Detecting the Sybil Attack in Wireless Sensor Network

A. V. PRAMOD
JJIIT, Maheshwaram

Md. Abdul Azeem
MVSR Engg college

M. OM PRAKASH
JJIIT, Maheshwaram

Abstract

Mobility is frequently a problem for providing security services in ad hoc networks. In this paper, we render that mobility can alsobe used to enhance security. Specifically, we render that nodes which are in passively monitor traffic in the network can able to detect a Sybil attacker which uses a number of network identities simultaneously. We can do through simulation that this detection can be done by a single node, or multiple trusted nodes can join to improve the accuracy of detection. We then show that although the detection mechanism will falsely identify groups of nodes traveling together as a Sybil attacker, we can extend the protocol to monitor collisions at the MAC level to differentiate between a single attacker spoofing many addresses and a group of nodes traveling in close proximity.

Keywords

WSN,PKI,PASID,PASID-GD

1 Introduction

Wireless Sensor Networks (WSNs) are obtainable into view as a innovative part in wireless and mobile computing research. Sensor networks are predicting innovative thriftily viable solutions to a range of applications Sensor networks are exceedingly distributed networks with small, lightweight wireless nodes and deployed in magnanimous numbers for supervise the environment by the facet of physical parameters such as temperature, pressure, or relative humidity. The sensor nodes are much likewise to that of a computer with components such as processing unit, limited memory, limited computational power source inform of a battery, and sensors. In a classic application, a WSN is garbled in a region where it is signified for collecting data through its sensor nodes. For defending or monitoring critical infrastructures a sensor network applications requires security. Security in sensor networks is refined due to broadcast nature of the wireless communication and be short of tamper resistant hardware. There are abundant protocols exist for imprinting ad hoc networks among cooperative mobile, radio-equipped nodes. Many of this protocol have been secured using reputation schemes [1] that rely on there being a limited number of attackers in the group and that assume each radio exemplifies a different individual. However, the broadcast life of radio allows a single node to make-believe to be many nodes at the same time by using many different addresses while transmitting. This attack is called the Sybil attack [2], can easily defeat repute [3] and threshold [2] protocols intended to protect against it. Douceur has shown that there is no practical defense against the attack even a PKI must ensure that each identity is actually one entity, this requires costly manual disturbance, that curbs the number of identities that can be carried off. In demarcation, protocols for detection do not suffer from such limitations. Moreover, detection is complemental to any method that attempts .In this paper, we exhibit that the mobility of nodes in a wireless network can be used to detect and identify nodes that are region of a Sybil attack. We rely on the fact that while individual nodes are free to move independently, all identities of a single Sybil attacker are bound to a single physical node and must move together. We propose two initial methods, which runs on standard, inexpensive equipment without any special transmitting aerial or hardware and with only very loose clock synchronization.

In the first method, called Passive Ad hoc Sybil Identity Detection (PASID), a node can detect Sybil attacks by reading the identities, of other nodes namely the IP or MAC, which discovers transmitting. Then the node builds a profile of all the nodes which are heard together, by which it helps in exposing Sybil attackers. We demonstrate through simulation that in networks with sufficient connectivity and mobility PASID can bring forth close to 100% accuracy in identifying the various attacker identities mean while avoiding any fake positives. As the network becomes more dense, with more nodes in less space, the fake positive rate increases as it becomes more thin, the accuracy rate declines as each node has fewer chances to hear its neighbors. Which will render multiple trusted nodes can share their observations to increase the accuracy of detection over a shorter time or in a more-sparsely connected network.

On other hand second method, is called PASID with Group Detection (PASID-GD), widens approach and reduces fake positives that can occur when a group of nodes moving together is falsely identified as a single Sybil attacker. This approach is successful by monitoring collisions at the MAC level we can show that can differentiate these cases. Because an attacker operating over a single channel can transmit only serially, whereas independent nodes can transmit in parallel.

