Research Article

A Replica Detection Scheme Based on the Deviation in Distance Traveled Sliding Window for Wireless Sensor Networks

Alekha Kumar Mishra, Asis Kumar Tripathy, Arun Kumar, and Ashok Kumar Turuk

1School of Computer Science and Engineering, VIT University, Tamil Nadu, India
2School of Information Technology and Engineering, VIT University, Tamil Nadu, India
3Department of Electrical & Computer Engineering, National University of Singapore, Singapore
4Department of Computer Science and Engineering, National Institute of Technology, Rourkela, India

Correspondence should be addressed to Alekha Kumar Mishra; alekha.mishra@vit.ac.in

Received 29 July 2016; Accepted 24 October 2016; Published 15 January 2017

1. Introduction

Sensor networks have been a topic of interest among the academia and industry due to their wide range applicability in various network environments. WSNs consist of a large number of tiny sensors, usually deployed densely in the target area to collect relevant data [1, 2]. Sensor nodes are resource-constrained due to their small size and this limits the ability of computation and communication. The most common applications of WSNs include habitat monitoring, border patrolling in military, traffic monitoring, and patient monitoring in healthcare [3, 4].

In MWSNs [5, 6], sensor nodes have additional mobility function to wander inside a target area. Mobile sensor nodes can provide accurate data compared to static nodes. The number of sensors required in MWSNs for covering a given area is quite lesser than static WSNs. Nevertheless, dense deployment of mobile nodes supports high reliability and load balancing. Despite all advantages, MWSNs possess a highly dynamic network topology due to mobility. Therefore, the challenges are multiplied compared to their static counterparts. The major challenges include communication, coverage, distributive cooperative control, and security [7, 8]. Security has always been a critical issue of concern in WSNs [9]. Taking the advantage of flexibility in deployment, an outsider can launch attacks at various levels of communication layers in WSNs. The list of attacks of high threat includes wormhole, sinkhole, desynchronization, hello flooding, and Sybil [10]. Node replica attack is a serious attack and difficult to handle in WSNs [11]. In this attack, an adversary captures a node and deploys one or more replicas of it in the network. The aim behind such attack is to take control over the...
network activities and launch insider attacks with the help of replicas. Therefore, detection of node replication attack should be done as early as possible. The replica detection schemes in MWSNs as reported in the literature are very few in numbers [12]. The schemes proposed by Ho et al. [13] and Deng et al. [14] rely on node’s location claim to detect replica. In these schemes, the received location claims are used to compute the node’s speed in the network. A replica is detected, when its speed exceeds the predefined maximum speed limit. However, in these schemes, BS performs the task of replica detection and therefore is exposed to single point of failure. These schemes also incur an additional communication overhead of sending node’s location claim to the BS.

In this paper, we propose a distributed replica detection scheme for MWSNs that uses time-distance broadcast of nodes to detect a replica. The proposed detection scheme is based on the fact that when a node $i$ and its replica coexist in the network, it would lead to deviation in the distance traveled by $i$ at any instance of time. In the proposed scheme, every node performs the role of an observer node for a number of nodes in the network. Observers are responsible for maintaining a sliding window of recent time-distance broadcast of nodes under observation. When a deviation is detected in the distance traveled by a node, this is recorded by the observer nodes. The degree of violation of a node is computed from the recorded deviations. When the degree of violation of a node is measured above a predefined threshold, it is detected as replica. The threshold measured using simulation experiments. The proposed scheme is fully distributed and found to have higher detection probability and an average communication overhead compared to the existing schemes such as EDD and MTLSD.

The rest of the paper is organized as follows: Section 2 provides an overview of the existing node replica detection schemes. Assumptions about the adversary and network are enlisted in Section 3. The proposed scheme is explained in Section 4. Analysis of the proposed scheme is presented in Section 5. Simulation results are detailed in Section 6, and Section 7 summarizes the concluding remark.

## 2. Related Works

The detection schemes for static WSNs are not applicable to MWSNs due to dynamic network topology. A detection mechanism in MWSNs must take the mobility of a node into account in order to detect replica [15]. The replica detection mechanisms in MWSNs as reported in the literature are described below.

The detection schemes Unary-Time-Location Storage and Exchange (UTLSE) and Multi-Time-Location Storage and Diffusion (MTLSD) proposed by Deng et al. [14] adopt time-location claim approach. Each node in UTLSE and MTLSD stores multiple instances of time-location claim of the tracked nodes. On meeting, the time-location schemes are exchanged between two trackers to verify the feasibility of location claims. When a conflict arises in the time-location verification process of a node, it is detected as replica.

