Modeling Node Capture Attacks in Wireless Sensor Networks

Invited Paper

Patrick Tague and Radha Poovendran
Network Security Lab (NSL), Electrical Engineering Department,
University of Washington, Seattle, Washington
Email: {tague, rp3}@u.washington.edu

***Abstract*—We formalize a model for node capture attacks in which an adversary collects information about the network via eavesdropping on the wireless medium and captures nodes based on the learned information. We show that attacks in this adversary model correspond to NP-hard optimization problems and discuss the behavior of a reasonable heuristic algorithm. We show that the goals of node capture attacks can be decomposed into a collection of primitive events, the impact of which can be evaluated and recombined to yield an overall evaluation of the attack. We demonstrate the use of the attack decomposition model for derivation of attack metrics and discuss the potential use of this decomposition technique for the purposes of defense against node capture attacks.**

I. Introduction

Two of the fundamental features of a wireless sensor network (WSN) are the ad-hoc nature of the system which is expected to set up and maintain its own communication architecture without the help of a centralized authority [1] and the fact that the sensor nodes are expected to operate for long periods of time without attention from a higher-level presence (e.g. a human). These features are due largely to the proliferation of wireless networking and sensor hardware technologies, allowing cheap wireless devices to make wireless sensor networking a viable option for personal, commercial, industrial, and military uses. However, with these rapid advances in wireless networking and sensor hardware technologies come inherent vulnerabilities, creating potential for a wide variety of attacks on network services.

Use of the open, shared wireless medium allows any nearby device to overhear messages transmissions (eavesdropping), interfere with or block message reception (jamming), insert messages into the network, or replay previously transmitted messages. Furthermore, even if the message payload is encrypted, the exchange of messages and the presence of message headers potentially allow an eavesdropper to learn about the network operation and protocol state. A large body of research exists on the problem of information privacy, aiming to prevent information leakage to eavesdroppers and protect the users' identities, credentials, and data [2], [3]. Such approaches typically rely on cryptographic primitives to create a security architecture within the network and prevent external eavesdroppers or adversaries from gaining access to the information held by valid network users.

In a WSN, an external adversary can, however, gain access into the network as a valid user by physically attacking, or capturing, sensor nodes and extracting the desired information from the nodes' memories [4], [5]. Even if the adversary does not need to extract secret information from the nodes to gain access into the network's security architecture, capturing nodes once the network is deployed provides a cheap and effective approach to establish an adversarial presence in the WSN and influence the outcome of network protocols. Once the adversary controls a set of sensor nodes, these nodes can be arbitrarily altered in terms of both software programming and hardware configuration, including attacks such as node cloning [6]. This access into the network and control of sensor nodes can then be used as a foundation for further attack on the network.

In order to eventually design network protocols and primitives that are robust to such *node capture attacks*, we must first understand these attacks at a fundamental level. Hence, in this work, we investigate the problem of modeling node capture attacks in WSNs in terms of the impact of the attacks on the network protocols and security by decomposing the attacks into a primitive set of events. Once the attacks are decomposed into primitive events that are well-understood, approaches to defend against node capture attacks can be made by defending the decomposed events.

In this work, we make the following contributions.

- We formalize node capture attacks from the adversary perspective as an NP-hard optimization problem and discuss the behavior of a reasonable heuristic algorithm.
- We discuss the event-based decomposition of attacks into primitive events and the ability to define attack evaluation metrics with respect to the decomposition.
- We demonstrate the use of the event-based attack decomposition and evaluation of attacks and illustrate example

This work was supported in part by the following grants: ONR YIP, N00014-04-1-0479; ARO PECASE, W911NF-05-1-0491; ARL CTA; and ARO MURI, #W 911 NF 0710287.

This document was prepared through collaborative participation in the Communications and Networks Consortium sponsored by the US Army Research Laboratory under the Collaborative Technology Alliance Program, DAAD19-01-2-0011. The US Government is authorized to reproduce and distribute reprints for Government purposes notwithstanding any copyright notation thereon. The views and conclusions contained in this document are those of the author and should not be interpreted as representing the official policies, either expressed or implied, of the Army Research Laboratory or the US Government.

node capture attacks.

- We discuss the potential use of the event-based decomposition and evaluation metrics from the network perspective for the purposes of defense against node capture attacks.

The remainder of this paper is organized is follows. In Section II, we characterize the metrics used in node capture attacks and formalize the node capture attack model in terms of an optimization problem. In Section III, we discuss the decomposition of attacks into a set of events and describe the derivation of attack metrics from the event-based decomposition. In Section IV, we illustrate the use of the proposed modeling and evaluation techniques with a variety of example attacks. In Section V, we discuss defense mechanisms based on the decomposition, and we conclude in Section VI.

