CHAPTER - 2

REVIEW OF RELATED WORK

2.1 INTRODUCTION

This chapter gives a survey of node replica detection schemes reported in the literature. It also classifies the existing schemes and analyzed their communication, storage, and message overhead. A few of the replica detection schemes are simulated and their performance is compared. The metrics considered for comparison are: detection probability, detection time, average number of packets sent/received, and energy consumption.

2.2 CLASSIFICATION OF REPLICA DETECTION METHODS

The existing mechanisms for replica detection can be broadly classified into the following categories:

a) Classification based on the detection mechanism

1. Centralized: In a centralized scheme, there exists a centralized trusted entity such as BS, which is responsible for replica detection.

2. Partially distributed: In this scheme, replica detection mechanism is distributed in nature. However, the involvement of BS is necessary for certain activities such as broadcasting a random number, revocation of message and global keys, etc.

3. Fully Distributed: This scheme does not require the involvement of any centralized entity. Nodes cooperate among themselves to detect replica.
b) **Classification based on the need of geographical information**

i) **Location dependent:** This scheme uses node’s location information to detect replica. In this scheme, the location-claim of a node is forwarded to a set of randomly generated witness nodes/location, or cells, by its neighboring nodes. Two or more different location-claims for the same node result in a location conflict, and a replica is detected. Location dependent replica detection is depicted in Figure 2.1. The nodes denoted as W in the figure are witness nodes. On receiving the location claim for node A from two different locations, the witness node W detects it as a replica. A node can compute its location information either using GPS [4] or using the mechanisms mentioned in [5, 19].

![FIG 2.3: Classification of replica detection schemes in WSN.](image)

ii) **Location independent:** A location independent scheme does not use location information for replica detection. This scheme relies on various parameters such as group membership of a node, synchronization timer/counter value etc. to detect a replica. Figure 2.2 depicts the group membership based replica detection. In this figure, node is detected as replica, since witness node W receives different group membership list for the same node A.
c) Classification based on the claim forwarding strategy

1. **Deterministic forwarding:** In this scheme, a claim received by a node is always forwarded to its destination.

2. **Probabilistic forwarding:** In this scheme, a claim received by a node is forwarded with a certain probability depending on the claim forwarding strategy.

d) Classification based on the message routing

1. **Link-based routing:** This scheme uses Link-based routing protocols [20] for routing messages in the network.

2. **Geographical routing:** This scheme uses Geographical Routing protocols [21] for routing messages in the network. Figure 2.3 shows the proposed classification of replica detection schemes.

### 2.3 EXISTING REPLICA DETECTION SCHEMES

This section, briefly discus the existing schemes for replica detection, with their relative merits and demerits. Notations used in this chapter are enlisted in the Table 2.1.

![FIG: 2.2 WIRELESS SENSOR NETWORK WITH CLONE NODES](image)

**Normal Node**  **Clone Node**
2.3.1 CENTRALIZED SCHEMES

SET

Choi et al. [25] proposed a scheme called SET, in which the network is divided into a number of non-overlapping sub-regions. Nodes in each sub-region are one-hop neighbor of each other. A leader is elected within each sub-region. Then, a tree is constructed with the BS as root, and the leaders at different levels in the tree. Each sub-region leader sends their members identity (ID) to its parent node. A parent node performs intersection operation between all child sub-regions before forwarding them to its parent. A non-empty intersection at any level of the tree leads to a conflict and is reported to the BS for further action. SET has lower communication overhead in replica detection. However, SET is a centralized detection mechanism, which is vulnerable to single-point of failure. Size of the member-ID message grows rapidly at higher levels of the tree close to BS.

