Behaviour Analysis of Node in Wireless Ad hoc Network using Genetic Algorithm

P.C. Kishore Raja[1], Suganthi.M[2]

Abstract

The security threats for wireless ad hoc networks have been increased in two-fold way in recent years. It keeps pace with technology. One of the techniques to thwart the security threat is behavior based wireless intrusion detection. This paper deals about behavior based wireless intrusion detection and it defines a feature set that characterizes wireless node behavior. The feature set is constructed from lower layers of WLAN to profile the normal behavior of wireless node. Observing a deviation from normal or expected behavior of wireless node identifies the misbehavior. The wireless node behavior is learned by genetic algorithm and current wireless node behavior can be predicted by genetic algorithm based on the past behavior. A 3-tuple value i.e. entropy index, newness index, mismatch index is calculated for constructed feature set in a session. The 3-tuple values of a wireless node behavior in a session are compared with expected non-intrusive behavior 3-tuple value to find intrusions. The performance of wireless intrusion detection is evaluated using detection probability and false alarm probability.

Keywords: Wireless Network, Intrusion Detection, Genetic Algorithm

[1] Assistant Professor, Department of Computer Science Engineering, Sri Venkateswara College of Engineering, Pennalur, Sriperumbudur, Tamilnadu, India – 602 105 Tel:91-44-27162321, email: kisraj@svce.ac.in, pckishoreraja@gmail.com.

[2] Professor, Department of Electronics and Communication Engineering, Thiagarajar College of Engineering, Madurai, India - 625 015. email: msece@tce.edu

1. Introduction

Wireless networking applications continue to proliferate at an incredible pace as wireless features, functions, security and throughput improve. At the same time, vulnerability of wireless networks keeps with technology. In a wireless network, one cannot make the assumption that wireless nodes are trusted. 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. There is a need of security measures for wireless networks. One of security measure is intrusion detection.

In this paper, we propose to use a behavior based intrusion detection technique using genetic algorithm to detect intrusions on wireless ad hoc networks. It is in contrast to signature based intrusion detection techniques, which may be impractical for ad hoc networks due to difficulties of specifying, distributing and updating signatures of attacks. Another challenge to behavior based intrusion detection in ad hoc networks is resource constraints. Our approach is to specify a reduced feature set of MAC layer to profile normal wireless node behaviors.

2. Related Work

Most of current work [Liu et al., 2005] on IDS for wireless networks employs either distributed and ical

architecture. Zhang and Lee [Zhang 2000] proposed the first distributed and cooperative anomaly based IDS framework. In this framework, local anomaly detection engine is built on a rule based classification algorithm RIPPER and local response is activated when a node locally detects an 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 [Huang 2003] 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. [Tseng 2003] 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.

3. 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. We extract feature set from MAC layer [Liu et al., 2005] to characterize wireless node behavior in wireless ad hoc network. Table I describes extracted wireless feature set from MAC.

Table 1 : Shows Statistical Wireless Feature Set and Its Values

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	[0,3] [3,5] [5,7] [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]
2 BIT STATE (Unauthenticated & Associated)	COUN T	[0] [1] [2] [3]
TIME OUT FOR RTS	COUNT	[0] [1] [2] [3]
SIGNAL TO NOISE RATIO	COUNT	[0] [1] [2] [3]

4. 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. 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 and knowledge base training module.

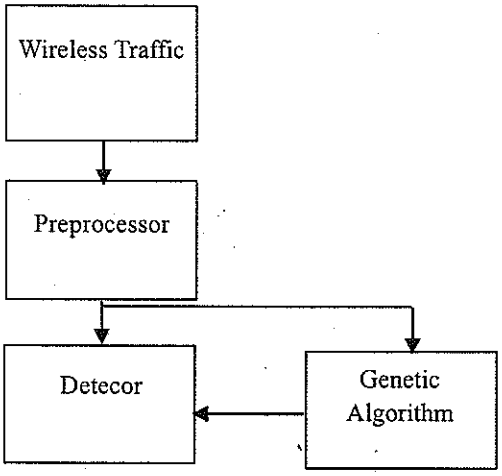

Figure 1 : Shows Wireless Intrusion Detection Modules

The simulation is conducted on the platform of Network Simulator (ns-2) [Liu et al., 2005] to collect the wireless traffic in wireless traffic capturing module. 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].

