Enhancing Resiliency Against Routing Layer Attacks in Wireless Sensor Networks: Gradient-based Routing in Focus

Ochirkhand Erdene-Ochir, Apostolos Kountouris
Orange Labs
France Telecom Group
38243 Meylan, France
Email: {ochirkhand.erdeneochir,apostolos.kountouris}@orange-ftgroup.com

Marine Minier, Fabrice Valois
Universite de Lyon, INRIA
INSA-Lyon, CITI
F-69621 Lyon, France
Email: {marine.minier,fabrice.valois}@insa-lyon.fr

Abstract—This paper focuses on the resiliency of wireless sensor network routing protocols against selective forwarding attacks by compromised nodes. Informally, resiliency should be understood as the capacity of the routing protocol to endure and mitigate the presence of a certain number of compromised nodes seeking to disturb the routing process. To provide for security when nodes may be compromised, cryptographic solutions must be completed by algorithmic solutions considering “beyond cryptography” approaches. In this article, after discussing the shortcomings of existing routing protocols against packet-dropping malicious nodes we describe some protocol behaviors enhancing routing resiliency under several combined routing attacks. These behaviors are mainly based on traffic redundancy and probabilistic selection for the next hop candidates, which permit to exploit and benefit from the inherent structural redundancy of densely deployed sensor networks. We consider the case that compromised nodes, prior to selective forwarding, and seeking to increase its impact, may perform several well known routing attacks such as Sinkhole, Sybil and Wormhole. Several variants of the well known gradient-based routing protocol were tested and simulation results show that using the proposed techniques resiliency can be improved. Nevertheless, as expected, resiliency comes at a cost and our results also shed some light on the resiliency-energy consumption trade-off. We propose in this paper the behaviors enhancing the resiliency of routing protocols under several combined routing attacks.

Keywords—wireless sensor networks, routing, security, attacks, resiliency, reliability.

I. INTRODUCTION

In typical Wireless Sensor Network (WSN) applications, a large number of resource constrained sensor nodes are deployed over a geographic area in order to collect physical world data and route them towards one or few destinations (data sinks). The rapid deployment capabilities, due to the lack of infrastructure, as well as the self organized and potentially fault-tolerant nature of WSNs make them attractive for multiple applications spanning from environmental monitoring (temperature, pollution, etc.) to building-industrial automation (electricity/gas/water metering, event detection, home automation etc.). In recent years WSNs have emerged as a very active as well as challenging research area in search for solutions to the open problems of scalability, adaptability, low energy consumption and security. In WSNs the difficulty of all these problems is exacerbated by the large numbers and the resource constrained nature of sensor nodes.

Security is particularly challenging in WSNs. Because of their open and unattended deployment, in possibly hostile environments, adversaries can easily launch Denial-of-Service (Dos) attacks, cause physical damage to sensors, or even capture them to extract sensitive information like for instance encryption keys, identities, addresses etc. Consequently node compromise poses severe security and reliability concerns since it allows an adversary to be considered as a legitimate node inside the network. After node compromise, an adversary can seek to disrupt the functionality of routing layer by launching attacks such as node replication, Sybil, Selective forwarding, Sinkhole or Wormhole. To cope with these “insider” attacks, stemming from node compromise, “beyond cryptography” algorithmic solutions must be envisaged to complement the cryptographic solutions for secure routing proposed in [1], [2], [3], [4]. The work presented in this paper is an extension of our first exploratory work [5] in this direction.

In the existing literature, papers often focus on a particular attack proposing ways to detect and to defend against it mainly by excluding malicious nodes [6], [7], [8], [9]. In this paper, we have chosen to follow a different path; we believe the difference is significant enough to justify the need for further study. Our main goal is not to detect attacks and eliminate malicious nodes, but rather to make the routing protocol capable to continue routing packets, at acceptable success rates, in the presence of malicious nodes. This routing protocol capability will be referred to as resiliency. Also, it is worth mentioning that even though routing resiliency is studied using the Selective forwarding attack as basis of our attack model, interestingly this attack is combined with several other well known routing attacks such as Sinkhole, Sybil and Wormhole. Such combinations represent more realistic attack situations than simply...
considering each attack separately. Finally, since we deal with “insider” attacks, malicious compromised nodes have access to the same information as honest nodes in agreement with Shannon’s maxim: “The enemy knows the system”. Therefore, malicious nodes, aware of defensive strategies against attacks, are expected to adapt their own strategies to optimize the impact of their attacks. From this standpoint our goal is also to dissuade an adversary from creating adapted attack strategies and just settle for basic (random) Selective forwarding.

It should be noted that we believe that if an attacker has decided to break down the network he will succeed by assuming the necessary cost. The required investment depends on cost-benefit analysis considerations quantifying the adversary’s interest in breaking down the network. Under such a worst case scenario protocol resilience will not be effective. However, this is also the case of other approaches, like for instance detecting and isolating malicious nodes. Even if a source node is capable of detecting and isolating malicious neighbors, the packet will not reach the sink if most of its neighbors are compromised. We also show in our simulation results that under Sinkhole attacks where most of the compromised nodes are located around the sink, the sink becomes almost completely disconnected from the rest of the network which in practice is equivalent to the sink being compromised. In what follows we assume that an adversary can compromise only a limited number of sensor nodes, since compromising a node has some cost. In other words, mass attacks, i.e., a large number of both insider and outsider attackers, are out of the scope of this paper. Our main goal is to render a network inherently resilient in the presence of a few malicious nodes, we therefore require that the network performance degrades gracefully as the number of compromised nodes increases. Numerous business applications such as periodic monitoring of electricity, gas, water metering, and environmental monitoring, manipulate some important but not highly sensitive data. In these non-mission critical cases, we assume that an adversary has limited power.