A detection scheme using node’s speed is proposed by Ho et al. [13]. This scheme uses Sequential Probability Ratio Test to compute node’s speed. When a node arrives at a new location, it broadcasts its time-location claim to the neighbors. Neighbors forward the received claim to BS after successfully verifying the authenticity of the message. The BS is responsible for gathering time-location claims of the nodes and measures their speed. A node with a speed higher than the predefined speed limit is detected as replica by the BS.

A pairwise key establishment process to detect existence of replica is presented by Deng and Xiong [16]. The total number of pairwise keys established by a node is stored using Counting Bloom filter. The count of number of keys established is periodically sent to the BS by each node. The received Counting Bloom filters are updated at BS for each node in the network. When the number of keys established for a node exceeds the predefined threshold value, the node is detected as replica by the BS.

A single-hop based replica detection scheme is proposed by Sindhuja and Padmavathi [17]. In their work, the witness node selection method of single-hop replica detection is improved by using clonal selection algorithm. The best suitable witness for a node in its single-hop neighbor is selected using the clonal selection algorithm. The neighborhood fingerprint sharing and verification method is used to detect replica.

The Extremely Efficient Detection (XED) and Efficient Distributed Detection (EDD) schemes are proposed by Yu et al. [18, 19]. The detection of replica in XED is based on the exchange of a random number between each pair of nodes, which is also called a challenge. When the same pair of nodes meet each other at a later point of time, the challenge verification is performed. A node that fails the challenge verification process is detected as a replica. In EDD, a replica is detected based on the count of number of meetings between a pair of nodes. If the number of meetings of a node over a time interval exceeds the predefined threshold, then it is detected as replica.

In the detection schemes by Ho et al. [13] and Deng and Xiong [16], the replicas are detected by the BS. In these schemes, BS is overburdened with computation and replica detection tasks. There is also an additional overhead in the network for communicating between all the nodes and the BS. The XED mechanism is not resilient to stealing of challenge from a captured node. In EDD scheme, the performance of detection mechanism depends on the number of meetings threshold, which is difficult to estimate in MWSNs. This is because the number of meetings with a node over a time interval depends on the network size, the area of deployment, node’s speed, and the mobility model of the nodes. It changes with the variation of any of these parameters in the network. For example, if the network size is increased, then the number of meetings with a particular node decreases, as $\text{prob}(\text{meeting}) \propto 1/\text{network size}$. Estimating a threshold on number of meetings with a node without considering the variation of the abovementioned parameters may result in false detection. In the scheme proposed by Deng et al. [14], the replica is detected solely based on a single case of conflict with the measured speed. The speed is computed.
using the Euclidean distance between reported locations of a node over a time interval. This may not compute the actual speed of a node in the network with random waypoint mobility model. When a node moves faster by changing the direction frequently, it never follows a straight path. Moreover, since the locations of replica and the original node are used to measure the deviating speed, the adversary may deploy the replicas to move within closer premises of the original node to keep the measured speed within the accepted range. In the scheme proposed by Sindhuja and Padmavathi [17], the neighborhood fingerprint mechanism is not suitable for MWSN, due to the dynamic topology of the network. Moreover, selection of a single witness node is not suitable for MWSN, due to the dynamic topology of the network. Moreover, since the locations of replica and the original node in MWSNs are distributed in nature. In this scheme, each node performs the role of observer for a set of nodes in the network. Each observer maintains a sliding window of size \( w \) to keep track of recent distance traveled values at various time instances. The degree of violation is computed based on the frequency of the deviation recorded in the distance traveled by a node. A replica is detected by an observer, when the degree of violation of a node is higher than a specified threshold value. Section 4.1 explains the process of sharing the time-distance values by the nodes and updating the sliding window by the observers, and Section 4.2 details the process of replica detection.

4.1. Time-Distance Pair Sharing. Each node computes and maintains the distance traveled at the end of each epoch and locally broadcasts a time-distance pair at regular intervals. The format of the broadcast message from a node, say, \( A \), is given below:

\[
A \longrightarrow * : [A, T^A, D, \text{Signature}_A].
\] (1)

Here, \( A \) is the identity of the node, \( T^A \) is the timestamp of \( A \), \( D \) is the distance traveled by \( A \) at time \( T^A \), and \( \text{Signature}_A \) is the signature, signed by \( A \).

