Visualization Assisted Detection of Sybil Attacks in Wireless Networks

Weichao Wang*
ITTC and Department of EECS
University of Kansas
weichaow@ittc.ku.edu

Aidong Lu
CS Department
University of North Carolina at Charlotte
alu1@uncc.edu

ABSTRACT

In wireless networks, the authenticity and uniqueness of node identities are essential to the fundamental operations such as routing, resource allocation, and intrusion detection. In this paper, we investigate Sybil attack, an attack in which a malicious node illegitimately acquires multiple identities and performs as these nodes simultaneously. We propose an effective approach to monitoring and detecting such attacks by integrating network security and visualization methods. The security component explores the time-varying network topology and its statistical and geometry information to detect the existence of Sybil attacks. The visualization component incorporates the detection results and provides an effective mechanism to illustrate abnormal topology patterns and locate fake identities. These two components are integrated into a practical system that takes advantage of both interactive visualization and intelligent security methods. Experimental studies are conducted to investigate the impacts of the network parameters such as node connectivity on the detection capability of the proposed mechanism.

Categories and Subject Descriptors

C.2.0 [**Computer-Communication Networks**]: General—*Security and protection*; H.5.2 [**Information Systems**]: Information Interfaces and Presentation—*User interfaces*

General Terms

Algorithms, Security

Keywords

Interactive Detection, Sybil Attacks, Visualization on Network Security, Wireless Networks, Topology Visualization

*Contact: weichaow@ittc.ku.edu

VizSEC'06, November 3, 2006, Alexandria, Virginia, USA.

1. INTRODUCTION

As wireless networks are widely adopted in various environments and applications, security has become one of their top priorities. Among many features that need to be protected, the uniqueness and authenticity of the node identity must be enforced to enable the fundamental operations such as routing, resource allocation, and misbehavior detection. For example, Intrusion Detection System (IDS) in wireless networks [36] detects the attacks and isolates the malicious nodes by matching the patterns of the known intrusions [3] or discovering the anomalies [6, 14, 20] in the network activities. If the attackers can easily generate a fake identity and participate into the network operations, the effectiveness of IDS will be drastically weakened. Therefore, mechanisms must be properly designed to prevent or detect such attacks.

In this paper, we focus on the *Sybil attack* [11], an attack that specifically targets the node identities in wireless networks. In a Sybil attack, a single malicious node plays the roles of multiple legitimate members of the network by impersonating their identities or claiming fake IDs. If there are a group of collusive attackers, each of them can pretend to be the whole group simultaneously at different places in the network, thus manipulating the results of localized voting or data aggregation [30]. Sybil attacks will also enable the malicious nodes to take over the control of the whole network by compromising a limited number of physical devices, and defeat the replication mechanisms in distributed systems [11].

Noticing the serious harm that Sybil attacks can make, researchers have proposed several approaches to defend against such attacks and encouraging results have been collected [5, 10, 11, 30]. Existing approaches usually detect Sybil attacks by verifying whether a pair of nodes has distinct resources, distinct knowledge, or distinct positions. Since the verifications are conducted in a localized manner, the approaches are especially effective in environments with a relatively stable topology such as sensor networks or when the nodes move slowly. However, under the scenarios that a fake identity dynamically switches among multiple collusive attackers, global network topology must be monitored. Since the ever-increasing network size and lengthened network lifetime will drastically increase the amount of information, we need more powerful techniques such as scientific visualization to assist the representation of the data and the discovery of the hidden connections among them.

In this paper, we present an approach to detecting Sybil attacks in wireless networks that integrates security and visualization methods. The mechanism monitors the neigh-

bor relations among wireless nodes and the network topology changes, and identifies the suspicious Sybil nodes by visualizing the anomalies introduced by fake identities. A comprehensive visualization interface is designed to provide a global view of the network topology evolution so that the attackers can be located even when they dynamically switch among multiple compromised physical devices.

The contributions of the proposed approach include: (1) Since the proposed approach is based on the evolution of global network topology, it provides an effective method to detect the Sybil attacks that cannot be identified in a localized manner. (2) The integration of the visualization techniques and security methods provides an intuitive and scalable vehicle to information representation and attack detection. (3) Since the proposed approach detects Sybil nodes solely based on neighbor relationships among wireless nodes, it can be applied to more dynamic environments such as mobile ad hoc networks.

As we will demonstrate, the proposed mechanism can effectively identify the Sybil nodes. Moreover, since many attacks on wireless networks will lead to anomalies in the network topology, the proposed analysis methods and visualization tools can be adapted to detect other attacks.

The remainder of the paper is organized as follows: Section 2 provides the background of wireless networking, how Sybil attacks are conducted, and the research challenges. Section 3 reviews the previous research efforts that contribute to our approach. Section 4 provides an overview of the approach. In section 5, we describe the design and development of the visualization tools. The details of the mechanism to detect Sybil attacks are presented in section 6. Section 7 presents the experimental results. Finally, section 8 concludes the paper and discusses future extensions.

2. BACKGROUND

While the Sybil attacks can be summarized as a malicious device presenting multiple identities to the network, a more detailed classification of the attacks will help improve our understanding and illustrate the detection capability of the proposed approach. Here we borrow the taxonomy defined in [30] and describe two dimensions of the attacks.

The first dimension focuses on the connections among the Sybil nodes and the legitimate nodes. If they can communicate directly, this is a direct Sybil attack. On the contrary, in an indirect Sybil attack, a malicious device will claim to have the paths to reach the Sybil nodes and all messages have to go through it. In the second dimension, the attacks are divided into two groups based on whether multiple fake identities participate into the network activities simultaneously. For example, in a non-simultaneous Sybil attack, only after a fake identity "leaves" the network, the next one will "join".

While it is more difficult to link together multiple fake identities that appear in different periods of network lifetime and detect non-simultaneous Sybil attacks, their impacts on network security are also limited. For example, a Sybil node that is not a member of the network cannot cast a vote during the leader election procedure. Therefore, in this paper, we focus on the simultaneous Sybil attacks. To evaluate the proposed mechanism in a more realistic environment, we assume that both direct and indirect Sybil attacks exist in the network and a malicious node can dynamically switch between the two kinds. We also assume that multiple malicious physical devices co-exist in the network and a Sybil node can switch among them. This assumption leads to the need to monitor the global network topology. Considering the large amount of information to be processed that is caused by the ever-increasing network size and lengthened network lifetime, we adopt visualization techniques to assist the representation of the data and the discovery of the implicit connections among them.