The rest of the paper is organized as follows. Section II, provides an overview of previous work insisting on information, which is relevant in the context of this paper; for instance, routing resiliency and its relationship to other similar notions such as survivability and robustness are discussed and it is explained why classical shortest path routing protocols are not resilient against insider attacks. In Section III, we illustrate our adversary model including network assumptions, adversarial definitions and several routing attacks considered in this paper. We then propose, in Section IV, several probabilistic node selection and packet replication strategies, which improve resiliency by making protocol behavior dynamic and redundant in order to exploit the inherent structural redundancy in the topology of densely deployed WSNs. In Section V we present our approach by mixing and applying these strategies to the well known Gradient-based routing protocol (GBR) [10]; simulations were performed for a basic Selective forwarding attack and for its combination with three other routing attacks, namely Sinkhole, Sybil and Wormhole attacks. Finally, Section VI concludes this paper and outlines future work directions.

II. SCOPe AND RELATED WORK

In this paper, we focus on the Selective forwarding attack where compromised nodes drop data packets. This attack is not only simple but it can be very effective as well. When multi-hop packet routing is considered even a small number of packet-dropping nodes can significantly deteriorate the packet delivery rate of the network. Furthermore, when several routing attacks such as Sinkhole, Sybil and Wormhole are considered in combination with Selective forwarding, this enables adversary nodes to attract more traffic and so amplify the impact of malicious packet dropping.

In this Section a rather rapid overview of previous work is given with the purpose of introducing relevant terminology, concepts and open issues. The vastness of the literature from one side and space limitation from the other do not permit to be more exhaustive but hopefully this brief discussion will help the reader situate our proposal within this research context.

A. Routing layer attacks and countermeasures

Attacks at the network layer were summarized in [11] as follows: (a) spoofed, altered or replayed routing information; (b) Selective forwarding, node replication, Sybil, Sinkhole or Wormhole and HELLO packets flooding. HELLO packets are special control packets sent by each node for neighborhood discovery. We are mainly interested on the attacks of the second type targeting the routing layer. After node compromise an adversary can extract all sensitive information stored in the node. Other attacks such as radio jamming, exhaustion, collisions, link layer jamming or attacks against data aggregation are out of the scope of this paper since they do not directly target the network layer.

In Selective forwarding, malicious nodes simply drop some packets (Greyhole) or all of them (Blackhole) instead of forwarding them as they are supposed to. The main principle of the Sinkhole attack is exactly the same except that the compromised nodes are, or pretend to be, near the sink to attract most of the traffic. After a successful Sinkhole attack the adversary performs Selective forwarding. One possible solution is to use traffic monitoring to ensure that neighbor nodes forward the messages. In [7], the authors propose to use a Watchdog scheme that identifies selfish nodes and a Pathrater scheme that helps routing protocols avoid such nodes. The Watchdog scheme is further extended by a reputation based scheme, [12], where the neighbors of any single node collectively rate the node according to how well the node executes the functions requested.
upon it. Alternatively an acknowledgment based scheme was proposed in [6] to detect maliciously packet dropping nodes.

In the Sybil attack, [8], a malicious device illegitimately takes on multiple identities. Doing so the malicious nodes can fill up their neighbors’ buffers with non existing neighbors and thus create a false topology. Another way to exploit node capture is the node replication/cloning attack [13], where an adversary replicates several nodes with the same identity at different places in the network. To defend against the Sybil attack, the network needs some mechanism to validate that a particular identity is the only identity being held by a given physical node. In [8] the authors describe resource tests and in [13] two distributed algorithms are proposed: randomized multicast and line-selected multicast exploiting the birthday “paradox” to defend against node replication.

Finally, the Wormhole attack, [9], occurs when an attacker receives packets at one location in the network and tunnels them to another, via an out-of-band connection, in order to replay them at this other location. This attack creates a totally false network topology. A Wormhole detection mechanism, called packet leashes, is introduced in [9] and is based on distance estimation; it consists in two mechanisms: geographical leashes and temporal leashes. Another technique to defend against Wormhole attacks consists in using directional antennas [14].

As a conclusion it can be said that most of the proposed solutions strive to defend against a single attack adopting a two stage approach: first, actively detect malicious nodes and second defend against the attack by avoiding routing traffic through the detected malicious nodes. In contrast our aim is not to detect and defend against a single attack, as previously done, but rather to limit damages when several routing attacks are combined together. We propose to do so by enhancing the resiliency of the routing protocols.

B. Resiliency and related notions

According to Webster [15] in mechanics, resiliency is the capability of a strained body to recover its size and shape after deformation caused especially by compressive stress. Hinging upon the general dictionary definition and after reviewing the multiple definitions of resiliency and other similar notions in networking, we define the resiliency in [16] as the ability of a network to “continue to operate” in presence of \( k \) compromised nodes, or in other words, the capacity of a network to endure and overcome internal attacks. Simply put, resiliency is a means to achieve a “graceful degradation” in packet delivery rate with increasing number of compromised nodes.

In the literature, several conceptually similar properties such as survivability [17] and robustness [18], have been discussed but mainly focus on system failures from causes of pure statistical nature contrary to attacks where there is some behind-the-scenes entity with malicious intention. Furthermore, the notion of enduring and overcoming an attack (failure) is not explicitly considered. Finally, resiliency as discussed in [19], [20], [21] is not applied in secure routing but in contexts like robust data aggregation, fault-tolerant routing and key distribution schemes respectively. Nevertheless it should be noted that as [19] compares the resiliency of aggregation functions, our aim is to compare the resiliency of several versions of a given routing protocol.

C. Deterministic routing and its limitations

Insofar minimizing power consumption has been considered a top priority in WSNs research. For increased efficiency, most of the routing protocols use a shortest path criterion to route DATA packets the goal being to reach the sink as quickly as possible. Reactive routing, such as Dynamic Source Routing (DSR) [22], geographical routing, such as Greedy Forwarding (GF) [23] and gradient-based, such as Gradient-Based Routing (GBR) [10], all employ a shortest path principle (with some appropriate definition of “short”). Unfortunately this underlying shortest-path optimization philosophy is responsible for the severe limitations of deterministic routing protocols when attacks involving compromised nodes are considered.