Let the node \( B \) be an observer of \( A \). Node \( B \) maintains a sliding window \( W^A_\Delta \) of size \( w \) containing recent time-distance broadcasts of \( A \). Let \( W^A_\Delta = \{ (d^A_1, t_1), (d^A_2, t_2), \ldots, (d^A_j, t_j) \} \), where \( (d^A_j, t_j) \) is the \( j \)th time-distance pair broadcast received by \( B \) from \( A \), and \( j > i \). Node \( B \) upon receiving a new time-distance pair from the node \( A \) verifies its signature and updates \( W^A_\Delta \) by removing the earliest pair from it. For example, if \( w = 10 \) and \( W^A_\Delta = \{ (d^A_1, t_1), (d^A_2, t_2), \ldots, (d^A_{16}, t_{16}) \} \), then after receiving \( (d^A_{16}, t_{16}) \) from \( A \) the updated \( W^A_\Delta \) is \( \{ (d^A_1, t_1), (d^A_2, t_2), \ldots, (d^A_{15}, t_{15}) \} \). Every time the sliding window of an observer is updated, it measures the deviation and the degree of violation of the node to detect replica. The following section defines the degree of violation and describes the process of replica detection.

4.2. Replica Detection. When an adversary deploys replica of a captured node, say, \( C \) in the network, the replicas will compute distance traveled and broadcast it at regular interval, despite their other malicious activities. In this circumstance, one or more observers of \( C \) are going to detect either of the following conditions while updating \( W^C_\Delta \) for any pair of time-distance pair instances \( (d^C_j, t_j) \) and \( (d^C_k, t_k) \):

\[
d^C_j \geq d^C_k \quad \text{if } t_j < t_k,
\] (2)

\[
d^C_k - d^C_j \leq [v^C_{\min}, v^C_{\max}] \quad \text{if } t_j < t_k, \quad d^C_j < d^C_k.
\] (3)

An observer marks the condition satisfied by either (2) or (3) as a deviation. This deviation of the node \( C \) is recorded.
using a binary deviation array, \( D \) of size \( w \). At the beginning of detection process, the \( D \) array of each node is set to zero. The deviation of a node \( C \) is updated using the following operations:

\[
D_C \ll 1, \\
\text{where } \ll \text{ is the left shift operation of bits,}
\]

\[
D_C [w-1] = \begin{cases} 
1 & \text{if deviation is marked} \\
0 & \text{otherwise.}
\end{cases}
\]

Using \( D_C \), the observers compute the degree of violation (deg \( V \)) for \( C \), which is defined by the following equation:

\[
\text{deg} V = k \cdot \frac{d_V}{n(d_V)}, \\
\text{where } d_V \text{ is given by}
\]

\[
d_V = \sum_{p \neq q \text{ s.t. } D_C[p] = D_C[q] = 1} \frac{1}{\sqrt{(p-q)^2}},
\]

\[
n(d_V) = \sum_{p \neq q \text{ s.t. } D_C[p] = D_C[q] = 1} .
\]

Here \( k \) is a constant. A node \( C \) is detected as replica when \( \text{deg} V \) is greater than a given threshold \( \eta \). The value of \( \eta \) is assigned using a set of replica-deployment scenarios during the simulation.

5. Analysis

**Claim 1.** For any pair of time-distance broadcasts \((d_j, t_j)\) and \((d_k, t_k)\), either of the following conditions holds:

(I) \( d_j \geq d_k \) if \( t_j < t_k \),

(II) \( \frac{d_k - d_j}{t_k - t_j} \notin [v_{\min}, v_{\max}] \) if \( t_j < t_k, \) \( d_j < d_k \).

**Proof.** Here, we show that when an observer receives time-distance broadcast from a node and its replica, either of the above conditions holds true. Let us consider two nodes \( n_a \) and \( n_b \) moving in a rectangular target field of size \( x \times y \). According to random waypoint mobility model, the node selects a target location and a speed at random from \([v_{\min}, v_{\max}]\) at the beginning of each epoch. Let the range of distance of an epoch be \([1, \sqrt{x^2 + y^2}]\). The distance traveled by a node and its replica will be the same for a given time interval only if the speed and target distance for all epochs are same. The probability of choosing a speed and distance by a node follows uniform distribution, that is, for a speed, \( v \in [v_{\min}, v_{\max}] \), and an epoch distance, \( d \in [1, \sqrt{x^2 + y^2}] \):

\[
P(v, d) = \frac{1}{(v_{\max} - v_{\min}) \cdot (\sqrt{x^2 + y^2} - 1)}. \\
\]

