IOSR Journal of Computer Engineering (IOSR-JCE)
e-ISSN: 2278-0661, p- ISSN: 2278-8727Volume 16, Issue 1, Ver. II (Jan. 2014), PP 01-11
www.iosrjournals.org

Protocols for detection of node replication attack on wireless sensor network

Mrs: Suvarna Game, Mr. Chandrashekhar Raut
Pursuing M.E. in computer science from Datta Meghe College of Engineering Airoli, Navi Mumbai
Prof. Datta Meghe College Of Engineering Airoli, Navi Mumbai

Abstract: Wireless sensor network has many small sensor nodes that work in collaborative manner to achieve a specific task. But it is deployed in unattended environment and that is why it is prone to attacks. These attacks mainly fall into two categories that is application dependent and application independent .In this paper the focus is on the node replication attack which falls under application independent attacks. In this paper a survey has been done related to node replication attack and existing techniques for solving this issue has been studied. The paper mainly focuses on the types of attacks on wireless sensor network and the two techniques centralized and distributed detection for detection of the node replication attack. Defending against this node replication attack is recently become a research topic in the security of wireless sensor network. The applications and advantages of centralized detection and distributed detection and their respective limitations has been studied.

Keywords: *Wireless sensor network, node replication attack, centralized detection, distributed detection, adversary*

I. Introduction

1.1 About Wireless Sensor Network

WSN has many small sensor nodes. These nodes vary from several hundreds to thousands. These sensor nodes work in a collaborative manner to achieve a common goal. Sensor network is used for interaction between computer system and there environment. These autonomous sensors are used to monitor physical or environmental conditions, such as temperature, pressure, sound etc and cooperatively pass their data through the network to a main location. The more modern networks are bi-directional, also enabling control of sensor activity. The development of wireless sensor networks was motivated by military applications such as battlefield surveillance; today such networks are used in many industrial and consumer applications, such as industrial process monitoring and control, machine health monitoring, and so on.

Basic components of WSN nodes are Sensors, memory, processor, GPS, Radio transceiver and power source and major components of WSN are sensor node and base station. Sensor nodes are known as sensing cells and base station as brain of WSN.WSN consists of spatially distributed autonomous sensors to monitor physical or environmental conditions. The sensor node vary in size .It may be as large as shoebox or size of grain of dust.

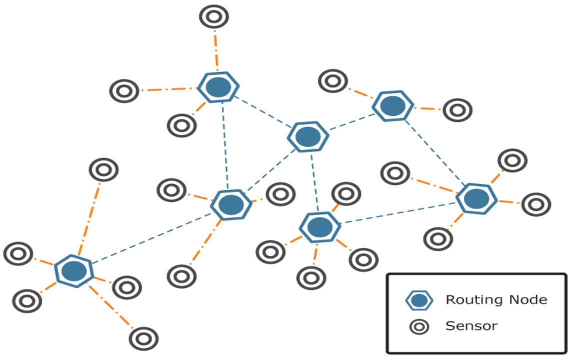

Fig: Representative Wireless Sensor Network

WSN is of two types that is Stationary and Mobile WSN. In stationary WSN nodes are stationary while in mobile WSN nodes can move and after deployment can interact with physical environment. Depending on the type of WSN attacks can vary.

1.2 Attacks on Wireless Sensor Network

Attacks on stationary WSN can be classified in to three categories Identity attacks, routing attacks and Network intrusion.

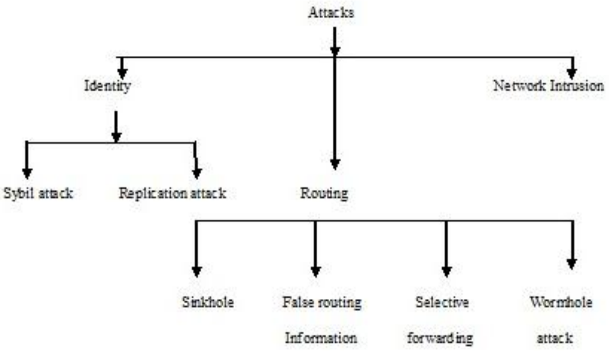

Fig: Types of attacks in WSN

1.2.1 Identity Attack: This is classified into two types as

1.2.1.1 Sybil attack: superficially similar to the node replication attack where one physical sensor node gains an unfair advantage by claiming multiple ids .The Sybil attack is also application-independent and enables one malicious node to multiply its inputs to subvert many protocols like distributed storage, routing, data aggregation, voting/agreement, resource allocation, and so on.

1.2.1.2 Replication attack: also called as node replication attack where one logical node id is reused by multiple physical sensor nodes. The Sybil attack is On a high level, the detection of sensor node replication attacks can be either network-based or not. A typical (but perhaps the only) instance in the latter case is found in Hussain and Rahman (2009), where radio signal strength is utilized at a receiver node to detect node replication we only consider network-based detections. A WSN can be either stationary (which is the prevalent case) or mobile, and replication detection scenarios in stationary and mobile WSNs can be substantially different The detection philosophy for stationary WSNs, on a coarse level, is based on the ‘‘exclusiveness’’ of node location (Boukerche etal., 2007). That is, a sensor node should be related to a unique deployment position; if one logical node id is found to be associated with two or more physical locations, node replication is detected. Clearly, this principle is inapplicable to the emerging mobile WSNs, where sensor nodes may roam in the deployment field all the time

1.2.2 Routing Attacks: In this type of attack a rouge node is placed on a routing path from source to base station. This may attempt to tamper legitimate data packets and this may discard the data packets.

1.2.2.1 Sinkhole attack: In this case adversary creates large sphere of influence because of which all traffic destined for the base station will get attracted from nodes which are several miles away from compromised node.

1.2.2.2 False Routing: This means injecting false routing control packets in to the network.

1.2.2.3 Selective Forwarding: In this case compromised node may refuse to forward the packets or forward only selective packets.

1.2.3 Network Intrusion: It is an unauthorized access to a system. This can be done either an external perpetrator or by an insider having lesser privileges. Among these various attacks in this paper we are focussing on Node replication attack which comes under Identity Attack. The attacks on WSN can be classified as application dependent and application independent attacks. The Node replication attack is application-independent Attack. In this attack an adversary prepares its own sensor node and induces that node in the network in such a way that a network will accept it as a legitimate node. An attacker is able to obtain compromised node’s memory and data within one minute after discovering it.