To facilitate discussion lets suppose that some insider attacker has compromised a number of nodes, which are uniformly distributed across the network and which drop all DATA packets they receive. If \( l \) denotes the path length in number of hops from source to destination; \( p_c \) denotes the probability that a node is compromised and \( p_n \) is the probability that a packet is delivered (i.e., all forwarding nodes on the route are legitimate), we have \( p_n = (1-p_c)^l \). In this case, the probability to find a “safe” route exponentially decreases with route length; essentially the same applies for Selective forwarding attacks where only part of the traffic is dropped.

In the presence of such attacks, the routing protocols using shortest paths have better overall delivery ratio. However, they are not resilient. First, as the routes are static all DATA packets from a source node take always the same route to reach a sink. Therefore, if at least one intermediate node is compromised along a route, all DATA packets will be lost and the source node will be completely disconnected from the sink. Second, if a source node has at least one malicious neighbor who will try to attract the traffic (best delay, best gradient, geographically closest to the sink etc.), all DATA packets will be engulfed by such a compromised node. Thus, the routing protocol as is will not be able to overcome this situation since the compromised node will always seem the best routing choice to make.

In previous work, a configurable secure routing protocol (SIGF) has been proposed in [24], extending geographical routing (IGF) [25] with the intention of adding security and protection against outsider attacks. It advocates for an incremental approach to security. As a basis SIGF uses
nondeterminism in neighbor selection, then it adds a reputation scheme and finally it considers cryptography. In a sense SIGF strives for a layered resistance to attacks. However, reputation schemes cannot defend against colluding malicious nodes and cryptographic primitives cannot defend against node compromise. Our aim is to contribute in a similar way by considering, as in SIGF, nondeterminism as a basis of protocol behavior but in our case striving for resiliency, instead of resistance, to several combined attacks, which is more appropriate when compromised possibly colluding nodes are considered.

III. NETWORK ASSUMPTIONS AND ADVERSARY MODELS

In this section we state the network assumptions and several adversarial definitions and we describe the implemented routing attacks.

A. Network assumptions

In the following two types of network device nodes are considered: ordinary sensors and data sinks. Sensor nodes sense and transmit data of the physical world to a single data collector, the sink. Here we deal with WSNs where all sensor nodes are physically identical in terms of transmission range, power, etc. Sensor nodes are densely deployed in a square region of size $N \times N$ and the physical topology of the network is represented by a connected graph. The packets are routed from the source (sensors) to the destination (sink) on this topology.

A common and practical graph model proposed for modeling WSNs is the fixed radius random graph. Let us consider a graph $G(\Omega, E)$ where $\Omega$ is a set of nodes wirelessly connected pairwise by a set of $E$ of undirected edges to represent communication links between nodes. In this model, the nodes are randomly placed in a $N \times N$ region according to a uniform distribution. A link exists between two nodes $i$ and $j$ if the Euclidean distance between these two nodes is less than the communication range $r$. We assume that the wireless links in our graph are bi-directional, i.e., if node $i$ hears node $j$ then node $j$ also hears node $i$.

In addition, from the network security standpoint we use the following, traditionally made, assumptions:

- the “sink” is considered robust, having enough resources in terms of memory, computational power and battery to support the cryptographic and routing requirements of the WSN. Thus, adversaries cannot compromise the sink in limited time.
- the “sensor” has limited resources in terms of memory, computational power and battery. Thus, sensor nodes are non-trustworthy since they are vulnerable to physical attacks and an adversary can compromise them.

B. Adversarial definitions

According to [26] an attack is an intentional act by which an entity attempts to evade security services and violate the security policy of a system; that is, an actual assault on system security that derives from an intelligent threat.

According to its capabilities an attacker can be characterized as:

- A laptop class attacker: It may have access to powerful devices with more computational resources, such as laptops or their equivalent. A single laptop-class attacker might be able to eavesdrop and/or jam the entire network.
- A mote class attacker: It has access to a few motes with the same capabilities as other ordinary sensor nodes. They have no resource advantages over legitimate nodes.

Attacks can also be characterized according to intent as:

- A passive attack: In this attack, the adversary attempts to learn or make use of information from a system but does not affect system resources. For example, passive eavesdropping that simply gathers information, can compromise privacy and confidentiality.
- An active attack: It attempts to alter system resources or affect system operations. Compared to the passive attack, here the goal of the adversary is to produce DoS attacks to disrupt communication by destroying links or exhaust available resources such as bandwidth or energy.

Finally, attacks can be characterized according to point of initiation as:

- An outsider attack: It is initiated from outside the security perimeter by an unauthorized or illegitimate user of the system. Examples are external attacks such as jamming, eavesdropping as well as injecting replayed or fabricated messages.
- An insider attack: It is one that is initiated by an entity inside the security perimeter, i.e., an entity that is authorized to access system resources but uses them in a way not approved by the party that granted the authorization. Selective forwarding, Sybil, Sinkhole or Wormhole attacks being notable examples.

With respect to this classification and given our network assumptions our adversary model considers: “mote-class”, “active”, and, “insider” attackers.

C. Implemented routing attacks

An adversary will try to disrupt communication and cause as much as possible damage to routing protocols. To compare the resiliency of the different protocols, firstly, we modeled the basic Selective forwarding attack, and secondly, we combined it with three other routing attacks; Sybil, Wormhole and Sinkhole. In this sense our attack model is a two-stage combination of simpler attacks. At the first stage, the attacker will launch some attacks in order to enable compromised nodes to attract a lot of traffic. Subsequently at the second stage the compromised nodes will launch the
Figure 1. Basic Selective forwarding attack

Figure 2. Combined Sinkhole attack

Figure 3. Combined Sybil attack

routing attack per se by performing selective forwarding on the attracted packets. We have considered the selective forwarding attack as a basis of our attack model not only because it is common to all protocols but also because this simple attack has a direct impact on reliable data delivery, which characterizes the success of routing protocols.

