Future Network and MobileSummit 2010 Conference Proceedings
Paul Cunningham and Miriam Cunningham (Eds)
IIMC International Information Management Corporation, 2010
ISBN: 978-1-905824-18-2

Effective and Robust Detection of Jamming Attacks [1]

Alexandros G. Fragkiadakis, Vasilios A. Siris[2], Apostolos P. Traganitis[3]
Institute of Computer Science, Foundation for Research and Technology - Hellas (FORTH)
P.O. Box 1385, GR 711 10 Heraklion, Crete, Greece
Tel: +3 2810 391567 Email:(alfrag,vsiris,tragani)@ics.forth.gr

Abstract: In this paper we present and evaluate anomaly-based intrusion detection algorithms for detecting attacks at the physical layer of wireless networks, by seeking for changes in the Signal-to-Noise ratio statistical characteristics. Two types of algorithms are proposed: simple threshold algorithms and cumulative sum (cusum) algorithms. Performance evaluation is performed in terms of the detection probability, false alarm rate, detection delay and the robustness of the algorithms to different detection threshold values. The algorithms are applied locally to measurements collected from three locations of an experimental network and under two attack intensities. The results show that the cumulative sum algorithms are more robust and achieve higher performance under both attack intensities. Next, we use the Dempster-Shafer algorithm to fuse the outputs provided by the above locally executed algorithms at different nodes, thus forming a collaborative intrusion detection system. The evaluation shows that the robustness substantially increases while the performance remains high, for both types of attacks.

Keywords: collaborative intrusion detection, signal-to-noise ratio, jamming, simple threshold algorithms, cumulative sum algorithms, performance evaluation, Dempster-Shafer

1. Introduction

The broadcast nature of wireless networks makes them more susceptible to attacks. Adversaries can exploit vulnerabilities in the Medium Access Control and Physical layers and heavily disrupt the services between the network nodes. This is feasible simply by using commodity hardware; therefore, the existence of an intrusion detection mechanism at the physical layer is necessary. Generally, intrusion detection falls into two categories: (i) misuse detection and (ii) anomaly detection. The former is based on known signature attacks, so it lacks the ability to detect new types of attacks, while the latter is more promising because of its potential ability to detect unknown types of intrusions. A primary assumption of intrusion detection is that a network's normal behavior is distinct from abnormal or intrusive behavior that can be the result of an attack. In this paper, we present and evaluate anomaly-based intrusion detection algorithms for detecting attacks at the physical layer. These attacks are referred as jamming and the attackers as jammers. The algorithms we investigate are of two types: (i) locally executed algorithms and (ii) fusion algorithms. Local algorithms execute independently on a number of monitors seeking for changes in the statistical

[1]This work is supported in part by the European Commission in the 7th Framework Programme through project EU-MESH, ICT-215320, www.eu-mesh.eu
[2]V. A. Siris is also with the Department of Informatics at the Athens University of Economics and Business
[3]A. P. Traganitis is also with the Department of Computer Science at the University of Crete

characteristics of the Signal-to-Noise ratio (SNR). SNR measurements are collected through the Ath5k driver [1]. Local algorithms are further divided into simple threshold algorithms and algorithms based on the cumulative sum (cusum) change point detection procedure. The algorithms consider SNR-based metrics which include the average SNR, minimum SNR, and max-minus-min SNR, in a short window. Our motivation for using SNR rather some other metric (e.g. number of PHY or CRC errors) is that hardware radio interfaces and wireless device drivers typically provide values of the SNR for the received packets. In general, fusion algorithms are used to fuse the output of the locally executed algorithms so as to form a collaborative intrusion detection system. In this work, we investigate the performance of the Dempster-Shafer fusion algorithm that successfully combines the outputs produced by the local algorithms, executing at all nodes.

