Detecting Sybil Attacks in Wireless and Sensor Networks Using Cluster Analysis

Jie Yang, Yingying Chen
Dept. of ECE, Stevens Institute of Technology
Hoboken, NJ 07030
{jyang, yingying.chen}@stevens.edu

Wade Trappe
WINLAB, Rutgers University
110 Frelinghuysen Rd, Piscataway, NJ 08854
trappe@winlab.rutgers.edu

Abstract

Wireless networks are vulnerable to sybil attacks, in which a sybil node forges multiple identifications to trick the system and conduct harmful attacks. The traditional approach to address sybil attacks is to employ cryptographic-related methods. However, conventional security approaches may not always desirable due to their infrastructural overhead. In this paper, we propose to utilize K-means cluster analysis for detecting sybil attacks based on the spatial correlation between the signal strength and physical locations. Our approach requires minimal overhead to wireless devices. We have evaluated our methods through experimentation using both an 802.11 (WiFi) network as well as an 802.15.4 (ZigBee) network in two office buildings. Our results show that the proposed sybil attack detector is highly effective with over 95% detection rates and under 5% false positive rates.

1 Introduction

Because of the flexibility and openness of wireless and sensor networks, they are especially vulnerable to sybil attacks where an attacker, a sybil node, can forge different identities to trick the network and conduct harmful attacks. The sybil attack can significantly reduce the network performance by defeating group-based voting techniques, fault-tolerant schemes (e.g., redundancy mechanisms [15], distributed storage [2], dispersity and multipath routing [1, 8]), and geographic routing protocols [9]. For example, the sybil attack can undermine geographic routing by misleading the routing scheme to go through fake nodes with different identities.

Existing approaches of defending against sybil attacks in wireless and sensor networks employ key distribution methods and identity-based encryption schemes. However, the application of conventional security methods requires reliable key distribution, management, and maintenance mechanisms. Thus, the traditional security approaches are not always possible due to high computational requirements and infrastructure maintenance on nodes. Further, another serious concern is that cryptographic methods are susceptible to node compromise where wireless nodes are easily accessible and allowing their memory to be scanned and altered.

In this paper, we take a different approach by using spatial correlation between the signal strength and physical locations to detect sybil attacks. Specifically, our proposed attack detector utilizes the Received Signal Strength (RSS) and K-means cluster algorithm to perform sybil attack detection. By analyzing the RSS from each location using K-means cluster algorithm, we have found that the distance between the centroids in signal space is a good test statistic for effective attack detection. Our scheme does not add any infrastructural, computational, or management overhead to the wireless devices and sensor nodes.

Further, we conducted experiments using both an 802.11 network as well as an 802.15.4 network in two real office buildings to evaluate the effectiveness of our sybil attack detector. We used detection rate and receiver operating characteristic curve to evaluate the performance of our attack detector. We have found that our K-means sybil attack detector is highly effective with over 95% detection rates and under 5% false positive rates.

The rest of the paper is organized as follows. Section 2 presents related research in sybil attack detection. In Section 3, we formulate the statistical model for attack detection and present the test statistic based on cluster analysis. Section 4 shows our experimen-

1-4244-2575-4/08/$20.00 ©2008 IEEE

tal methodology and our evaluation results. Finally, we conclude our work in Section 5.

2 Related Work

Employing cryptographic-related methods [6, 12, 9, 11] are the traditional approaches to prevent sybil attacks. To address the issues of computational constraints, low power consumption, and small memory footprint on wireless and sensor nodes, [15] proposed schemes based on symmetric key cryptography to satisfy the resource requirements and [10] used unique random pairwise key establishment schemes based on $t - degree$ polynomials. However, the choice of the threshold t is a challenge. If t is too small, the attacker only needs to compromise a small percent of nodes to compromise the whole network. Whereas if t is too large, the storage overhead for each node will grow very high.