In the following the main constituents of our attack model will be described in more detail.

1) Basic attack: In multi-hop routing, messages may cross many hops before reaching their final destination. However, a malicious node in the path of data transmission can refuse to forward messages. Selective forwarding is a simple and basic routing attack easy for an insider adversary to launch. After node compromise, malicious nodes instead of forwarding messages with probability 1 they do so with some lower probability. For instance, they can drop all messages (probability to forward = 0) or they can selectively drop some of them in order to avoid detection of their malicious activity (Fig. 1).

2) Combined attacks: For more efficiency, an adversary can exploit its “insider” knowledge to first try to attract traffic and then drop it. Selective forwarding is effective when malicious nodes are on the routes of packet transfer so it is logical to consider it as the final stage of more complex attack behavior where malicious nodes firstly employ some other attack to advantageously place themselves on the routes of heavy traffic and then effect Selective forwarding. Hence, well known routing attacks such as Sybil, Wormhole, Sinkhole could be combined with basic Selective forwarding. This type of combined attacks is explicitly considered within our model.

For instance, to create a Sinkhole, an adversary will try to compromise nodes closer to the sink, exploiting knowledge of location information, to attract most of the traffic (Fig. 2). After a successful Sinkhole attack, the adversary will perform Selective forwarding. The nature of sensor networks where all the traffic flows towards one (or few) sink node(s) makes this type of attacks highly relevant.

Sybil attack is defined by malicious nodes illegitimately taking on multiple identities (Fig. 3) thus compromising the neighborhood discovery process. For instance, a malicious node taking two or more identities will increase the probability of being selected by legitimate nodes as their next hop and then produce Selective forwarding to disrupt routing.

In Wormhole, a malicious node receives packets at one point in the network and tunnels them to another point via an out-of-band connection (Fig. 4). Thus, two malicious nodes can make believe that they are neighbors even if they are physically distant. Well placed Wormholes, for instance an adversary closer to the sink, make possible to attract the traffic of the two hop neighborhood. Wormhole attack is particularly dangerous against routing protocols not only because it creates false topologies but also it permits to attract effectively the traffic. It should be noted that Wormholes is also an effective means to create Sybil identities using existing identities in case legitimate nodes can detect fabricated or duplicated identities.

In the remainder of this paper we propose some routing
behaviors, which could make protocols inherently resilient to such attacks. Our goal is not to detect and to eliminate attacks, but rather to enhance the routing protocols resiliency in order to limit damages.

IV. PROTOCOL BEHAVIORS ENHANCING RESILIENCY

Deterministic protocol behavior forces traffic to flow on a subset of “best” routes, in the quest of optimization (see the discussion in Section II-C). As a result of this, packet delivery success and failure are not fairly distributed among the network nodes; some nodes will have a good delivery ratio and others very bad ones. This is a limitation of the protocol since the network structural (i.e., network topology) redundancy is not exploited to benefit from physically existing alternative routes. In this Section, the techniques that can be employed in order to circumvent this limitation are described.

In this respect resiliency will permit: first, to avoid complete disconnection of nodes; second, graceful degradation of the delivery ratio as the number of compromised nodes increases; and third, obtain packet delivery ratios higher than those achieved by the standard protocols under the same conditions.

The complexity (overhead) of our proposal compared to the deterministic protocol is provided in terms of energy consumption.

Our goal then is to make resiliency emerge through modified protocol behavior. To this end, inspired by previous work, we believe that techniques enabling both dynamic and redundant behavior at the protocol level are needed.

A. Random selection of the next hop

A dynamic (random) behavior can be introduced in different ways according to the routing protocol features. In protocols that require a route discovery process, such as DSR, multiple routes can be discovered once and for each DATA packet the source node can each time select randomly a different route among the discovered ones. In a protocol without route discovery, such as GF, each node can determine a subset of direct neighbors that are closest to the sink compared to itself and choose the next hop randomly in this subset. Depending on how “greedily” a DATA packet should be forwarded, several neighborhood subsets can be constructed. For instance, in a GBR, each node can randomly choose a next hop among those who have a “height” strictly less than itself.

Generally speaking implementing this behavior requires two things. First, the set of selection candidates needs to be defined; it can be of arbitrary size constrained by some maximum allowed distance from the sink. Second, a selection probability law on this set needs to be specified; for instance, it may be desirable that the network node chooses neighbors closer to the sink with higher probabilities. The network node has thus the opportunity to make a random choice for the next hop with a probability to choose the nodes more or less close to the sink.

With this method, the structural redundancy of a physical topology can be effectively exploited in making the protocol fairer in terms of packet loss per node and thus more resilient since the overall packet delivery success can be attributed to a larger population of nodes. Furthermore the energy dissipation at the network is also fairer since the most solicited nodes under a deterministic scheme, i.e., those along the shortest routes, are relieved. Yet another advantage is that attacks targeting state information become less effective since now a single compromised node is not enough to compromise an entire neighborhood. However, this method may decrease packet delivery ratio and increase power consumption due to the lengthening of routes to the sink. There is thus a resiliency-power trade-off that needs to be evaluated. It is possible that by varying the parameters, of candidate set size and selection probability law, this trade-off can be controlled and kept to acceptable levels.

B. Traffic redundancy

Another means to effectively exploit the structural redundancy of the network is to enforce some degree of replication of sent packets. Each replica should then follow its own path to reach the sink.

Here two packet replication schemes to achieve redundancy are considered:

- Nodes replicate their own packets a number of times and send them to an equal number of appropriately selected neighbors. The forwarding nodes do not replicate packets and discard duplicates.
- Packets are replicated both at the source and at each intermediate node along the route. Intermediate nodes discard duplicates of already forwarded packets.