3. RELATED WORK

3.1 Sybil Attack Detection in Networks

Sybil attack is one particularly harmful attack on distributed systems and wireless networks [11]. This attack has been demonstrated to be detrimental to many important network functions. For example, Sybil attack is discussed in an architecture for secure resource peering in an Internet-scale computing infrastructure [16]. The problem of Sybil attack has been formalized [10] to show that there is no symmetric sybilproof reputation function and a collection of flow-based asymmetric reputation functions can be given under some conditions. Newsome et al. have systematically classified these attacks into several types and analyzed their threats to wireless sensor networks [30].

Based on the detection mechanisms, we roughly divide previous approaches into two categories: identity-based or location-based methods. Identity-based approaches usually mitigate the Sybil attacks by limiting the generation of valid node information, such as the pre-distributed secret keys [30]. To secure routing for structured peer-to-peer overlay networks, Sybil attacks are decreased through charging expensive fees to each newly created nodeID or binding nodeIDs to real-world identities [8]. In evaluating the admission control framework, a special purpose certificate asserting group membership is issued to each peer upon admission [33]. Another detecting approach is proposed for vehicular ad hoc networks through possible explanations for collected data of each node [18]. The method of radio resource testing [30] is based on the assumption that a node cannot send out two signals at different frequencies at the same time.

Location-based approaches utilize the fact that each node can only be at one position at a specific moment. Localization algorithms, such as SeRLoc [24], are proposed to allow sensors to passively determine their locations under known attacks including Sybil attack. The geometric properties of message transmission delay are also explored to reduce the impacts of Sybil attacks [5]. This technique is based on several assumptions which may limit the attack complexities. Our approach utilizes the network topology information. With the integration of visualization and security techniques, our approach can be used to detect Sybil attacks under more sophisticated scenarios.

3.2 Visualization for Network Security

With the fast development of computer security mechanisms, the scale and complexity of the security data put ever-increasing challenges to the representation and understanding of the information. Visualization techniques have been adopted by the researchers to bridge the gap. For example, researchers have designed mechanisms that can provide an overview of the traffic patterns of thousands of hosts [4]. Mechanisms have been developed to provide a scalable rep-

resentation of the intrusion alarms in multiple class-B IP address ranges [2, 22, 23]. Researchers have also developed a visualization methodology to characterize the most common and versatile intrusions, network scans, based on their patterns and wavelet scalograms [29]. Another approach uses IP address and port number histographs to detect and analyze the scan attacks [32]. VisFlowConnect-IP [35] allows anomalous traffic detection through a link-based network flow visualization tool.

Under many conditions, the security data acquired from multiple methods must be investigated jointly to improve the detection accuracy and efficiency. The research efforts in [15] provide a visual correlation between the host processes and network traffic. In both [29] and [32], the approaches can identify the similarity among different scan attacks or NetFlow signatures.

While many visualization approaches to network security require a large amount of high-dimensional data, several mechanisms focus on the big picture. For example, the mechanism in [28] takes very coarsely detailed data to help uncover interesting security events. The mechanism in [31] overcomes the scalability issues inherent in visualizing massive networks through sampling. In [19], low level textual data are provided in the context of a high-level, aggregated graphical display. Disparate logs are also visualized to show the correlation of network alerts based on what, when, and where [26].

4. SYSTEM OVERVIEW

4.1 Network Assumptions

We assume that the links among wireless nodes in the network are bidirectional and two neighboring nodes can always send packets to each other. This assumption will hold under most conditions when the power of the nodes has not been exhausted. We assume that two nodes are neighbors when the distance between them is shorter than r, where r is defined as the radio range.

We assume that a special node exists in the network, which is called the "controller". It will integrate, process, and visualize the topology information that is collected by the wireless nodes and detect the Sybils. We assume that the controller has the storage and computation resources that are required by the proposed mechanism. For example, if the network topology is viewed as a graph, the controller can locate the cut points of it in a short period of time. In our experimental studies, we use a PC with 1.8GHz CPU as the controller and it can process the information of a network containing several hundreds of nodes in real time.

For the wireless networks that have an infrastructure, the controller can be chosen from the special nodes. For example, in a multi-hop cellular network [27], a base station can play the role of controller. In those pure ad hoc networks, leader election mechanisms [34] can be adopted to determine the controller based on the trustworthiness of the mobile nodes and the available resources.

For each wireless node in the network, we assume that it has established a pair-wise key with the controller. This task can be accomplished during the network initiation procedure or based on some pre-distributed information [9, 12, 13, 25]. We also assume that every node moves independently in the network [21]. The impacts of group movement on the proposed approach will be investigated in future extensions.

4.2 System Overview

To monitor and detect Sybil attacks, we have integrated security and visualization methods into a convenient detection tool. Our system concentrates on visualizing and observing significant patterns from time-dependent network topology information. Intelligent algorithms are embedded in the system to identify potential abnormal events and provide additional validation methods.

As shown in Figure 1, the information of network topology is first collected (section 5.1), processed, and visualized at the controller (section 5.2). To simplify user interaction in the visualization process, we use statistical topology information to identify a suspicious node list (section 6.1.1). The users can adjust this suspicious node list with their expertise and change the visualization easily to better reveal the event correlations (section 5.2). To provide additional validation of the user decisions, two algorithm components are designed for the detection of direct Sybil attacks (section 6.1.2) and indirect Sybil attacks (section 6.1.3) respectively. Details of the system design will be discussed in section 6.2.

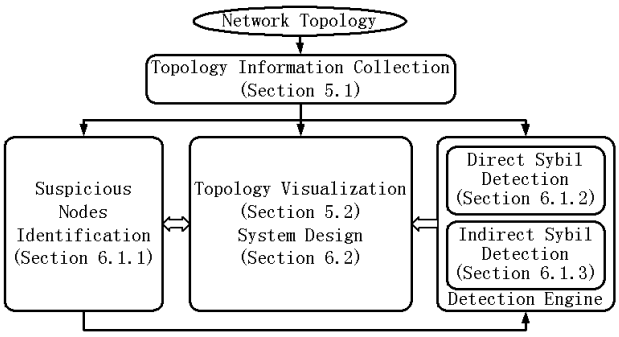


Figure 1: The system architecture.

5. TOPOLOGY VISUALIZATION

5.1 Collection of Topology Information

Since the proposed mechanism detects the Sybil attacks by monitoring the changes and anomalies in the network topology, in this section, we describe how the information can be correctly collected and integrated by the controller.

We know that the neighbor relations among wireless nodes may change because of various reasons such as node movement, device malfunction, battery exhaustion, and unreliable transmission medium. Therefore, a node must be able to detect its active neighbors dynamically. A widely adopted approach is to let the mobile node periodically broadcast a short message containing its identity (called 'beacon' packet) and the neighbors receiving this packet will add the node into their neighbor lists. In the proposed approach, every node will periodically send its neighbor list to the controller. To prevent the list from being altered during transmission, it is protected by the pair-wise key between the controller and the node.