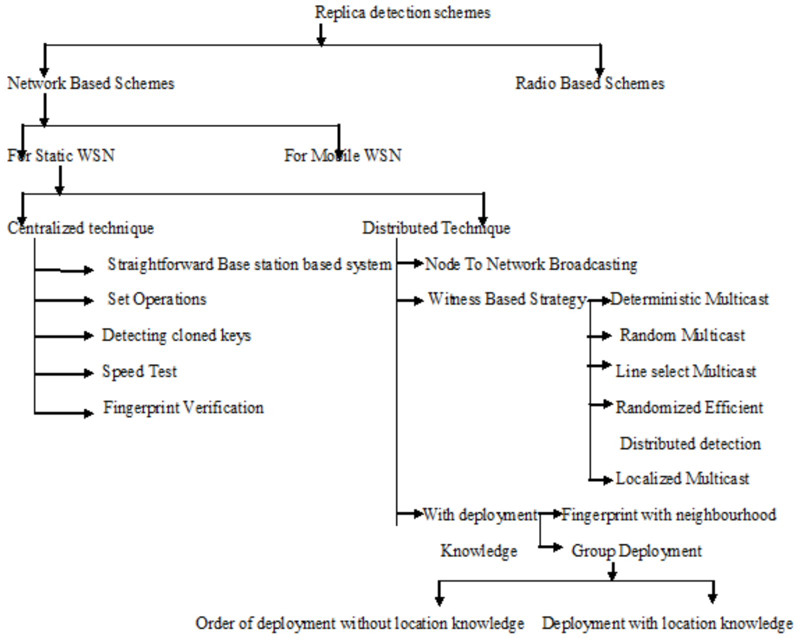

Fig: Node replication attack detection schemes

There are two types of detection techniques that are described below:

II. Centralized Detection Technique

1.1 Straightforward scheme: It requires each node to send a list of its neighbours (more specifically, a list of their ids) and the positions claimed by these neighbours (and signed by them, e.g., with an id-based signature scheme) to the base station, which then examines every neighbour list to look for replicated sensor nodes. In a stationary WSN, conflicting position claims for one node id indicates a replication. In a stationary WSN, conflicting position claims for one node id indicates a replication. Once the base station spots one or more replicas, it can revoke the replicated nodes by flooding the network with an authenticated revocation message for broadcast authentication. While conceptually simple, this approach suffers from several drawbacks (Parno et al., 2005) inherent in a centralized system. First, the base station introduces a single point of failure and can become a significant bottleneck. Nodes close to the base station are known as hotspots will receive the huge routing load and thus there power supply get depleted .Thus network connectivity get depleted. Beside this base station hotspots are also target for attacks.

1.2 Set Operations: This is a centralized detection scheme known as SET. This technique attempts to reduce the overhead by computing set operations such as Union, Intersection of exclusive subsets in the network. SET partitions the network logically in to non-overlapping regions called clusters. These clusters are managed by leaders known as cluster heads. These leaders report all the ids of the nodes in the region in the form of subset to base station. The intersection of an two subset of reports is empty that means no replication has occurred otherwise replication is detected. The actual set protocol is complex because of its complicated components. Which contribute to overload. SET protocol has to perform multiple rounds for countering the colluding replicas. One drawback is there for revoking an honest node an adversary can use SET protocol.

1.3 Detecting cloned keys: It is used for addressing cloned cryptographic keys rather than cloned sensor nodes. There is a random key distribution and key of a genuine nod should follow certain pattern. This helps for monitoring the key usage. Key usage refers to number of times key is used for setting the secure connection between neighbouring nodes. In this technique cloned keys are detected by analyzing node authentication statistics those keys whose usage exceeds a certain threshold (determined by the false positive rate) are considered cloned and erased from the network. Now each node will report its preloaded keys o base station. Now it is the task of base station to discover cloned keys. For collecting the key usage data bloom filter is used. Now this detection is effective when i) size of keys pre-distributed to each node is small, (ii) more clones exist in (i.e., are inserted into) the network, and (iii) a high false positive rate is set. Practically any WSN sensor node can communicate with only limited number of its neighbours within finite radius. This number is d known as average node degree or node density. The value of d can be adjusted by selecting appropriate transmission range. There is one drawback that e.g., how to ensure that the participating clones report their keys honestly (and exactly) to the base station. All the Centralized solution bear the deficiencies such as any compromise of base station will make the solution useless. For another example, even if there are no attacks, the nodes surrounding the base station (i.e., the hotspots) will suffer an undue communication burden that may shorten the life expectancy of the WSN. Because of this distributed solution become necessity.

III. Distributed Detection of Node Replication Attacks

In all four protocols, we assume that nodes know their own geographic positions. We also assume that the nodes in the network
remain relatively stationary, at least for the time it takes to perform one round of replication detection. If the network designers anticipate occasional mobility, they can schedule regular detection rounds. As long as a node successfully participates in a round,
it can continue to communicate until the next round, even if its position changes in the interim.
Distributed detection of node replication attack can be done using following three techniques. These are:

- Node to network broadcasting
- witness based strategy
- with deployment knowledge.

2.1 Node-To-Network Broadcasting: N2NB where each node floods the entire network with authenticated broadcast to claim its own location(instead of its neighbours') .Each node stores the location information for its neighbours, incurring a storage cost of O(d). Each node upon receiving a conflicting claim invokes a revocation procedure against the offending nodes , and eventually any replica will be cut off by all its neighbours (thus isolated from the WSN). Each node stores the location information for its neighbors and if it receives a conflicting claim, revokes the offending node. This protocol achieves 100% detection of all duplicate location claims under the assumption that the broadcasts reach every node. This assumption may not hold if the adversary can jam key areas or otherwise interfere with communication paths through the network.The challenge for detecting replication attacks has roots in the resource scarcity of sensor nodes. For stationary WSNs , such detection essentially requires network-wide comparison of location-dependent authentication information, and the limited memory capacity and energy supply places ever constraints on how much authentication information can be stored per node and exchanged in the network(Zhang etal.,2009). Hence it is reasonable to trade the detection rate(e.g.,100%forN2NB)for other major performance criteria like energy efficiency and memory efficiency . Note that the wireless transceiver is the biggest energy consumer for sensor nodes, while the communication cost for N2NB is Oðn2Þ.