2 Related Work

Sybil attack can occur in a broadcast system that engages without a central authority to verify the identities of each communicating entity [2]. Sybil attacker can acquire many different identities by sending messages with different identifiers. an entity in the system can endeavor to influence if some set of entities are distinct by testing their resource limits, but this is tough because each entity is only aware of others through messages over a communication channel. If a single Sybil attacker pretends to be multiple entities, then it may not have the same computational, storage, and bandwidth capabilities as multiple independent entities. A Sybil attacker that has more resources than expected can pose a number of entities proportional to the amount its resources are underestimated. However, testing based on such an assumption requires an accurate model of the attacker's resources Similarly, a set of entities that are more resource-constrained than expected may fail to prove their independence. The testing entity might also attempt to verify identity and independency circuitously by asking entities to guarantee for each other. This strategy is prone to the Sybil attack because multiple entities can be the multiple identities of one or more Sybil attackers.

Newsome, et al [4] proposed several methods for detecting Sybil entities in a sensor network. They present an excellent discussion on threats that Sybil attack poses to sensor networks. In counterpoint to these methods, the detection techniques are proposed active tests which require the engagement of all neighboring nodes by inquiring them to respond to queries on specified channels or to carry pre-distributed keys. Such type of queries/response resource tests are a challenge to undertake in a mobile environment where neighbors legitimately may change with great frequency and without notice. To detect or prevent a Sybil attack is based on a significant assumption that each entity has been assigned exactly one key, which is difficult to ensure in practice in general.

Our methods of detecting Sybil attackers are related to malicious attacks against anonymous routing protocols [5]. Which allows an identity to remain identical from other nodes in the system. An attacker who wishes to influence the identity of an initiator can track the membership of the group over time. Every time it identifies a message, and records the group membership. As when there is a changes in membership owed to nodes joining or leaving the group deliberately or as of network failures, then the intersection of all the recorded memberships meets to only the initiator.

We are trying to work out in this paper an application of the intersection attack applied to geographic location in an ad hoc network. Likewise, a Sybil attacker wishes to hold on her multiple identities are identical from others in the system. However, there are differences between a Sybil attacker and legitimize nodes in a mobile wireless scenario, particularly in self-governing nodes are mobile but the identities of a Sybil node move together.

2.1 Sybil Attacks in Sensor Networks

An advantage of sensor network is that no fixed infrastructure is required a network for routing data can be formed from whatever nodes are available. Nodes forward messages for each other to provide connectivity to nodes. Routing protocols are used to find a path end-toend through the cooperative network [6, 7]. In unguaranteed routing protocols, such as DSR or AODV, the address-based identifiers can be easily faked by malicious nodes, which presents an opportunity for a Sybil attack. Morever, allowing unauthenticated address presents a series of other attacks, including spoofing, route direction and error fabrication [8].Sybil attacks may not be the most significant problem present. Our methods work whether addresses are authenticated or not, though given the wide range of attacks possible against unauthenticated networks.

PKI-based protocols. Much of the initial work in ad hoc network security focuses on secure routing [8]. To counter these type of routing attacks, a multifariousness of protocols have been proposed. Some of which require a central authority for less flexible and some other mechanism to distribute cryptographic material to nodes in the system prior or during deployment.. Allowing nodes to join without pre-distributing keys leaves a potential Sybil attack.

Reputation Schemes. Security mechanisms include protocols for determining and maintaining reputation information about nodes for ad hoc networks. Every node capable of developing trust in the other nodes that it conceives correctly routed. The Sybil attack weakens these protocols because it can use multiple identities to falsely vouch for or otherwise support an identity that would otherwise gain a bad reputation. A reliance on cryptographic certificates or keys does not prevent the Sybil attack in general because one entity may be in possession of multiple keys. For example, if PKI credentials are simply purchased, then the PKI is reduced to a resource test of each identity's wealth, which can be without bound. Unfortunately, implementing a stronger approach is problematic. This is because in practice it is untenable to create a foolproof system that can scale to a significant number of users to check identities for independence before the keys are issued. Deploying a fool proof systems touches on issues including physical security and attacks involving social engineering or physical force. It would require checking a person against some set of unforgeable documents; but even government issued documents are forged regularly.

Threshold-based protocols. These type of protocols are used to avoid the indefensible requirement of a PKI. Where group of trusted nodes distributes cryptographic material only if a subset of that group agrees on the trustworthiness of new members. Sybil attackers can additionally defeat schemes that rely on threshold cryptography because verifying the true number and independence of nodes in the network is difficult. If a Sybil attacker can generate identities to meet the threshold requirements it can effectively control the routing of the network.

