Increasing Attack Resiliency of Wireless Ad Hoc and Sensor Networks

Harald Vogt
Department of Computer Science
ETH Zurich, Switzerland
vogt@inf.ethz.ch

Abstract

An ad hoc or sensor network that is employed for security sensitive applications is expected to tolerate a certain quantity of maliciously behaving nodes. Algorithms must be designed in a way to withstand attacks such as manipulation and injection of messages. End-to-end security mechanisms, albeit desirable, are generally too costly with regard to the limited computational resources available in sensor networks. We propose and evaluate a robust, scalable scheme for interleaved message authentication, which approximates security guarantees of end-to-end schemes.

1 Introduction

In this paper, we consider large-scale ad hoc and sensor networks that are comprised of resource-deprived devices and engage mostly in ad hoc communication relationships. Communication takes place over short-range wireless connections, and devices relay messages for each other, thus creating a multi-hop environment. Due to network dynamics and load balancing, roles are frequently changed, which means that it is impractical that certain nodes permanently take over special responsibilities, for example regarding security. Creating an end-to-end security relationship between communication end-points is generally prohibitive due to the overhead involved for key establishment and storage. We aim at providing mechanisms that, despite all these constraints, increase the robustness of such networks against certain types of attacks, for example one that tries to partition the network into disconnected parts.

Routing in such networks is most likely based on characteristics of nodes and messages, not mainly on the identity of nodes as in classical networks such as the Internet. Routing mechanisms that are based on message content [6] and on geographic positions of nodes [7] have been proposed. In the following, we assume geographic routing due to its simplicity.

We also assume a *synchronous* system, by which we mean that no unexpected messages are being sent. That means that whenever a message arrives, the receiver knows that the legitimate sender has previously created a message, although not necessarily of the same content as the one that has been received. This property guarantees that the attacker cannot inject extra messages. The only feasible attack is manipulating the content of existing messages.

Our main focus is the protection of the *integrity* of messages that are sent in such a network. Integrity is an important quality of messages, since the results obtained from processing messages depends on their content[1]. Integrity protection is usually achieved through checksums (against accidental faults) or message authentication (against intentional faults). When end-to-end security mechanisms are not available, other means are necessary to strengthen the confidence that a message is genuine, such as multipath transmission and interleaved authentication [9, 10].

2 Adversary Model

We assume that an adversary is able to break into devices that are part of the network and that such an attack requires a fixed effort for each node. Such attacks are possible due to likely weaknesses in the physical layout and protection of nodes, which are physically accessible. Making them tamper-proof would be in general impossible due to the required costs. Once the attacker has gained control over a device, he fully determines its operation. He can block, examine and retransmit all incoming messages, and perform any computational function the device is capable of. In particular, he can make compromised devices cooperate.

We assume that the adversary is not able to make his own devices participate in the network, since all legitimate devices must be certified in some way. This can be guaranteed, for example, by a key-predistribution scheme (see Sect. 3.1). This ensures that only nodes that were present in the pre-distribution phase are able to participate in the network.

[1] Not exclusively, though. Also traffic characteristicts, such as message frequency, could bear important information.

We will not consider attacks that emanate from outside of the network, for example from a powerful device that has a wide radio range (compared to devices in the network) and superior computational power. We assume that all direct communication links are secured (authenticated and encrypted) through locally shared keys.

We will also rule out simple denial of service attacks, since their value is limited to an adversary and they can be easily recognized. We assume that a major goal of the attacker is to prevent the user (i.e. consumer of reports issued by the sensor network) from learning that an attack is going on. Only then it is possible to fool the user into accepting reports from the network that are manipulated by the adversary.

3 Interleaved Authentication

In this section, we describe the *Canvas* communication protocol for protecting the integrity of messages travelling through a network. The core idea of the protocol is to restrain malicious nodes from manipulating messages by implicitly monitoring their actions. Interleaved authentication can handle sets of collaborating nodes up to a certain size, depending on a parameter k. This parameter determines the neighbourhood of a node, i.e. a node must share keys with all nodes within a radius of k hops. For example, for $k = 2$, every node shares a distinct secret key with all other nodes that are one or two hops away, i.e. direct neighbours and their neighbours.