2.2 Witness based strategy has following different schemes:

1. Deterministic multicast(DM)
2. Randomized multicast(RM)
3. Line select multicast(LSM)
4. Randomized efficient distributed detection (RED)
5. Localized multicast(LM)

2.2.1 Deterministic Multicast: DM is based on claimer–reporter–witness framework. Main goal behind the design of DM is reduced communication cost and the main idea is to only send a node's location claim to a limited set of deterministically chosen nodes serving as witnesses.For improving the communication cost of the N2NB protocol, we describe a detection protocol that only shares a node's location claim with a limited subset of deterministically chosen "witness" nodes. When a node, referred to as the claimer, locally broadcasts its location claim to its neighbours, each neighbour, serving as a reporter, employs a function to map the claimer id to a witness. Then the neighbour forwards the claim to the witness, which will receive two different location claims for the same node id if the adversary has replicated anode. Now there is a immediate issue the adversary can also employ the function to know about the witness for a given claimer id, and may locate and compromise the witness node before she inserts the replicas into the WSN so as to evade the detection. To solve this problem DM employs n instances of a function so that one claimer id is mapped to n different witnesses. Nevertheless, each of the d neighbours does not necessarily need to forward the location claim to each of all g witnesses. If t is assumed that each reporter behaves independently. The well- studied coupon collector's problem (Cormen etal.,2001) tells us that if the reporters randomly select in all g∑gi ¼ 1 1=i _ gðln gþ0:58Þ repeatable destinations from all g witnesses for the claimer, then More formally, in this protocol, whenever node γ hears a location claim lα from node α, it computes F(α) = {ω1, ω2, . . . , ωg}, where F maps each node ID in the set of possible node IDs, S, to a set of g node IDs: $F : S \rightarrow \{\sigma : \sigma \in 2^S, |\sigma| = g\}$ (1) The nodes with IDs in the set {ω1, ω2, . . . , ωg} constitute the witnesses for node α. Node γ forwards lα to each of these witnesses. If α claims to be at more than one location, the witnesses will receive conflicting location claims, which they can flood through the network ,discrediting α. In this protocol, each node in the network stores g location claims on average. For communication, assuming α's neighbors do not collaborate, we will need each of α's neighbors to probabilistically decide which of the ωi to inform. If each node selects (g ln g) /d random destinations from the set of possible ωi, then the coupon collector's problem [7] assures us that each of the ωi's will receive at least one of the location claims. Assuming an average network path length of O(√n) nodes, this results in O((g ln g√n)/d) messages. Unfortunately, this cost does not provide much security. Since F is a deterministic function, an adversary can also determine the ωis. Thus, they become targets for subversion. If the adversary can capture or jam all g of the messages destined to the ωis, then she can create as many replicas of α as she desires (limited only by the requirement that no two replicas share a neighbor). Since the communication costs of this protocol grow as O(g ln g), we cannot afford a large value for g, and yet a small value for g allows the adversary almost unlimited replication abilities after compromising a fixed number of nodes; in other words, if the adversary controls the g witnesses for α, she can create unlimited replicas of α and suppress the conflicting reports arriving at the witness nodes. These disadvantages make this protocol unappealing.

2.2.2 Randomized multicast (RM): RM also follows claimer–reporter–witness framework .But the witnesses become unpredictable for the adversary Therefore, both schemes can be regarded as improvements of the above DM. One major difference between RM and LSM lies in that in the former protocol the reporters randomly select several witnesses, while in the latter protocol nodes forwarding a location claim (i.e., on the path from a reporter to the corresponding witness) also save the claim for inspection, serving as additional witnesses.To improve the resiliency of the deterministic multicast protocol discussed in Section 4.2, we propose a new protocol that randomizes the witnesses for a given node's location claim, so that the adversary cannot anticipate their identities. When a node announces its location, each of its neighbors sends a copy of the location claim to a set of randomly selected witness nodes. If the adversary replicates a node, then two sets of witnesses will be selected. In a network of n nodes, if each location produces √n witnesses, then the birthday paradox predicts at least one collision with high probability, i.e., at least one witness will receive a pair of conflicting location claims. The two conflicting locations claims form sufficient evidence to revoke the node, so the witness can flood the pair of locations claims through the network, and each node can independently confirm the revocation decision. At a high level, the protocol has each node broadcast its location claim, along with a signature authenticating the claim. Each of the node's neighbors probabilistically forwards the claim to a randomly selected set of witness nodes. If any witness receives two different location claims for the same node ID, it can revoke the replicated node. The birthday paradox ensures that we detect replication with high probability using a relatively limited number of witnesses. More formally, each node α broadcasts a location claim to its neighbors, β1, β2, . . . , βd. The location claim has the format _IDα, lα, {H(IDα, lα)} where lα represents α's location (e.g., geographic coordinates(x, y)). Upon hearing this announcement, each neighbor, βi, verifies α's signature and the plausibility of lα (for example, if each node knows its own position and has some knowledge of the maximum propagation radius of the communication layer, then it can loosely bound α's set of potential locations). Then, with probability p, each neighbor selects g random locations within the network and uses geographic routing (e.g.,GPSR [19]) to forward α's claim to the nodes closest to the chosen locations (as in GHT [30]). Since we have assumed the nodes are distributed randomly, this should produce a random selection from the nodes in the network. In Section 5.3, we show that the probability of selecting the same node more than once is generally negligible. Collectively, the nodes chosen by the neighbors constitute the witnesses for α. Each witness that receives a location claim, first verifies the signature. Then, it checks the ID against all of the location claims it has received thus far. If it ever receives two different locations claims for the same node ID, then it has detected a node replication attack, and these two location claims serve as evidence to revoke the node. It blacklists α from further communication by immediately flooding the network with the pair of conflicting location claims, lα and l_α. Each node receiving this pair can independently verify the signatures and agree with the revocation decision. Thus, the sensor network both detects and defeats the node replication attack in a fully distributed manner. Furthermore, the randomization prevents the adversary from predicting which node will detect the replication.

Security Analysis Let malicious node α claim to be at L locations, l1,l2,. . . , lL. We would like to determine the probability of a collision using the randomized multicast protocol outlined above, since a collision at a witness corresponds to detection of α's replication. At each location li, p · d nodes randomly select g witnesses. If the neighbors coordinated perfectly, this would store α's location claim at exactly p · d · g locations. However, since we prefer to have each neighbor act independently, there may be some amount of overlap between the it nesses each neighbor selects. To determine the impact of this overlap, Nreceive, that will receive the location claim assuming the neighbors choose witnesses independently. If Pclaimis the probability that a node hears at least one claim and Pnone is the probability that a node hears no location claims, then we have:

