon node conceals itself as legitimate node claiming multiple identities and locations resulting in wrong location estimation by sensor nodes.

Replay: Malicious node first jams the signal from transmitter and then replays the same message to the receiver impersonating transmitter’s I.D and location.

Wormhole attack: Attack is launched by two or more colluding malicious nodes. The location information collected at one end is tunneled to other end of network, ensuing erroneous location information.

Link-spoofing Attack: The compromised beacon node can publicize fake connectivity link with non-neighbors (Y. C. Hu et al., 2003). Thus in range free localization techniques, where communication is multi-hop, the beacon signal can be routed to this compromised node which can then corrupt the data. Table II presents the Type & Proposed solutions for some of localization attacks.

Table II. Localization attacks and proposed solution

<table>
<thead>
<tr>
<th>Attack</th>
<th>Type</th>
<th>Proposed solution</th>
</tr>
</thead>
<tbody>
<tr>
<td>Message corruption/modification</td>
<td>External</td>
<td>Encryption e.g. (G. Yang et al., 2007)</td>
</tr>
<tr>
<td>Falsifying routing information to mislead location information</td>
<td>External</td>
<td>Cryptography e.g. SPINS (F. Hu et al., 2003)</td>
</tr>
<tr>
<td>False node</td>
<td>External</td>
<td>Trust and reputation mechanism(F. Mahmood, 2009)</td>
</tr>
<tr>
<td>Sybil</td>
<td>Internal</td>
<td>e.g. (D. Liu et al., 2005)</td>
</tr>
<tr>
<td>Wormhole</td>
<td>Internal</td>
<td>e.g. (Y. C. Hu et al., 2003)</td>
</tr>
<tr>
<td>Replay</td>
<td>Internal</td>
<td>e.g. (S. Misra et al., 2007)</td>
</tr>
<tr>
<td>Link spoofing</td>
<td>Internal</td>
<td>e.g. DRBTS (A. Srinivasan et al., 2006)</td>
</tr>
</tbody>
</table>

Since false external node will have no previous reputation value so its communication attempt will be rejected by the node. (Y. C. Hu et al., 2003) is resilient to wormhole attack. Link spoofing attacks can be detected by DRBTS because in DBRTS beacon node maintains reputation for neighbors and then exchange it with each other. The claim of malicious node being neighbor will be detected since it will not have any reputation in neighbor reputation table.

Secure localization

As elaborated in section IV localization process in WSN is liable to be attacked by the adversary in hostile environment and cryptographic methods are not enough to ensure correct localization. Therefore secure localization being a crucial issue has been searched over the past few years and researchers have proposed several solutions to improve it. However, research is continued to find optimum solution. To address the problem of localization researchers have come up with different approaches some of which are briefed as follows.

Robust Positioning: This approach assumes that there are malicious nodes in network but positioning
scheme should be such that it may tolerate this malicious behavior and still work for correct location discovery. This scheme comprises of statistical methods and separates out the erroneous information. Li et al in (Z. Li et al., 2005) introduced the theme of robust localization. They proposed two statistical methods for robust localization against attack in Wireless Sensor Networks. However these methods hold an assumption that benign nodes outnumber compromised nodes.

Liu et al presented a method in (D. Liu, 2008) to key out and bump off malicious information. The technique uses minimum mean squared estimation to calculate the location of sensor nodes. They presented a second technique for calculation of location which is voting based and is iteratively refined for resource-constrained sensor networks. The scheme is dependent on majority of benign nodes, so it would fail if the adversary launches more malicious nodes than benign ones.

Lazos and Capkun have designed a robust localization system (L. Lazos et al., 2005) which is robust to Sybil and wormhole attacks. It is a decentralized localization method that requires less no of beacons (locators). Although this method does not involve statistical methods, it accomplishes robust localization in presence of malicious nodes.

**Verification**: Another approach is based on the verification of the estimated position. Capkun et al in (S. Capkuan et al., 2005) have designed an algorithm that computes position of a sensor using verifiable multilateration process. Lazos and Capkun in (L. Lazos et al., 2005) have presented hybrid algorithm which utilizes the cryptographic methods for secure communication between sensors and locators and then devises an algorithm for location determination using verification. Locators verify the location of claimant by checking whether it claims to be outside the transmission range of locator. However authors of ROPE and SPINE have neglected the scenario where a number of locators themselves are compromised and that may collude. In (N. Sastry et al., 2003) N.Sastry et al present an algorithm named Echo which verifies the area of location of claimant. However the verification process sometimes may also be attacked by the adversaries. Authors in (Y. Zengt et al., 2009) provide brief overview of problems and solutions of secure location verification.