While a legitimate node will faithfully report its neighbors, an attacker will manipulate the list to avoid being detected. For example, a malicious physical device may claim to have a route to an indirect Sybil node so that more traffic will be attracted to it. However, it will not report the Sybil node as its neighbor to the controller. To prevent the

manipulation and its impacts on Sybil detection, we require every neighbor list that is transmitted to the controller to be authenticated by the Message Authentication Code (MAC) of the nodes in the list. The nodes that are not in the list will not be adopted by the neighbors in routing or other network activities. Therefore, a Sybil node cannot be hidden from the controller.

Because of the topology changes, the routes from the mobile nodes to the controller need to be updated. In the proposed mechanism, the controller will periodically broadcast a route discovery packet to the nodes within the radio range and mark the path length to itself as 0. The nodes receiving the packet will increase the path length by one and re-broadcast it. With every node remembers the previous hop, increases the path length by one, and re-broadcasts the packet, the routes to the controller will be established. The frequency to broadcast the route discovery packets can be determined by the radio range and the node movement patterns [17].

Using the received neighbor lists, the controller can regenerate the network topology. For example, a matrix representing the connectivity and shortest paths among the wireless nodes can be calculated. When multiple neighbor matrices are sorted by their sampling time, the changes in network topology can be illustrated as a volumetric data. Since the amount of information will increase fast with the size of the network and the number of topology snapshots, we adopt visualization techniques to represent the data and assist the Sybil detection.

5.2 Visualization of Network Topology

In a wireless network, the node mobility has created severe challenges to defending against malicious attacks. Since the network topology is common information in many networking applications, we concentrate on building a visualization tool for this data that can be easily extended to detect multiple attacks.

The network topology data often contains enough information to monitor and detect intrusions. However, it is very difficult to visualize this information in a manner that can be easily understood by users. Since the topology information acquired from multiple time steps composes a regular volumetric data, we first try to use general 2D and 3D visualization approaches to look at this data, including various 2D cut views, statistical views, and 3D direct volume rendering techniques. As shown in Figure 2 (a)(b) and Figure 3 (a), it is difficult to obtain much useful information directly from these visualizations.

An interesting feature we find is that when we sort the node sequence according to certain criteria, we may see some obvious patterns in both 2D and 3D visualizations. For example, the 2D statistical view shows a grid pattern in Figure 2 (c). Therefore, our basic idea is to develop an approach to reveal the significant patterns in the topology information by grouping the nodes based on the similarity among their topology features, thus enabling the detection of malicious attacks.

We have explored the typical patterns under Sybil attacks, including both direct and indirect simultaneous attacks. Figure 2 (c) and Figure 3 (b) show a grid structure from the indirect attacks which is caused by the lack of direct communication among the fake identities and legitimate nodes. The bright square in the bottom left corner in Figure 4 (c)(d) indicates direct Sybil attacks. On the contrary, there does not exist any obvious pattern under normal network operations.

Scalability is a practical issue for the topology visualization, since the rendering resolution is limited by the screen size and human perception capability. To preserve the significant features in the topology information across different scales, we need to enlarge the range-of-interest in the previous patterns, such as the highlighted regions in Figure 5 (a). One simple solution is to assign the range-of-interest according to the total neighbor numbers of each node in the specified time period. As shown in Figure 5, this scaling method preserves more significant information than general zoom out function.

It will be time consuming if a user is asked to manually adjust the node sequence. Therefore, we integrate an automatic computation process to assist in determining the suspicious nodes and to accelerate the Sybil attack detection. This computation mechanism will be discussed in the next section and the details about the interaction will be described in the system design section.

6. SYBIL ATTACKS DETECTION

6.1 Detection Algorithms

6.1.1 *Determining the Suspicious Node List*

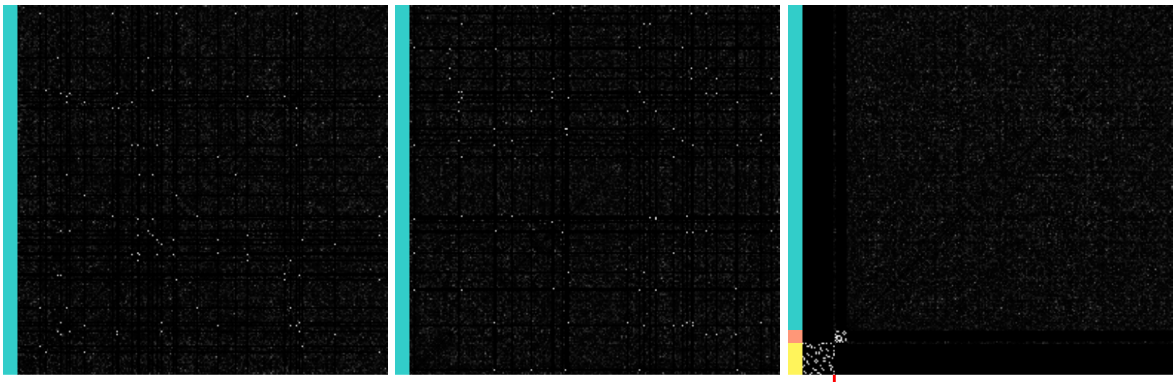
As we will demonstrate in the later parts, the detection operations of the Sybil attacks will focus on the neighbor relations among the wireless nodes. Although the operations on every single node are not computationally intensive (e.g. determine whether removing a node will disconnect the network), when the network consists of hundreds or even thousands of nodes, the controller will be overwhelmed by the processing overhead. Therefore, an efficient mechanism must be designed to filter out a suspicious node list.

Studying the scenarios of direct and indirect Sybil attacks, we find that the connectivity among the Sybil nodes attaching to the same physical device presents a "locality". For example, for two fake identities under direct Sybil attacks, although they can pretend not to be neighbors, a two-hop path must exist through a legitimate neighbor. Similarly, for indirect Sybil attacks, the path length among the fake identities is solely determined by the number of Sybil nodes and their claimed organizations. On the contrary, for the legitimate nodes, since we assume that they move independently, the average path length depends on the network scale and node density.

Therefore, we can efficiently calculate the distribution of the path length between every pair of nodes based on the collected topology information and identify the group of suspicious nodes. Figure 6 illustrates the cumulative distribution function (CDF) of the path length between a pair of legitimate nodes and Sybil nodes. The identities under direct Sybil attack pretend not to be neighbors, and a two-hop path exists between them. The malicious device claims that a three-hop path exists between the two identities under indirect Sybil attack. The anomalies can be easily identified.

