WIRELESS NODE BEHAVIOR BASED INTRUSION DETECTION USING GENETIC ALGORITHM

P C Kishore Raja , Dr.M.Suganthi.M, R.Sunder
Department of Electronics and Communication Engineering
Sri Venkateswara Colleg of Enginneing ,Sriperumbudur India
Thiagarajar College of Engineering ,Madurai India
kisraj@svce.ac.in , rsunder@svce.ac.in, msece@tce.edu

ABSTRACT

Ad-hoc networks are facing increased number of security threats in recent years. Despite numerous technological advancements in wireless network security, it is still very difficult to protect the wireless ad-hoc networks because of lack of centralized traffic concentration, thus requiring monitoring the behavior of individual wireless nodes. This paper presents a behavior-based wireless network intrusion detection using genetic algorithm which assumes misbehavior identification by observing a deviation from normal or expected behavior of wireless node's event sequence. The features set are constructed from MAC layer to profile the normal behavior of wireless node. If any deviations from the normal behavior pattern of wireless node can be used to detect the intrusions in the wireless ad-hoc network. The wireless node behavior is learnt by using genetic algorithm. Current wireless node behavior can be predicted by genetic algorithm based on the past behavior. A 3-tuple value is calculated for constructed feature in a network session. The 3-tuple value of a wireless node behavior in a session are compared with expected non-intrusive behavior 3-tuple value to find intrusions.

Keywords: Wireless Network, Intrusion Detection, Genetic algorithm
.

1 INTRODUCTION

Wireless networking applications continue to proliferate at an incredible pace as wireless features, functions, security and throughput improve. Wireless local area network (WLAN) offer much needed flexibility in a world of high-speed data networks. This flexibility provides an easy way to install a new wireless network, whether an extension of a wired network or pure WLAN. At the same time, vulnerability of wireless networks keeps with technology. In a wireless network, one cannot make the assumption that wireless users are trusted. The first wireless security solution for 802.11-based networks was wired equivalent privacy that received a great deal of coverage due to various technical failures in the protocol. Wireless networks are more vulnerable due to the nature of mobility that does not exist in wired networks. There is a need of security measures for wireless networks. Intrusion prevention such as encryption, authentication is not guaranteed to work all the time. This clearly underscores the need for wireless intrusion detection mechanism. Wireless Intrusion detection mechanism enables the collection of information about intrusion techniques that can be used to strengthen the intrusion prevention facility.

Section II describes vulnerabilities of wireless network. Section III describes related work. Section IV describes feature of interest. Section V describes wireless intrusion detection architecture. Section VI describes experiment results and performance evaluation.

2 VULNERABILITIES OF WIRELESS NETWORK

Wireless LAN exists in either infrastructure network or ad-hoc network. In infrastructure network, wireless nodes associate themselves with an access point, which is connected to wire line network that solves centralized network management function. In case of ad-hoc networks, network does not have a centralized network management function. All leads to increase in vulnerability that ranges from passive eavesdropping to active interfering. So wireless nodes are autonomous units that are capable of

roaming independently. Nodes are having with inadequate physical protection. It has been easily captured, compromised and hijacked. Even many of the current MAC protocols for wireless channel access are vulnerable. Wireless network is vulnerable due to its features of open medium, dynamic changing network topology, lack of centralized monitoring and management point, and lack of a clear line of defense. The vulnerabilities of wireless network are MAC Address Spoofing, Default Configuration, Service Coverage Area and Public WLAN Concerns. The internal attacks are Distributed false route request and Denial of service Impersonation.

.

3   RELATED WORK

Most of current works on IDS for wireless networks employ either distributed and cooperative architecture or distributed and hierarchical architecture. Zhang and Lee proposed the first distributed and cooperative anomaly based IDS framework. In this framework, local anomaly detection engine [8] is built on a rule based classification algorithm RIPPER and local response is activated when a node locally detects a anomaly or intrusion with high confidence. When a node detects an anomaly or intrusion with weak confidence, if then initiates a global intrusion detection procedure through a cooperative detection engine. Huang and Lee extended their previous work on local anomaly detection and developed a cross feature analysis technique to explore the correlations between features using classification decision tree induction algorithm C4.5. Their detection engine uses features extracted from routing table and also they incorporated statistical features. However, system is unable to localize the attack.
Tseng at al. developed distributed IDS using specification based detection techniques to detect on attacks on AODV routing protocol. Generally specification based detection a technique of any kind has to balance trade off between model complexity and accuracy.