$E[N_{receive}] = n \cdot P_{claim}$ (2)
$P_{claim} = 1 - P_{none}$ (3)
Since each neighbor is assumed to select g random, unique witness locations, the probability (Pf) that a

$$P_{f} = 1 - \frac{g}{n} \qquad (4)$$

Since each neighbor decides independently whether to send out location claims, the number of nodes that actually send out location claims is distributed binomially, with mean p · d and variance d · p(1 −p). For a network with d = 20 and p = 1 10 , the variance will be less than0.005, so we will approximate the number of neighbors that send out locations claims as p · d. Since the neighbors choose their destinations independently, we have:

$$P_{none} = \left(1 - \frac{g}{n}\right)^{2} \qquad (5)$$

Combining equations 2, 3 and 5, the number of witness nodes that receive at least one location claim is:

$$E[N_{receive}] = n.(1- \left(1 - \frac{g}{n}\right)^{p-d}) \quad (6)$$

The Binomial Theorem allows us to approximate (1 − x)y as (1 − xy) for small x, so as long as g _ n,we have Nreceive ≈ p·d·g, so overlapping witness locations should not impact the security of the protocol. As an example, in a network with n = 10, 000, g = 100,d = 20, and p = 0.1, perfect coordination would tell200 nodes, while independent selection would tell 199.Thus, for the remainder of the analysis, we will assume that p · d · g nodes receive each location claim. If the adversary inserts L replicas of α, we would like to determine the probability that two conflicting location reports collide at some witness node, since this corresponds to the probability that a witness detects the node replication. Note that even if there are more than two replicas of α, we still only need two location claims to collide in order to completely revoke all L of the replicas, since one collision will prompt a network-wide lood of the duplicate claims (li and lj) and any other node that has heard location claim lk for k = i, j will also revoke α. Following the standard derivation of the birthday paradox [7], the probability Pnc1 that the p · d · g recipients of claim l1 do not receive any of the p · d · g copies of claim l2 are given by:

$$P_{nc1} = \left(1 - \frac{p \cdot d \cdot g}{n}\right)^{p \cdot d \cdot g} \qquad (7)$$

Similarly, the probability Pnc2 that the 2·p·d·g recipients of claims l1 and l2 do not receive any of the p·d·g copies
of claim l3 is given by:

$$P_{nc2} = \left(1 - \frac{2 \cdot p \cdot d \cdot g}{n}\right)^{p \cdot d \cdot g} \qquad (8)$$

$$P_{nc} = \pi_{i-1}^{L-1} \left(1 - \frac{i \cdot p \cdot d \cdot g}{n}\right)^{p \cdot d \cdot g}$$

Thus, the probability Pnc of no collisions at all is given by:

Using the standard approximation that $(1+x) \leq e^{x}$, we have:

$$\begin{aligned} P_{nc} &\leq \prod_{i=1}^{L-1} e^{\frac{-i \cdot p2 \cdot d2.g2}{n}} && (10) \\ &\leq e^{\frac{-p2 \cdot d2.g2}{n}} \sum_{i=1}^{L-1} i && (11) \\ &\leq e^{\frac{-p2 \cdot d2.g2}{n}} \frac{L(L-1)}{2} && (12) \end{aligned}$$

Unlike the deterministic proposal (Section 4.2), we no longer need to worry about the adversary using a limited number of captured nodes to enable an unlimited amount of replication. If the adversary captures neighboring nodes α and β, then the total number of claims about either node is reduced by 1d , essentially reducing the neighbor count of each node by one. Since all of the protocol decisions are made locally by individual nodes, the adversary has only two options remaining: it can disrupt the routing of packets from the remaining legitimate neighbors or it can capture all of the legitimate neighbors. Routing disruptions create tell-tale signs of the adversary's presence in the network and will be avoided by a prudent adversary

2.2.3 Line-Selected Multicast

If we interpret RM as a randomized version of DM, LSM can then be regarded as a less expensive version of RM to decrease the communication cost. On a high level, when node a's location claim propagates from a reporter to a witness, all the intermediate nodes on the forwarding path also learn about the information, and can serve as additional witnesses, as shown in Fig. 2. Hence whenever a conflicting location claim by a replica a0 crosses the forwarding path for a, the intermediate node at the intersection of the two path scan detect the conflict, i.e., an intersection corresponds to a detection of the replication attack. This idea can be compared to the following geometric theorem: for x randomly drawn lines within a circle, the expected number of intersections is about 0:339xðx_1Þ, and thus we only need a few such lines to insure an intersection(e.g. ,with only x¼3 random lines weexpect2collisions). Each location claim from a node a is forwarded to r random nodes following the claimer–reporter– witness framework, but the intermediate nodes on the r forward-ing paths also save a copy of the claim. Once another location claim is received by a witness(either selected by a reporter, or more likely, an additional one), it floods the network with the unforgivable evidence to exclude both a and a0 . One may notice that LSM actually ‘‘draws’’ r line segments (paths) that originate from reporters around a central point(the claimer node) and radiate out in random directions(to r random witnesses), instead of random lines (as in the aforementioned geometric theorem).However, even so, simulations indicate that even if there is only one replica(i.e., r random paths radiate from a, another r paths from a0), setting r¼2 assures that the probability for generating at least one intersection (i.e., the detection rate Pd) is above 56%, and r¼ 5 leads to 95%.Similar reliability holds for realistic WSN deployment fields with irregular topologies far different froma circular domain. Compared with RM, LSM has the network communication costs scaling as (O($\sqrt{n}$ ·n· r) and an average node storage cost of (1/n) ·r$\sqrt{n}$ ·n location claims scaling as (O $\sqrt{n}$) .

2.2.4 Single deterministic cell (SDC) and parallel multiple probabilistic cells (P-MPC)

Zhu etal.(2007) proposed two schemes SDC and P-MPC under the brand ‘‘localized multicast’’. Essentially both are variants of DM (Section 4.2), and can be parsed as network-wide deterministic multicast, followed by in cell broadcast and probabilistic storage. In both schemes, the WSN deployment field is considered as a geographic grid of cells, and a location claim from node a is sent by its reporters to g¼1 (SDC)or g41 (P-MPC)cells for in-cell broadcast, the cell id(s)of which is/are deterministically mapped from ID a; each node in the destination cell(s) then probabilistically chooses to be a witness by saving the claim. If there is a replica a0 , its location claim is sent to the same cell(s) for in-cell broadcast, and thus the witnesses can spot the conflict. One may remark that the concept ‘‘localized multicast’’ advocated in Zhu etal.(2007) is not very exact. Both schemes also bear a similar dilemma with DM(Section 4.2) that if the cell size s is too large,theyincurexpensivecommunicationcostlikeN2NB (Section 4.1); if s is too small, they degenerate back to DM, and an adversary can defeat both schemes by compromising all nodes in the g deterministic ‘‘tiny’’ cells. Note that in the latter case(a very small s), all prospective witnesses in one cell are deployed close to each other with in a geographically limited region instead of sparsely spreading throughout the deployment field, and thus it is easy for an adversary to physically approach and compromise them once for all. Therefore, the practicality of SDC and P-MPC relies on careful selection of s. Unfortunately, in Zhu etal.(2007) the critical issue of choosing an appropriate cell size s is over- looked; for all provided examples, s is setto100nodeswithout any explanation/discussion. In practice, one needs to choose s carefully to find an appropriate trade off between efficiency and security. Another problem omitted in Zhu etal.(2007) is what weterm the indistinguishable dilemma. Take SDC for example. Once a location claim by node a arrives at the destination cell, it should be flooded within the cell so that each node in the cell independently stores the claim(i.e., becomes a witness)with probability ps. To reduce the in-cell broadcast to overhead, SDC requires that the