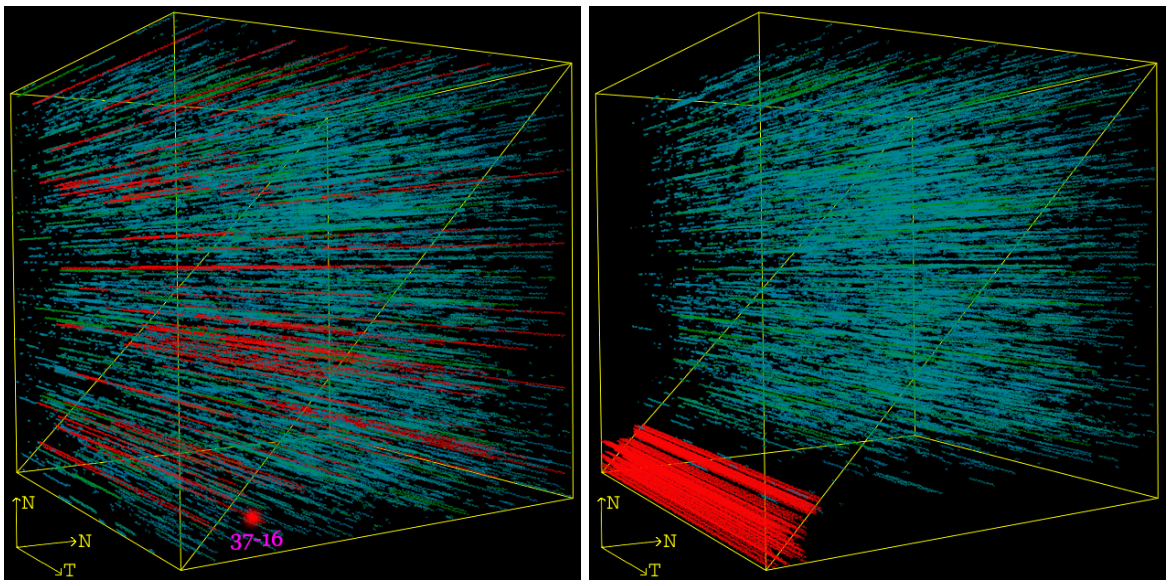
6.1.2 *Locating Direct Sybil Node Pairs*

After the group of suspicious nodes are determined, more computationally expensive detection operations can be conducted. In this part we introduce the detection of direct



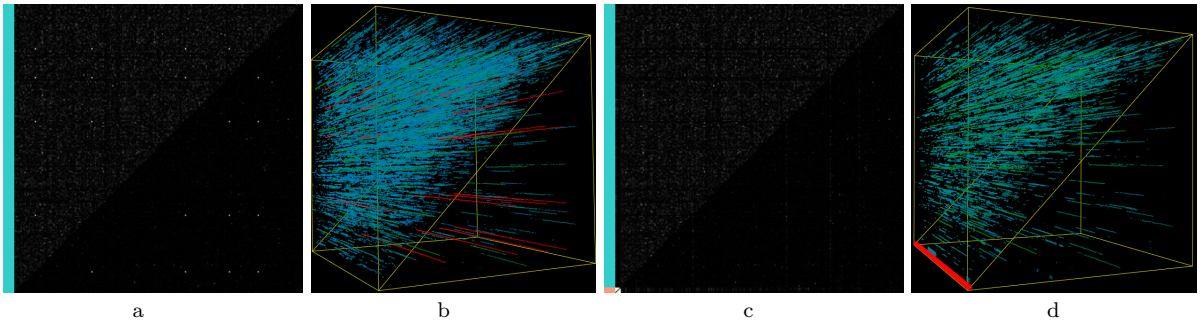
a b c

Figure 2: Three statistical neighbor relation views of the nodes, with bright colors indicating high connection values. (a) It is very difficult for users to directly obtain useful information from the topology, although it contains enough information to detect anomalies. (b) Changing the node sequence does not necessarily show additional information. (c) Only when arranged appropriately, the topology information can demonstrate significant features for users to detect intrusions. The left colormap beside each 2D view shows the node group information and the bottom red point suggests the main role in the attack.



a b

Figure 3: The 3D visualization of the original node sequence (a) and sorted sequence according to a suspicious node list (b). The three axes are two node sequences and one time series. A blue to red colormap is used in the rendering according to the statistical neighbor relationships between two nodes, with red indicating high connections. A label function is used to inquire the node sequence numbers.



a b c d

Figure 4: (a)(b) 2D and 3D views from the original node sequence. (c)(d) An obvious pattern, located at the left bottom corner, is revealed and may indicate a direct Sybil attack. Combined statistical neighbor relationship (left top) and similarity information (right bottom) are used in both 2D and 3D views.

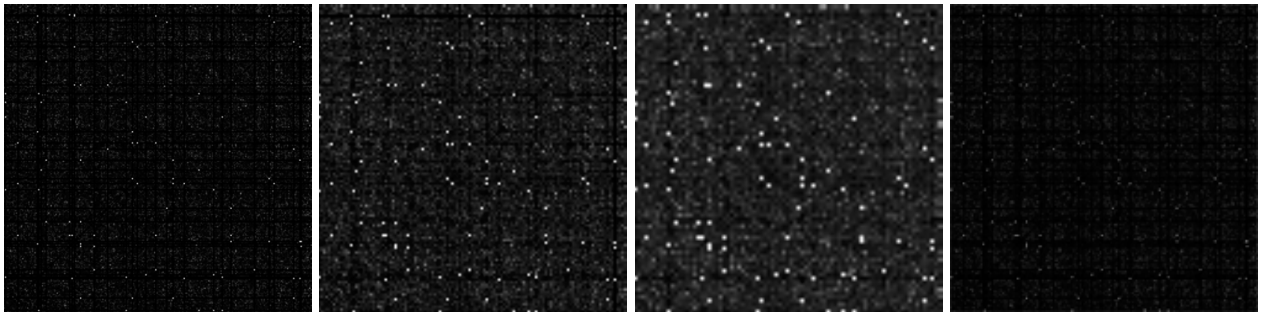


Figure 5: A 2D pattern is continuously zoomed out from (a) to (c), noticing that (c) preserves more significant patterns than general zoom out result (d).