**Compromised Node Detection**: In (D. Liu et al., 2005) Liu et al have come up with more genuine approach to detect and revoke the malicious nodes. Three solutions have been presented in this regard. In (D. Liu et al., 2005) Liu at el propose a technique for identification and removal of malicious beacon nodes. This algorithm assumes key-distribution scheme for secure communication between the nodes. The detecting node dissembles target node by using a different detecting I.D and thus appears to be regular node. The detecting node on receiving the location reference from target node calculates distance between them using the signal received and location information sent by the target node. If the two are consistent with difference less than maximum distance error, it is considered benign otherwise malicious. Authors also proposed a technique to filter out locally replayed beacon signals by calculating RTT. They then developed revocation scheme which is based on the suspiciousness of a beacon node calculated at the base station through alert based scheme. However the detection scheme has a shortcoming. It has not considered the environmental affect. WSN are mostly deployed in an unattended environment, if there were worse weather changes, there would be much measurement error in distance and two distances may differ much larger than maximum measurement error. In this way many alerts would be reported to the base station which may increase suspiciousness of benign node and as a result, number of benign nodes would be revoked from the network. Srinivasan et al in (A. Srinivasan et al., 2006) have extended the work in (D. Liu et al., 2005). They introduced for the first time the concept of trust and reputation for removal of malicious beacons. They have proposed distributed reputation beacon trust system in which beacon nodes help sensor nodes to decide whether to use given beacon’s location information or not by maintaining reputation of their neighbor beacon nodes. This reputation is maintained in neighbor reputation table (NRT) at each beacon node by monitoring its neighborhood for misbehavior. In this work a decentralized approach is adopted i.e. the sensor nodes decide about beacon node’s trustworthiness based on local neighbor’s experience. Authors have shown the robustness of method through simulation. However, since WSNs pose unique challenges in terms of memory, computational capability and battery the reputation based mechanisms usually can add overhead. Also without being globally informed, it is usually hard to cope with local majority and collusion of nodes that have been compromised (Q. Zhang et al., 2008). Satyajayant Misra and associates in (S. Misra et al., 2007) proposed a scheme to detect and remove malicious beacon nodes. This technique uses a mobile verifier, which is sent to sensor network by the base station iteratively. For every iteration it obtains a number of location estimations from locators and performs hypothesis test to find the mean and variance of location estimations. If the mean and variance of error values fall near to 0 and \( \sigma^2 \), the locator is a benign one otherwise malicious. The technique is inherently resilient to replay and wormhole attacks. Authors have shown through simulations that their proposed method detects almost 80% of malicious locators while the false positives percentage is almost zero. However, this scheme has
some drawbacks. Mobile verifier needs preprogrammed paths to be stored to follow through the network. Due to restricted number, there is repetition of paths which weakens the algorithm’s effectiveness. An additional requirement is that the mobile verifier needs to be charged after every iteration which makes it unsuitable for most of the unattended wireless sensor network applications. Above all the whole scheme is dependent on single mobile verifier which is assumed to be uncompromised failing which, the whole scheme will collapse.

SWARM INTELLIGENCE

The expression swarm intelligence (SI) denotes the collective behavior of self regulating, decentralized systems. These systems consist of population of simple agents with no central authority to prescribe their behavior. The simple agents are un-intelligent and interact locally with each other and environment to produce an intelligent and complex behavior. Swarm intelligence exemplifies meta-heuristic approach to solve a problem.

A number of optimization algorithms have been modeled incorporating swarm behavior. Well known examples are ant colony optimization (ACO), particle swarm optimization (PSO), gravitational search algorithm (GSA) and recently explored intelligent water drops (IWD). Some of the applications of SI algorithm are described below.

- ACO: Ant colony optimization algorithm is modeled on the ants’ action-mechanism to find their food. Algorithm can be applied to the problem of finding the optimal path to achieve the goals. E.g. ACO is employed in (R. Muraleedharan and L.A. Osadciw, citeseex)
- PSO: Particle Swarm optimization is inspired by social behavior of bird flocking or fish schooling. It is a particle based search algorithm to find global best solution in an n-dimensional space. Each particle has its own best solution and global best solution among the particles then travel through search space and updates its best and global best valves by its position and velocity information. Thus through global search in space final best solution is obtained. One of applications of PSO is (K. S. Low et al., 2005),
- GSA: Gravitational search Algorithm is based on gravitational law and use of laws of Newtonian physics. The search space consists of masses as agents. The gravitational force is responsible for the moments of these agents towards the global heavier mass which corresponds to the best global solution of problem (E. Rashedi et al., 2009)
- IWD: Intelligent water drops algorithm introduced in 2007 by Shah Hosseini is swarm based optimization approach. The algorithm is nature inspired and mimics the procedural behavior of water drops and soils of the riverbed (H. Shah-Hosseini, 2009). In IWD algorithm collection of artificial water drops interact to find optimal way with lowest soil. IWD has been used to solve traveling salesman Problem (TSP), the n-queen puzzle (H. Shah-Hosseini, 2009), the multiple knapsack problem (H. Shah-Hosseini, 2009), Automatic Multilevel Thresholding (AMT) (H. Shah-Hosseini, 2009) and Air Robot Path planning (H. Duan et al., 2008).