II. Node Capture Attacks

Node capture attacks result from the combination of passive, active, and physical attacks by an intelligent adversary. In order to initialize or set up an attack, the adversary will collect information about the WSN by eavesdropping on message exchanges, either local to a single adversarial device or throughout the network with the aid of a number of adversarial devices deployed throughout the network. Even if message payloads are encrypted, the adversary can extract information about the network operation and state, effectively learning about the network structure and function. In addition to passive learning, the adversary can actively participate in network protocols, probing the network for information and maliciously injecting information into the network. Once a sufficient amount of passive and active learning has taken place, the adversary can physically capture nodes. The gathered information can be used to help the adversary make an informed decision of which sensor nodes to capture in order to optimize the performance of the attack with respect to a specific attack goal.

To serve as a basis for attack optimization and in order to measure the performance of an on-going attack, the adversary must develop an attack performance metric for the specified attack goal. In addition to the evaluation metric, the adversary must also be able to associate a value with each potential node capture in order to optimize the impact of the attack. Because of the purpose of the node value metric, it will be intimately related to the attack performance metric. A notable difference between the performance metric and the node value metric, however, is that the adversary may incorporate the heterogeneity of the WSN nodes by normalizing the node value with respect to the associated attack cost. This attack cost is not an immediate factor in the attack performance metric.

We suppose that the adversary is interested in performing the node capture attack which optimizes the derived attack performance metric with minimum total attack cost. The optimization variable in this formulation is the subset $\mathcal{C}$ of nodes to capture. Since the choice of the subset $\mathcal{C}$ from the set of nodes $\mathcal{N}$ in the network is equivalent to choice of an optimal $|\mathcal{N}|$-dimensional binary vector of weight $|\mathcal{C}|$, the optimization formulation is that of an integer programming problem. Due to inherent computational complexity of determining optimal solutions to integer programming problems, we assume the attack will be carried out using a suitable heuristic. In particular, we suppose that nodes are iteratively added to the set $\mathcal{C}$ by choosing the node $n \in \mathcal{N} \setminus \mathcal{C}$ with the maximum normalized incremental value $v_{\mathcal{C}}(n)$ with respect to the previously captured nodes in $\mathcal{C}$ given by

$$v_{\mathcal{C}}(n) = \frac{value(\mathcal{C} \cup \{n\}) - value(\mathcal{C})}{cost(n)}, \tag{1}$$

which is a non-negative quantity, as adding a node to the captured node set $\mathcal{C}$ cannot decrease the value of the attack.

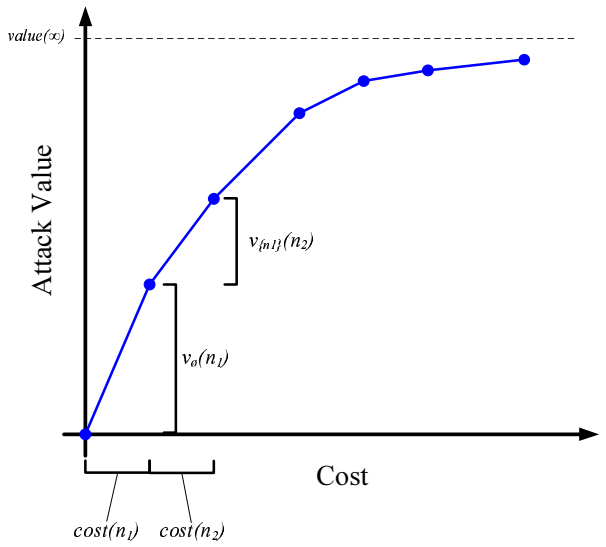


Fig. 1. The function $value(\cdot)$ plotted as a function of total attack cost, itself an increasing function of time, increases to the value $value(\infty)$ with a decreasing rate.

Using the iterative heuristic according to the node value function in (1), the performance metric $value(\cdot)$ can be plotted as a function of total attack cost, which is itself an increasing function of time. Graphically, the goal of the heuristic attack at each iteration according to the normalized value function is to cause this function to increase with the greatest slope at each iteration, as illustrated in Figure 1. By construction, the slope will be decreasing as time progresses, meaning that the value of the attack approaches the final value $value(\infty)$ with a decreasing rate.

Since the basic goal of the attack is to optimize the performance of the attack, one of the most valuable assets to the adversary in performing node capture attacks is a suitable metric. As described above, these metrics can be developed by decomposing the attack goal into a set of events. This decomposition is described in what follows.

III. Event-Based Attack Decomposition

In this section, we propose a method for the development of suitable performance metrics for node capture attacks by decomposing the attack goal into a collection of events. By arranging the attack tasks into a graphical structure, the value

of certain events can be computed via graph composition as a function of the corresponding sub-event values.