Further, because of the resource constraints on sensor nodes, it may not always possible to deploy cryptographic-based schemes. Radio resource testing and registration approaches are two methods that deviates from the conventional security approaches. Radio resource testing [11] assumes each node only has one radio, and cannot send and receive simultaneously on more than one channel. This testing process may consume a lot of battery power. Whereas registration alone cannot prevent sybil attacks, because a malicious attacker may get multiple identities by non-technical means such as stealing. The most closely related work to our paper is [5], which proposed the use of RSSI for sybil attack detection, whereas our approach utilizes K-means cluster analysis to detect sybil attacks. [14] utilized signal strength distributions to detect and localize sybil nodes in Vehicular Ad Hoc Networks (VANETs). Their statistical algorithms are closely associated with VANETs.

Our work differs from most previous research in that we use statistical significance testing and cluster analysis to serve as our mathematical foundation of sybil attack detection and our scheme does not add any overhead to the wireless devices and sensor nodes. In addition, we experimentally validated our approaches using different networks in real office building environment.

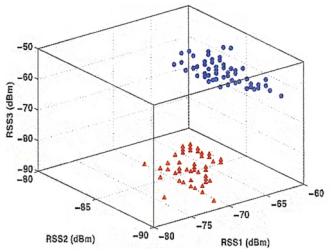


Figure 1. Illustration of RSS readings from two physical locations.

3 Sybil Attack Detector

In this section, we first formulate the sybil attack detection problem as statistical significance testing. We then derive the test statistic based on cluster analysis techniques and evaluate the feasibility of the derived test statistic.

3.1 Statistical Model for Attack Detection

We propose to formulate sybil attack detection as a statistical significance testing problem, where the null hypothesis is

$$\mathcal{H}_0 : \text{normal (no attack).}$$

In significance testing, a test statistic $\mathbf{T}$ is used to evaluate whether the observed data belongs to the null-hypothesis or not. For a particular significance level, α (defined as the probability of rejecting the hypothesis if it is true), there is a corresponding *acceptance region* Ω such that we declare the null hypothesis valid if an observed value of the test statistic $\mathbf{T^{obs}} \in \Omega$, and reject the null hypothesis if $\mathbf{T^{obs}} \notin \Omega$ (i.e. declare an attack is present if $\mathbf{T^{obs}} \in \Omega^c$, where Ω^c is the *critical region* of the test). In our sybil attack detection problem, the region Ω and decision rule is specified according to the form of the detection statistic $\mathbf{T}$ (for example, when using distance in signal strength space for $\mathbf{T}$ as presented below, the decision rule becomes comparison against a threshold), and rejection of the null hypothesis corresponds to declaring the presence of an attack.

3.2 Spatial Correlation of Signal Strength to Location

The RSS readings at different locations in physical space are distinctive and are determined by the distance to landmarks [3]. We define the RSS vector as $\mathbf{s} = \{s_1, s_2, ...s_n\}$ (n is the number of landmarks/access

points(APs) that are monitoring the RSS of the wireless nodes). The sequence of RSS sample vectors from one physical location will be close to each other, and will fluctuate around a mean vector due to random noise, environmental bias, and multipah effects [16].

Since each RSS reading vector corresponds to a point in a n-dimensional signal space [4], the RSS readings from the same physical location will belong to the same cluster points in the n-dimensional signal space, whereas the RSS readings from different locations should form different clusters in signal space as shown in Figure 1, which presents RSS reading vectors of three landmarks from two different physical locations. This observation suggests that we may conduct cluster analysis to perform attack detection.