When communicating, a node relaying a message generates k authentication codes, for the next k nodes on the path towards the target. Such a path is shown in Figure 1. A receiving node expects k authentication codes from different nodes in order to accept a message. If at least one of them doesn't match the message content, the message is rejected, so k nodes have to agree on the message content to convince the receiver that a message is genuine. This means that sets of up to $k - 1$ collaborating malicious nodes are prevented from manipulating messages without being detected.

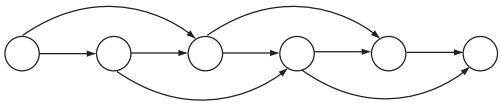


Figure 1. A communication path with interleaved message authentication ($k = 2$)

We note that interleaved authentication paths are essentially equivalent to multiple communication paths on which node-to-node authentication is employed. Their properties can be expected to be similar. However, interleaved paths are cheaper with regard to communication cost, since each message travels only on one physical path. They also impose easier requirements on node deployment, since interleaved authentication is possible even if only one physical communication path exists.

3.1 Local Key Agreement

We assume that any two nodes are in principle able to agree on a shared secret key. Generally, this is achieved through the help of a central, trusted party or a key predistribution scheme [2, 4]. If the context of network deployment allows for it, we can assume a time frame within which no attack is possible and therefore all communication can be considered safe from eavesdropping and manipulation. Within this time frame, nodes could exchange keys in plaintext, but after it has been closed, no further key agreement would be possible.

Pre-distribution of keys enables each pair of nodes to establish a common key (at least, with high probability). Generally, this is affordable for two arbitrary nodes only if they are either close to each other (which induces only a small communication overhead) or, for distant nodes, if the overhead pays off due to special requirements. Shared keys between very distant nodes are thus likely to be scarce.

Before the *Canvas* protocol is used for communication, nodes agree with their neighbours on pairwise shared secret keys. The number of keys a node has to store depends on the topology of the network. In a grid network, for $k = 2$, a node would have to store 24 such keys.

In a static network, this key negotiation phase would take place only once. In networks where nodes can be replaced or are mobile, key agreement would take place whenever a node appears in another node's range. This event is then propagated to all nodes within a k-hop distance, which would then also initate a key agreement with the new node.

3.2 Communication

When a message transmission is initiated, the message source selects the first two hops on the path towards the target. How exactly these hops are chosen depends on the routing algorithm. In a static network using geographic routing, the source can choose the next k hops independently, since the positions of all neighbours within a k-radius can be assumed to be known (e.g., exchanged during key setup). Under other routing regimes or under different network dynamics, it might be necessary to negotiate the next hops step-by-step. However, we will not elaborate on this procedure in this paper. For the presentation of the *Canvas* protocol we will fix a parameter $k = 2$, but the extension to other values is straightforward.

For $k = 2$, the source, let's call it S_1, computes two authentication codes, one targeted at each of the follow-

ing hops. Thus, the source authenticates the message towards two distinct nodes. The message will then be passed on along the communication path only if both these nodes agree on the content of the message.

In order to send a message M on its way, the message source S_1 performs the following steps:

> Compute two authentication codes: $a_{1,2} = \mathsf{MAC}(K_{1,2}, M||S_1)$ and $a_{1,3} = \mathsf{MAC}(K_{1,3}, M||S_1)$. Send to S_2: $\langle \textsc{Init}, M, S_1, (S_1, a_{1,2}), (S_1, a_{1,3}) \rangle$

First, two authentication codes are computed using the keys shared between S_1 and the next nodes on the path, S_2 and S_3. Here, the authentication code comprises the message content M and the identity of S_1. Next, S_1 sends a message to S_2, containing all information allowing S_2 and S_3 to verify the authenticity of M. The marker INIT denotes the message as a non-relayed message. This is important since such messages are accepted by a receiving node if they are authenticated only once, but they have to be authenticated by the original source itself.

The second node on the path, S_2, receiving the message constructed by S_1, performs the following steps:

> Check if $\mathsf{MAC}(K_{1,2}, M||S_1)$ matches the value contained in the message. If it does, compute $a_{2,3} = \mathsf{MAC}(K_{2,3}, M||S_1)$ and $a_{2,4} = \mathsf{MAC}(K_{2,4}, M||S_1)$. Send to S_3: $\langle \textsc{Pass}, M, S_1, (S_1, a_{1,3}), (S_2, a_{2,3}), (S_2, a_{2,4}) \rangle$

If the verification in the first step fails, S_2 discards the message (possibly issuing a warning). Otherwise, it chooses the next hop on the path and creates two more authentication codes. It then continues to send a message with a PASS marker that indicates that this is a relayed message. Note that the last entry in the received message, the pair $(S_2, a_{2,4})$, is simply copied to another position in the new message, so that it is simply passed through to the next node without any processing, since it contains no information useful to S_2.