Suppose that the adversary is interested in achieving a particular attack goal. This goal is most likely to cause some sort of noticeable effect on the network and it is likely to be a composition of a number of attack events. By decomposing the goal into these individual events, the adversary is better able to gauge the progress of the attack toward the desired goal. To further simplify the attack evaluation, we suggest a further decomposition of attack events into simpler sub-events, until a collection of easily described primitive attack events is obtained, noting that such a decomposition need not be unique. The decomposition of the attack into these primitive events similarly allows for decomposition of the attack evaluation metric into quantities that measure the value of accomplishing individual events. Once a set of nodes $\mathcal{C}$ has been captured, the attack performance metric can be evaluated by recombining the values of the achieved primitive events by reversing the original decomposition. This decomposition and recombination process is formalized as follows.

A. Graph Decomposition for Attack Evaluation

Suppose the desired attack goal is given by a particular event $\mathcal{G}$ which can be decomposed into a collection of events $\mathcal{E}$, including a subset $\mathcal{P} \subseteq \mathcal{E}$ of primitive events. We assume that the event $\mathcal{G}$ is contained in the union of events

$$\mathcal{G} \subseteq \bigcup_{e \in \mathcal{E}} e \tag{2}$$

and that at least one event e must be accomplished in order to satisfy the event $\mathcal{G}$. We let $\mathcal{D}(\mathcal{G})$ denote the directed graph of the decomposition for the attack goal corresponding to the event $\mathcal{G}$. The vertex set V of $\mathcal{D}(\mathcal{G})$ is given by the set of events $\mathcal{G} \cup \mathcal{E}$, and the edge set E is given by the collection of pairs (e_i, e_j) where e_j is a sub-event of e_i. For every non-primitive event $e \in \mathcal{E} \setminus \mathcal{P}$, we assume a condition similar to that in (2) given by

$$e \subseteq \bigcup_{(e,e_i) \in E} e_i. \tag{3}$$

Note that the graph $\mathcal{D}(\mathcal{G})$ is necessarily acyclic. An example decomposition and the corresponding decomposition graph are given in Figure 2.

B. Event Evaluation in Graph Decomposition

Once the graph decomposition is constructed, the relationships between the event values $v(e)$ for each $e \in \mathcal{E}$ can be determined. In order to relate event values, we make use of rules analogous to those of discrete probability distributions by normalizing the event values such that $v(\mathcal{G}) = 1$. We can thus make use of results such as the inclusion-exclusion principle as

$$v(e_i \cup e_j) = v(e_i) + v(e_j) - v(e_i \cap e_j), \tag{4}$$

and the implied union bound

$$v\left(\bigcup_{e \in \mathcal{A}} e\right) \leq \sum_{e \in \mathcal{A}} v(e). \tag{5}$$

Further, conditional probability rules can be used to recombine the achieved primitive event values $v(p)$ for $p \in \mathcal{P}$ into the event values at higher levels of the graph.

IV. Examples

In this section, we present an example of node capture attacks to illustrate the use of the attack model in Section II and the event-based decomposition in Section III. We illustrate an example in which the adversary's goal is solely to gain access into the security architecture of the WSN.

A. Example Attack: Access to Key Management

Suppose the adversary's goal in the node capture attack is to recover a set $\mathcal{K}$ of cryptographic keys collectively held by nodes in the network. The set $\mathcal{K}$ can represent the entire space of keys assigned to sensor nodes or a particular subset required for a continued attack. Due to the absence of authorities in the WSN, each node n is required to share information about the keys it holds. Suppose first that each node broadcasts a list of identifiers for the keys it holds. We assume that the adversary can recover this information by eavesdropping on the identifier exchange.

In this example, the goal $\mathcal{G} = \{\text{recovery of } \mathcal{K}\}$ can be directly decomposed in to the set of primitive events $\mathcal{P} = \{p_1, \ldots, p_{|\mathcal{K}|}\}$ with $p_i = \{\text{recovery of } i^{th} \text{ key in } \mathcal{K}\}$. Here, the $|\mathcal{K}|$ primitive events are independent, and, assuming the keys are equally valuable to the adversary, each primitive events has value $v(p_i) = 1/|\mathcal{K}|$. Letting $\mathcal{K}_n$ denote the subset of keys in $\mathcal{K}$ held by node n, the initial value $v_\varnothing(n)$ of each node n is equal to the fraction