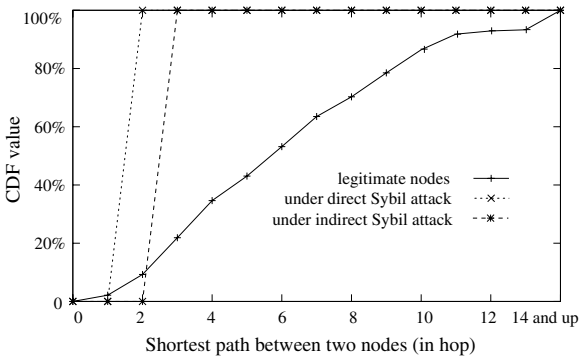


Figure 6: Abnormal patterns of Sybil attacks: distribution of the shortest path length between node pairs.

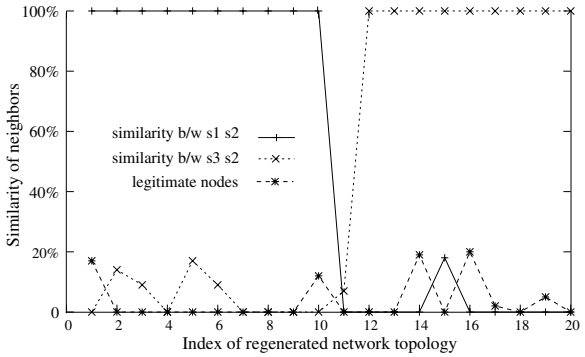


Figure 7: Abnormal patterns of direct Sybil attacks: similarity of neighbor relations.

Sybil nodes based on the similarity of their neighbor relations. The detection of indirect Sybil attacks will be discussed in section 6.1.3.

Let us consider a pair of fake identities under direct Sybil attacks. Since their neighbor discovery packets are transmitted by the same physical device, the group of legitimate nodes that can receive the packets and list them as neighbors are almost the same. On the contrary, a pair of legitimate nodes will not have this similarity when each of them moves independently.

Based on the collected network topology information, we can calculate the similarity of the neighbor relations between two suspicious nodes s_1 and s_2. If the neighbor set of a node s_1 in a regenerated network topology G is represented as $N_{s_1}^G$, a normalized value to describe the similarity between the neighbor sets of two nodes can be calculated as:

$$\frac{N_{s_1}^G \cap N_{s_2}^G}{N_{s_1}^G \cup N_{s_2}^G} \quad (1)$$

The controller can estimate the similarity of the neighbor relations between two nodes in each regenerated network topology. Since a Sybil node can dynamically switch its attached physical device, different pairs of nodes may demonstrate the similarity in different periods of the network lifetime. For example, Figure 7 illustrates the scenario when s_2 switches from node s_1 to s_3 at round 11. As a comparison, we also present the similarity value between a pair of legitimate nodes.

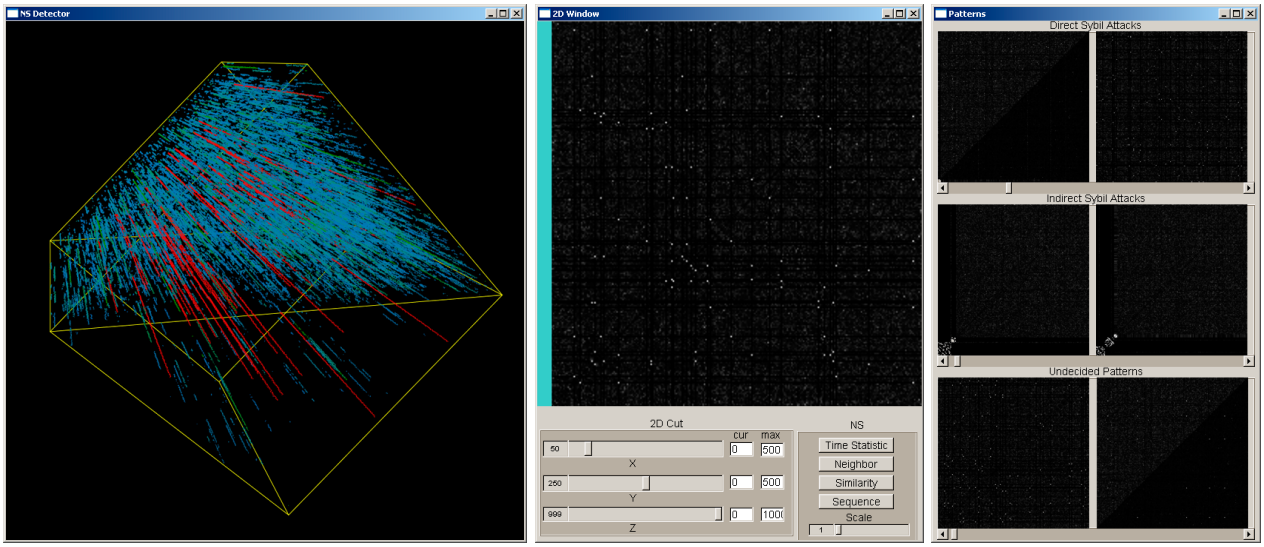
6.1.3 *Locating Anchor Nodes for Indirect Sybils*

In an indirect Sybil attack, only through the malicious node that claims to have paths to the fake identities can other legitimate members reach them. Therefore, if the network topology is viewed as a graph, this "anchor" node is a cut point in it: removing this node and the associated links will disconnect the graph.

Based on this observation, we examine the frequency that a node is a cut point in the topology. Since multiple fake identities can attach to the same physical device in an indirect Sybil attack, a special method must be adopted to mitigate the impacts on detection accuracy when the attackers rotate the identities of the "anchor" node. Therefore, we will also count the frequency that a node is isolated from the majority of the network when the anchor node is removed. Figure 8 illustrates an example of indirect Sybil attacks and the accumulated count values for each node. The differences between the four fake identities and the legitimate nodes can be easily identified.

6.2 System Design

To provide a robust Sybil attack monitoring and detection tool, we have developed a detection algorithm and a topology visualization approach as two essential components of our security system. The integration of these two components can benefit each other from multiple aspects. Our basic idea is to use the visualization component to intuitively understand the network topology and security results. This also allows users to adjust and interact with the network information according to their expertise. The security com-



a b c

Figure 9: Our system interface is composed of a 3D view (a), a 2D view (b), and a pattern organization window (c). Both 2D and 3D views are provided to visualize data correlations from multiple aspects. The pattern organization window helps users to visualize similar typed patterns and locate attackers quickly.