3 Detecting the Sybil Attack

The malicious node established by a Sybil attacker (i.e having The multiple identities) differ from those of an honest node in several ways, irrespective of their representation either by their IP addresses, MAC addresses, or public keys . This is because of resources used by single node for simulating multiple identities, any exceptional accepted identity will be resource constrained in computation, storage, or in bandwidth. Douecer has shown that a Sybil attacker cannot be prevented by tests of finite resources [2]. All identities of a Sybil attacker must share the same set of resources, and this sharing can be detected in some scenarios [4]. In the mobile environment, a single entity posing multiple identities has an important constraint that can be detected easily as all identities are part of the same physical device, therefore they must move in unison, while other nodes are absolve to move at their will. When we consider the nodes move geographically, all the Sybil identities may appear or disappear simultaneously as it moves in and out of range. Assuming an attacker uses a single-channel radio, multiple Sybil identities must transmit serially, whereas multiple independent nodes can transmit in parallel.

3.1 Overview

In this scheme, time intervals plays vital which captures behavior from all the Sybil identities of an attacker, individual nodes that wish to detect Sybil attackers have to monitor all transmissions that they receive over many time intervals. These intervals are chosen long enough for capturing behavior of attacker, which includes data transmissions, like HELLO, keep alive messages, routing requests and replies. The node keeps track of the different identities heard during this time interval. After many observations, the node analyzes the data to find identities that appear together often and that appear apart rarely. These identities likely comprise a Sybil attack.

If the Sybil attacker does not transmit using all its different identities within an interval the results will be skewed. Sybil attackers may actively foil detection by changing identities frequently. However, doing so limits the effectiveness of an attack when false identities would best be long-lived, for example to foil a reputation scheme or to defeat threshold cryptography. False positives can be caused by using Sybil identities that belong to other nodes in the network. An attacker can corrupt trust in legitimate nodes in this way. False positives can also occur if a collection of nodes moves together in unison in close proximity, either accidentally or intentionally. For example, a military unit with many members channelizing information, each of whom has a wireless device, will appear as a Sybil attacker based on their physical proximity. That can reduce the rate of false positives in such cases by analyzing the rate of packet collision at the MAC layer. The first is that it should run on any normal node without any unusual hardware, nor does it require any directional antennas or specialized clocks. This protocol requires only that a node be able to receive transmissions. Later, when protocol is protracted to multiple nodes, it also requires that they be able to share data among them by forwarding it, and each node have a similar notion of what time it is to within a few seconds, both of which are part of the normal operation of the network. Because of the hardware simplicity is low, the protocol can be widely deployed . We assume that the Sybil attacker maintains its identities over time rather than disposing of them and creating new ones that is, it is a simultaneous Sybil attacker [4]. This assumption is reasonable for Sybil attackers wishing to foil long-lived protocols, such as those with threshold cryptography or maintain reputation information.

3.2 Detection Protocol

Here, we describe two versions of our first detection protocol i.e a single observer case and a multi-observer case. Which include information from the MAC network layer.

3.2.1 Single Node Observer

The protocol, Passive Ad hoc Sybil Identity Detection (PASID), is strictly passive it does not require any type of active probing of suspected Sybil nodes, though the techniques are complimentary to active methods [4]. Instead, it operates effectively on a single node that records the identities of nodes that it hears broadcasting. Protocol performance will be improved by sharing this data among a set of trusted nodes. A node that wishes to detect the presence of a Sybil attacker in the network starts by recording the identities of all other nodes it hears broadcasting over a series of intervals. A complete record of all data transfered may not needed. The observation period, referred to as the time bucket, is long enough to capture the likely behavior of all normal node which also includes normal data flow, regular HELLO and keep-alive messages and other periodic route requests for nodes that have data to send but have no current route for the destination. The length of this time period depends on the underlying protocol engaged in our simulations 30 seconds was adequate as it far exceeded the period between routing updates and requests. A more thorough investigation of bucket times would reveal the advantages of longer or dynamically chosen bucket times; however, we reserve this topic for future work.