Related work includes several important contributions. The authors in [2] present two types of detection algorithms considering metrics such as the packet delivery ratio, bad packet ratio and energy consumption amount. The evaluation shows high detection rates but trade-off points regarding the false alarm rate versus detection probability or detection delay are not presented. The authors in [3] use the Dempster-Shafer algorithm to fuse data provided by heterogeneous monitors. Their intrusion detection system considers metrics for the detection of UDP and ICMP flooding attacks as well as SYN attacks. The work presented in [4] evaluates two types of algorithms for the detection of SYN attacks. The evaluation shows that the simple detection algorithm has satisfactory performance for high intensity attacks but deteriorates for the low intensity attacks. The cumulative sum algorithm, on the other hand, has robust performance for different types of attacks. The authors in [5] present a distributed anomaly detection system based on simple thresholds. A method for combining measurements using the Pearson's Product Moment correlation coefficient is also presented. A disadvantage of this method is that "raw" RSSI measurements by several sniffers are needed. This could generate a high volume of traffic flowing from the sniffers to a main node where the algorithm executes. We propose to use the outputs of several local detection algorithms without the need of transmitting SNR values in a per-packet basis. Several adversarial models are presented in [6], all focusing on RF jamming attacks. One of the proposed algorithms, applies ***high order crossings***, a spectral discrimination mechanism that distinguishes normal scenarios from two specific types of jammers. The authors introduce two detection algorithms, based on thresholds that use signal strength and location information as a consistency check to avoid false alarms. The authors in [7] present a cross layer approach to detect jamming attacks. They consider jamming performed at the physical layer by using RF signals, and at the MAC layer by targeting the RTS/CTS and NAV mechanisms of the IEEE 802.11 protocol. Although significant, none of these works is referred to robustness of the detection algorithms.

Our main contributions are as follows: (i) we consider different metrics for the local algorithms based on the SNR: average, minimum, and max-minus-min SNR, (ii) we consider the Dempster-Shafer algorithm for combining information from a number of monitors, (iii) we investigate the performance of the local and fusion algorithms in terms of the detection probability, false alarm rate, detection delay, and their robustness to different detection threshold values and (iv) we present the performance of the local and fusion algorithms for measurements from a real network, under two attack intensities, collected from locations whose distance from a jammer varies.

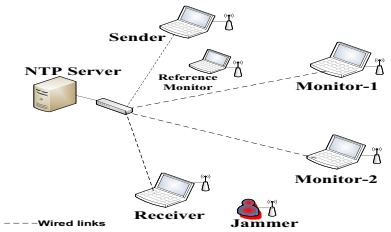


Fig. 1: Network layout for collecting SNR measurements

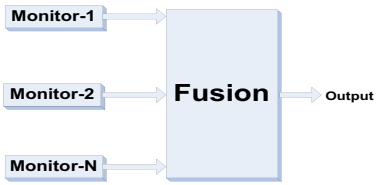


Fig. 2: Data fusion for collaborative intrusion detection

The rest of the paper is organised as follows. Section 2. presents the experimental network layout deployed for the collection of SNR measurements and the jamming model used to launch the attacks. The local detection algorithms and their evaluation is described in Section 3. In Section 4. we discuss the Dempster-Shafer algorithm for data fusion and its evaluation. Finally, conclusions and further work appear in Section 5.

2. Experimental Setup and Jamming Model

The experimental network layout shown in Fig. 1, consisting of off-the-shelf IEEE 802.11a devices, has been deployed for launching the jamming attacks and collecting SNR measurements. UDP traffic is sent from Sender to Receiver at a constant rate of 18 Mbps. In addition to Receiver, Monitor-1 and Monitor-2 also collect SNR measurements. Synchronisation is achieved through the use of an NTP server and a wired backbone. SNR measurements are collected along with their corresponding timestamps, through a modified version of the Ath5k driver. Timestamps are necessary to time align measurements at all nodes participating in the experiment and to investigate the performance evaluation of the algorithms. All nodes shown in Fig. 1, except Jammer, operate on the same channel. Furthermore, the network interface cards of both monitors are set to monitor mode, hence receive all packets sent in the channel. Jammer operates on a different channel broadcasting UDP packets for a period of 30 seconds, after which it remains inactive for 30 seconds. We have also modified the values of several hardware registers (through Ath5k) that are part of the IEEE 802.11 wireless card (Atheros), disabling the backoff and CCA (clear channel assessment) mechanisms of Jammer. With these mechanisms disabled, Jammer is now a non-compliant IEEE 802.11 node immune to the energy radiated by the legitimate nodes and thus it can perform jamming independently. The reference monitor (RM) depicted in Fig. 1 measures the noise radiated by Jammer and assists to the classification of two attack intensities as follows: (i) high intensity attack, where the packet loss is over 50%, the throughput degradation over 80% and the noise reported by RM is over -55 dbm, (ii) low intensity attack, where the packet loss is less than 15%, the throughput degradation less than 30% and the noise reported by RM is below -75 dbm (throughput and packet loss are measured at the Receiver).