3.3 Test Statistic for Attack Detection

The K-means algorithm is one of the most popular iterative descent clustering methods [7]. The squared Euclidean distance is chosen as the dissimilarity measure. If there are N RSS sample readings for a node, the K-means clustering algorithm partitions N sample points into K disjoint subsets S_i containing N_i sample points so as to minimize the sum-of-squares criterion:

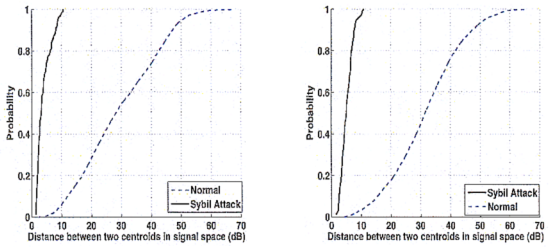
$$I_{min} = \sum_{i=1}^{K} \sum_{\mathbf{s_n} \in S_i} \|\mathbf{s_n} - \alpha_\mathbf{i}\|^2 \quad (1)$$

where $\mathbf{s_n}$ is a RSS vector representing the nth sample point and $\alpha_\mathbf{i}$ is the geometric centroid of the sample points for S_i in signal space. We further propose to choose the distance between two centroids as the test statistic $\mathbf{T}$ for sybil attack detection,

$$H_c = ||\alpha_\mathbf{i} - \alpha_\mathbf{j}|| \quad (2)$$

with $i, j \in \{1, 2..K\}$. We note that different from most existing works [13], the advantage of our proposed attack detection approach is that it does not assume that RSS readings follow any distributions (e.g. Gaussian distribution).

The basic idea behind using the proposed K-means cluster analysis to detect sybil attacks relies on physical locations of nodes. When examining the RSS readings from two different identities, we can apply K-means cluster analysis to the *mixture* of these two RSS streams. Under normal conditions without a sybil attack, the observed value of the test statistic $\mathbf{T^{obs}}$ (i.e., H_c^{obs}) should be large, since there are basically two



(a) 802.11 network (b) 802.15.4 network

Figure 2. Cumulative Distribution Function (CDF) of the test statistic H_c in the signal space.

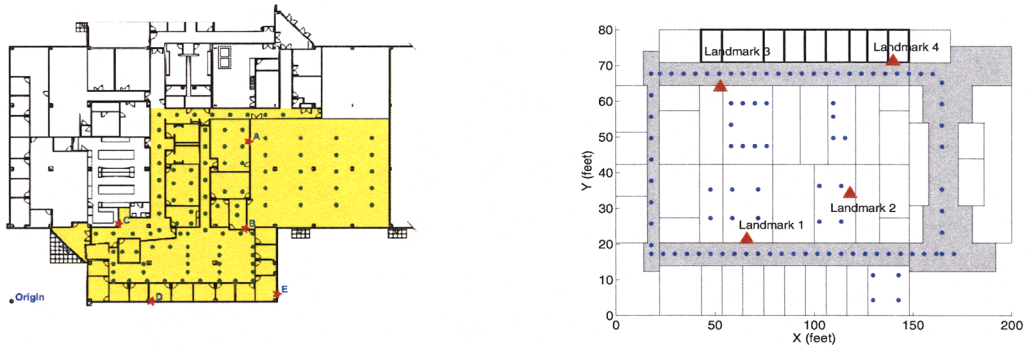
different RSS clusters coming from two physical locations. Whereas when a sybil attack is present, the $\mathbf{T^{obs}}$ is small and satisfies $\mathbf{T^{obs}} < \tau$, because the RSS readings are originated from one physical location (i.e., the location of a sybil node) and thus there is basically only one cluster in the signal space.

3.4 Determining Thresholds

The thresholds define the critical region for the significance testing. Appropriately choosing a threshold τ will allow the attack detector to be robust to false detections. We show how we determine the thresholds through empirical training for K-means algorithm. During the off line phase, we can collect the RSS readings across a set of locations over the experimental area and obtain the distance between two centroids in signal space for each node. We then use the distribution of the training information to determine the threshold τ. At run time, based on the RSS sample readings for each node pair (i.e. the mixture of two RSS streams), we can calculate the observed value H_c^{obs}. Our condition for declaring the presence of a sybil attack is:

$$H_c^{obs} < \tau. \quad (3)$$

Figure 2 (a) and (b) present the Cumulative Distribution Function (CDF) of the H_c in signal space for an 802.11 network as well as an 802.15.4 network. We found that the the value of H_c is small under sybil attacks, whereas the value of H_c is large under normal situations. This suggests that using H_c as a test statistic is effective for detecting sybil attacks.