This probability is significantly small for even a small range of speed and distance. For example, the common speed range of mobile sensor is taken as \([2,8] \text{ m/s}\), and a small target area size is \(100 \times 100 \text{ m}^2\). Using these values in the above equation, the probability is given by

\[
P(v, d) = \frac{1}{8 \times 140.4} = 0.0009.
\]

Hence, the distance traveled by a node and its replica will never be same. Therefore, at any time, \( d_j \neq d_k \). Based on the broadcast time of the pair \((d_j, t_j)\) and \((d_k, t_k)\), either \( d_j > d_k \) or \( d_k > d_j \) will hold true. In case (II), the distance \( d_j \) is always received from a replica and \( d_k \) from the original node. This is because the replicas are deployed after the deployment of original node. In this case, the \( d_j - d_k \) will be greater than the traveling limit of the nodes. Hence, in either case, the above conditions will hold true.

5.1. Security Analysis. A malicious node may attempt to forge a time-distance pair. However, on receiving the time-distance broadcast, the node’s identity and the signature are verified by the observers. Therefore, malicious node cannot forge time-distance broadcast. A node may skip the process of broadcasting time-distance pair. In this case, when observers do not get an update from a node over a longer period of time, they blacklist the node for violating security protocol. An adversary will not gain much benefit if replicas move together and stay close enough so that all could travel the same distance. This is because these nodes would essentially have the same set of neighbors and an observer would always receive a single time-distance broadcast from the same node at a time. An adversary may also modify replicas to broadcast a fake or manipulated time-distance pair that is close to the legitimate node. However, due to random waypoint mobility model, a node cannot predict the movement of another node in the network. In other words, a replica cannot guess the distance traveled by its legitimate node by any means. In this case, the intended replica would broadcast a random distance instead of the actual ones. As per the given conditions to measure the deviation, any random distance may lead to a deviation with a high probability and can be handled by an observer.

5.2. Communication and Storage Overhead. The communication in the proposed scheme involves local broadcast in the neighborhood. Therefore, the communication complexity of the proposed scheme for a network of size \( N \) is \( O(N) \). The storage overhead of the proposed scheme depends on \( w \). According to Birthday Paradox [23], \( \sqrt{N} \) number of
Table 1: Comparison of communication and storage overhead.

<table>
<thead>
<tr>
<th>Schemes</th>
<th>Type</th>
<th>Communication</th>
<th>Storage</th>
</tr>
</thead>
<tbody>
<tr>
<td>UTLSE &amp; MTLSD [14]</td>
<td>Distributed</td>
<td>$O(N)$</td>
<td>$O(\sqrt{N})$</td>
</tr>
<tr>
<td>Ho et al. [13]</td>
<td>Centralized</td>
<td>$O(N \sqrt{N})$</td>
<td>$O(N)$</td>
</tr>
<tr>
<td>Deng and Xiong [16]</td>
<td>Centralized</td>
<td>$O(N \sqrt{N})$</td>
<td>$O(k)$</td>
</tr>
<tr>
<td>XED [19]</td>
<td>Distributed</td>
<td>$O(1)$</td>
<td>$O(N)$</td>
</tr>
<tr>
<td>EDD [19]</td>
<td>Distributed</td>
<td>$O(1)$</td>
<td>$O(N)$</td>
</tr>
<tr>
<td>Proposed</td>
<td>Distributed</td>
<td>$O(N)$</td>
<td>$O(w \sqrt{N})$</td>
</tr>
</tbody>
</table>

Table 2: Simulation parameters.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Simulation area</td>
<td>$1000 \times 1000 \text{ m}^2$</td>
</tr>
<tr>
<td>Network size</td>
<td>100–1000 nodes</td>
</tr>
<tr>
<td>Deployment type</td>
<td>Uniformly random</td>
</tr>
<tr>
<td>Communication range</td>
<td>30 meter</td>
</tr>
<tr>
<td>Node movement</td>
<td>Random way-point mobility model</td>
</tr>
<tr>
<td>Node’s speed</td>
<td>2–8 meters/sec</td>
</tr>
<tr>
<td>Simulation time</td>
<td>800 sec</td>
</tr>
</tbody>
</table>

observers is sufficient to detect replica with a high probability. Therefore, the storage overhead associated with a node in the proposed scheme is $O(w \sqrt{N})$, where $w \ll N$. The comparison of communication and storage overhead of the proposed scheme with the existing schemes is shown in Table 1.

$$P_{\text{det}} = P \left( \text{probability of meeting of a pair of nodes in the network, such that one of them is a replica, and other is its observer} \right)$$

$$= r \cdot \sqrt{N} \left( \frac{2 \cdot r \cdot \sqrt{N}}{N(N-1)} \right).$$

From the above equation, it can be inferred that the detection probability increases with the number of replicas and observers in the network.