AN AREA-BASED APPROACH

An area based scheme is proposed by Naruephiphat et al. [26]. In their scheme, a node having maximum number of neighbors is selected as the central node. Then, the network is divided into sub-areas. Each sub-area has equal angle around the central node. A witness node is selected in each sub-area using the method similar to the one used to select the central node. A node sends its location-claim to the witness node of its area. If a witness node detects a location conflict, then it broadcasts a conflict notification to all nodes in the network. On reception of claims without conflict, the witness node sends these claims to the central node for further detection of replica at network level.
Table: 2.1: Notations

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>N</td>
<td>Network Size</td>
</tr>
<tr>
<td>D</td>
<td>Average Degree of a node</td>
</tr>
<tr>
<td>P</td>
<td>Probability of forwarding a message</td>
</tr>
<tr>
<td>W</td>
<td>Number of nodes (witnesses) that store a location-claim</td>
</tr>
<tr>
<td>S</td>
<td>Number of sensor node in a cell</td>
</tr>
<tr>
<td>T</td>
<td>Number of clusters in the network</td>
</tr>
<tr>
<td>G</td>
<td>Number of witnesses generated by a neighboring node</td>
</tr>
<tr>
<td>Vp</td>
<td>Verification point used in the network</td>
</tr>
</tbody>
</table>

In this scheme, sub-areas can be scaled easily by dividing their angles. However, this scheme is subject to single-point of failure. Moreover, the communication overhead is quite high in inter sub-area replica detection.

2.3.2 PARTIALLY DISTRIBUTED SCHEMES

REAL-TIME DETECTION SCHEME

A fingerprint based scheme is proposed by Xing et al. [27], where the fingerprint of a node is computed using the neighborhood information. In this scheme, each node is preloaded with a codeword generated from a superimposed s-disjunct code [28] prior to their deployment. A node computes its unique fingerprint as well as the fingerprints of its neighboring nodes using their codeword. The fingerprint of a node is included in every message that is sent to the BS. Neighboring nodes verify the genuineness of a node by comparing the fingerprint in the message with their own record. A conflict in the fingerprint of any node is reported to the BS, and is detected
as a replica. In this scheme, a replica is detected either by the neighboring nodes or the BS. The generation of codeword from superimposed s-disjunct code is computationally expensive.

**NEIGHBOR-BASED DETECTION SCHEME (NBDS)**

A neighbor-based detection scheme was proposed by Ko et al. [29]. According to their scheme, a node moving to a new location must broadcast a rejoining claim to its new neighbors. The rejoining claim consists of the list of its old neighbors. A new neighbor on receiving the rejoining claim forwards it to a randomly selected node from the list of old neighbors. An old neighbor on receiving the rejoining claim checks for the node ID in its neighbor table. If the node ID is present, then the old neighbor reports to the BS, otherwise it concludes that the node has moved to another location. The old neighbor then sends a challenge to the node to verify its existence in its neighborhood. If the node is present in the old position, then it is detected as a replica. On the other hand, if the new neighbors do not receive any revocation message within a stipulated time period, then, they accept the newly joined node as their neighbor. This scheme allows occasional mobility of nodes in the network. However, an intelligent adversary may bypass the replica detection process. The verification process of the scheme increases the communication overhead.

**LOCALIZED MULTICAST**

Zhu et al. [22] proposed a method called Localized Multicast where the witness nodes are randomly selected and localized to a restricted geographical region. There are two variations of Localized Multicast: i) Single Deterministic Cell (SDC), and ii) Parallel Multiple Probabilistic Cells (P-MPC). In their scheme, Zhu et al. have
considered a geographical grid network and divided it into a number of rectangular cells. In SDC, neighboring nodes forward the location claim of a node to its target cell with a probability \( p \). A geographic hash function [30] is used to map the node’s ID to a target cell. The authenticity of location claim is verified by all nodes within this target cell. On successful verification, each node in the cell caches the location claim with a certain probability. The claimer of two conflicting location claims is detected as a replica by a node within the target cell. P-MPC is an extension of SDC. In P-MPC, the location claim is mapped to a number of destination cells instead of one. This improves the detection probability in P-MPC. This scheme is simple and its communication overhead is low. However, the strength of SDC and P-MPC in replica detection depends on the probability of forwarding the location claim to a cell and storing the claim by the cell members.