The ith node on the communication path, starting with $i = 3$, will perform the following steps:

> Check if $\mathsf{MAC}(K_{i-2,i}, M||S_1)$ and $\mathsf{MAC}(K_i - 1, M||S_1)$ match their respective counterparts in the message. If they do, compute $a_{i,i+1} = \mathsf{MAC}(K_{i,i+1}, M||S_1)$ and $a_{i,i+2} = \mathsf{MAC}(K_{i,i+2}, M||S_1)$. Sent to S_{i+1}: $\langle \textsc{Pass}, M, S_1, (S_{i-1}, a_{i-1,i+1}), (S_i, a_{i,i+1}), (S_i, a_{i,i+2}) \rangle$

Here, we make no assumptions about when the protocol terminates. In fact, the protocol could be seen as broadcasting the message along the path, since all nodes have access to the content of M. The protocol would then terminate simply when there are no more nodes left to which the message could be sent. Otherwise, a destination address, such as a geographic location, could be included in M. If a node close enough to this location receives the message, it would terminate the protocol and start processing M.

3.3 Limits of Interleaved Authentication

Given a parameter k, interleaved authentication can tolerate up to $k - 1$ consecutive nodes on a communication path. That means that for $k = 2$, single, isolated malicious nodes without malicious, cooperating neighbours, have no effect on the network's operation. However, larger groups of compromised nodes could manipulate messages.

Since there is no direct authentication of a message between its source and its destination, there is no immediate way of knowing whether a message is genuine or not. There might be k or more malicious cooperating nodes on the communication path, manipulating or injecting messages, but the receiving node has no way of knowing about them. Other means are required to determine if there are compromised nodes acting in the network. For example, the receiver of a message could occasionally send an acknowledgement back to the source, including a hash code of a recent message, deliberately choosing a path physically different from the path over which the original message arrived. If such a path exists and can be used, and unless this path has been compromised, the sender will learn whether his message has arrived correctly or not.

A receiver could also create a (temporary) secret key between him and the sender for the purpose of giving the sender feedback about recent messages. If the sender learns that some of his messages have been tampered with, he could emit a warning.

Generally, the attacker could try to misroute messages such that they pass through compromised nodes. For example, a compromised node could pretend that the only connection available is one that goes through another compromised node. It is generally not possible to verify if this claim is true. We propose to design the network in such a way that multiple paths do exist so that such claims are not plausible and would lead to an intrusion alert. A malicous node could also lie about its neighbours. However, it cannot introduce faked nodes, which is prevented by the key pre-distribution scheme.

A powerful attack on a network that is equipped with the *Canvas* protocol is a partitioning attack. An attacker would have to break into a very limited number of nodes only, comprising a narrow band of width k that separates the two parts, in order control all messages being sent between the two parts. All paths between the partitions would lead through this band and would be subject to manipulation. A node located in one part of the network could receive

no authentic messages from the other part any more. We therefore propose the shortcuts as an extension to the basic *Canvas* protocol, described in the following subsection.

3.4 Shortcuts

Shortcuts are links that are established between distant nodes, i.e. nodes that are not in their mutual neighbourhood. Each node stores a certain, small, number of these keys that "re-enforce" the *Canvas* path from a message source to its destination. When a message is sent, it is first routed to the shortcut node that is closest to the message target. On its way from the shortcut node to the destination, the message is authenticated with the basic *Canvas* protocol. This "last mile" is therefore much more weakly protected, but the impact of such a short compromised path is much more localized than with basic *Canvas* only.

We can also build interleaved paths from shortcuts, as for example shown in Fig. 2, to span very large distances. They can be built by local information only: Each node on a segment from A to B (A and B sharing a key) looks if it finds an own shortcut node closer to the target than B. The MAC for the "best" such pair is then passed on by B. If no further shortcut node can be found, the last segment, which is very short in most cases, can be bridged by basic *Canvas*.

Figure 2. An long-distance interleaved authentication path

The assignment of shortcut nodes can be made randomly before the network is deployed, or nodes select their shortcuts during operation, performing normal key agreement with distant nodes they choose randomly, possibly based on the traffic in the network. Pre-assignment, however, has clear advantages since it requires no message exchange.