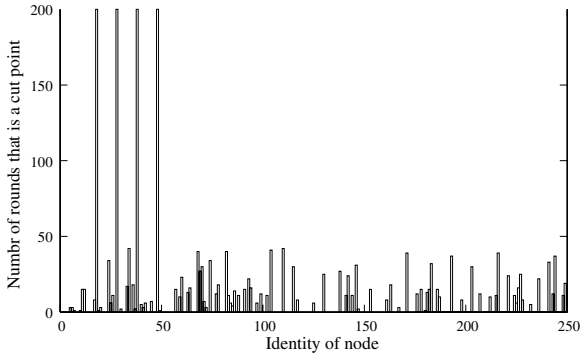


Figure 8: Abnormal patterns of indirect Sybil attacks: frequency to be cut points.

ponent is used as an identification and validation tool to assist users to draw final decisions and reduce their interaction overhead. The following will discuss our interface design and a case study respectively.

As shown in Figure 9, we have arranged three parallel windows for visualizing the event correlations: a 3D view, a 2D view, and a pattern organization window. The 3D view is mainly designed for displaying neighbor relationships within a time sequence. The 2D view alternatively displays cut views from the volumetric topology data (neighbor relationships at a time step or neighbors of a node through the time sequence) and statistical topology matrices (node connections or similarities). We also add a pattern organization window to store typical abnormal patterns and undecided patterns for users to compare with. These three windows allow us to observe the network topology information through multiple aspects and reveal the data correlations.

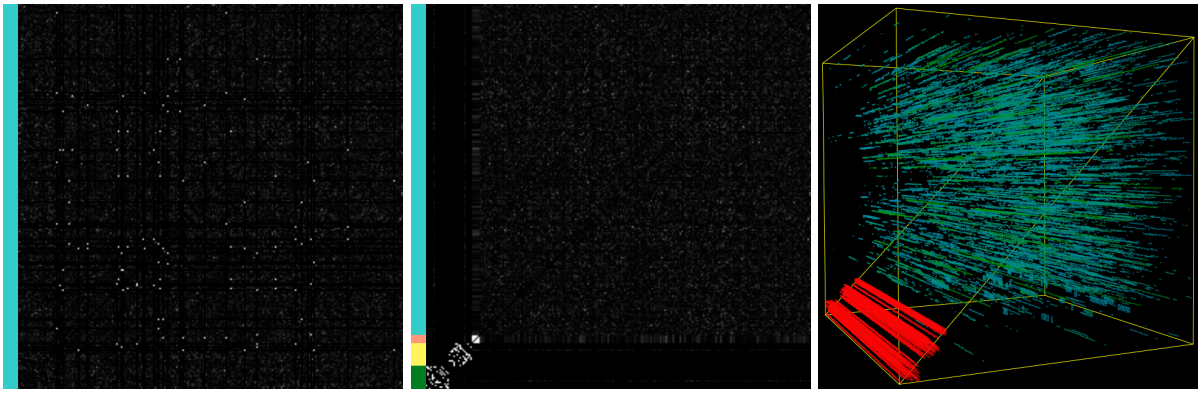
The interaction of node sequence in the 2D view is achieved through picking and dragging into the corresponding node group. The system automatically adjusts the node sequence according to their topology features, such as total neighbor numbers. The interaction of the 3D view is managed from the 2D view, which is used as a transfer function. The users can interactively adjust the 3D view through common 1D transfer functions of the connection values to reveal the focus-of-interest and reduce the overlapping in space. As shown in Figure 3(a), we use an easy pick function by the movement of the mouse to display the node IDs.

We use one case study to demonstrate this interaction process. As shown in Figure 10, the network topology data is first collected and the 2D view shows the statistical neighbor relationship in Figure 10 (a). The security component calculates the suspicious node list and guides users to adjust the 2D view to reveal an abnormal pattern (Figure 10 (b) and (c)). These patterns are compared with typical direct and indirect Sybil attack patterns and suggest that this is a hybrid attack. The suspicious node list is sent to the security validation component to justify node behaviors, which can be run through the data of a longer period. The users can make final decisions based on both visualization and security results. The located fake identities will be removed from the network.

7. EXPERIMENTAL RESULTS

The proposed mechanism is examined through simulation. The experiments are conducted in two phases. In the first phase, we use ns2 [1] to simulate the neighbor discovery procedures and the report of the topology to the controller. In the second phase, the proposed mechanism tries to detect Sybil attacks and locate the fake identities. The mobile nodes are deployed in a square area with the edge length of $1500m$. The radio range r of the nodes is $170m$, and any two nodes that have a distance shorter than r can directly communicate to each other.

Within the simulated area, 300 nodes are randomly and uniformly distributed and the average degree of connectivity is 12.0. We adopt the random trip movement model that is



a b c

Figure 10: A case study: a hybrid direct and indirect Sybil attack is discovered through abnormal patterns when arranged according to the identified suspicious node list. (a) The original statistical neighbor relationship pattern; (b) The adjusted neighbor pattern; (c) The adjusted 3D topology pattern.

proposed in [7] and the highest moving speed of the nodes is $17m/s$. The controller collects the network topology every 10 seconds, which is a rough estimation of the lifetime of a link based on the radio range r and the highest node moving speed. In every simulation, 200 rounds of topology will be collected.

We will investigate both direct and indirect Sybil attacks. The number of compromised physical devices and fake identities will be described in detail in each group of experiments. When a Sybil node switches its attached physical device, it will temporarily leave the network and rejoin later so that the moving speed restriction will not be violated.

Since the proposed mechanism uses the network topology information to detect Sybil attacks, we will investigate the impacts of the parameters such as the degree of connectivity on the detection capability. Every data point in the following figures represents the average value over 15 trials under different network setups.

7.1 Direct Sybil Attacks

The proposed mechanism detects the direct Sybil attacks in two steps: it first determines the suspicious node list based on the distribution of the path length between two nodes, then uses the similarity of the neighbors to locate the Sybil pairs. This mechanism will be impacted by the length of the duration that the neighbor relations are monitored. For example, under the extreme condition, when only one snapshot of the network topology is available, the distribution of the path length between two nodes cannot be derived. Therefore, in this group of experiments, we examine the impacts of the length of the monitored duration on the detection capability.

Since the frequency of network topology changes heavily depends on the radio range r and the movement model of the mobile nodes, we use the ratio between the radio range and the highest moving speed as a time unit to measure the length of the monitored duration. The experiment results on the two steps of the mechanism are presented and discussed respectively.

The cumulative distribution function (CDF) of the path length between two nodes is illustrated in Figure 11. For the pair of fake identities under direct Sybil attacks, we assume that they pretend not to be neighbors. Therefore, their distance is always two hops during the monitored duration. As the comparison, we randomly choose ten pairs of legitimate nodes that are two hops away at the beginning of the monitored duration and the average values are illustrated in the figure. From the figure, we find that when more than 10 time units are monitored, the differences between legitimate nodes and fake identities can be easily identified.