A NOTE-BASED RANDOMIZED AND DISTRIBUTIVE PROTOCOL

Meng et al. [31] proposed a mechanism that uses a note containing the subset of neighbors as a claim. The claim-forwarding process starts with the selection of a reporter node among the neighbors. Once a reporter node is selected, the claimer requests for a signature note from the reporter node. Upon receiving the signature note, claimer first verifies and then forwards the claim. The reporter node on receiving a claim, generates a pre-defined number of witness nodes and forwards the claim to those nodes. A conflict in the claim of a node at any of its witness node is detected as a replica. This scheme has an additional overhead of selecting reporter nodes. This is because, the trust worthiness of a claim-forwarding node can easily be verified by listening to its broadcast message. Moreover, the replica detection depends on the uniqueness of the subset member-list which may not be ensured in all cases.
RANDOMIZED, EFFICIENT AND DISTRIBUTED SCHEME (RED)

Conti et al. [32] proposed a scheme called randomized, efficient and
distributed protocol (RED) in order to improve the performance of location-claim
based schemes. In their scheme, a random value is broadcasted to all nodes. Then the
nodes digitally sign their location-claim, and broadcast to their neighboring nodes.
Each neighbor generates a number of pseudo-random witness locations using the
random value, their own ID, and the number of witnesses. Then they send the
location-claim to the witness locations with a probability p. Nodes located at a distance
less than a pre-defined value from these witness locations store the location-claim. A
conflict in the location claim is detected as a replica. RED is secured as witness
locations are generated randomly. However, it has limitations in sparse as well as
dense sensor networks. In a sparsely distributed sensor network, a case may arise
where there will be no or limited witness nodes within the predefined range of a
randomly generated location. In such a scenario, the claim may or may not reach a
witness node. This will limit detection probability. In a highly dense network, the
number of witness nodes can be significantly higher. In this scenario, the average
storage overhead of a network increases.

HIERARCHICAL NODE REPLICATION DETECTION SCHEME

Znaidi et al. [23] proposed a mechanism for three-tier hierarchical network
structure. Their mechanism is based on the use of Bloom filter [33, 34]. The replica
detection process is divided into the following three phases:

1. **Pre-distribution Phase**: In this phase, nodes are equipped with cryptographic
   keying materials and other parameters required for Bloom filter operation.
2. **Election Phase:** This phase is performed periodically to select cluster-head using Local Negotiated Clustering Algorithm (LNCA) protocol [35].

3. **Detection Phase:** Elected cluster-heads exchange their member IDs among themselves using Bloom filter. A node ID that is found to be a member of more than one cluster is detected as a replica.

   The memory overhead of this scheme is significantly lower. However, it has an additional communication overhead in cluster formation. For a small number of hash functions the false detection probability is higher and requires relatively larger number of bits in Bloom filter.

### 2.3.3 FULLY DISTRIBUTED SCHEMES

**DISTRIBUTED DETECTION SCHEME**

Parno et al. [36] proposed two mechanisms which they called: i) Randomized Multicast (RM), and ii) Line-Selected Multicast (LSM). In RM, witness nodes are randomly chosen. The location claim broadcast of a node is forwarded to a set of randomly selected witness nodes by its neighbors. If a clone is present in the network, then the location claim of the clone is also forwarded to a set of witness nodes. According to birthday paradox [37], if each location claim is forwarded to $\sqrt{N}$ number of witnesses, then at least one witness node will receive two different location claims, one from the original node and the other from its replica with a higher probability. This leads to a location conflict and the node is detected as a replica. LSM has lower communication cost than RM. In LSM, the location claim of a node is cached at each intermediate node before forwarding to the next-hop node on the path to the witness node. This creates a line across the cached intermediate nodes. When
the line of a clone’s location-claim crosses the line of its legitimate node, the intermediate node at the crossing point detects it as a replica. In RM, the broadcasting of location claim of a node to all of its witness nodes is expensive in terms of communication overhead. In LSM, nodes that are common to multiple location-claim lines suffer from higher storage overhead.

**SYMMETRIC PAIR-WISE KEY ESTABLISHMENT SCHEME**

Bekara et al. [38] proposed a scheme based on symmetric pair-wise key establishment. In their scheme, each node is associated with a unique generation or group. The generation of a node can be computed using its ID and a symmetric polynomial. According to their protocol, only a newly deployed node that belongs to a newly deployed generation is able to establish a pair-wise key with their neighbors. Therefore, when a clone that belongs to an old generation tries to establish a pair-wise key with its new neighbors, it is detected as a replica. This scheme is simple and incurs less communication overhead. The duration of a nodes’ generation is a critical parameter. An inaccurate duration may detect the genuine nodes of previous generation as replicas.