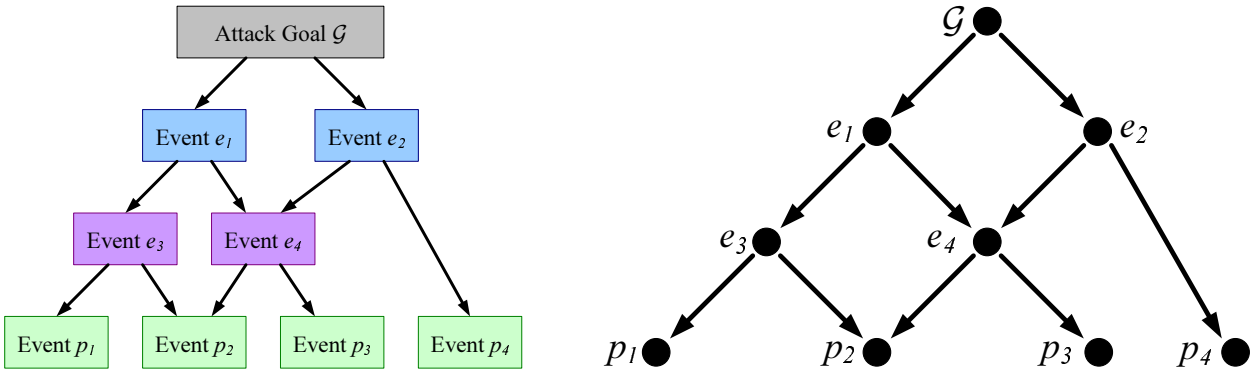
$$v_\varnothing(n) = \frac{|\mathcal{K}_n|}{|\mathcal{K}|}. \tag{6}$$

After a set $\mathcal{C}$ of nodes have been captured, the value $v_\mathcal{C}(n)$ of each node is given by

$$v_\mathcal{C}(n) = \frac{|\mathcal{K}_n \setminus \bigcup_{c \in \mathcal{C}} \mathcal{K}_c|}{|\mathcal{K}|}. \tag{7}$$

V. Future Work: Attack Defense

In order to defend against the debilitating effects of node capture attacks, it is necessary to understand these attacks at a primitive level. The ability to decompose attack goals into primitive events depends on the abilities of network researchers to discover the necessary primitives. As these primitive attack events are identified, they can be evaluated relative to other primitive events and worked into the attack model from the network perspective, allowing the network designer to focus defense resources on the primitive events that have the highest impact on the operation of the sensor network. As the identification of attack primitives is largely an open problem in sensor networks, and in wireless networks in general, this is an area of interest in future research.



(a) (b)

Fig. 2. In (a), an example event-based decomposition of the attack goal $\mathcal{G}$ is given. In (b), the corresponding directed graph $\mathcal{D}(\mathcal{G})$ is illustrated.

VI. Conclusion

In this work, we presented a model for node capture attacks in wireless sensor networks, incorporating attack evaluation metrics into a framework for optimal attacks. We discussed the ability to decompose the goal of a node capture attack into primitive attack events and use the event-based decomposition to evaluate the impact of an attack with respect to the attack primitives. We illustrated the use of the event-based decomposition with an example node capture attack. Finally, we highlighted the potential for future research in the identification of primitive attack events with respect to node capture attacks in sensor networks.

References

[1] I. F. Akyildiz, W. Su, Y. Sankarasubramaniam, and E. Cayirci, "A survey on sensor networks," *IEEE Communications Magazine*, vol. 40, no. 8, pp. 102–114, Aug. 2002.

[2] A. Perrig, R. Szewczyk, J. D. Tygar, V. Wen, and D. Culler, "SPINS: Security protocols for sensor networks," in *Proc. 7th Annual International Conference on Mobile Computing and Networking (MobiCom'01)*, Rome, Italy, Jul. 2001, pp. 189–199.

[3] S. Pai, S. Bermudez, S. B. Wicker, M. Meingast, T. Roosta, S. Sastry, and D. K. Mulligan, "Transactional confidentiality in sensor networks," *IEEE Security & Privacy*, vol. 6, no. 4, pp. 28–35, Jul./Aug. 2008.

[4] L. Eschenauer and V. D. Gligor, "A key-management scheme for distributed sensor networks," in *Proc. 9th ACM Conference on Computer and Communications Security (CCS'02)*, Washington, DC, USA, Nov. 2002, pp. 41–47.

[5] P. Tague and R. Poovendran, "Modeling adaptive node capture attacks in multi-hop wireless networks," *Ad Hoc Networks*, vol. 5, no. 6, pp. 801–814, Aug. 2007.

[6] B. Parno, A. Perrig, and V. Gligor, "Distributed detection of node replication attacks in sensor networks," in *Proc. 2005 IEEE Symposium on Security and Privacy*, Oakland, CA, USA, May 2005, pp. 49–63.