4   FEATURE OF INTEREST

In wireless networks, MAC layer manages and maintains communication between mobile nodes by coordinating access to a shared radio channel and utilizing protocols that enhance communications over a wireless medium. The proactive mechanisms are employed in wireless networks before any data communication. These mechanisms cannot give prefect prevention. This work concentrates on reactive mechanism, which detects intrusion or anomaly behavior in wireless networks. This works assumes that wireless networks use both CSMA/CD

and CSMA/CA. To characterize wireless node behavior in wireless network, we extract feature set from MAC layer.

| FEATURE	UNIT	RANGE
NAV	SECOND	[0,1] [1,3] [3,5] [5,∞]
TRANSMIT-TRAFFIC-RATE	BYTE	[0,102.4k] [102.4k,204.8k] [204.8k,3027.2k] [3027.2k,∞]
RECEIVE-TRAFFIC-RATE	BYTE	[0,102.4k] [102.4k,204.8k] [204.8k,3027.2k] [3027.2k,∞]
RETRANSMIT RTS	COUNT	[0,3] [3,5] [5,7] [7,∞]
RETRANSMIT DATA	COUNT	[0,7] [7,∞]
NEIGHBOR-NODE-COUNT	COUNT	[0,7] [7,15] [15,23] [23,29]
FORWARD -NODE-COUNT	COUNT	[0,0] [1,1] [2,2] [3,3] [3,29]

Table I: Wireless Feature Set

Network Allocation Vector (NAV) is a counter resides at each node that represents the amount of time that channel will be occupied by the current sending node. RTS/CTS/DATA/ACK are four features defined over these four types of packets. The transmission and reception traffic rate (transmit-traffic-rate and receive-traffic-rate) are busy/idle indicators. The retransmission rates of RTS (Retransmit RTS) packets and DATA (Retransmit DATA) packets are channel congestion indicators. Active Neighbor node count indicates the number of active neighbor nodes of the monitoring node.The remaining extracted features from network layer are tabulated in TABLE I

5   WIRELESS INTRUSION DETECTION ARCHITECTURE

The goal of intrusion detection is seemingly simple: to detect intrusions and also to identify unauthorized use, misuse and abuse of wireless nodes by both internal attackers and external penetrations. In other words, Intrusion detection is a process of identifying and responding to malicious activity targeted at computing and network resources.

A network intrusion is a sequence of activities by a malicious individual that results in unauthorized security threats to a target network. Generally, Intrusion detection is classified as 1.Profile based intrusion detection 2.Signature based detection. The designing an IDS in wireless networks is tougher challenge due to vulnerabilities and lack of physical infrastructure. Without centralized audit point such as routers and gateways, an IDS for wireless networks is limited to using only the current traffic coming in and out of the node an audit data

This paper describes wireless intrusion detection architecture to monitor and detect the malicious activity of wireless node. The entire architecture consists of wireless traffic capturing module, preprocessor module, detector module, knowledge base training module and decision module. The first step is to collect the wireless feature set using NS2 in a session. This feature set is fed into preprocessing module. In the preprocessing module, wireless feature set is encoded into alphabets. The detection module has two jobs. First job is to learn the encoded wireless feature set and second job is to detect the intrusions from learned wireless feature set. The detection module uses genetic algorithm to perform training and detection.
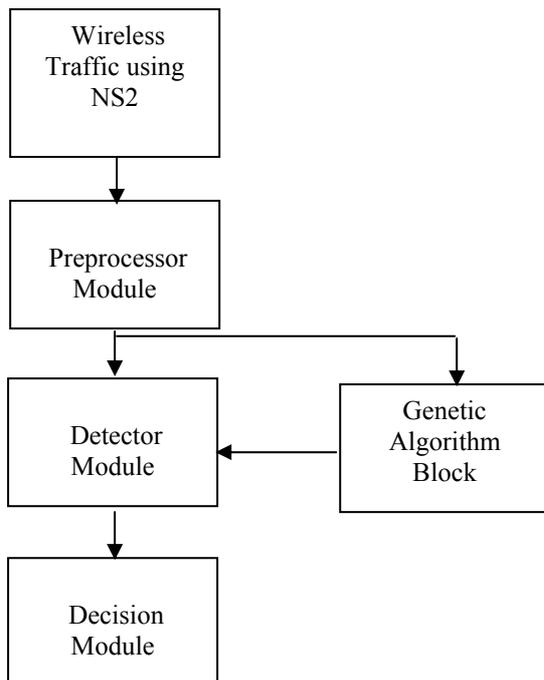
Figure 1: Wireless Intrusion Detection Blocks