**DISTRIBUTED DETECTION SCHEME RESILIENT TO MANY COMPROMISED NODES**

Sei et al. [39] have proposed a resilient replica detection scheme. This scheme does not require any trusted entity, and is resilient to a number of compromised nodes in the network. Each node is pre-loaded with detection process start time. A node in its turn sends a one-time seed, its ID, and location with a signature to all other nodes. If a node fails to start the detection process in its turn within a pre-defined interval of
time, then the next node starts its process. To improve resiliency against node capture, nodes are divided into groups, and each node starts its detection process using a role ID assigned to it. The neighboring nodes responsible for forwarding the location claim of a node to its witness nodes are called reporter nodes. A reporter node forwards the location claim to a number of witness nodes that are responsible for detecting replica. This scheme incurs significantly higher communication overhead because, the detection process is based on message broadcast over the network. Moreover, the scheme does not explain the procedure to ensure that at least one neighbor would voluntarily become a reporter node.

MEMORY EFFICIENT PROTOCOLS

Zhang et al. [40] have identified two problems associated with LSM protocol; which they called crowded-center problem, and cross-over problem. To overcome the above two problems, they proposed the following four protocols: i) B-MEM, ii) BC-MEM, iii) C-MEM, and iv) CC-MEM. In B-MEM, the location claim of a node is sent to a random location with a probability of p. A node located closer to this location, stores the claim. An intermediate node on the path, also known as watcher node stores the ID and location of the claimer using Bloom filter. Any location conflict with a stored ID is detected as a replica by the watcher node. In BC-MEM protocol, the deployment area is divided into a number of virtual cells. Each node in the cell is associated with an anchor point and an anchor node. The anchor point is determined using the node ID, whereas the anchor node of a node is the node, which is closer to its anchor point. In BC-MEM, a location claim is forwarded from an anchor point of one cell to another cell, which are intersected by the line segments
until it reaches the last cell. The intermediate anchor nodes act as a watcher to detect clone. This overcomes the cross-over problem, and reduces the storage overhead. C-MEM protocol overcomes the crowded-center problem by forwarding the location claim to a random point called cross point. The claim is again forwarded along the horizontal and vertical lines from the cross point. Nodes on these lines act as watcher, and those closer to the cross point are witness node. CC-MEM uses the concept of both cross-forwarding and cell-forwarding used in CMEM and BC-MEM respectively. The detection probability is higher in CC-MEM. Above memory efficient schemes are able to achieve lower storage overhead using Bloom filter. However, their communication overhead increases significantly with improvement in detection probability in C-MEM and CC-MEM.

**DISTRIBUTED DETECTION WITH GROUP DEPLOYMENT KNOWLEDGE**

Group deployment knowledge is used in the schemes proposed by Ho et al. [41]. They have proposed three schemes. In their first scheme, a node is preloaded with its pre-determined group deployment point. The nodes within the same group are expected to be deployed closer to each other. Nodes closer to their group deployment point are considered to be trusted, whereas the nodes that are far away from the group deployment point are considered to be untrusted. A node ignores the message received from the untrusted nodes. Nodes belonging to different groups can communicate with each other, only if the distance between their group deployment point is less than a threshold value. In their second scheme, they have relaxed the criteria to communicate with untrusted neighbor. In this scheme, a node can communicate with an untrusted node, only if the untrusted node provides sufficient
evidence in terms of location claim that it is not a replica. In their third scheme, a node forwards the location claim of an untrusted neighbor to multiple groups instead of the untrusted neighbor’s home group. Since, the above schemes use deployment knowledge, their memory requirements are significantly lower. Verification of untrusted nodes may require large number of message exchanges. This will increase the communication overhead of the network. Moreover, the above schemes may not be suitable for the applications, where it is difficult to determine or compute the group deployment point well in advance.