Table 2 : Ns-2 Simulation Environment

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

The wireless feature set is extracted from wireless traffic. It is fed into preprocessing module. In the preprocessing module, wireless feature set is encoded into alphanumeric value, which forms a gene to learn regularities of wireless node. Wireless node feature set in a session is divided into string of alphabets of size 'n' called behavior gene. We have used 3 alphanumeric values for normal behavior and 4 & above for abnormal behavior [Raghavan 2001] to represent MAC layer feature set. The alphanumeric values forms input population to genetic algorithm. 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. In the genetic algorithm, Fitness is calculated for each individual of genes. 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

$$\text{Fitness} = 1 - | \sigma x - \phi | \quad (1)$$

where σx is wireless node behavior entropy of predicated gene and ϕ is the average wireless node behavior entropy of 'm' previous behavior 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. 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.

Entropy is defined as

$$\text{Entropy} = \Sigma I - (P(i) * \log(P(i)) / \log(n)) \quad (2)$$

Where p(i) is fraction of number of time alphanumeric value 'I' occurred to size of behavior gene and n is number of alphabets in the behavior genes. Three indices have been used for detecting the intrusion. They are 1. Match Index 2. Newness Index 3. Entropy Index. The match index, entropy index and newness index is used to characterize the wireless node behavior. These 3-tuple values determine the threshold of current behavior gene. The 3-tuple values is described as follows

Match Index

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

Match Index = Number of each alphanumeric value predicated correctly in a wireless node feature set / Size of the wireless node feature sample

Entropy Index

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

$$\text{Entropy Index} = \Sigma I - (P(i) * \log(P(i)) / \log(N)) \quad (3)$$

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

Newness Index

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

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

New index is created from match index called mismatch index. This is because of the direction of inequality operators. These 3-tuple value is calculated for the current behavior gene and then it is compared with the past behavior gene in order to detect the intrusions.

Mismatch index = 1 – match index.

Threshold = $\alpha 1 * MMI + \alpha 2 * EI + \alpha 3 * NI$

The weights a1, a2, a3 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.

5. Experiment Results And Discussion

The 30 wireless node feature set is collected. Each wireless feature set has 10 values to form one behavior gene. The genetic algorithm for finding the normal behavior of a wireless node receives the initial 500 wireless feature set as initial population. The wireless feature set that occurs after the first 500 are used for testing. Three intrusive samples are appended in beginning, middle and end to normal behavior sample set. For the genetic algorithm, the cross probability is set as 0.6, the mutation probability is set as 0.001 and the number of generations is set as 5. These values have been obtained after experimental analysis. Different values for these parameters have been tried on known no intrusive behavioral set of samples. As can be seen from the Table 2 the set with the least probability of false alarm rate from amongst these has been used here. The weights a1, a2, a3, 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. The following figure shows that entropy index, network index, newness index, sequence index with intrusive samples. From figures, three feature set sample have crossed threshold value and treated as intrusive sample. It can be seen that mismatch index and newness index are significantly higher for intrusive behavior. The intrusive behavior does not evoke any change in the entropy index.

Table 2 : Shows Least Probability of False Alarm

Cross over Probability	Mutation Probability	No. of Generations Probability	False alarm
0.1	0.03	10	0.1
0.2	0.03	5	0.1
0.3	0.001	10	0.12
0.4	0.005	10	0.08
0.6	0.0001	5	0.10
0.6	0.001	5	0.04
0.7	0.001	5	0.06