(a) 802.11 network (b) 802.15.4 network

Figure 3. Experimental setup across two networks in two office buildings.

4 Performance Evaluation

In this section, we first describe the experimental methodology and metrics that we use to evaluate our approach of sybil attack detection. We then present our experimental results.

4.1 Experimental Methodology

In order to evaluate the effectiveness of our sybil attack detection mechanisms, we have conducted experiments using two networks: an 802.11 (WiFi) network at the Wireless Information Network Laboratory (WINLAB) and an 802.15.4 (ZigBee) network on the 3rd floor of the Computer Science Department at Rutgers University. The size of these two floors are 219x169ft and 200x80ft respectively. Figure 3 (a) shows the 802.11 (WiFi) network with 5 landmarks deployed to maximize signal strength coverage shown in red stars. Whereas the 802.15.4 (ZigBee) network is presented in Figure 3 (b) with 4 landmarks distributed in a squared setup in order to achieve optimal landmark placement [3] as shown in red triangles.

The small dots in floor maps are the locations used for testing. For the 802.11 network, there are 101 locations and we used 300 packet-level RSS samples at each location, while for the 802.15.4 network, there are 94 locations and 300 packet-level RSS samples are used at each location.

Based on the measured data, we then apply K-means algorithm to RSS samples from one location (corresponding to cases of sybil attacks)and the mixture of RSS samples from two different locations (corresponding to normal situations) respectively by setting $K = 2$. We ran sybil attack detection test through all the possible combinations of location pairs on the experimental floor using all the locations in both networks. There are total 5050 location pairs for the 802.11 network and 4371 location pairs for the 802.15.4 network.

4.2 Performance Metrics

We use the following metrics to evaluate the performance of our sybil attack detector.

Detection Rate and False Positive Rate: A sybil attack will cause the significance test to reject $\mathcal{H}_0$. We are thus interested in the statistical characterization of the attack detection attempts. The detection rate is defined as the percentage of sybil attack attempts that are determined to be under attack. Note that, when the sybil attack is present, the detection rate corresponds to the probability of detection P_d, while under normal conditions it corresponds to the probability of false positive P_{fa}. The detection rate and false positive rate vary under different thresholds τ.

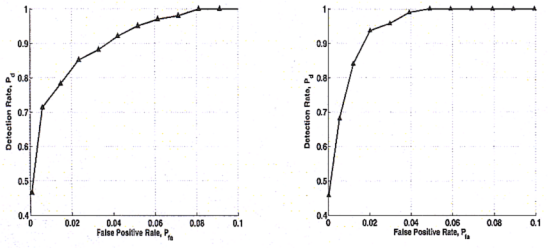
Receiver Operating Characteristic (ROC) curve: To evaluate the effectiveness of an attack detection scheme we want to study the probability of detection P_d and the false positive rate P_{fa} together. The ROC curve is a plot of attack detection accuracy against the false positive rate. It can be obtained by varying the detection thresholds τ. The ROC curve provides a direct means to measure the trade off between false-positives and correct detections.

4.3 Experimental Evaluation

In this section we present the evaluation results of the effectiveness of the sybil attack detector. Table 1 presents the detection rate and false positive rate for

Network, Threshold	Detection Rate	False Positive Rate
802.11, τ = 10.6dB	1.0000	0.0736
802.11, τ = 9.1dB	0.9505	0.0485
802.11, τ = 8.2dB	0.9010	0.0346
802.15.4, τ = 10.9dB	1.0000	0.0409
802.15.4, τ = 9.1dB	0.9574	0.0269
802.15.4, τ = 7.9dB	0.9043	0.0173

Table 1. Detection rate and false positive rate of the sybil attack detector under different thresholds τ in two networks.