**RANDOMLY DIRECTED EXPLORATION**

Li et al. [42] proposed a claim broadcast based technique called randomly directed exploration. In this scheme, nodes generate a number of claim messages and each message containing a Time-to-leave (TTL) value is forwarded to a randomly selected neighbor. An intermediate node computes its angle with the witness location. To forward the location claim, an intermediate node selects a node that is closer to the computed angle plus \( \pi \). When two or more conflicting location claims of a node is received by a witness node, it is detected as a replica. This scheme is similar to RM, but differs from RM in the claim forwarding mechanism. In RDE, a TTL field is used to prevent indefinite routing. This scheme suffers from higher communication overhead.

**DISTRIBUTIVE, DETERMINISTIC AND RESILIENT SCHEME (DDR)**

Kim et al. [24] proposed a distributive, deterministic, and resilient (DDR) scheme. In this scheme, nodes are associated with a verification point in the network. Verification point is generated by the BS prior to the node deployment. The
The verification point is the destination for the location-claim of a node. A node generates the location-claim, estimates the expected hop-count to the verification point, and then sends the location claim to the verification point through a randomly chosen neighboring node. An intermediate node on receiving the location claim decreases the hop-count by one, and caches the location claim based on some probability, before forwarding it to the verification point. A node that receives two different location-claims for the same ID detects it as a replica. DDR is similar to LSM. Both schemes cache the location-before claim forwarding to next node. Common intermediate nodes for a large number of forwarding paths are overloaded with the task of storing and forwarding location claim.

Table 2.2: Categorization of different replica detection schemes

<table>
<thead>
<tr>
<th>Scheme</th>
<th>Detection Mechanism</th>
<th>Geographical Information</th>
<th>Claim forwarding</th>
<th>Routing Type</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Centralized</td>
<td>Partially Distributed</td>
<td>Location Dependent</td>
<td>Deterministic</td>
</tr>
<tr>
<td>RM &amp; LSM [36]</td>
<td>-</td>
<td>√</td>
<td>√</td>
<td>-</td>
</tr>
<tr>
<td>SET [25]</td>
<td>√</td>
<td>-</td>
<td>-</td>
<td>√</td>
</tr>
<tr>
<td>Bekara et al. [38]</td>
<td>-</td>
<td>√</td>
<td>-</td>
<td>√</td>
</tr>
<tr>
<td>Xing et al. [27]</td>
<td>-</td>
<td>√</td>
<td>√</td>
<td>-</td>
</tr>
<tr>
<td>Sei et al. [39]</td>
<td>-</td>
<td>√</td>
<td>√</td>
<td>-</td>
</tr>
<tr>
<td>NBDS [29]</td>
<td>-</td>
<td>√</td>
<td>√</td>
<td>-</td>
</tr>
<tr>
<td>B-MEM, BC-MEM</td>
<td>-</td>
<td>√</td>
<td>√</td>
<td>-</td>
</tr>
<tr>
<td>C-MEM, CC-MEM</td>
<td>-</td>
<td>√</td>
<td>√</td>
<td>-</td>
</tr>
<tr>
<td>Ho et al. [41]</td>
<td>-</td>
<td>√</td>
<td>√</td>
<td>-</td>
</tr>
<tr>
<td>RDE [42]</td>
<td>-</td>
<td>√</td>
<td>√</td>
<td>-</td>
</tr>
</tbody>
</table>
EARLY AND LIGHTWEIGHT DISTRIBUTED DETECTION PROTOCOL

Tran et al. [43] proposed two light-weight protocols which they called: i) LANCE, and ii) SACRED. In LANCE protocol, a network-wide counter is maintained which is incremented by one, after a pre-defined time interval. A clone will have the counter value different from others. When a clone broadcasts its counter value, it is considered as a replayed counter value and is detected as a replica. The SACRED protocol is a secured version of LANCE. In this protocol, nodes take the help of location claim and authentication mechanism similar to LSM [36]. LANCE assumes that nodes are time-synchronized. It is difficult to ensure time-synchronization in a sensor network. Moreover, a powerful adversary can forge the original counter value.

2.4 ANALYSIS OF EXISTING APPROACHES

In this section, brief quantitative comparison of different replica detection schemes was made. The metrics used for comparison are communication and storage overhead associated in detecting a replica. Communication and storage overhead of different schemes are analyzed in sub-section 2.3.1. Number of nodes required for detecting a replica also affects replica detection ability of a scheme. In sub-section 2.3.2, we have analyzed the number of nodes required for replica detection. Communication cost is the number of messages exchanged in the network.
for replica detection. Storage overhead is the additional information stored per node in replica detection. The analysis of communication and storage overhead of different schemes is given below.