By construction deterministic protocols such as DSR, GF, GBR, cannot take advantage of redundant sends to increase their delivery ratio. If at least one node is compromised along the route, all redundant packets are lost, as they take
always the same route. Such protocols need to be modified to be able to construct alternative shortest routes for each replica but even then their static nature does not allow them to be resilient. In this respect the discovery, construction and maintenance of alternative routes becomes an important consideration. In the literature, most of the multi-path routing protocols use multiple node (or link) disjoint paths to send redundant packets as shown in [27], [28] for example. A packet delivery rate can be increased significantly by using node disjoint multi-path routing. However, as the protocol gets more complex the energy required to discover and to maintain such multiple node disjoint paths is high.

C. Probabilistic routing with traffic redundancy

Finally, we can mix all presented strategies to obtain a random probabilistic routing with traffic redundancy. In this case, the structural redundancy of a physical topology is effectively exploited with some probability to choose longer routes. It will be shown that the random choice of a next hop candidate combined to packet replication naturally implements efficient enough route diversity even though for protocol simplicity node disjoint multiple paths are not guaranteed.

V. SIMULATIONS AND RESULTS

As a first attempt to better understand routing resiliency as well as the associated cost in terms of power consumption, these techniques were applied on the conventional GBR protocol to analyze through extensive simulation if and how its resiliency to attacks is improved. Simulations were performed using WSNet [29], an event-driven simulator for wireless networks.

A. Simulation environment

In our simulations, a unique sink is assumed at the center of the field. The deployed nodes have fixed positions during each simulation. The simulations are averaged over 100 trials for each case with a 95% confidence interval. Table I sums up the simulation parameters.

At this stage we configure WSNet for ideal MACPHY layers (e.g., no interference, no path-loss and no collisions) in order to isolate the impact of the defined attacks on routing and conceptually validate our approach before engaging into more resource consuming simulations and pilot deployments, which ultimately will be necessary.

B. Protocol under study

GBR [10] is a flooding based routing protocol, which is suitable for routing DATA packets from all source nodes to a sink. GBR uses two types of packets: INTEREST and DATA packets. The sink floods an INTEREST packet in order to setup a gradient. The INTEREST packet records the number of hops taken from the sink. Then a node can discover its minimum number of hops from the sink, called the node “height”. The height difference between a node and each of its neighbors is the gradient on that link. The gradient setup process is executed only once at the beginning of the simulation. The following variants of GBR are considered:

1) Deterministic GBR: A given network node $s$ sends DATA packets to a forwarding candidate with the minimum “height” in order to make maximum progress toward the sink. The next hop candidate, $v_i$, is chosen in $V(s)$, $1 \leq i \leq m_s$ (Table II). If several neighbors have the same “height”, we choose the first one registered.

2) Random GBR: A given network node $s$ sends DATA packets to a randomly chosen forwarding candidate with strictly lower “height” than itself. The next hop candidate, $v_i$, is chosen randomly in $V(s)$, $1 \leq i \leq m_s$ (Table II). The nondeterminism introduced by a random selection of the next hop is conceptually similar to SIGF [24]. However, we used a gradient value instead of a geographic distance.

3) Random probabilistic GBR: We have considered two cases according to the probability to select the next hop candidate. Let $p_i$ and $\tilde{p}_i$ be real numbers such that $p_i + \tilde{p}_i = 1$. The considered cases are $p_i = \{0.8, 0.6\}$ and $\tilde{p}_i = \{0.2, 0.4\}$. For a network node $s$, $p_i$ is the probability to choose the next hop candidate $v_i \in V(s)$, $1 \leq i \leq m_s$ (Table II) in the subset of neighbors closer to the sink and $\tilde{p}_i$ is the probability to choose the next hop candidate $w_j \in W(s)$, $1 \leq j \leq l_s$ (Table II) in the subset of neighbors with the same height as itself.

4) Random probabilistic GBR with redundancy: Two cases of redundancy are considered; DATA packets are replicated twice (i) at the source node and (ii) by each node along a full path. In those two cases, duplicate copies of a packet are dropped by forwarding nodes.

C. Implemented attacks

In the following, we assume a unique trustworthy sink. Sensor nodes are assumed untrustworthy since they are

<table>
<thead>
<tr>
<th>Table I</th>
<th>SUMMARY OF THE SIMULATION PARAMETERS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Parameter</td>
<td>Value</td>
</tr>
<tr>
<td>Number of nodes</td>
<td>300</td>
</tr>
<tr>
<td>Area size</td>
<td>100 x 100m</td>
</tr>
<tr>
<td>Transmission range</td>
<td>20m</td>
</tr>
<tr>
<td>Topology</td>
<td>uniformly distributed</td>
</tr>
<tr>
<td>Traffic generation</td>
<td>Poisson distribution $\lambda = 1 \text{ p/s}$</td>
</tr>
<tr>
<td>Simulation time</td>
<td>100s</td>
</tr>
<tr>
<td>Number of packets</td>
<td>30000</td>
</tr>
<tr>
<td>Number of runs</td>
<td>100</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Table II</th>
<th>SUMMARY OF NOTATIONS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Notation</td>
<td>Description</td>
</tr>
<tr>
<td>$s$</td>
<td>a network node</td>
</tr>
<tr>
<td>$h_s$</td>
<td>height of $s$</td>
</tr>
<tr>
<td>$U(s)$</td>
<td>neighbors of $s$</td>
</tr>
<tr>
<td>$V(s)$</td>
<td>neighbors of $s$ with height $&lt; h_s$</td>
</tr>
<tr>
<td>$W(s)$</td>
<td>neighbors of $s$ with height $= h_s$</td>
</tr>
</tbody>
</table>

2011, © Copyright by authors, Published under agreement with IARIA - www.iaria.org
vulnerable to physical attacks and can be compromised. With respect to definition described in Section III-B, malicious nodes may belong to one of the following adversarial categories: “mote-class”, “active”, and, “insider” attackers.

We implemented Selective forwarding as a basic attack and further we considered combining this basic attack with Sybil, Wormhole and Sinkhole attacks.