3. Local Detection Algorithms

Local detection algorithms execute independently at both monitors and the Receiver and fall into two categories: (i) simple threshold algorithms and (ii) cumulative sum (cusum) change point detection algorithms. Both types are applied to different metrics

all based on SNR: average SNR, minimum SNR, and max-minus-min SNR. The values of these metrics are measured over a small time window and then compared to another metric over a large time window. The simple threshold algorithms trigger an alarm when the metric that the algorithm considers deviates from its normal (expected) value by some amount. The normal value is given by the value of the metric, estimated in a long time window, whereas the amount of deviation which activates an alarm is determined by the detection threshold. If K is the number of samples in the small window, M the number of samples in the long window and N the total number of the collected samples, then $\forall n \in [M+1, N]$, we define as D_n^S the metric of the simple algorithm measured in the small window and D_n^L the metric in the large window. D_n^L for each algorithm is shown in Table 1 (the formulas for D_n^S are omitted since they are straightforward explained by the definitions that follow). Different algorithms consider different metrics within the small and large windows. The Simple average algorithm (S_{avg}) considers the average value of SNR both in the small and large windows. The Simple min algorithm (S_{min}) uses the minimum value of SNR in the small window and the average SNR in the long window, while Simple max-min (S_{mm}) uses the maximum-minus-minimum value of SNR in the small window and the average of the differences of maximum-minus-minimum values of SNR in the long window. The last column of Table 1 shows the combined metric that is compared to a detection threshold. If h is the detection threshold, an alarm is raised at the arrival of frame n, if $Z_n \geq h$.

Table 1: Metrics of the simple local algorithms

Algorithm	D_n^L	Z_n
S_{avg}	$\frac{\sum_{i=n-M+1}^{n} SNR_i}{M}$	$1 - \frac{D_n^S}{D_n^L}$
S_{min}	$\frac{\sum_{i=n-M+1}^{n} SNR_i}{M}$	$\frac{D_n^L}{D_n^S}$
S_{mm}	$\frac{\sum_{i=n-M+1}^{n} D_i^S}{M}$	$D_n^S - D_n^L$

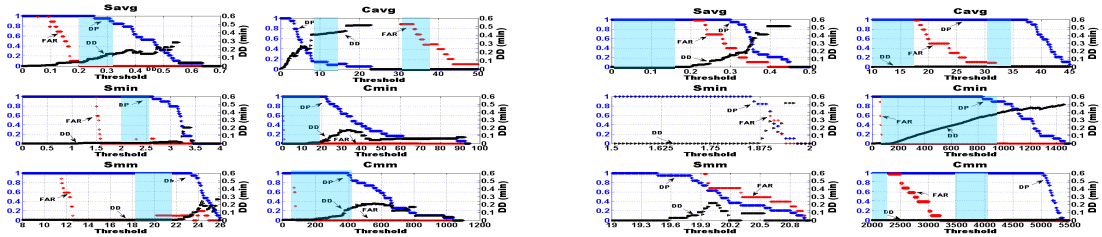
The second category of the local algorithms that we investigate in this paper are the cumulative sum (cusum) algorithms. This type of algorithm has been widely used in the literature [8, 9, 10, 11]. Cusum is a sequential change point detection procedure aiming to detect abrupt changes of a specific metric. If the probability distribution of the metric (e.g. SNR-based metrics) before an incident (e.g. attack) and after the incident are unknown, cusum is suitable for detecting such changes based on the assumption that the average value of the metric is negative before the change and becomes positive after the change. Generally, there are two main categories of cusum algorithms: (*i*) parametric and (*ii*) non-parametric. For the parametric cusum, a parametric model for $\{x\}$, where x is an independent and identically distributed random variable, is required which is not easy to obtain in the area of the communication networks and especially for the SNR, due to its variability. For this reason, we use non-parametric cusum algorithms where a model of $\{x\}$ is not required. The cusum algorithms considered in this work are the "cusum-versions" of the simple algorithms considered above, namely Cusum average (C_{avg}), Cusum min (C_{min}) and Cusum max-min (C_{mm}). The regression formula for each algorithm is given by: $y_n = max(0, y_{n-1} + Z_n - a)$, where Z_n is given in Table 1 and $a > 0$ is a tuning parameter.