As we will show in the next section, shortcuts are very effective in defeating partitioning attacks. If shortcut keys are randomly distributed, it is highly likely that at least one shortcut node is located close to the target, even if the number of shortcut keys per node is rather small.

4 Simulation Results

Baran used the fraction of hosts being able to communicate as a metric for the "survivability" of a network under attack [1]. Similarly, we use the fraction of *functional* paths as a metric for the security of a network. (A path is called *functional* if both endpoints are uncompromised and the endpoints can communicate securely.) The following simulation results are were obtained from graph-based simulations with the following parameters: 250 nodes, randomly placed on a 1000×500 units area, communcation range 100 units, 10 shortcuts per node (if applicable). The neighbourhood was given by $k = 2$. The simulations were run 10 times and the results averaged.

Fig. 3 compares the effectiveness of different authentication schemes under the partitioning attack. Each compromised node breaches a set of paths: all paths that end at that node, plus those paths that pass through the node and are controlled by the attacker. The left axis shows the fraction of end-point-breached paths to all breached paths. This means that under an end-to-end authentication scheme, the value would always be 1 (only end-point-breached paths are breached). To partition the network, the attacker compromises one node after the other in a band that goes vertically through the area in the middle of it. As one can see, the Interleaved (i.e., interleaved shortcuts) and the Shortcut/Canvas schemes come close to the optimum. The others are much worse, i.e. an additional single compromised node has much greater impact. Of these others, basic Canvas performs best for few compromised nodes. (Hop-to-hop authentication means that only traffic between neighbours is authenticated. This scheme performs worst of all and can be slightly improved through shortcuts.)

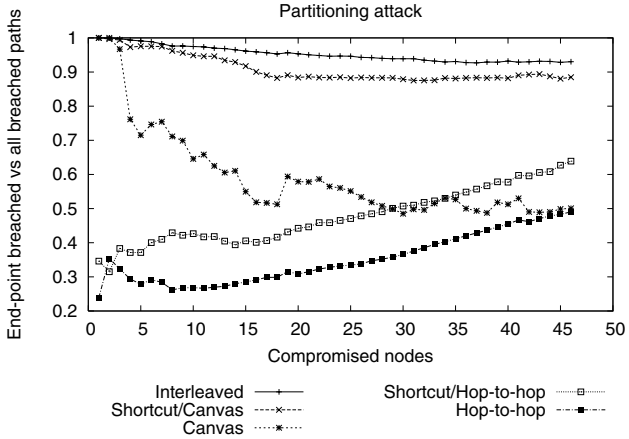


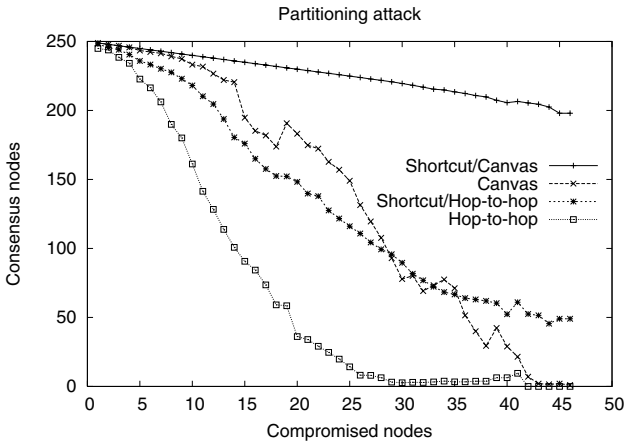
Figure 3. Impact of compromised nodes

One important feature of networks is the ability to agree on a common value, e.g. for taking a crucial decision. Thus, we evaluate the number of uncompromised nodes that are able to participate in a consensus protocol. Byzantine agreement is possible with at most one third of compromised nodes [8], so we are interested in those nodes that can securely communicate with at least $2/3$ of all non-compromised nodes. In Fig. 4(a) and 4(b), the number of consensus-enabled nodes is shown for the partitioning attack and a different attack type, the "concentrated" attack, where the attacker starts in the middle of the network and successively compromises nodes that are connected to al-