2.2.5 Randomized, efficient, and distributed (RED) detection

Conti etal.(2007) proposed a randomized, efficient, and distributed (RED)protocol, which combines both merits of DM (Section 4.2) and RM(Section 4.3.1). The major motivations from the fair nessor so called ‘‘quality’’ of the detection protocol (Conti etal., 2006): resilience to attack scan be improved by designs that associate individual sensor nodes with equal risk level. For example, a protocol where the likelihood for a genuine node to serve as a witness node (known as the ‘‘node appeal’’) is independent of the node’s geographical position is more favourable, because such an ‘‘area-oblivious’’ protocol actually associates sensor nodes with almost even responsibility. In RED, each of the d neighbours of a claimer a becomes a reporter with probability p, and each reporter sends a’s location claim to a set of g pseudo-randomly selected network locations (hence to g witnesses, like RM).The point is that these pseudo- random locations are computed from ID a with a1-to-g deterministic mapping(like DM), which is seeded with a non received from centralized broadcasting(e.g., from a satellite).Once the random seed is shared network-wide at the beginning of each protocol iteration, the g witnesses are actually deterministic, and the witness set selected by any reporter for a is actually the same. Compared with RM and LSM (Section 4.3), RED’ s philosophy lies in ‘‘just enough witnesses’’, which is inherited from DM. The product dpg can be merely very small constant (5 ffiffiffi n p); it is even enough to set g¼1. Clearly, then ode storage is dpg location claims.. Importantly, the incurred overhead almost evenly balanced among sensor nodes. The probability that a claimer has no reporter is (1-p^{d}), and thus the detection rate is ((1-(1-p^{d}))^{2} assuming there are only two nodes sharing the same id. The pseudo-random choice of witnesses leads to a uniform witness distribution (‘‘area-oblivious’’). On the contrary, in LSM (Parno etal.,2005) a very small central area(for a convex deployment field like a square)may accommodate a large portion of all the witnesses that spot non-coherent location claims, because two forwarding paths are more likely to intersect in the central area; these nodes are just another type of hotspots(recall Section 3.1), and may become appealing target so f attack and/or exhausted quickly. This is termed the crowded center problem in Zhang etal.(2009) to be reviewed in the next subsection. RED (Conti etal., 2007) following Conti etal.(2006) solves this problem justifiably, and an updated version is in Conti etal. (2011). we consider RED as one of the most promising replication detections in the state of the art. Never the less, in Zhang etal. (2009) it is also noted that the infrastructure for distributing RED’s random seed may not always be available. Moreover, since for each protocol iteration the witnesses set for any node is deterministic, there might exist a dilemma in selecting an appropriate g so as to balance between efficiency against node compromise (Zhu et al., 2010).

2.2.6 Memory efficient multicast: B-MEM, BC-MEM, C-MEM, and CC-MEM

Zhang etal.(2009) proposed four replication detection protocols in the name of memory efficient multicast(MEM).The first, B-MEM, is an extension of LSM(Parno etal.,2005), and is the basis of all other three schemes. It reduces the number of stored location claims per node by factor ffiffiffi n p through the use of two compact Bloom filters, which are maintained by semi witnesses (known as watchers) and are reset right before each detection round. However, additional memory consumption per node has to be incurred for storing the two filters(essentially ‘‘compressed’’ location claims),and the overall node storage still scales as Oð ffiffiffi n p Þ (i.e., of the same level with LSM).Moreover, simulations show that B-MEM may lower the detection rate of LSM due to so called false verifications(Zhang etal.,2009) (essentially the intrinsic false positives of Bloom filters).The second, BC-MEM, employs a technique called cell forwarding to solve the cross over problem that unlike geometric line segments intersecting at a common point, in LSM even when two forwarding paths cross they may not intersect at a common node. One can indeed verify the problem by reconsidering the geographic routing(Section 2.2). The third, C-MEM, employs a technique called cross forwarding to address the aforementioned crowded center problem(Section 4.5) that in LSM random forwarding paths end to pass the central area of the deployment field more frequently, where the nodes suffer far worse over heads. For each claimer, C-MEM first select sa random point called the cross point in the network, and forwards the location claim to that point. From there, the claim is then forwarded in four directions, along the horizontal and vertical lines that pass the cross point. Last, CC-MEM integrates cell forwarding and cross forwarding, and thus is a combination of BC-MEM and C-MEM. Simulation results show that the performance of C-MEM is comparable to BC-MEM, because two set so crossing lines have a very high probability to intersect at one or two locations. That is, C-MEM can also mitigate the crossover problem. However, cross forwarding achieves a high probability for intersection only for a convex deployment field, particularly a rectangle (the simulations in Zhang etal.,2009 actually employed a square).For the various irregular topologies considered by LSM such as ‘‘thin cross’’, ‘‘large H’’,etc.(Parno etal.,2005), the cross forwarding technique employed by both C-MEM and CC-MEM may work far poorer than in a rectangle; the detection rate may drop drastically. As to cell forwarding in BC-MEM, the basic idea is to divide the deployment field into virtual cells (like SDC and P-MPC Zhu etal., 2007). By employing a pseudo-random mapping similar to RED (Conti etal.,2007) but seeded with the detection round number W.T. Zhuetal./JournalofNetworkandComputerApplications35(2012)1022–1034 1030 in each cell an anchor node is assigned for each claimer in the network; one anchor node as a representative of the cell may ‘‘serve’’ different network-wide nodes. To solve the cross over problem, BC-MEM only chooses witnesses from these anchor nodes, which serve as definite intersections for forwarding paths. The price is increased energy expenditure, as any location claim is no longer forwarded along an approximately straight path but a zigzag path. A major problem with BC-MEM is that similar to Ho et al. (2009b), the cell division and anchor node selection ask for highly accurate localization, which may not be affordable for the current generation of WSNs. A less serious problem is that an adversary may circumvent BC-MEM by compromising certain deterministic anchor nodes, assuming the detection only runs for a few rounds. An unaddressed problem is the policy for cell size selection (as also observed in SDC and P-MPC Zhu et al., 2007, recall Section 4.4), which makes fair comparison with other schemes difficult. In all simulations the deployment field is always divided into 100 cells without further explanation/discussion.