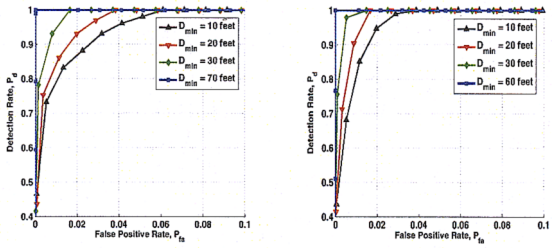


(a) 802.11 network (b) 802.15.4 network

Figure 4. Receiver Operating Characteristic (ROC) curves over all the testing points across the experimental floor.

both the 802.11 network and the 802.15.4 network under different threshold (τ) settings. The corresponding ROC curves are displayed in Figure 4. The results are encouraging showing that the attack detector can achieve detection rate over 95% with less than 10% false positive rate. Even when the detection rate reaches 100%, the false positive rate is only 7.4% for the 802.11 network and 4.1% for the 802.15.4 network respectively. In addition, from Table 1, we observed that the similar thresholds are achieved for both networks under detection rates of 90%, 95% and 100%. These results indicate that our sybil attack detector is generic across different networks and is highly effective in performing attack detection.

We further study how the detection rate and the false positive rate are affected by the distance between two wireless nodes in a network. We define a distance threshold D_{min}, which is the minimum distance between two nodes. Figure 5 shows the ROC curves under different thresholds of D_{min} for both the 802.11 and the 802.15.4 network. We note that the distances between two nodes in the data subset that used to generate each ROC curve in the figure are larger than D_{min}. We found that the ROC curves shift to the left when increasing the



(a) 802.11 network (b) 802.15.4 network

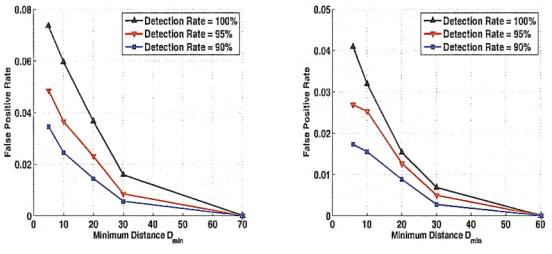
Figure 5. Receiver Operating Characteristic (ROC) curves when varying the node distance threshold D_{min}.

threshold D_{min}. This indicates that the farther away two nodes are from each other, the higher the detection rate and the lower false positive rate are achieved.

Since if two wireless devices are close to each other, the resulting test statistic H_c^{obs} will be small and may be less than the threshold (i.e., $H_c^{obs} < \tau$). Consequently the attack detector will claim a false positive (i.e., declaring the presence of a sybil attack). Thus, we further examine how likely the false positive rate of our detector can be reduced by varying the node distance threshold D_{min}. Figure 6 presents the false positive rate as a function of D_{min} under different detection rates for both the 802.11 network and the 802.15.4 network. First, the curves of false positive rate show that the higher detection rate usually results in higher false positive rate, which is inline with our observation when using ROC curves. Second, the results indicate that the false positive rate decreases as the D_{min} increases. For instance, by examining the curve under the detection rate of 95%, the false positive rate decreases from 3.66% to 0.85% in the 802.11 network and from 2.53% to 0.49% in the 802.15.4 network respectively, when D_{min} increases from 10 feet to 30 feet. Additionally, we observed that the detector can achieve 100% detection rate with 0% false positive rate when D_{min} reaches 68 feet in the 802.11 network and 56 feet in the 802.15.4 network respectively.

5 Conclusion

In this paper, we proposed to detect sybil attacks in wireless and sensor networks using statistical significance testing and cluster analysis. Compared with conventional cryptographic-based approaches, our cluster analysis based attack detector does not add additional



(a) 802.11 network (b) 802.15.4 network

Figure 6. False positive rate as a function of the node distance threshold D_{min}.

overhead to the wireless devices and sensor nodes. More specifically, we formulated the sybil attack detection problem as a statistical significance testing problem. We then use the distance between centroids derived from the K-means cluster analysis as the test statistic.