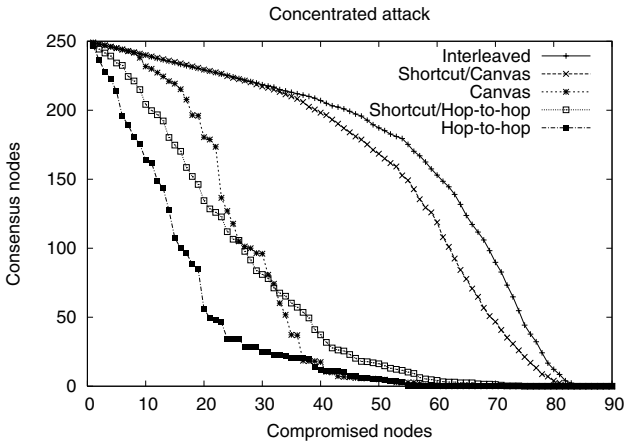
ready compromised nodes.

Under the partitioning attack, the Canvas/Shortcut performs nearly optimal (the optimum is all uncompromised nodes). The Interleaved scheme is not shown here, since its graph overlaps with the Canvas/Shortcut scheme almost completely. Thus, we have found an effective countermeasure against the partitioning attack.

The concentrated attack (Fig. 4(b)) was carried out until more than $1/3$ of all nodes were compromised. At that point, consensus is not possible anymore. Both Interleaved and Shortcut/Canvas can withstand such an attack very long.



(a) Consensus (partitioning attack)



(b) Consensus (concentrated attack)

Figure 4. Ability for consensus under attack

5 Related Work

Interleaved authentication was studied in [11] in a slightly different context than ours. There, groups of sensor node agree on a common value that is sent along a communication path towards a base station. Interleaved authentication is applied on the communication path, allowing it to drop messages that have been manipulated. Interleaved communication was, to our knowledge, first introduced in a paper by Craig and Reed [3] in order to increase network efficiency and reduce routing costs, but without any reference to security or reliability.

6 Conclusion and Future Work

We are looking for scalable security mechanisms for networks which cannot support costly end-to-end mechanisms. Interleaved authentication with shortcuts is a proposal for protecting the integrity of messages. It is suitable for resource-constrained devices and supports transient ad hoc communication.

Future work includes more and improved simulation scenarios, protocol verification, investigation of applications for interleaved authentication, and the general connection between small-world [5] properties and security.

References

[1] Paul Baran. On Distributed Communications Networks. *IEEE Transactions on Communications Systems*, 12(1):1–9, 1964.

[2] Haowen Chan, Adrian Perrig, and Dawn Song. Random Key Predistribution Schemes for Sensor Networks. In *Proceedings of the 2003 IEEE Symposium on Security and Privacy*, pages 197–213. IEEE, May 2003.

[3] L. J. Craig and I. S. Reed. Overlapping Tessellated Communications Networks. *IEEE Transactions on Communications Systems*, 10(1):125–129, March 1962.

[4] L. Eschenauer and V. D. Gligor. A Key-Management Scheme for Distributed Sensor Networks. In *CCS'02*. ACM, 2002.

[5] Brian Hayes. Graph Theory in Practice: Part II. *American Scientist*, 88(2):104–109, 2000.

[6] Chalermek Intanagonwiwat, Ramesh Govindan, and Deborah Estrin. Directed Diffusion: A Scalable and Robust Communication Paradigm for Sensor Networks. In *Proc. of 6th Ann. Int. Conf. on Mobile*

Computing and Networking (MobiCom), pages 56–67. ACM, 2000.

[7] B. Karp and H. T. Kung. GPSR: Greedy Perimeter Stateless Routing for Wireless Networks. In *Proc. of MobiCom*. ACM Press, 2000.

[8] Leslie Lamport, Robert Shostak, and Marshall Pease. The Byzantine Generals Problem. *ACM Trans. Program. Lang. Syst.*, 4(3):382–401, 1982.

[9] Harald Vogt. Exploring Message Authentication in Sensor Networks. In *Proc. of European Workshop on Security of Ad Hoc and Sensor Networks (ESAS)*, LNCS. Springer-Verlag, 2004.

[10] Harald Vogt. Integrity Preservation for Communication in Sensor Networks. Technical Report 434, ETH Zürich, Institute for Pervasive Computing, February 2004.

[11] Sencun Zhu, Sanjeev Setia, Sushil Jajodia, and Peng Ning. An Interleaved Hop-by-Hop Authentication Scheme for Filtering False Data Injection in Sensor Networks. In *IEEE Symposium on Security and Privacy*. IEEE, 2004.