3.1 Performance Evaluation of the Local Detection Algorithms

The performance of the algorithms is evaluated in terms of the detection probability (DP), false alarm rate (FAR), average detection delay (DD), and the robustness to different detection threshold values. Detection probability is defined as the ratio of the detected attacks over the total number of the attacks. The false alarm rate is the ratio of the number of false alarms over the total duration of the experiment in minutes. A false alarm occurs when there is no attack but an alarm is raised. The average detection delay is the mean time between the start of an attack and its detection (note that more than one alarms can be raised during a single attack), measured in minutes. The size of the sliding window used in the experiments is 1000 and 10 samples, for the long and short windows, respectively. Note that 1000 samples correspond to approximately 0.9 seconds, which is about one thirtieth of the attack duration (30 seconds). Traditionally, performance evaluation is presented by showing the trade-off points between FAR and DP [4, 12, 10, 13]. Although this is a significant method for performance evaluation, it is not complete, as it provides no information regarding the robustness of the algorithms. By robustness we mean the variation in the performance in terms of the detection probability and false alarm rate, when the detection threshold changes. We define a new metric that addresses this concern as: $M = \frac{1}{c+DP} + FAR$. M combines the detection probability DP and false alarm rate FAR. $c > 0$ is selected so as M approaches $\frac{1}{c}$ and not infinity, when $DP \to 0$. We define that a detection algorithm is (relatively) robust if the detection threshold needs to change by more than 20% for the metric M to change by more than 20%. Both these numbers and the specific form of function M are operator-controlled.

Fig. 3a shows DP, FAR and DD as function of the detection threshold, when the measurements at the Receiver are used, and for the high intensity attack. The values for DD are referred on the right vertical axis. The range of the detection threshold values for which an algorithm is robust is shown as a shaded area. In this figure we observe that all algorithms, except C_{avg}, reach maximum performance (DP=1 and FAR=0). Also, the threshold areas within which they reach the maximum performance are robust areas. Regarding detection delay, within the robust regions, it is less than 0.05 minutes for S_{min}, S_{mm} and C_{min}. For the S_{avg}, delay is up to 0.18 minutes while for C_{mm} is up to 0.1 minutes. For the low intensity attack we have observed (graph not shown due to space constraints) that the performance of the algorithms slightly deteriorates achieving $DP = 1$ and $FAR = 0.1$, except for C_{mm} that still can reach maximum performance. As in the high intensity attack scenario, the areas where the maximum or good performance is achieved, algorithms remain robust. The detection delay for all algorithms is less than 0.5 minutes.

Fig. 3b shows the performance evaluation of the local algorithms when the SNR traces collected at Monitor-1 are considered, and for the high intensity attack. Here we observe that S_{avg} has robust regions but within these regions the performance is low (high FAR). S_{min} and S_{mm} have no robust regions at all. On the other hand, all cusums are robust within threshold regions where the performance is high. The delay, within the robust regions, for C_{avg} and C_{mm}, is less than 0.05 minutes, while for C_{min} can vary up to 0.5 minutes. For the low intensity attack, we have observed that S_{avg} and S_{min} are robust only in regions where the performance is low. S_{mm} has no robust regions. The cusums again are robust within threshold areas where they achieve



(a) Receiver (b) Monitor-1

Fig. 3: Performance evaluation of the local algorithms when applied to measurements at the Receiver and Monitor-1 (high intensity attack)

high performance. Their detection delay can vary from 0.002 to 0.35 minutes. When considering the measurements at Monitor-2, we observe similar performance.