A wireless node behavior exhibits some regularity in wireless node's event sequence in wireless traffic. Any deviation from wireless normal node behavior treated as signing of intrusion. In this work, genetic algorithm is used to learn the wireless node behavior in a computing environment due to its robustness and adaptability to changes in the environment. The

genetic algorithm involves two steps. The first step involves coding the vectors with a string of bits, which form the input population of a genetic algorithm. The second step is finding a fitness function to test each individual of population against some evaluation criteria. In the learning process each event sequence of wireless node behavior forms a gene. Fitness is calculated for a collection of genes. If genes with required fitness cannot be found in the current generation, new set of genes are evolved through crossover and mutation. The process of evolution continued until genes with required fitness are found. The detection process involves defining vectors for event data and testing whether the vector indicates an intrusion or not.

5.1 Preprocessor Module

This section presents steps involved in mapping wireless feature set into genes to learn regularities in wireless node behavior profile. In preprocessing, each wireless node feature set forms a gene. Alphabets are assigned for each wireless feature set in a session. Wireless node feature set in a session is divided into string of alphabets of size 'n' called behavior gene. We have used 27 alphabets to represent MAC layer feature set of wireless node in evolution of behavior.

In order to the wireless intrusion detection system to recognize intrusive behavior, it must first learn encoded feature set extracted from MAC layer of wireless node and wireless node behavior profile is formed in each wireless node in order to describe its normal behavior. Detector involves a process of establishing profiles of normal wireless node behavior and past behavior with current one. Detection depends on an assumption that extracted features set of MAC layer of wireless node exhibit predicable, consistent patterns of usage. The approach also accommodates adaptations to changes in wireless node behavior profile over time. In this approach, fitness function required for reproduction is based on the observation that extracted features set of MAC layer of wireless node can be best captured by observing the trend in total behavior entropy. This total behavior entropy gives a measure of amount of randomness in the wireless node behavior profile. It gives the frequency of entries in wireless node behavior profile. Frequent change in the total behavior results in large entropy value and the entropy value remains approximately the same for normal behavior.

5.2 Fitness Function

Determination of appropriate fitness function to measure the fitness of behavior gene is important to improve the accuracy of the prediction. The fitness function a gene is given by

Fitness = 1- | σ_x - ϕ |

where σ x is wireless node behavior entropy of predicated gene and ϕ is the average wireless node behavior entropy of 'm' previous behavior genes. Entropy is defined as

Entropy = Σ I – (P (i) * log (P (i)) / log (n))

Where p(i) is fraction of number of time alphabet I occurred to size of behavior gene and n is number of alphabets in the behavior genes

5.3 Parameters for Wireless Node Behavior Characterization

The input of the genetic algorithm consists of 'm' behavior genes. A behavior gene is a set of n encoded feature set extracted from MAC layer of wireless node in the session .The three genetic operators are reproduction, crossover and mutation that are applied on the input population and output is a gene that describes the normal behavior of wireless node profile .The current gene is then taken and a 3-tuple value match index, entropy index and newness index is calculated from the actually occurred behavior and the predicated behavior. The 3 tuple value[10] is described as follows

5.3.1 Match Index

The match index is a measure of regularity in the wireless node behavior. It is given by

Match Index = Count of encoded feature predicated correctly in a wireless node feature set / Size of the feature set

5.3.2 Entropy Index

The entropy is a measure of wireless node behavior dynamics in the wireless feature set profile. It is given by

Entropy Index = Σ I – (P (i) * log (P (i)) / log (N))

Where p(i) give probability of occurrence of wireless node feature 'i' in the wireless node feature set.
N gives the number of unique wireless node feature in the wireless node feature set

5.3.3 Newness Index

The newness index is a measure of the number of new wireless node feature of wireless node, which have not occurred earlier in the feature set. It is given by

Newness index = 1- Number of new wireless node feature of wireless node as well as in feature set / length of wireless node feature set.

This 3 tuple value calculated for the current feature set extracted from network layer and MAC layer of wireless node is compared with the threshold values to determine whether feature set extracted from network layer and MAC layer of wireless node is intrusion or not. Having a single threshold instead of having three-threshold values .The three-index value are combined to form and checked against the single threshold performs the detection. New index is extracted from three index called New mismatch index. This is because of the direction of inequality operators. If the match index of feature set extracted from network layer and MAC layer of wireless node is less than threshold of the match index , sample is considered intrusive. Also if the entropy index or newness index of current sample is greater than the threshold of corresponding index, sample is declared intrusive.