1) Basic attack: Selective Forwarding. Assuming that the adversary has no information about the location of the sink, the \( k \) compromised nodes are randomly and uniformly distributed on a \( N \times N \) square field (Fig. 5). For simulations \( k \) varies between 10% and 50% of the node population. Malicious nodes do not disturb gradient setup phase and retransmit interest packets with correct hop count. They drop all data packets coming from their neighbors, however, they generate and send their own data packets to the sink.

2) Combined attack #1: Sinkhole with Selective forwarding. Assuming that the adversary has some information about the location of the sink, the \( k \) compromised nodes are randomly distributed on a \( M \times M \) (e.g., \( M = N/2 \)) square field around the sink (Fig. 6). For simulations \( k \) varies between 10% and 30% of the node population. Malicious nodes simply drop all data packets coming from their neighbors. However it is assumed that malicious nodes do not disturb the gradient setup phase, retransmit the interest packets used for gradient setup with with a correct hop count and finally, they normally generate and send their own data packets to the sink.

3) Combined attack #2: Sybil with Selective forwarding. The \( k \) compromised nodes are randomly and uniformly distributed on a \( N \times N \) square field (Fig. 5). For simulations \( k \) varies between 10% and 50% of the node population. According to Sybil attack taxonomy [8], our model corresponds to “direct communication” where Sybil nodes communicate directly with legitimate nodes, using “fabricated identities” where an attacker can simply create arbitrary new Sybil identities (not existing in the network) and it is of the “simultaneous” form where an attacker may participate all of his Sybil identities simultaneously in the network. In this adversary model, malicious nodes take two identities. A compromised node disturbs gradient setup phase by duplicating interest packets. A malicious node puts a false identity to the duplicated interest packet to make believe to their neighbors that there are two nodes, while physically there is only one node. The probability to be chosen for the next hop increases for a malicious node and it can attract more traffic. A malicious node does not lie about its gradient and the two identities take the same true gradient. We choose this particular strategy to separate the impact of Sinkhole attack (which will be the case if the Sybil node lies on its gradient) and of the Sybil attack itself. The false identity is chosen randomly in the large interval of non existing identities to avoid collisions. Once two identities are created, a malicious node drops all data packets coming from its neighbors for both its own and Sybil identities. We also assume only one Sybil identity to be convinced that a Sybil node will not be detected by simple mechanisms such as node degree comparison even if this strategy limits the impact of Sybil attack.

4) Combined attack #3: Wormhole with Selective forwarding. Two colluding malicious nodes can make believe that they are neighbors even if they are physically distant by tunneling messages via an out-of-band connection. Every pair of malicious nodes \((w_1; w_2)\) with a distance greater than two hops, creates a Wormhole link. An interest packet received by \(w_1\) is directly transmitted to \(w_2\) by using the out of band connection. Thus, tunneled interest packets arrive sooner than other packets transmitted over a normal multi-hop route. If \(w_1\) is placed near the sink, \(w_2\) obtain a gradient lesser than its neighbors and \(w_2\) can attract its neighbors’ traffic. The \( k \) malicious nodes are randomly distributed across the whole network, except in the border. The total number of Wormhole links is \(k/2\). For simulations \( k \) varies between 10% and 50% of the node population. Once a Wormhole link is created between two malicious
nodes \((w1; w2)\), they will drop all DATA packets coming from their neighbors. A given malicious node only belongs to one Wormhole link, the case of several Wormhole links coming from a single Wormhole node is not treated here. The Wormhole malicious nodes use legitimate traffic to perform their activity: falsify neighborhood information and attract traffic; collect node identities and use them as Sybil ones instead of having to fabricate false ones.

D. Evaluation metrics

To gain insight concerning the WSN routing resiliency some metrics are needed in order to meaningfully summarize the information collected by simulations. A single such metric is currently lacking and is an object of ongoing research. As a provisional substitute we have used the following metrics:

- **Average delivery ratio (ADR):**
  \[
  ADR = \frac{N_r}{N_s},
  \]
  where \(N_r\), \(N_s\) are respectively the total number of received and sent packets.

  \(ADR\) is an important metric to evaluate the overall success of routing functionality, i.e., packet delivery. To refine over the information provided by \(ADR\), we also measured the delivery ratio per node and we grouped the measurements into 5 classes.

  - **ADR classes:**
    - Class \(c1\): nodes with \(ADR = 100\%\)
      all DATA packets from these nodes are received by the sink
    - Class \(c2\): nodes with \(ADR \in [66\%; 100\%]\)
    - Class \(c3\): nodes with \(ADR \in [33\%; 66\%]\)
    - Class \(c4\): nodes with \(ADR \in [0\%; 33\%]\)
    - Class \(c5\): nodes with \(ADR = 0\%\)
      no DATA packet from these nodes is received by the sink and so they are totally disconnected from the sink

  This measure allows to determine the distribution of transmission success in the node population and the fracture of the network connectivity. In our point of view, the higher the number of connected source nodes (even if with a low \(ADR\)), the more the routing protocol is resilient.

  - **ADR per distance:** The delivery ratio per node is measured and grouped according to the distance (in number of hops) of nodes from the sink.

  To get the distance in number of hops, we take the geographical distance between the source nodes and the sink, and we divide it by the transmission range. All source nodes have the same transmission range. The routing protocols are more resilient if more distant nodes are able to still reach the sink and thus successfully transmit packets.

- **Average path length (APL):** The number of hops crossed by each received packet.

  The end-to-end delay is not explicitly measured in this paper since for our simulations we configure WSNet for ideal MAC/PHY layers which implies no retransmission and no propagation delays. However, the average path length (i.e., hop count) is directly proportional to the average end-to-end delay of the network (see Fig. 11a and Fig. 10) and in this sense it provides an indication of.

- **Normalized power consumption (NPC):**
  \[
  NPC = \frac{T_e}{\tilde{T}_e}
  \]
  where the total energy consumption \((T_e)\) is normalized by the energy consumption of the deterministic GBR without attack and without packet replication sent (\(\tilde{T}_e\)).