After a sufficient observation period, which consists of a number of buckets, the node attempts to determine if it has observed a Sybil node. The length of the observation period depends on the amount of mobility within the network; highly mobile networks need fewer intervals than more static networks. In our simulations, 200 observations over 6,000 seconds, or 100 minutes of simulated time, was sufficient. The node then determines which pairs of nodes are related. While correlation would be the most obvious candidate for doing so, it suffers from the fact that nodes that are never seen together will be highly correlated. In this case, however, nodes that are not seen together cannot be assumed to be related,they might be in separate parts of the network, unheard by both the observing node and each other. We therefore tried a number of different techniques to measure the relationship between nodes, including machine learning tools.

Our final and simple solution reflects the intuition that, during some observation period, seeing a pair together provides some evidence they are related, that seeing one but not the other of a pair provides stronger evidence they are not related; and that not seeing either of a pair nodes provides no evidence, because it is not possible to tell if they appear together elsewhere or separate elsewhere. Our solution also reflects that having more observations of the nodes in question provides more evidence than fewer observations.

After a period of observation, the detection algorithm then works in a series of simple steps:

1. We calculate Aij , the affinity between nodes i and j, as

$$Aij = \frac{Tij-2Lij}{N} Tij + Lij \qquad (1)$$

where Tij is the number of intervals in which nodes i and j were observed together, Lij is the number of intervals in which either i or j were observed alone, and N is total number of intervals in the observation period.

2. After the affinity between each pair of nodes has been computed, the observer constructs a graph in which the node identities are the vertices and the undirected edges are weighted with the affinity values between them. Only edges that are greater than a specific threshold parameter are included. Using our measure of affinity, we recorded our results using a threshold of 0.1.

3. Depth-first search (DFS) is then run over each vertex to discover the connected components. Each of the components found represents a possible Sybil attacker. While there can be several different connected components, we took only the largest to be a Sybil attacker, in line with the working assumption that there was only one per network. If there were more, they would appear as separate components. Note that this approach is clearly not optimal,DFS can have a long running time for large numbers of nodes. We will look to improve the scalability of the analysis algorithm in our future work. The justification for this affinity measure is that each identity of an attacking node must transmit often enough to participate in the protocols that operate the network, including routing. If the observation periods are long enough, the attacker will be forced to transmit within a single period to maintain the fiction of multiple identities. For example, in AODV, routes that are not used for 3 seconds are dropped, thus our observation period is set to 30 seconds in our evaluations to catch route re-formations. Accordingly, we expect that for most realistic scenarios, the attacker will find it difficult have identities transmit individually in separate periods. In situations where this is not the case, the period can be lengthened or the weights of observations together and apart can be adjusted to account for the change in difficulty.

3.2.2 Multiple Node Observers

While observations from a single node can be accurate in identifying a Sybil attacker, any single observer is inherently throttled in that area can be monitored easily. Collaborating observers might be able to determine that different identities are not related as they were seen in different areas at different times. Therefore PASID should increase in accuracy as we add observers to the network. We assume a subset of the legitimate nodes in the network can share observations periodically using the normal data transmission capabilities of the ad hoc network, and that these nodes can trust each other to perform this task honestly. Each node again tracks all other nodes that it hears over many time buckets. At the end of the observation period, it exchanges the information of what identities were heard during what time periods with the other nodes it trusts in the calculations. Note that this exchange does not have to occur often; in our simulations, it would only happen every 100 minutes. We do not simulate the exchange in our simulations, and assume that there is sufficient connectivity for all trusted nodes to reach each other; if this is not the case, detection will be delayed until node movement allows one or more nodes to accumulate the results of all observations. We will see, however, that the accuracy increases even if only some of the additional observations are received.

When using observations from more than one node, the counts are totals over all observing nodes and the last term of Equation 1 is at most 1. We let G represent the number of nodes sharing observations with one another. We modify our other variables to account for the multi-observer case. Now, Tij(n) is number of intervals in which nodes I and j were observed together by node n, defined as

$$Tij = \sum_{n \epsilon G} Tij(n)$$

We let Lij(n) is the number of intervals in which either I or j were observed alone by node n, defined as

$$Lij = \sum_{n \epsilon G} Lij(n)$$

N is still the total number of intervals in the observation period. Accordingly, the affinities for the multi-observer case are calculated as follows.