2.2.7 Randomly directed exploration (RDE)

In Li and Gong(2009a), a simplified version of N2NB(Section 4.1) known as randomly directed exploration(RDE)is proposed, where a location claim along with the claimer’s neighbour list is forwarded in such a manner that the each of the forwarding paths are approximately a straight line segment. We notice such a ‘‘directed’’ (i.e., oriented) forwarding approach is just a special (yet simple) implementation of geographic routing(Section 2.2) that only works for a convex deployment field(them or regular the better).There al interesting part lies in its motivation: RDE tries to mimic N2NB while suppressing broadcast flood. The underlying idea can be interpreted as follows: if the WSN is small-scale but very densely deployed, a ‘‘thin’’ forwarding path can become a ‘‘thick belt’’ to cover sufficient over hearing nodes. Hence it is plausible to substitute such an ‘‘any cast’’ (as suggested in Li andGong,2009a) for broadcast. RDE’s node storage cost remains the same with N2NB (i.e., O(d)), while the network communication overhead is reduced from O(n^2) to O(d· n$\sqrt{n}$)at the price of decreased detection rate. Note that we add the coefficient d to count for the cost of additionally forwarding claimer’s neighbour list(whichisnotforwardedinN2NB),whereas this cost is overlooked in the evaluation in Li and Gong(2009a). Actually, the communication reduction from O(n^2) to O(d· n$\sqrt{n}$)is not very beneficial. Moreover, RDE only seems feasible for an ideal network model, and the detection rate may not be very significant even for a convex deployment field.

2.3 With deployment knowledge scheme can be done using two techniqus.These are as follows:
1.Fingerprint with neighborhood: the best of our knowledge, the only approach that achieves real-time detection of clone attacks in WSN was proposed by Xing et al. [7]. In their approach, each sensor computes a fingerprint by incorporating the neighborhood information through a superimposed s-disjunctive code [16]. Each node stores the fingerprint of all neighbors. Whenever a node sends a message, the fingerprint should be included in the
message and thus neighbors can verify the fingerprint. The messages sent by clone nodes deployed in other locations will be detected and dropped since the fingerprint does not belong to the same “community”.
2.Group deployment
Group deployment technique is again divided into two strategies
1. Order of deployment without location knowledge.
2. Deployment with location knowledge.

REFERENCES

[1] M. Bawa, H. Garcia-Molina, A. Gionis, and R.Motwani. Estimating aggregates on a peer-to-peer network. Technical report, Stanford University, 2003.

[2] C. Blundo and A. Cresti. Space requirements for broadcast encryption. In Advances in Cryptology (EUROCRYPT), 1995.

[3] C. Blundo, L. Mattos, and D. Stinson. Trade-offs between communication and storage in unconditionally secure schemes for broadcast encryption and interactive key distribution. In Advances in Cryptology (CRYPTO), 1996.

[4] D. Braginsky and D. Estrin. Rumor routing algorithm for sensor networks. In Proceedings of ACM Workshop On Wireless Sensor Networks and Applications, 2002.

[5] N. Bulusu, J. Heidemann, and D. Estrin. GPS-less low cost outdoor localization for very small devices. IEEE Personal Communications Magazine, October 2000.

[6] H. Chan, A. Perrig, and D. Song. Random key pre distribution schemes for sensor networks. In Proceedings of IEEE Symposium on Security and Privacy, May 2003.

[7] T. Cormen, C. Leiserson, R. Rivest, and C. Stein. Introduction to Algorithms. MIT Press, 2001.

[8] L. Doherty, K. S. J. Pister, and L. E. Ghaoui. Convex position estimation in wireless sensor networks. In oceedings of IEEE Infocom, 2001.

[9] D. Dolev and A. C. Yao. On the security of public key protocols. IEEE Transactions on Information Theory,1983.

[10] J. R. Douceur. The Sybil attack. In Proceedings of Workshop on Peer-to-Peer Systems (IPTPS), Mar. 2002.

[11] J. Dyer, M. Lindemann, R. Perez, R. Sailer, L. van Doorn, S.W. Smith, and S. Weingart. Building the IBM 4758 Secure Coprocessor. IEEE Computer, 2001.

[12] J. Elson, L. Girod, and D. Estrin. Fine-grained network time synchronization using reference broadcasts.SIGOPS Oper. Syst. Rev., 2002.

[13] L. Eschenauer and V. Gligor. A key-management scheme for distributed sensor networks. In Proceedingsof the ACM Conference on Computer and Communication Security (CCS), Nov. 2002.

[14] D. Estrin, R. Govindan, J. S. Heidemann, and S. Kumar. Next century challenges: Scalable coordination in sensor networks. In Mobile Computing and Networking, 1999.

[15] A. Fiat and M. Naor. Broadcast encryption. In Advances in Cryptology (CRYPTO), 1994.

[16] J. Garay, J. Staddon, and A. Wool. Long-lived broadcast encryption. In Advances in Cryptology (CRYPTO), 2000.

[17] V. D. Gligor. Security of emergent properties in ad-hoc networks. In Proceedings of International Workshop on Security Protocols, Apr. 2004.

[18] A. Hu and S. D. Servetto. Asymptotically optimal time synchronization in dense sensor networks. In Proceedings of ACM International Conference on Wireless Sensor Networks and Applications, 2003.

[19] B. Karp and H. T. Kung. GPSR: Greedy perimeter stateless routing for wireless networks. In Proceedings of Conference on Mobile Computing and Networking (MobiCom), Aug. 2000.

[20] D. Liu and P. Ning. Efficient distribution of key chain commitments for broadcast authentication in distributed sensor networks. In Proceedings of Network and Distributed System Security Symposium (NDSS), Feb. 2003.

[21] M. Luby and J. Staddon. Combinatorial bounds for broadcast encryption. In Advances in Cryptology (EUROCRYPT), 1998.

[22] D. Malan, M.Welsh, and M. Smith. A public-key infrastructure for key distribution in Tiny OS based on elliptic curve cryptography. In Proceedings of IEEE Conference on Sensor and Ad hoc Communications and Networks (SECON), Oct. 2004.

[23] J. M. McCune, E. Shi, A. Perrig, and M. K. Reiter. Detection of denial-of-message attacks on sensor network Broadcasts. In Proceedings of IEEE Symposium on Security and Privacy, May 2005.

[24] R. Merkle. Protocols for public key cryptosystems. In Proceedings of the IEEE Symposium on Research in Security and Privacy, Apr. 1980.

[25] R. Merkle. A digital signature based on a conventional encryption function. In Advances in Cryptology (CRYPTO), 1988.