  \(NPC\) allows to objectively compare energy expenditure under attacks for each case (including redundancy) without having to enter at this time into low level considerations requiring power consumption modeling. The energy model of WSNet as detailed in the WSNet documentation (see in [30]) is linear: the sleep and idle modes of the MAC layer are not taken into account whereas the basic model considers that the cost for one bit sent is 1 and the cost for one bit received is 2. The total energy is thus computed taking into account the energy cost of each bit received or sent.

E. Results and analysis

The focus of our simulations is on comparing the four versions of GBR (Deterministic, Random, Random probabilistic \(p_t = 0.8\) and Random probabilistic \(p_t = 0.6\)) with a single and two types of redundant DATA packets under four implemented attacks discussed in Section V-C, in term of metrics discussed in Section V-D.

An example of the functional flow diagram with traffic redundancy (double sent full path) under a basic Selective forwarding attacks is presented in Fig. 7.

1) Results for the basic Selective Forwarding attack:

As expected the average delivery ratio (Fig. 8), the average path length (Fig. 11) and the total energy consumption (Fig. 12) decrease with increasing number of compromised nodes under the basic Selective forwarding attack. When a single DATA packet is considered, Deterministic and Random GBR have a higher delivery ratio than others (Fig. 8(a)). The path length is inversely proportional to the average delivery ratio. With probability \(p_t\) decreasing, the average path length (Fig. 11(a)) and the total energy consumption (Fig. 12(a)) increase. However, as the number of the next hop candidates is increased, the structural redundancy of the network is better exploited.

As shown in Fig. 9 in Deterministic GBR only two classes appear. For any source node \(s\) either all DATA packets will be successfully delivered \((ADR_s = 100\%)\), i.e., no malicious node is along the route, or all DATA packets will be lost \((ADR_s = 0\%)\), i.e., at least one forwarding node is compromised along the route. In last case, a source...
node $s$ is completely disconnected from the sink. Note also that the number of disconnected nodes ($c5$) is significantly important (15%) for Deterministic GBR. On the contrary with all variants of Random GBR four classes $c1$ to $c4$ appear. With Random GBR a low number of nodes are completely disconnected from the sink ($c5$). Note that since the network saves energy due to dropped packets by the compromised nodes, this energy gain can then be exploited by redundant DATA packets to further improve resiliency and ADR. In this way, the source nodes can reach the sink as long as possible, thus, enhancing the network connectivity (Fig. 9).

Resiliency and ADR over Deterministic GBR further improve when probabilistic behaviors are mixed with DATA packet replication at the source because DATA packets may take potentially different routes thanks to the random selection of next-hop neighbors. As shown in Fig. 8(b), all random versions exhibit higher delivery ratio performance, though their average path length is higher (Fig. 11(b)), than the Deterministic GBR whose performance remains unchanged. With traffic redundancy, in Fig. 9(b) and (c), we can observe that the number of nodes with higher delivery ratio ($c1$ and $c2$) is increased and the number of disconnected nodes from the sink ($c5$) is decreased for all Random GBR protocols, while for Deterministic GBR the situation remains unchanged. Network reliability is thus improved since most source nodes remain connected.

As expected (Fig. 9 (a)) with decreasing probability $p_t$, ADR decreases when the distance from the sink (in number of hops) increases due to the route length effect. However, with traffic redundancy, the ADR of distant nodes is increased for all random versions, while for Deterministic GBR it remains unchanged (Fig. 9 (b) and (c)). Resiliency is thus improved since distant nodes have better delivery ratio. Nevertheless this has a price, as shown in Fig. 12(b).
the resiliency. It appears that for uniformly distributed compromised nodes variation of the probability $p_t$ does not influence the delivery ratio. So, we may choose the value of $p_t$ that has lower energy consumption. In this respect Random GBR ($p_t = 1$) remains the better trade-off in term of energy-resiliency (Fig. 12(c)). However, it remains to be confirmed if for more realistic spatially distributed compromised nodes, the lower probability $p_t$ may allow better delivery rates as it increases the number of next hop candidates.

2) Results for the combined attacks: In this Section we illustrate results of four versions of GBR with combined attacks; Sybil, Wormhole and Sinkhole with traffic redundancy, where DATA packets are replicated at each intermediate node along a full path.

Sybil attack results: In Fig. 13(a), we observe that the impact of combined Sybil attack is more important than with basic Selective forwarding. When malicious nodes create two identities, they increase the probability to be chosen as the next hop by their neighbors, if they have smallest gradient. Once chosen as the next hop, they receive more packets for retransmission. With traffic redundancy all Random GBR variants have better delivery ratio than Deterministic GBR. The number of nodes in classes $c_1$ and $c_2$ is higher than in other classes for all Random GBR variants (Fig. 14(a)). As a result, with all Random GBRs, most of source nodes have ADR greater than 66% with 10% of compromised nodes and very few nodes are disconnected ($c_5$). In Deterministic GBR, 20% of source nodes are disconnected from the sink with 10% of compromised nodes, while with Random GBR
Figure 11. Basic Selective forwarding - Average path length (APL)

\( p_i = 1 \) only 0.01% are completely disconnected. Network reliability and resiliency are improved again with Random GBR, since most of the source nodes remain connected. The ADR of distant nodes is increased for all random versions, whereas for Deterministic GBR ADR remains unchanged (Fig. 14 (a)). Resiliency is improved with Random GBR under combined Sybil attack, since distant nodes have better delivery ratio. However, the energy consumption with traffic redundancy (Fig. 16(a)) is increased about 3 times.