**Table 2.3: Authentication technique used and the content of claim message in different scheme**

<table>
<thead>
<tr>
<th>Scheme</th>
<th>Content of the claim message</th>
<th>Authentication Technique</th>
</tr>
</thead>
<tbody>
<tr>
<td>RM, LSM [36]</td>
<td>(ID, Location, Signature)</td>
<td>Merkle-Winternitz signature [45]</td>
</tr>
<tr>
<td>SET [25]</td>
<td>(ID, Sets of Node IDs, MAC)</td>
<td>HMAC [46]</td>
</tr>
<tr>
<td>Bekara et al. [38]</td>
<td>(Generation ID, MAC)</td>
<td>Symmetric Bivariate Polynomial [47]</td>
</tr>
<tr>
<td>Xing et al. [27]</td>
<td>(ID, Node’s Fingerprint, data)</td>
<td>-</td>
</tr>
<tr>
<td>Sei et al. [39]</td>
<td>(ID, Location, Signature)</td>
<td>Identity-Based Signature Scheme (IBSS) [48, 49]</td>
</tr>
<tr>
<td>NBDS [29]</td>
<td>(ID, Neighbor ID List, Signature)</td>
<td>EIBSSBP [50]</td>
</tr>
<tr>
<td>B-MEM, BC-MEM, C-MEM, CC-MEM [40]</td>
<td>(ID, Location, Signature)</td>
<td>IBSS</td>
</tr>
<tr>
<td>Ho et al. [41]</td>
<td>(ID, Location, MAC)</td>
<td>TinyECC [51]</td>
</tr>
<tr>
<td>RDE [42]</td>
<td>(TTL, ID, Location, Neighbor ID List, Signature)</td>
<td>IBSS</td>
</tr>
<tr>
<td>DDR [24]</td>
<td>(ID, Location, hop-count, verification point, MAC)</td>
<td>Symmetric Key [52]</td>
</tr>
<tr>
<td>SACRED [43]</td>
<td>(ID, Location, Signature)</td>
<td>Symmetric Key</td>
</tr>
<tr>
<td>SDC, P-MPC [22]</td>
<td>(ID, Location, Signature)</td>
<td>EIBSSBP [50]</td>
</tr>
<tr>
<td>RAWL, TRAWL [44]</td>
<td>(ID, Location, Signature)</td>
<td>TinyECC</td>
</tr>
<tr>
<td>Meng et al. [31]</td>
<td>(ID, Neighbor ID sub-list, time, Signature)</td>
<td>EIBSSBP [50]</td>
</tr>
<tr>
<td>RED [32]</td>
<td>(ID, Location, Signature)</td>
<td>IBSS</td>
</tr>
<tr>
<td>Naruephiphat et al. [26]</td>
<td>(ID, Location, Signature)</td>
<td>-</td>
</tr>
<tr>
<td>Znaidi et al. [23]</td>
<td>(Encrypted Bloom Filter, Signature)</td>
<td>Elliptic-Curve Cryptography (ECC) [53]</td>
</tr>
</tbody>
</table>
SET [25]

In SET [25], each node participates in the subset leader election process by sending a broadcast message. In addition to leader election, each subset leader has an additional task of forwarding membership list to the BS. The average communication overhead of each node is $O(1)$ and therefore, communication cost of SET is $O(N)$. In SET, there is no storage overhead because none of the nodes including the subset leaders need to store the membership information.

SYMMETRIC PAIR-WISE KEY ESTABLISHMENT SCHEME [38]

In this scheme, each node establishes pair-wise key with its neighbors on deployment. Therefore, overall communication overhead is $O(N)$. There is no storage overhead associated with this scheme because the replica detection depends only on the pair-wise key establishment made by a node with its neighbors.

REAL-TIME DETECTION SCHEME [27]

In this scheme, each node needs to share its codeword with the neighboring node for the computation of fingerprint. This sharing requires communication overhead of $O(N)$ in the network. Each node needs to store the codeword of the neighboring nodes. Therefore, storage overhead of this scheme $O(d)$.