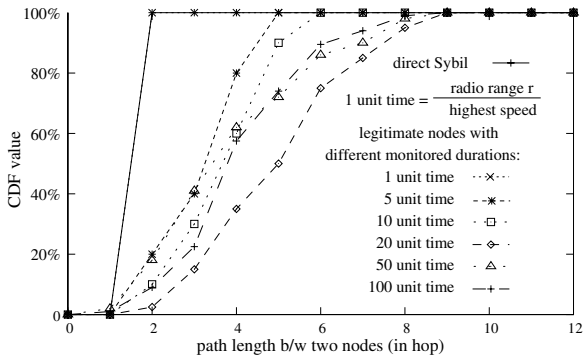


Figure 11: Relationship between the distribution of path length and the monitored duration.

The CDF values of the similarity between the neighbors of two nodes are illustrated in Figure 12. For the pair of Sybil nodes, their neighbors will always be the same. As the comparison, we randomly choose ten pairs of legitimate nodes that are neighbors at the beginning of the monitored duration and the average values are illustrated in the figure. When more than 5 time units are monitored, the differences can be easily identified.

Combining the results in Figure 11 and 12, we find that when the length of monitored duration is longer than 10 time units, the proposed approach can locate the fake identities. In real deployment of the mechanism, the length can be determined based on previous simulation results or field test results. A threshold of the differences in the CDF values can be adopted to determine the list of suspicious nodes.

7.2 Indirect Sybil Attacks

The proposed mechanism detects the indirect Sybil attacks based on the frequency that a node is a "cut point" in

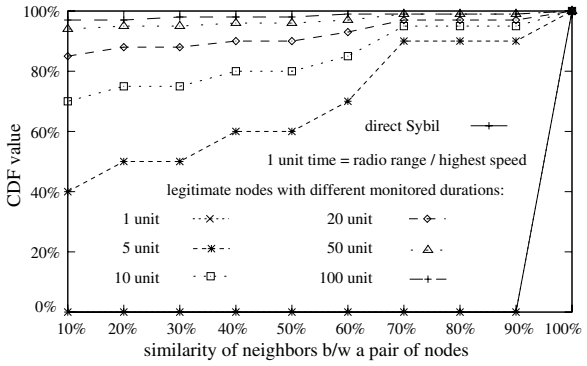


Figure 12: Relationship between similarity of neighbors and the monitored duration.

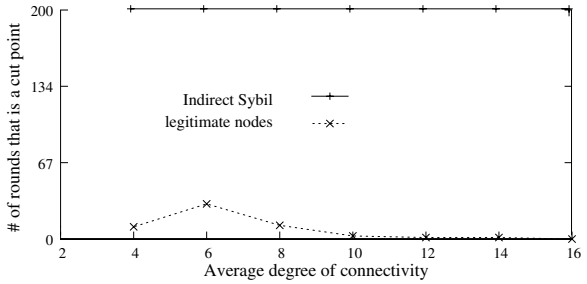


Figure 13: Relationship between the frequency to be a cut point and the degree of connectivity.

the network. This feature, however, is also impacted by several network parameters. In this group of experiments, we investigate the impacts of the average degree of connectivity on the detection capability of the proposed mechanism. We adjust the degree of connectivity by altering the radio range. The simulation results are presented in Figure 13.

For an indirect Sybil node, it will always be a cut point or in the isolated subnetwork when the cut point is removed. On the contrary, the frequency of a legitimate node will change with the degree of connectivity. As shown in Figure 13, when the average number of neighbors is very low, most of the nodes are isolated and they are not the cut points of the graph. As the node density increases, the nodes become connected but few pairs of nodes have disjoint paths. Therefore, more of them become the cut points. When the network becomes denser, most of the nodes have multiple disjoint paths among them and removing a single node will not disconnect the graph. Therefore, the frequency will decrease again.

From the results in Figure 13, we find that the proposed mechanism will have a better detection accuracy when the mobile nodes have a relatively large degree of connectivity. The detection of Sybil attacks in sparse networks remains an open problem and will be investigated in the future work.

8. CONCLUSIONS AND FUTURE WORK

In this paper, we propose an approach for detecting Sybil attack in wireless networks, which is a particular harmful attack for many network functions. Our approach concentrates on visualizing, organizing, and detecting significant abnormal patterns from network topology information. We have designed security methods to locate suspicious nodes and validate their behaviors using topology geometry information. By integrating these intelligent algorithms, a user-friendly visualization method is designed to reveal meaningful event correlations from the network topology. This approach allows users to monitor and detect simultaneous direct and indirect Sybil attacks effectively.

Because of the popularity of network topology information, our approach can be expanded to a common intrusion detection tool for many applications and attack types. In this paper, we integrate security and visualization methods to provide a robust Sybil attack monitor and detection tool. The intelligent detection component significantly accelerates and simplifies the user interaction in the system; while visualization component increases the accuracy and tolerance of the proposed security algorithm. This allows users to detect more complex attack scenarios than single security approaches.

In the future, we will design and perform a systematic user study for the proposed approach and incorporate the evaluation results to further simplify the system interface and user interaction. We will investigate methods to effectively distinguish topology patterns from group movements and malicious attacks. We are also interested in exploring detection methods for non-simultaneous Sybil attacks, which may create more complex scenarios when conducted by multiple attacker groups in a longer time duration. Finally, since topology visualization and organization is a common problem for network security, we will extend the proposed approach to detect more kinds of intrusions for administration and application purposes.

9. ACKNOWLEDGEMENTS

The authors would like to thank the anonymous reviewers for their valuable comments. This research is supported in part by KU New Faculty General Research Fund and Department of Energy under Award DE-FG02-06ER25733.

10. REFERENCES

[1] *Proceedings of IEC Workshop on Internet Simulations with the NS simulator*, 2000.

[2] K. Abdullah, C. Lee, G. Conti, J. Copeland, and J. Stasko. IDS RainStorm: Visualizing IDS Alarms. In *Proc. of VizSEC*, 2005.

[3] F. Anjum, D. Subhadrabandhu, and S. Sarkar. Signature based Intrusion Detection for Wireless Ad-Hoc Networks: A Comparative study of various routing protocols. In *Proc. of VTC*, 2003.

[4] R. Ball, G. Fink, A. Rathi, S. Shah, and C. North. Home-Centric Visualization of Network Traffic for Security Administration. In *Proc. of ACM VizSEC/DMSEC*, 2004.

[5] R. Bazzi and G. Konjevod. On the establishment of distinct identities in overlay networks. In *PODC '05: Proceedings of the twenty-fourth annual ACM SIGACT-SIGOPS symposium on Principles of distributed computing*, pages 312–320, 2005.