Wormhole attack results: Fig. 13(b) shows that the impact of combined Wormhole attack is more important than both basic Selective forwarding and combined Sybil attacks. If we consider a pair \( (w_1; w_2) \) of malicious nodes

(c) DATA packets replicated by all forwarding nodes along the route

Figure 12. Basic Selective forwarding - Norm. Power Consumption (NPC)
and if \( w_1 \) is placed near the sink, \( w_2 \) obtains a gradient lesser than its neighbors and the Wormhole can attract the traffic. Here again, all Random GBR protocols have better delivery ratio than Deterministic GBR. In Deterministic GBR, 25% of source nodes are disconnected from the sink with 10% of compromised nodes and with Random GBR \((p_t = 1)\) it is 0.06% (Fig. 14(b)). Network reliability and resiliency are also improved with all Random GBR variants, since the majority of source nodes remain connected (Fig. 14(b)) and the ADR of distant nodes is increased (Fig. 14(b)). Resiliency is improved with Random GBR under combined Wormhole attack and the energy consumption (Fig. 16(b)) due to traffic redundancy remains almost the same as combined Sybil attack.

**Sinkhole attack results:** In Fig. 13(c), we observe that the impact of combined Sinkhole attack is the most important
Figure 15. Combined attacks with DATA packets replicated by all forwarding nodes along the route - Average path length (APL).

Figure 16. Combined attacks with DATA packets replicated by all forwarding nodes along the route - Normalized Power Consumption (NPC).

compared to all other attacks. When the compromised nodes are close to the sink, they receive for retransmission more packets than other nodes: they naturally attract most of the traffic. It is worth noting the significant differences in terms of delivery ratio for all random versions compared to Deterministic GBR as well as among the different versions of Random GBR with traffic redundancy. As packets can take longer routes with Random GBR $p_t = 0.6$ (Fig. 15(c)), messages can find “unaffected” routes around the sink if exist. Hence, distant nodes have more chance to find those “healthy” routes near the sink. The source nodes close to the sink have lower ADR because of the important number
of compromised nodes in their neighborhood (Fig. 14(c)). When all nodes around the sink are compromised, the sink receives packets only from these malicious nodes and no DATA packets are received from the legitimate nodes. That is why we observe on Fig. 15(c) a path length that tends to 1. Resiliency is improved with a Random GBR under the combined Sinkhole attack and the energy consumption (Fig. 16(b)) due to traffic redundancy remains almost the same as with other attacks.

VI. CONCLUSION

In this article, we have considered the case of mote-class/active/insider attacks against WSN multi-hop routing protocols. In this specific context of node compromise cryptography needs to be complemented by algorithmic approaches. We have proposed WSN routing strategies enhancing the protocol resiliency in the presence of maliciously packet-dropping compromised nodes. The basic Selective forwarding attack as well as its combination with Sinkhole, Sybil and Wormhole attacks was thoroughly investigated in the context of the well established GBR.

We have started by analyzing the conditions required for resiliency at the routing layer. The two main findings were that, first, the shortest-path optimization principles though good for energy efficiency are not adapted at all from the routing layer security (i.e., resiliency to insider attacks) standpoint and, second, that the structural redundancy in the network topology should be effectively exploited by employing some form of redundant protocol behavior.

In accordance with these findings our proposal consists in combining random next-hop selection and packet replication; both are needed. A random and probabilistic choice of the next hop candidates allows a dynamic behavior in route selection exploiting thus the structural redundancy of the network. However, the packet delivery ratio may suffer since packets may take longer routes.

With increasing path length (in terms of packet hop count), the overall delay across the network increases as well. The overall delay is directly proportional to the average path length (ideal MAC/PHY layers). However, we observed that under worst attack scenario such as Sinkhole attacks, the average path length of successfully delivered packets tends to one. This can be explained by the fact that with increasing number of compromised nodes, the sink ends up receiving packets only from its direct neighbors. Similarly, in the worst case mass attack scenario (a large number of both insider and outsider attackers), the observed overall delay across the network will also decrease since most of the packets from distant nodes will be lost.

To counterbalance the longer route effect such dynamic (probabilistic) behavior needs to be combined with some form of packet replication. To validate our ideas we have extensively simulated the proposed techniques by modifying in various ways the well-known routing protocol GBR. The results show that the resiliency of routing protocols can be effectively enhanced.

The main merits of our proposal compared to the classical deterministic protocols are:

- the delivery ratio is improved; “graceful” degradation of the delivery ratio with increasing number of compromised nodes.
- the delivery success is fairly distributed; more sources transmit with a high delivery ratio and distant nodes have better delivery success.
- the connectivity is improved; more sources are remain connected to the sink with increasing number of compromised nodes.
- the structural redundancy of the physical topology is better exploited and the energy consumption is fairly distributed; more nodes participate to the routing.

From simulations, we found that traffic redundancy is extremely energy consuming when no attack, but energy efficiency of the protocol is improved when under attack. Hinging on this observation a future work perspective is the search of a mechanism to dynamically adapt the degree of dynamic/redundant behavior to equalize energy cost and so keep the energy consumption-resiliency trade-off at acceptable levels. It also seems that keeping the routes short (in terms of hop count) should be sought but there are some particular cases (e.g., combined Sinkhole attack) where longer routes should be permitted in order to get around obstacles. It is worth mentioning that in our simulation study we have gone beyond the simple Selective forwarding attack to consider combined attacks (such as Wormhole and Selective forwarding) concluding that these attacks have extreme impact on routing especially when Sinkhole and selected forwarding are combined together.

From our simulation analysis we conclude that an operational definition resiliency, in the context of network routing, should incorporate the notions of fairness, preservation of connectivity and graceful degradation of delivery ratio. Thus, our ongoing research especially concerns the definition of a metric of resiliency that includes all those notions. Such a metric will be a valuable tool in analyzing protocol resilience and will greatly simplify the process of protocol comparison.

Finally, in a near future, we also need to relax the ideal MAC/PHY assumption to validate the performance of resilient routing techniques when channel imperfections and medium access limitations are taken into account; to this end it would be interesting to consider modeling packet loss due to MAC/PHY limitations as a form of unintentional Selective forwarding.

ACKNOWLEDGMENT

Ochirkhand Erdene-Ochir is partially supported by a PhD graduate scholarship through a CIFRE contract between France Telecom and the French Ministry of Industry.
REFERENCES