RED [32]

In RED, each neighboring node generates a random number of witness locations for a location-claim and forwards with a probability $p$. For $g$ number of witnesses, communication overhead of RED is $O(N.g.d.p.(\sqrt{N}))$. Nodes that are closer to witness locations receive and store the location-claim of a node. Therefore, storage overhead of RED is $O(g.d.p)$. 
HIERARCHICAL NODE REPLICATION DETECTION SCHEME [23]

The major communication overhead of this scheme is in exchanging the member IDs among the cluster-heads. In a network of $t$ number of clusters, the communication overhead is $O(t^2)$. The cluster-heads keep the member IDs of its own and other clusters with the help of Bloom filter. Therefore, the storage overhead of this scheme is $O(t)$.

2.5 COMPARISON OF NUMBER OF NODES RESPONSIBLE FOR REPLICA DETECTION

The probability of replica detection also depends on the number of nodes responsible for detecting a replica. We call this as Number of Replica Detectors Per Node. In the detection schemes such as RM and LSM [36], Sei et al. [39], [40], Ho et al. [41], DDR [24], SACRED [43], SDC and P-MPC [22], RAWL and TRAWL [44], RED [32], and Naruephiphat et al. [26], Number of Replica Detectors Per Node is equal to number of witnesses per node. In SET [25], all subset leaders in the path from the node to the BS are responsible for replica detection. For $t$ number of subset leaders, Number of Replica Detectors Per Node is given by $t$. In the schemes such as Bekara et al. [38], Xing et al. [27], NBDS [29], RDE [42], and LANCE [43], the neighboring nodes are responsible for detecting replica. Therefore, Number of Replica Detectors Per Node for these schemes is equal to $d$. In hierarchical detection schemes such as Znaidi et al. [23], the cluster-heads are responsible for replica detection. Therefore, the Number of Replica Detectors Per Node for these schemes is equal to $t$. It can be observed from the Table that the Number of Replica Detectors Per Node is comparatively higher in location-based schemes.
2.6 PROBLEM STATEMENT

2.6.1 CLONE (REPLICA) NODE ATTACK

Sensor nodes are geographically organized in Wireless Network. In general, it is very difficult to keep track the changes happen in the sensor node because of security issues in the wireless network. A few kinds of the security issues arise on the sensor node during the transmission of data between the nodes. One of most important security attacks is clone node attack, which pretense serious threat in Wireless Sensor Network. The challenger or opponent aimed to launch the attack in sensor network by capturing the sensor node manually and gain the access of node, replicate them in to several locations and easily take over control of the entire networks. To overcome these issues different solutions were proposed, which was not adaptive to the dynamic changes in the number of nodes in the network also network performance get reduced. Most of the replica detection schemes reported in the literature are centralized and location dependent. Centralized schemes are vulnerable to a single point of failure. Forwarding location information incurs additional overhead in location dependent schemes. Most of the replica detection mechanisms require the exchange of membership information among the nodes. Therefore, a lightweight mechanism for exchanging information among the nodes is required in replica detection.

2.7 SUMMARY

In this chapter, a brief discussion of the various schemes for replica detection as reported in the literature. It contains the existing schemes based on the detection mechanism, need of geographical information, claim forwarding strategy, and message routing type. A quantitative analysis of different schemes is presented.
Simulation of a few popular replica detection schemes such as SET [25], LSM [36], SDC [22], Sei et al. [39], and RED [32] were compared. The metrics considered for comparison are detection probability, detection time, average number of packets sent/received, and energy consumption. It is observed that the RED has higher detection probability, and relatively higher detection time. The SET incurs lower communication overhead than other schemes. So it is necessary to develop a scheme that not only have higher detection probability but also lower detection time. Since, energy is an important issue in WSN, a replica detection scheme should also consume lesser energy as well. In a nutshell it is desirable to have a replica detection scheme that has higher detection probability, lower detection time, communication overhead and energy consumption. Most of the replica detection mechanisms require the exchange of membership information among the nodes. Therefore, a lightweight mechanism for exchanging information among the nodes is required in replica detection. Next chapter presents a technique called H-RAWL which is compared with the previous techniques and it finds the clone node in effective manner.