[6] V. Bhuse and A. Gupta. Anomaly intrusion detection in wireless sensor networks. *Journal of High Speed Networks*, 1(15), 2006.

[7] J.-Y. Le Boudec and M. Vojnovic. Perfect Simulation and Stationarity of a Class of Mobility Models. In *Proc. of IEEE Infocom*, 2005.

[8] M. Castro, P. Druschel, A. Ganesh, A. Rowstron, and D. S. Wallach. Secure routing for structured peer-to-peer overlay networks. In *OSDI '02: Proceedings of the 5th symposium on Operating systems design and implementation*, pages 299–314, 2002.

[9] H. Chan, A. Perrig, and D. Song. Random key predistribution schemes for sensor networks. In *Proc. of IEEE Symposium on Security and Privacy*, 2003.

[10] A. Cheng and E. Friedman. Sybilproof reputation mechanisms. In *P2PECON '05: Proceeding of the 2005 ACM SIGCOMM workshop on Economics of peer-to-peer systems*, pages 128–132, 2005.

[11] J. Douceur. The Sybil Attack. In *IPTPS '01: Revised Papers from the First International Workshop on Peer-to-Peer Systems*, pages 251–260, 2002.

[12] W. Du, J. Deng, Y. S. Han, and P. K. Varshney. A pairwise key pre-distribution scheme for wireless sensor networks. In *Proc. of ACM CCS*, pages 42–51, 2003.

[13] L. Eschenauer and V. D. Gligor. A key-management scheme for distributed sensor networks. In *Proceedings of ACM CCS*, pages 41–47, 2002.

[14] W. Fan, M. Miller, S. Stolfo, W. Lee, and P. Chan. Using Artificial Anomalies to Detect Unknown and Known Network Intrusions. *Knowledge and Information Systems*, 6(5), 2004.

[15] G. Fink, P. Muessig, and C. North. Visual Correlation of Host Processes and Traffic. In *Proc. of VizSEC*, 2005.

[16] Y. Fu, J. Chase, B. Chun, S. Schwab, and A. Vahdat. SHARP: an architecture for secure resource peering. In *SOSP '03: Proceedings of the nineteenth ACM symposium on Operating systems principles*, pages 133–148, 2003.

[17] M. Gerharz, C. de Waal, M. Frank, and P. Martini. Link Stability in Mobile Wireless Ad Hoc Networks. In *Proceedings of the IEEE Conference on Local Computer Networks (LCN)*, pages 30–39, 2002.

[18] P. Golle, D. Greene, and J. Staddon. Detecting and correcting malicious data in VANETs. In *VANET '04: Proceedings of the 1st ACM international workshop on Vehicular ad hoc networks*, pages 29–37, 2004.

[19] J. Goodall, P. Rheingans, W. Lutters, and A. Komlodi. Preserving the Big Picture: Visual Network Traffic Analysis with TNV. In *Proc. of VizSEC*, 2005.

[20] J. Hall, M. Barbeau, and E. Kranakis. Using mobility profiles for anomaly based intrusion detection in mobile networks. In *Proc. of Wireless and Mobile Security Workshop*, 2005.

[21] W. J. Hsu, K. Merchant, H. Shu, C. Hsu, and A. Helmy. Weighted waypoint mobility model and its impact on ad hoc networks. *SIGMOBILE Mob. Comput. Commun. Rev.*, 9(1):59–63, 2005.

[22] A. Komlodi, P. Rheingans, U. Ayachit, J. Goodall, and A. Joshi. A User-Centered Look at Glyph-Based Security Visualization. In *Proc. of VizSEC*, 2005.

[23] K. Lakkaraju, W. Yurcik, A. Lee, R. Bearavolu, Y. Li, and X. Yin. NVisionIP: NetFlow Visualizations of System State for Security Situational Awareness. In *Proceedings of ACM VizSEC/DMSEC*, 2004.

[24] L. Lazos and R. Poovendran. SeRLoc: Robust localization for wireless sensor networks. *ACM Trans. Sen. Netw.*, 1(1):73–100, 2005.

[25] D. Liu, P. Ning, and R. Li. Establishing Pairwise Keys in Distributed Sensor Networks. *ACM Transactions on Information and System Security*, 8(1):41–77, 2005.

[26] Y. Livnat, J. Agutter, S. Moon, R. Erbacher, and S. Foresti. A Visualization Paradigm for Network Intrusion Detection. In *Proceedings of the IEEE Information Asssurance Workshop*, pages 92-99, 2005.

[27] B. S. Manoj, R. Ananthapadmanabha, and C. Siva Ram Murthy. Multi-hop cellular networks: architecture and protocols for best-effort and real-time communication. *J. Parallel Distrib. Comput.*, 65(6):767–791, 2005.

[28] J. McPherson, K. Ma, P. Krystosk, T. Bartoletti, and M. Christensen. PortVis: A Tool for Port-Based Detection of Security Events. In *Proc. of ACM VizSEC/DMSEC*, 2004.

[29] C. Muelder, K. Ma, and T. Bartoletti. A Visualization Methodology for Characterization of Network Scans. In *Proc. of VizSEC*, 2005.

[30] J. Newsome, E. Shi, D. Song, and A. Perrig. The sybil attack in sensor networks: analysis & defenses. In *IPSN '04: Proceedings of the third international symposium on Information processing in sensor networks*, pages 259–268, 2004.

[31] D. Rafiei and S. Curial. Effectively Visualizing Large Networks Through Sampling. In *Proc. of IEEE Visualization*, 2005.

[32] P. Ren, Y. Gao, Z. Li, Y. Chen, and B. Watson. IDGraphs: Intrusion Detection and Analysis Using Histographs. In *Proc. of VizSEC*, 2005.

[33] N. Saxena, G. Tsudik, and J. Yi. Admission control in Peer-to-Peer: design and performance evaluation. In *SASN '03: Proceedings of the 1st ACM workshop on Security of ad hoc and sensor networks*, pages 104–113, 2003.

[34] S. Vasudevan, B. DeCleene, N. Immerman, J. Kurose, and D. Towsley. Secure Leader Election Algorithms for Wireless Ad Hoc Networks. In *Proc. of IEEE DARPA Information Survivability Conference and Exposition (DISCEX)*, 2003.

[35] W. Yurcik. VisFlowConnect-IP: A Link-Based Visualization of NetFlows for Security Monitoring. In *18th Annual FIRST Conference on Computer Security Incident Handling*, 2006.

[36] Y. Zhang and W. Lee. Intrusion detection in wireless ad-hoc networks. In *Proc. of ACM MobiCom*, pages 275–283, 2000.