Mismatch index = 1 – match index.

Threshold = $\alpha1$ * MMI + $\alpha2$* EI + $\alpha3$* NI.

The weights $\alpha1$, $\alpha2$, $\alpha3$, need to be chosen for the finding the threshold. These weights are fixed by observing the values of the three indices that determine whether total behavior gene is intrusive or not.

6 SIMULATION ENVIRONMENT AND FEATURE SET CONSTRUCTION

The simulation is conducted on the platform of Network Simulator (ns-2) [21]. Table II lists the ns-2 parameters in our simulation. In the simulation, each node starts its move from a random location to a random destination with a randomly selected speed that uniformed distributed between [0, maxspeed]. Once the destination is reached, the node stays there for as long as specified by pause time. then another destination location is chosen. Dynamic network topology and different mobility scenarios are modeled by varying the maxspeed and the pause time. To prevent all flows start from the beginning at the same time , each source node chooses its starting time for sending packets from the range of [0,stime].

Parameter	Value /Choice
Topology	500m X 500m
Node Movement	Random waypoint model
Max movement speed	10 m/s
Radio range	250m
Node set count	30

Total number of flows	25
Average transmission rate per flow	2 packets /s , 512 b /packet
Training Execution Time	2000s
Testing Execution Time	200s
Feature sampling interval	5s

Table II : Ns-2 simulation environment

We use the feature set described in Table I to construct feature vectors. These feature set is encoded and fed into intrusion detection module. The performance of wireless intrusion detection was tested using 30 wireless nodes. In experimental study, effect of wireless node feature set in the accuracy of predication, selection of length of initial observation period to learn wireless node behavior, evaluation of performance using false alarm rate and accuracy of intrusion detection are studied. Accuracy of behavior gene prediction decides the value of 3-tuple which in turn affects the accuracy of intrusion detection.

7 OPERATION OF WIRELESS NODE BEHAVIOR BASED IDS

The Wireless node feature set is extracted from simulated wireless networks. It is encoded and used as input stream to genetic algorithm based intrusion detection module. Learning module learns the trend in entropy value across wireless node feature set. The 3-tuple value < match index, entropy index, newness index> is calculated in a session . Expected normal behavior 3-tuple value is compared with calculated 3-tuple value to calculate the deviation in wireless node behavior. Intrusion detection module computes probability of current wireless node feature set being intrusive from deviation in wireless node behavior.

7.1 Performance of Wireless Intrusion Detector

The performance of wireless intrusion detector is evaluated through two parameters

7.1.1 Accuracy of Intrusion

In this approach, intrusion is a set of actions which are deviating from the normal wireless node behavior. Probability of a feature set being intrusive is same as accuracy of detection.

Accuracy = [1 – n / N] * 100

n: count of feature value that are in total feature set .
N: Initial size of total feature set

7.1.2 False Alarm Rate

False alarm rate is a measure of count of instances in which a genuine wireless node is classified as an intruder. False alarm rate = [n / N] * 100

n: count of feature value that are in total feature set .
N: Initial size of total feature set

## 8 EXPERIMENT RESULTS AND DISCUSSION

The 30 wireless node feature set is collected. Intrusion is artificially created by appending random encoded alphabets in a wireless feature set. Initial values are assigned to 3 tuple parameter. The probability of current wireless feature set was computed. The following figure shows that entropy index, network index , newness index , sequence index with intrusive samples.
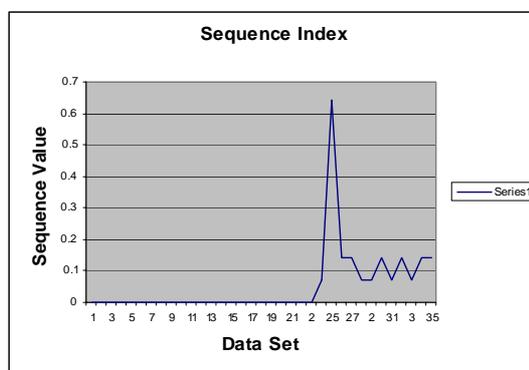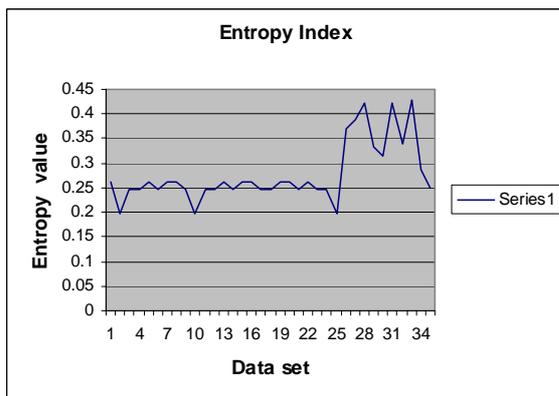