5.3. Detection Probability. Detection of replica depends on the probability of successful reception of time-distance broadcast from a replica by an observer. Let $r$ be the number of replicas deployed in the network, where $r \ll N$. Using Birthday Paradox [23], let $\sqrt{N}$ be the number of observers per node. Then, the probability of detecting a replica is expressed as the probability of meeting a replica by its observer. That is,

6. Simulation and Results

The Castalia 3.2 [24] simulator is used for simulation of the proposed scheme in Omnet++ simulation environment [25]. An area of $1000 \times 1000 \text{ m}^2$ is considered for simulation, where nodes are uniformly deployed within the simulation area. The network size is varied from 100 to 1000 nodes during simulation. Nodes use random waypoint mobility model for movement. Replicas are randomly deployed by compromising multiple nodes. We observed during the simulation that the replica detection is high when $\eta$ is 0.5. The summary of simulation environment parameters is provided in Table 2. The performance of the proposed scheme is evaluated using the following metrics: (i) detection probability, (ii) first replica detection time, (iii) number of packets sent/received, (iv) energy consumed per node, and (v) false detection rate. The proposed scheme is compared with two popular distributed replica detection mechanisms: MTLSD [14] and EDD [19].

Figure 1 shows the comparison of detection probability versus network size for number of replicas equal to 10. It is observed that proposed scheme has higher detection probability in comparison to EDD and MTLSD. This is because in the proposed scheme the deviation in distance traveled by the replica is successfully recorded and detected by its observer nodes. However, in the schemes EDD and MTLSD the replica detection relies upon the number of meetings with the nodes and the time-location claims exchanged between a pair of nodes, respectively. It is found that it is difficult to achieve higher detection probability by considering these parameters and as a result the detection probability is lower in these schemes.

The comparison of detection probability versus number of replicas for the network size of 500 is shown in Figure 2.
It is observed that the detection probability of proposed scheme increases with the number of replicas in the network. The increase in the number of replicas contributes to higher meeting probability of replica with its observer nodes. It is also observed that the detection probability of proposed scheme is higher than other schemes with different number of replicas. The sliding window size is one of the parameters that influence the efficiency of recording deviation. The plot for detection probability versus window size by varying the number of replicas is shown in Figure 3. It is observed that the detection probability is high when window size is 30. The detection probability remains unchanged, when the window size is greater than 30.

The plot for comparison of first replica detection time versus number of replicas deployed is shown in Figure 4. It is observed that the first replica detection time of the proposed scheme lies between EDD and MTLSD. This is because the detection of replica is done based on the degree of violation, which takes a marginally higher amount of time than time-distance claim diffusion process in MTLSD. EDD have higher detection time than the rest of the schemes, because EDD checks for number of meetings with other nodes over a predefined time interval and this takes relatively higher time for detection of replica.

Figure 5 shows the comparison of average number of packets sent/received by a node per epoch versus network size. It is observed that average number of packets sent/received per epoch of the proposed scheme is higher than EDD and lower than MTLSD. This is because the only communication in the proposed scheme involves the
broadcasting of time-distance pair in the neighborhood. This value is quite less than the overhead of MTLSD, where all time-location claims in common are shared among the trackers. EDD has lower number of packets sent/received per node because it uses only hello broadcast message.

The energy consumed per node is compared in Figure 6. Energy consumed by the proposed scheme is higher than EDD and lower than MTLSD. This is because the energy consumption in a node mostly depends on the transmission and reception of messages. Since the communication complexity of the proposed scheme is higher than EDD and lower than MTLSD, the energy consumption per node of the proposed scheme lies between EDD and MTLSD. Finally, the plot for false positive versus network size is shown in Figure 7. The schemes considered for comparison have marginally higher false positive for network size more than 700. The false positive rate is higher in EDD compared to others since it relies on number of meetings to detect replica.

7. Conclusion

In this paper, we proposed a distributed replica detection scheme for MWSNs based on sliding window approach. The scheme uses the time-distance pair to compute the degree of violation, deg \( V \). Each node acts as an observer for a number of nodes in the network. The observer nodes maintain a sliding window of recent time-distance pair of these nodes. Every mark of deviation of a replica contributes to a higher value of deg \( V \). When the degree of violation of a node is higher than \( \eta \), it is detected as replica. The simulation results show that the proposed scheme has detection probability nearly equal to one compared to EDD and MTLSD schemes.

Competing Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

References


Submit your manuscripts at https://www.hindawi.com