[26] D. Naor, M. Naor, and J. Lotspiech. Revocation and tracing schemes for stateless receivers. In Advances in Cryptology (CRYPTO), 2001.

[27] J. Newsome, E. Shi, D. Song, and A. Perrig. The Sybil attack in sensor networks: Analysis and defenses. In Proceedings of IEEE Conference on Information Processing in Sensor Networks (IPSN), Apr. 2004.

[28] J. Newsome and D. Song. GEM: Graph embedding for routing and data-centric storage in sensor networks without geographic information. In ACM Conference on Embedded Networked Sensor Systems (SenSys), Nov. 2003.

[29] A. Perrig, R. Szewczyk, V. Wen, D. Culler, and J. D.Tygar. SPINS: Security protocols for sensor networks. In ACM Conference on Mobile Computing and Networks (MobiCom), July 2001.

[30] S. Ratnasamy, B. Karp, L. Yin, F. Yu, D. Estrin,R. Govindan, and S. Shenker. GHT: A geographic hash Bekara C, Laurent-Maknavicius M. A new protocol for securing wireless sensor networks against nodes replication attacks. In: Proceedings of the 3rd IEEE international conference on wireless and mobile computing, networking and communications (WiMob'07); 2007. October. Bonaci T, Bushnell L, Poovendran R. Node capture attacks in wireless sensor networks: a system theoretic approach. In: Proceedings of the 49th IEEE conference on decision and control (CDC'10); 2010. p. 6765–72, December. Boukerche A, Oliveira HABF, Nakamura EF, Loureiro AAF. Localization systems for wireless sensor networks. IEEE Wireless Communications 2007;14(December): 6–12. Brooks R, Govindaraju PY, Pirretti M, Vijaykrishnan N, Kandemir MT. On the detection of clones in sensor networks using random key predistribution. IEEE Transactions on Systems, Man, and Cybernetics, Part C: Applications and Reviews 2007;37(November):1246–58. Chan H, Perrig A. Security and privacy in sensor networks. Computer 2003;36(October):103–5. Choi H, Zhu S, La porta TF. SET: detecting node clones in sensor networks. In: Proceedings of the 3rd international conference on security and privacy in communications networks and the workshops (SecureComm'07); 2007. p. 341–50, December. Conti M, Di Pietro R, Mancini LV, Mei A. Requirements and open issues in distributed detection of node identity replicas in WSN. In: Proceedings of the 2006 IEEE international conference on systems, man, and cybernetics (SMC'06); 2006. p. 1468–73, October. Conti M, Di Petro R, Mancini LV, Mei A. A randomized, efficient, distributed protocol for the detection of node replication attacks in wireless sensor network. In: Proceedings of the 8th ACM international symposium on mobile Ad Hoc networking and computing (MobiHoc'07); 2007. p. 80–9, September.

[31] Conti M, Di Pietro R, MAncini LV, Mei A. Emergent properties: detection of the node-capture attack in mobile wireless sensor networks. In: Proceedings of the 1st ACM conference on wireless network security (WiSec'08); 2008. p. 214–19, March. Conti M, Di Pietro R, Mancini LV, Mei A. Mobility and cooperation to thwart node capture attacks in MANETs. EURASIP Journal on Wireless Communications and Networking 2009: 13 (Article ID 945943). Conti M, Di Pietro R, Mancini LV, Mei A. Distributed detection of clone attacks in wireless sensor networks. IEEE Transactions on Dependable and Secure Computing 2011(September/October):685–98. Cormen TH, Leiserson CE, Rivest RL, Stein C. Introduction to algorithms. MIT Press; 2001. Deng J, Hartung C, Han R, Mishra S. A practical study of transitory master key establishment for wireless sensor networks. In: Proceedings of the 1st international conference on security and privacy for emerging areas in communication networks (SecureComm'05); 2005. p. 289–99. September. Dolev D, Yao AC. On the security of public key protocols. IEEE Transactions on Information Theory 1983;29(March):198–208. Duan M-J, Xu J. An efficient location-based compromise-tolerant key management scheme for sensor networks. Information Processing Letters 2011;111(May): 503–7. Gligor V. Security of emergent properties in ad-hoc networks. In: Proceedings of the 12th international workshop on security protocols; 2004. p. 256–66. April. He W, Liu X, Nguyen H, Nahrstedt K, Abdelzaher T. PDA: privacy-preserving data aggregation in wireless sensor networks. In: Proceedings of the 26th IEEE conference on computer communications (INFOCOM'07); 2007. p. 2045–53 May. Ho J-W, Liu D, Wright M, Das SK. Distributed detection of replica node attacks with group deployment knowledge in wireless sensor networks. Ad Hoc Networks 2009;7(November):1476–88. Ho J-W, Wright M, Das SK. Fast detection of replica node attacks in mobile sensor networks using sequential analysis. In: Proceedings of the 28th IEEE conference on computer communications (INFOCOM'09); 2009b. p. 1773–81. April. Hussain S, Rahman MS. Using received signal strength indicator to detect node replacement and replication attacks in wireless sensor networks. In: SPIE Proceedings of the data mining, intrusion detection, information assurance, and data networks security; 2009. April. Karlof C, Wagner D. Secure routing in wireless sensor networks: attacks and countermeasures. Ad Hoc Networks 2003;1(September):293–315. Karp B, Kung HT. GPSR: greedy perimeter stateless routing for wireless networks. In: Proceedings of the 6th international conference on mobile computing and networking (MobiCom'00); 2000. p. 243–54. August. Kim C, Park C, Hur J, Lee H, Yoon H. A distributed deterministic and resilient replication attack detection protocol in wireless sensor networks. Commu-nications in Computer and Information Science 2009a;56(December):405–12. Kim C, Shin S, Park C, Yoon H. A resilient and efficient replication attack detection scheme for wireless sensor networks. IEICE Transactions on Information and Systems 2009b;E92-D(July):1479–83. Ko L-C, Chen H-Y, Lin G-R. A neighbor-based detection scheme for wireless sensor networks against node replication attacks. In: Proceedings of the 2009 international conference on ultra modern telecommunications and workshops (ICUMT'09); 2009. October. Li Z, Gong G. Randomly directed exploration: an efficient node clone detection protocol in wireless sensor networks. In: Proceedings of the 6th IEEE interna- tional conference on mobile adhoc and sensor systems (MASS'09); 2009a. p. 1030–5. October. Li Z, Gong G. DHT-based detection of node clone in wireless sensor networks. In: Proceedings of the 1st international conference on ad hoc networks (ADHOC- NETS'09); 2009b. p. 240–55. September. Liu J, Baek J, Zhou J, Yang Y, Wong J-W. Efficient online/offline identity-based signature for wireless sensor network. International Journal of Information Security 2010;9(August):287–96. Mathur S, Reznik A, Ye C, Mukherjee R, Rahman A, Shah Y, et al. Exploiting the physical layer for enhanced security. IEEE Wireless Communications 2010;17(October):63–70. Meng X, Lin K, Li K. A note-based randomized and distributed protocol for detecting node replication attacks in wireless sensor networks. In: Proceedings of the 10th international conference on algorithms and architectures for parallel processing (ICA3PP'10); 2010. p. 559–70. May. Newsome J, Shi E, Song D, Perrig A. The Sybil attack in sensor networks: analysis & defenses. In: Proceedings of the 3rd international symposium on information processing in sensor networks (IPSN'04); 2004. p. 259–68. April. Parno B, Perrig A, Gligor V. Distributed detection of node replication attacks in sensor networks. In: Proceedings of the 26th IEEE symposium on security and privacy (S&P'05); 2005. p. 49–63. May. Perrig A, Szewczyk R, Tygar JD, Wen V, Culler DE. SPINS: security protocols for sensor networks. Wireless Networks 2002;8(September):521–34. Poovendran R, Wang C, Roy S. Secure localization and time synchronization for wireless sensor and ad hoc networks.New York Inc: Springer-Verlag; 2007. R¨ uhrup S. Theory and practice of geographic routing. In: Liu H, Leung Y-W, Chu X, editors. Ad hoc and sensor wireless networks: architectures, algorithms and protocols. Bentham Science Publishers; 2009. Sei Y, Honiden S. Reporter node determination of replicated node detection in wireless sensor networks. In: Proceedings of the 3rd international conference on ubiquitous information management and communication (ICUIMC'09); 2009. p. 566–73. January. Table 3 Comparison between the scenarios in replica detections under the claimer– reporter–witness framework. Protocol Assumed deployment model N2NB (Parno etal.,2005) Arbitrary network very small in size DM, RM,LSM(Parno etal.,2005) Arbitrary network SDC, P-MPC(Zhu etal.,2007) A(preferably rectangle)grid of cells RED (Conti etal.,2007) Arbitrary network, preferably rectangle B-MEM (Zhang etal.,2009) Arbitrary network BC-MEM (Zhang etal.,2009) A(preferably rectangle)grid of cells C-MEM (Zhang etal.,2009) Rectangle network CC-MEM (Zhang etal.,2009) A Rectangle grid of cells RDE (Li andGong,2009a) Convex, small-scale, and dense network W.T. Zhuetal./JournalofNetworkandComputerApplications35(2012)1022–1034 1033