4. Collaborative Intrusion Detection

A collaborative intrusion detection system (CIDS) is a system that collects and fusses information provided by the monitors, taking the final decision about a possible attack (Fig. 2). Depending on the output of the locally executed algorithms, several types of fusion algorithms can be used (e.g. average, product, majority vote, Dempster-Shafer etc). The performance evaluation of the local algorithms, presented in Section 3.1, was based on a binary output format (attack, no attack). In this work, we evaluate the Dempster-Shafer (D-S) fusion algorithm that is used when the output of the local algorithms is continuous. The basic Dempster-Shafer Theory of Evidence (D-S) is a mathematical framework for the representation of uncertainty and its main advantage is that no *a priori* knowledge of the system is required; thus making it suitable for anomaly detection of previously unseen information [14]. This advantage becomes more useful for our work, where the anomaly detection algorithms are based on SNR metrics. It is well known that SNR exhibits high variability; therefore no models exist to describe its fluctuation under different network conditions. D-S has several essential concepts: (i) the *frame of discernment* (Θ), is the set of all possible mutually exclusive and complete states of a system, here we have selected $(\Theta) = \{attack, normal, (attack \text{ or } normal)\}$, (ii) the *probability assignment function* (or mass function), expresses a belief based on some evidence. Initially, the monitors (including Receiver) that comprise the collaborative intrusion detection system express a belief for each state that is described in (Θ). (iii) the *belief function* measures the belief of a hypothesis A and it computes the sum of all non-empty subsets of A and is given by $Bel(A) = \sum_{B \subseteq A} m(B)$, where $m(B)$ is the mass function of B, subset of A. We focus only on detecting a single type of attack (jamming at the physical layer), so the belief function and the mass function are equivalent (there is only one subset B of proposition A). D-S has the ability to combine evidence from different information sources. Assuming there are two information sources (in our case two monitors) then, if Monitor-1 believes that hypothesis A is true with confidence $m_1(A)$, and Monitor-2 believes that hypothesis A is true with confidence $m_2(A)$, D-S can combine these two separate beliefs into a single (combined) belief: $m_{12}(A) = \frac{\sum_{B \cap C = A} m_1(B) m_2(C)}{1-K}$, where $K = \sum_{B \cap C = \varnothing} m_1(B) m_2(C)$ represents a basic probability mass related to conflict. Conflict appears, when sources

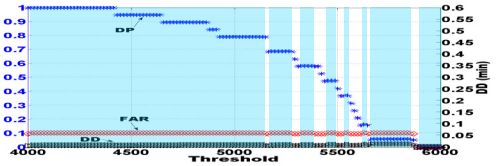


Fig. 4: Performance evaluation of Dempster-Shafer when using C_{mm}, for the high intensity attack

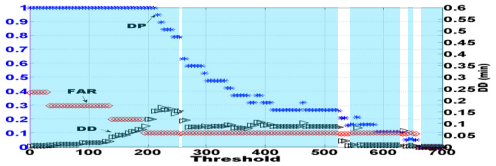


Fig. 5: Performance evaluation of Dempster-Shafer when using C_{mm}, for the low intensity attack

express contradictory beliefs. If information from a third source is available, D-S due to its associative characteristic, updates the combined mass function; therefore, D-S can combine the information provided by the local algorithms executing at Monitor-1, Monitor-2 and Receiver (Fig. 1) and produce a combined mass function m_a about a possible attack . If h is the detection threshold, an alarm is raised if $m_a \geq h$. The output of the local algorithms is transformed to a belief through the use of a linear function that is proportional to the difference $Z_n - h$ for the simple algorithms, and $y_n - h$ for the cusum algorithms (Z_n is computed according to Table 1).

4.1 Performance Evaluation of the Dempster-Shafer Algorithm

Fig. 4, shows the performance evaluation of D-S when fusing information provided by Monitor-1, Monitor-2 and Receiver, and for the high intensity attack. The local algorithm that executes in all monitors is C_{mm}. Observe that D-S has extended robust regions achieving a performance of DP=1 and FAR=0.1 that corresponds to 100% detection probability with a single false alarm. This performance is achieved within areas where the algorithm is robust. The detection delay is up to 0.01 minutes. For the low intensity attack (Fig. 5) D-S has a performance of DP=1 and FAR=0.1, within robust threshold areas. The detection delay is up to 0.17 minutes.

As we observed in Fig. 3a and Fig. 3b the performance of the algorithms depends not only on their type (simple or cusum) but it also depends on the relative distance between the jammer and a monitor. This is because the metrics we consider are based on the SNR and SNR decreases as the distance from the jammer decreases because the level of noise is higher close to Jammer. Nevertheless, when fusion is used, the combined output is less dependent on the distance from the jammer as it combines the output of all monitors. This is more obvious from the low intensity attack results.