Aij = (Tij − 2Lij) wij

Where

$$\begin{cases} \frac{Tij + Lij}{N} & if\ Tij + Lij < N, \\ 1 & other\ wise \end{cases}$$

4 Conclusion

To manifest the effectiveness of our first detection protocol, Kotz, et al [13] have conveyed concerns about using ns2 to simulate mobile networking ,we simulate a series of ad hoc networks using the ns2 network simulator [14]. We pass judgment about the single observer and multi-observer cases.

We also introduced a limiting to the protocol that micturates use of MAC layer information. We tried to show that a single node can accurately identify the Sybil attacker with various identities, and that cooperating nodes can increase the accuracy of the process.

However, we believe the evaluations we presented are sufficient as a preliminary exploration of our method. This is because our sybil detection approach relies on simple observations but not dependent on the actual throughput rate of a channel or the particular efficacy of some routing or MAC protocol. In other words, at a minimal, our simulations clearly show the feasibility of our approach, more over a realistic models or evaluations over real traces would complicate performance numbers. Additionally, we model mobility using the random way point model, which come under scrutiny. This model is reasonable here because it does not restrict mobility along some path or sub area for any particular node. Changing the mobility model will not impact operation of the protocol, though it would again refine the results.

References

[1] S. Buchegger and J. Le Boudec. A Robust Reputation System for P2P and Mobile Ad hoc Networks. In Proc. Wkshp Economics of Peer-to-Peer Systems, June 2004.

[2] J. R. Douceur. The Sybil Attack. In Intl Wkshp on Peer-to-Peer Systems, March 2002.

[3] A. Cheng and E. Friedman. Sybilproof Reputation Mechanisms In ACM Wkshp on the Economics of Peer-to-Peer Systems, August 2005.

[4] J. Newsome, E. Shi, D. Song, and A. Perrig. The Sybil Attack in Sensor Networks: Analysis & Defenses. In Proc. Intl Symp on Information Processing in Sensor Networks, 2004.

[5] N. Mathewson, P. Syverson, and R. Dingledine. TOR: The Second-Generation Onion Router. In Proc. USENIX Security Symp, August 2004.

[6] C. E. Perkins and E. M. Royer. Ad hoc On-Demand Distance Vector Routing. In Proc. WMCSA, Feb. 1999.

[7] D. Johnson and D. Maltz. Dynamic Source Routing in Ad hoc Wireless Networks. In Mobile Computing, volume 353. Kluwer Academic Publishers, 1996.

[8] Y. Hu and A. Perrig. A Survey of Secure Wireless Ad hoc Routing. IEEE Security & Privacy, 2(3):28–39, May/June 2004.

[9] K. Sanzgiri, B. Dahill, D. LaFlamme, B. N. Levine, C. Shields, and E. Belding-Royer. A Secure Routing Protocol for Ad hoc Networks. JSAC Special Issue on Ad hoc Networks, March 2005 .

[10] Y. Hu, D. Johnson, and A. Perrig. SEAD: Secure Efficient Distance Vector Routing for Mobile Wireless Ad hoc Networks. In Proc. Wkshp on Mobile Computing Systems and Applications, Jun. 2002.

[11] P. Papadimitratos and Z. Haas. Secure Link State Routing for Mobile Ad hoc Networks. In Proc. Symp on Applications and the Internet Wkshps, January 2003.

[12] P. Papadimitratos and Z. Haas. Secure Routing for Mobile Ad hoc Networks. In Proc. Communication Networks and Distributed Systems Modeling and Simulation Conference, Jan. 2002.

[13] D. Kotz, C. Newport, R. Gray, J. Liu, Y. Yuan, and C. Elliott. Experimental evaluation of wireless simulation assumptions. In Proc. ACM/IEEE Intl Symp on Modeling, Analysis and Simulation of Wireless and Mobile Systems (MSWiM), pages 78–82, October 2004.

[14] S. McCanne and S. Floyd. Network Simulator Version 2. http://www.isi.edu/nsnam/ns.