To evaluate the effectiveness and generality of our sybil attack detector, we conducted experiments in both an 802.11 (WiFi) network as well as an 802.15.4 (ZigBee) network in two real office building environments. The performance of the sybil attack detector is evaluated in terms of detection rates and receiver operating characteristic curves. Our attack detector has achieved high detection rates, over 95% and low false positive rates, below 5%. Therefore, our experimental results provide strong evidence of the effectiveness of our approach in detecting sybil attacks in wireless and sensor networks.

References

[1] A. Banerjea. A taxonomy of dispersity routing schemes for fault tolerant real-tiem channels. In *Proceedings of ECMAST*, volume 26, pages 129–148, May 1999.

[2] M. Castro and B. Liskov. Practical byzantine fault tolerance. In *Proceedings of the Symposium on Operating System Design and Implementation (OSDI)*, 1999.

[3] Y. Chen, J. Francisco, W. Trappe, and R. P. Martin. A practical approach to landmark deployment for indoor localization. In *Proceedings of the Third Annual IEEE Communications Society Conference on Sensor, Mesh and Ad Hoc Communications and Networks (SECON)*, September 2006.

[4] Y. Chen, K. Kleisouris, X. Li, W. Trappe, and R. P. Martin. The robustness of localization algorithms to signal strength attacks: a comparative study. In *Proceedings of the International Conference on Distributed Computing in Sensor Systems (DCOSS)*, pages 546–563, June 2006.

[5] M. Demirbas and Y. Song. An rssi-based scheme for sybil attack detection in wireless sensor networks. In *Proceedings of the International Workshop on Advanced Experimental Activities on Wireless Networks and Systems*, 2006.

[6] L. Eschenauer and V. D. Gligor. A key-management scheme for distributed sensor networks. In *Proceedings of the 9th ACM conference on Computer and communications security (CCS)*, November 2002.

[7] T. Hastie, R. Tibshirani, and J. Friedman. *The Elements of Statistical Learning, Data Mining Inference, and Prediction*. Springer, 2001.

[8] K. Ishida, Y. Kakuda, and T. Kikuno. A routing protocol for finding two node-disjoint paths in computer networks. In *Proceedings of the International Conference on Network Protocols*, pages 340–347, November 1992.

[9] C. Karlof and D. Wagner. Secure routing in wireless sensor networks: Attacks and countermeasures. In *Proceedings of the First IEEE International Workshop on Sensor Network Protocols and Applications*, May 2003.

[10] D. Liu and P. Ning. Establishing pairwise keys in distributed sensor networks. In *Proceedings of the 10th ACM Conference on Computer and Communications Security (CCS)*, October 2003.

[11] J. Newsome, E. Shi, D. Song, and A. Perrig. The sybil attack in sensor networks: analysis and defenses. In *Proceedings of 3rd International Symposium on Information Processing in Sensor Networks (IPSN)*, April 2004.

[12] A. Perrig, R. Szewczyk, V. Wen, D. Culler, and J. Tygar. Spins: Security protocols for sensor networks. *Wireless Networks*, 8(5):521–534, September 2002.

[13] Y. Sheng, K. Tan, G. Chen, D. Kotz, and A. Campbell. Detecting 802.11 mac layer spoofing using received signal strength. In *Proceedings of the IEEE International Conference on Computer Communications (INFOCOM)*, April 2008.

[14] B. Xiao, B. Yu, and C. Gao. Detection and localization of sybil nodes in vanets. In *Proceedings of the Workshop on Dependability Issues in Wireless Ad Hoc Networks and Sensor Networks (DIWANS)*, 2006.

[15] Q. Zhang, P. Wang, and D. P. N. Reeves. Defending against sybil attacks in sensor networks. In *Proceedings of the 25th IEEE International Conference on Distributed Computing Systems Workshops (ICDCSW)*, June 2005.

[16] G. Zhou, T. He, S. Krishnamurthy, and J. A. Stankovic. Models and solutions for radio irregularity in wireless sensor networks. *ACM Transactions on Sensor Networks*, 2:221–262, 2006.