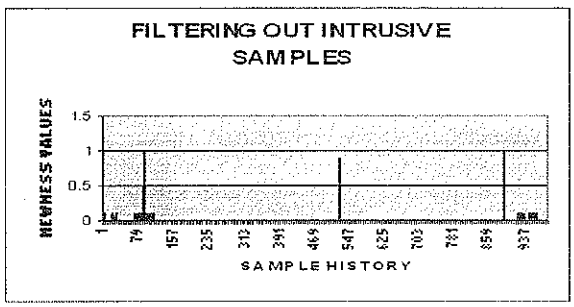

Figure 2 : Shows Filtering Out Intrusive Samples Using Newness Index

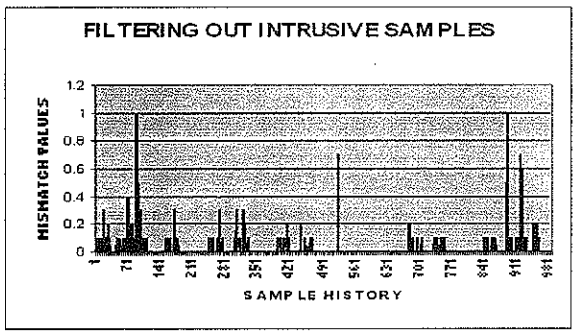

Figure 3 : Shows Filtering Out Intrusive Samples Using Mismatch Index

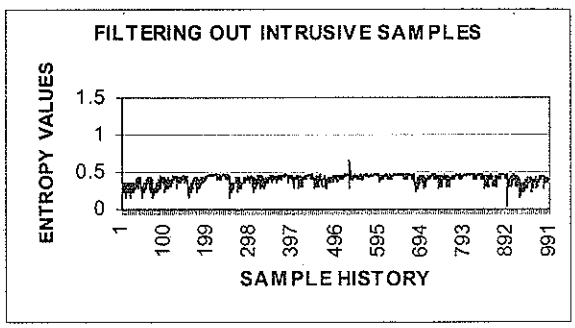

Figure 4 : Filtering Out Intrusive Samples Using Entropy Index

6. Performance of Wireless Intrusion Detector

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. The performance of wireless intrusion detector is evaluated through two parameters

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.

$$\text{Accuracy} = [\,1 - n / N\,] * 100 \qquad (5)$$

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

False Alarm Rate

False alarm rate is a measure of count of instances in which a genuine wireless node is classified as an intruder.

$$\text{False alarm rate} = [\,n / N\,] * 100 \qquad (6)$$

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

The performance of wireless intrusion detection is specified in terms of detection probabilities and false alarm probability. The figure 5 shows that detection probability against false alarm probability for different threshold values

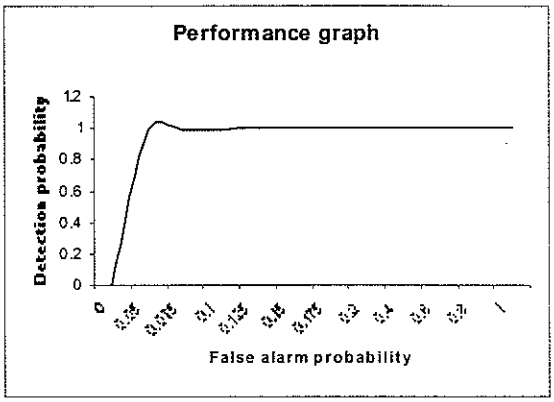

Figure 5 : Shows Performance Evaluation Of Wireless Intrusion Detection

7. 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.

References

[1] G.Y.Zhang, W.Lee and Y.A Huang, *" Intrusion Detection Techniques for Mobile Wireless Networks"*, ACM Journal. Wireless Networks, Vol 9, No.5, PP. 545-56, 2003.

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

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

[4] 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, 2006.

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

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

[7] Y. Huang and W. Lee, *"A Cooperative Intrusion Detection System for Ad Hoc Networks"*, in Proceedings of the 1st ACM workshop on Security of ad hoc and sensor networks, PP. 135-147, October 2003.

[8] C. Tseng, P. Balasubramanyam, C. Ko, R. Limprasittiporn, J. Rowe and K. Levitt, *"A Specification-based Intrusion Detection System for AODV"*, in Proceedings of the 1st ACM workshop on Security of ad hoc and sensor networks, PP. 125-134, October 2003.

Author's Biography

P C Kishore Raja is working as an assistant professor at Sri Venkateswara College of Engineering, Sriperumbudur. He received a B.E in Electronics and Communication Engineering from Kongu Engineering College, Bharatiyar University and M.E in Communication Systems from Thiagarajar College of Engineering. His research interests are Wireless Network Security, Intrusion Detection Systems and low power VLSI design.

Dr. Suganthi .M is working as a professor at Thiagarajar College of Engineering. She received her Ph.D. from Madurai Kamaraj University . Her research interests include Mobile Ad-hoc Networking, VLSI, and Security.