[33] Song H, Xie L, Zhu S, Cao G. Sensor node compromise detection: the location perspective. In: Proceedings of the 3rd international conference on wireless communications and mobile computing (IWCMC'07); 2007. p. 242–7. August. Sun B, Osborne L, Xiao Y, Guizani S. Intrusion detection techniques in mobile ad hoc and wireless sensor networks. IEEE Wireless Communications 2007;14(October):56–63. Xie M, Han S, Tian B, Parvin S. Anomaly detection in wireless sensor networks: a survey. Journal of Network and Computer Applications 2011;34(July): 1302–25. Xing K, Cheng X. From time domain to space domain: detecting replica attacks in mobile ad hoc networks. In: Proceedings of the 29th IEEE conference on computer communications (INFOCOM'10); 2010. March. Xing K, Liu F, Cheng X, Du DHC. Real-time detection of clone attacks in wireless sensor networks. In: Proceedings of the 28th international conference on distributed computing systems (ICDCS'08); 2008. p. 3–10. June. Yu C-M, Lu C-S, Kuo S-Y. Mobile sensor network resilient against node replication attacks. In: Proceedings of the 5th IEEE communications society conference on sensor, mesh and ad hoc communications and networks (SECON'08); 2008. p. 597–9. June. Yu C-M, Lu C-S, Kuo S-Y. Efficient and distributed detection of node replication attacks in mobile sensor networks. In: Proceedings of the 70th IEEE vehicular technology conference (VTC'09-Fall); 2009. September. Zeng K, Govindan K, Mohapatra P. Non-cryptographic authentication and identi- fication in wireless networks. IEEE Wireless Communications 2010;17(October): 56–62. Zhang Y, Liu W, Lou W, Fang Y. Location-based compromise-tolerant security mechanisms for wireless sensor networks. IEEE Journal on Selected Areas in Communications 2006;24(February):247–60. Zhang Q, Yu T, Ning P. A framework for identifying compromised nodes in wireless sensor networks. ACM Transactions on Information and Systems Security

[34]Zhang M, Khanapure V, Chen S, Xiao X. Memory efficient protocols for detecting node replication attacks in wireless sensor networks. In: Proceedings of the 17th IEEE international conference on network protocols (ICNP'09); 2009. p. 284–93. October. Zhou J, Das TK, Lopez J. An asynchronous node replication attack in wireless sensor networks. In: Proceedings of the 23rd international information security conference (SEC'08); 2008. p. 125–39. September. Zhu WT. Analysis of a replication attack detection protocol for wireless sensor networks. In: Proceedings of the 3rd international conference on networks security, wireless communications and trusted computing (NSWCTC'11); 2011a. p. 593–6. April. Zhu WT. Node replication attacks in wireless sensor networks: bypassing the neighbor-based detection scheme. In: Proceedings of the international con- ference on network computing and information security (NCIS); 2011b. p. 156–60. May. Zhu S, Setia S, Jajodia S, Ning P. An interleaved hop-by-hop authentication scheme for filtering of injected false data in sensor networks. In: Proceedings of the 25th IEEE symposium on security and privacy (S&P'04); 2004. p. 259–71. May. Zhu B, Addada VGK, Setia S, Jajodia S, Roy S. Efficient distributed detection of node replication attacks in sensor networks. In: Proceedings of the 23rd annual com- puter security applications conference (ACSAC'07); 2007. p. 257–66. December. Zhu B, Setia S, Jajodia S, Roy S, Wang L. Localized multicast: efficient and distributed replica detection in large-scale sensor networks. IEEE Transactions on Mobile Computing 2010;9(July):913–26. Zhu WT, Zhou J, Deng R, Bao F. Detecting node replication attacks in mobile sensor networks: theory and approaches. Security and Communication Networks, available online May 2011. Znaidi W, Minier M, Ubeda S. Hierarchical node replication attacks detection in wireless sensors networks. In: Proceedings of the 20th IEEE international symposium on personal, indoor and mobile radio communications (PIMRC'09) 2009. p. 82–6. September