Swarm Intelligence (IWD) based Detection of Malicious Beacon Nodes

Swarm Intelligence (IWD) based detection assumes same base model as in (D. Liu et al., 2005) & (A. Srinivasan et al., 2006). However incorporating SI, the detection method can be made more affective considering environmental factor. Previously different solutions were proposed to either tolerate the presence of malicious nodes or to detect and revocate them. Also individual solutions have been proposed against particular attacks.

IWD algorithm based approach can perform well even in worse environmental conditions because algorithm is based on relative comparison of errors. The distance-error values of neighbor beacons are compared to find the probability of trustworthiness or goodness of a node. IWD algorithm incorporates the method of natural water drops to the select the next location. Velocity of an IWD increases inversely to soil between its current and next location so the drop will gain velocity on the path with low soil. IWD prefers the path with low soil, so the probability of selecting a path with low soil is higher.

Probability of selecting a next position by IWD is calculated by

$$P(j) = \frac{P_{best}(i,j)}{\sum_{i} P_{best}(i,j)}$$  \hspace{1cm} (1)

This formula can be used in improving security and localization in WSN for the detection of malicious beacon node. Function of soil between two points i & j can be considered as the error of distance between two nodes. Every node can be considered as intelligent water drop which calculates the goodness of the other node. According to the formula, the node producing more error will give less ‘P’ value challenging its trustworthiness. The error between distances can be found by

$$Error = |d_{(ij)} \text{calculated} - d_{(ij)\text{actual}}|$$  \hspace{1cm} (2)

Where $d_{(ij)}$ calculated is the distance, determined by any range estimation technique (e.g. TOA) between two nodes i & j and $d_{(ij)\text{actual}}$ is the one which is found by estimating Euclidian distance between two locations.
Node i receives locations from nodes X, Y, Z 

\{
  for \ j = x \ y \ z 
    Estimate distance d(ij)calculated 
    Estimate distance d(ij)actual 
    Calculate Distance Error 
    Find Probability of trust = P 
    If 
      P < Threshold 
      Consider Node as malicious
\}

Fig 1. IWD based algorithm

If the P value is below carefully obtained threshold value, the trustworthiness of a node can be challenged.

The model above is based on optimum point selection by IWD, which is in our case the node with most malicious behavior. With this model some security attacks associated with WSN localization can be detected.

- **Sybil Attack:** If a node is compromised and it launches “Sybil attack” i.e. spoofs the identity of some other beacon node and its location, a considerable error will be generated between the actual and calculated distances. This error will be compared by the errors produced by other neighbor nodes. Since other nodes are benign (assumption), the error produced by this will be large as compared to others resulting in reducing the probability of its trustworthiness below threshold and thus will be caught.

- **Replay Attack:** If the internal adversary has launched replay attack, distance estimated by any range techniques will not be same as calculated by using its claimed location. The error produced will be further checked with its neighbors. If error is comparably higher than others, its P value generated will be low showing its maliciousness.

- **Wormhole Attack:** If the ID and location of a node has been tunneled to some other end, the range estimation distance method will point out the difference between calculated and actual distances. The error would be large as compared to neighbors, so the P value will be reduced.

- **False Node:** An external false node striking valid signature trying to disguise about its location, will be caught by error generated and compared to its neighbors.

Although the distances-estimated using TOA, TDOA, RSS, AOA are not accurate and have some errors, it seems as if all the nodes will produce lower P value but this is not the case. Since the group of neighbor nodes are in same environment facing same obstacles and weather, their error values will be close to each other, however the malicious node error will be larger than these and thus will be caught with lowest P value. Moreover, if there is more than one malicious node their P values will also be lower than threshold so there is potential & ability in algorithm to detect colluding compromised nodes as well. However, the identifiable number of malicious nodes is a future research direction.

In this model each node is able to judge other node based on its own personal local experience, this personal experience can be converted to the global solution using a center authority. The future work of this research will focus on following points

- Using P value to either count for reputation of a node or vote (positive or negative) for that node.
- Using a central authority to authenticate the maliciousness of a node.
- Development of a model to revocate malicious nodes.
- Developing simulation model and interpreting results.

**CONCLUSION**

In this paper, IWD based algorithm for detection of malicious beacon node has been presented. IWD has been used for the first time in area of Wireless Sensor Networks. Also the algorithm for the first time incorporates the environmental effects in wireless communication while detecting compromised beacon nodes. The algorithm has space and potential for development and improvement about which few notions are mentioned in future work. A brief and critical overview and comparison of existing secure localization schemes have also been presented. More over we have described the algorithm to cope with some threats associated with localization.

**REFERENCES**