5. Conclusions-Further work

In this paper we have described and evaluated anomaly-based intrusion detection algorithms for the detection of jamming attacks at the physical layer. Our algorithms seek for changes in the statistical characteristics of SNR and are of two types: (i) simple threshold algorithms and (ii) cusum-type algorithms. We collected SNR traces from three locations of a real network and evaluated the algorithms in terms of the detection probability, false alarm rate, detection delay and their robustness to different detection threshold values and under two attack intensities (high and low). When the measurements at Receiver are used, and for both attack intensities, all algorithms (except C_{avg}) have good performance within threshold regions where they are robust (Receiver is located very close to Jammer, so SNR fluctuations during the attacks are higher). When

the measurements at Monitor-1, that is in a larger distance from Jammer, are considered, and for both attack intensities, cusums have good performance while being robust. The simple threshold algorithms can be robust but their performance deteriorates when considering the measurements at Monitor-1. This is because Monitor-1 is located in a larger distance from Jammer, so the small SNR fluctuations do not necessarily lead to alarm triggering. We obtain the same results for the traces at Monitor-2. Next, we used the Dempster-Shafer algorithm to fuse the information provided by one of the local algorithms that executes at Receiver, Monitor-1 and Monitor-2. The performance evaluation shows that under both attack intensities, the robust areas have been increased compared to the areas of the local algorithm at the separate (single) locations. Furthermore, D-S achieves good performance regardless of the attack intensity with 100% detection probability and a single false alarm.

D-S has been criticised as giving counterintuitive results when there is high conflict between the information sources. The next step in our work will include the investigation of conflict between the monitors and the use of alternative algorithms for data fusion. These algorithms can include fusion rules such as the majority voting, average, product and more advanced algorithms that consider conflict solving.

References

[1] "Linux wireless drivers, ath5k, `http://linuxwireless.org/en/users/Drivers/ath5k`.".

[2] M. Cakiroglou and T. Ozcerit, "Jamming detection mechanisms for wireless sensor networks," in *Proc. of 3rd Int. Conference on Scalable Information Systems*, (Napoli, Italy), June 2008.

[3] V. Chatzigiannakis, G. Androulidakis, K. Pelechrinis, S. Papavassiliou, and V. Maglaris, "Data fusion algorithms for network anomaly detection: classification and evaluation," in *ICNS '07: Proceedings of the Third International Conference on Networking and Services*, (Washington, DC, USA), p. 50, IEEE Computer Society, 2007.

[4] V. Siris and F. Papagalou, "Application of anomaly detection algorithms for detecting syn flooding attacks," *Computer Communications*, vol. 29, no. 9, pp. 1433–1442, 2006.

[5] A. Sheth, C. Doerr, D. Grunwald, R. Han, and D. Sicker, "MOJO: a distributed physical layer anomaly detection system for 802.11 WLANs," in *ACM MobiSys*, 2006.

[6] W. Xu, W. Trappe, Y. Zhang, and T. Wood, "The Feasibility of Launching and Detecting Jamming Attacks in Wireless Networks," in *Proc. of ACM MobiHoc*, May 2005.

[7] M. Thamilarasu, S. Mishra, and R. Sridhar, "A cross-layer approach to detect jamming attacks in wireless ad hoc networks," in *Proc. of Milcom 2006*, (Washington DC, USA), pp. 1–7, October 2006.

[8] A. Cardenas, S. Radosavac, and J. Baras, "Evaluation of detection algorithms for mac layer misbehavior: Theory and experiments," *To appear in IEEE/ACM Transactions on Networking*, 2009.

[9] G. Yan, Z. Xiao, and S. Eidenbenz, "Catching instant messaging worms with change-point detection techniques," in *Proc. of the 1st Usenix Workshop on Large-Scale Exploits and Emergent Threats*, April 2008.

[10] Y. Chen, K. Hwang, and W. Ku, "Distributed change-point detection of ddos attacks: Experimental results on deter testbed," in *Proc. of USENIX Security Symposium*, (Boston, USA), August 2007.

[11] G. Verdier, N. Hilgert, and J. Vila, "Adaptive threshold computation for cusum-type procedures in change detection and isolation problems," *Elsevier, Computational statistics and data analysis*, pp. 4161–4174, 2008.

[12] Y. Chen, K. Hwang, and W.-S. Ku, "Collaborative detection of ddos attacks over multiple network domains.," *IEEE Trans. Parallel Distrib. Syst.*, vol. 18, no. 12, pp. 1649–1662, 2007.

[13] Y. Sheng, K. Tan, G. Chen, D. Kotz, and A. Campbell, "Detecting 802.11 mac layer spoofing using received signal strength.," in *Proc. of INFOCOM'08*, pp. 1768–1776, IEEE, 2008.

[14] Q. Chen and U. Aickelin, "Anomaly detection using the dempster-shafer method," in *Proc. of the 2006 International Conference on Data Mining, DMIN 2006*, pp. 232–240, 2006.