Figure 1: Sequence Index with Intrusive Sample
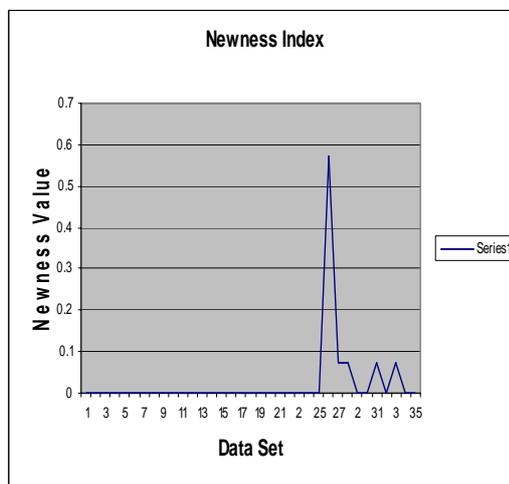
Figure 2: Newness Index with Intrusive Sample
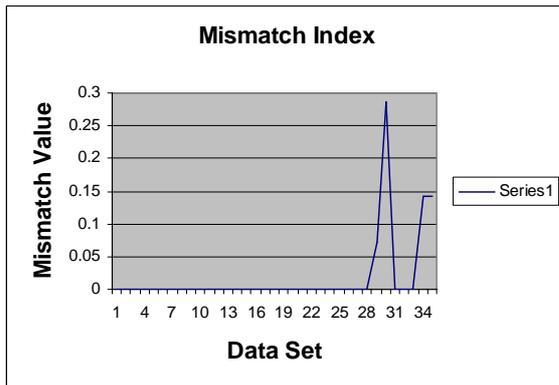
Figure 3: Mismatch Index with Intrusive Sample

Figure 4: Entropy Index with Intrusive Sample

9 CONCLUSION

We described a novel idea of wireless intrusion detection architecture. As we know that wireless security vulnerabilities will keep pace with the technology. Since the traditional perimeter defenses are inadequate for wireless network. The proposed work offers new kind of defense against intrusion.

10 REFERENCES

[1] J. M. Davis: Simplified Diaphragm Analysis, Journal of Structural Div., ASCE, Vol. 103, pp. 2098-2109 (1977).

[2] G.Y.Zhang, W.Lee and Y.A Huang, "Intrusion Detection Techniques for Mobile Wireless Networks", ACM J.Wireless Networks, vol 9, no.5, September 2003 pp.545-56

[3] G.Y.Zhang, W.Lee," Intrusion Detection in Wireless Ad-hoc Networks ", 6th international conference on Mobile computing and Networking, August 2000 pp.275-83

[4] S. Forrest S. A. Hofmeyr, A. Somayaji, and T. A. Longstaf. "A sense of self for Unix processes", In Proceedings of the 1996 IEEE Symposium on Security and Privacy, pages 120-128, Los Alamitos, CA, 1996. IEEE Computer Society Press

[5] IEEE Std. 802.11, "Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications," 1997

[6] Air Defense Inc, " Wireless LAN Security for the Enterprise," Air Defense, [Website] Available HTTP: http://www.airdefense.net/

[7] T.S Rappaport. Wireless Communication: Principles and Practice, Prentice Hall, 2002

[8] Dorothy E. Denning, "An intrusion-detection model. IEEE Transactions on Software Engineering", Vol. SE-13 (No. 2): 222–232, February 1987

[9] Yu Liuy, Yang Liy, Hong Many, "MAC Layer Anomaly Detection in Ad Hoc Networks", 6th IEEE Information Assurance Workshop, USA, 15-17 June 2005

[10] S. Balachandran, D. Dasgupta, L. Wang, "A Hybrid Approach for Misbehavior Detection in Wireless Ad-Hoc Networks," in Proceedings of the Symposium on Information Assurance, June 2006

[11] S. V. Raghavan and B. Balajinath, "Intrusion Detection Through Learning Behavior Model", Appeared in the International Journal Of Computer Communications, Vol 24, No 12, July 2001, pp. 1202-1212