Revealing of Reducing Manners in Ad Hoc Networks with Crosslayer Approach Using SVM and FDA in Distributed Architecture

S. Jasmin Salma and B.Aysha Banu

Department of Computer Science and Engineering,
Mohamed Sathak Engineering College, Kilakarai, Tamil Nadu, India

jasminsalma786@gmail.com

***Abstract* - Ad hoc network is a structure less network with independent nodes. In the ad hoc network, the nodes have to cooperate for services like routing and data forwarding. The routing attacks in ad hoc networks have given rise to the need for designing novel intrusion detection algorithms, different from those present in conventional networks. In this work, distributed intrusion detection system (IDS) have proposed for detecting malicious sinking behavior in ad hoc network. Detection process of that sinking behavior node is very important to do the further forwarding process in network. Intrusion detection system use linear classifiers for training the intrusion detection model. Cross -layer approach is involved to increase the accuracy of intrusion detection process in ad hoc network. A machine learning algorithm in non linear manner named as Support Vector Machine (SVM) involved for training the detection system and used together with Fisher Discriminant Analysis (FDA). The proposed cross-layer approach aided by a combination of SVM and FDA reduces the feature set of MAC layer without reducing information content.**

Keywords - Cross-layer designs, routing attacks, Ad hoc networks, intrusion detection, sinking

I. Introduction

Wireless ad-hoc networks are vulnerable to various kinds of security threats and attacks due to relative ease of access to wireless medium and lack of a centralized infrastructure. Thus, the functioning of the network highly depends on the cooperativeness of nodes in the network. This unprecedented cooperative nature of the routing and data forwarding mechanism has spawned a unique security vulnerability called routing attacks. Thus, the conventional IDS architecture and algorithms have to be redesigned to suit the ad hoc network technology. Therefore, the goal of the IDS is to distinguish packet dropping induced by network conditions and from those caused by malicious sinking.

Sinking is a malicious behavior of nodes, where nodes do not cooperate in the routing and forwarding operations of the network. In this work, the distributed IDS is proposed for detecting sinking behavior in an ad hoc network. The proposed detection system uses a cross-layer approach to maximize detection accuracy. Node routing behavior is defined by collecting statistics from the protocol communication at network, MAC, and physical layers. To further maximize the detection accuracy, a nonlinear machine learning algorithm, namely Support Vector Machines is used for training the detection model. The proposed IDS preprocess the training data for reducing the computational overhead incurred by SVM. Number of features in the training data is reduced using predefined association functions. Also, the proposed IDS use a linear classification algorithm, namely Fischer Discriminant Analysis to remove data with low-information content.

A. Objective

The scope of the proposed IDS is to distinguish packet dropping induced by network conditions and from those caused by malicious sinking. In this work, the distributed IDS is proposed for detecting sinking behavior in an ad hoc network.

B. Routing Attacks and Networks Vulnerabilities

The threat model for the proposed IDS consists of three entities and their characteristics. The entities are the network, the attack, and the attacker. The characteristics of these entities define the threat model of the proposed IDS. Characteristics of the network include factors that help to camouflage the malicious behavior. These characteristics of the network cause nodes to benignly drop packets. This kind of dropping behavior due to the network conditions resembles the behavior of malicious sinking.

Possible factors which can induce benign dropping include the following:

- mobility of nodes,
- network/traffic density,
- traffic type, and
- channel and fading conditions.

II. Literature Review

A. Existing System

An autonomous host-based intrusion detection system existed for detecting malicious sinking behavior. Various experiments are conducted with varying network conditions and malicious node behavior. The effects of factors such as mobility, traffic density, and the packet drop ratios of the malicious nodes are analyzed. Experiments based on simulation show that

the proposed cross-layer approach aided by a combination of SVM and FDA performs significantly better than other existing approaches. If any attacking behavior node involved in corresponding network then the information will pass to the source node by cross layer detection and also the further analysis will handle by SVM and FDA.

B. Limitation of the Existing System

Here the detection process will predict only by the source node of the ad hoc network. Rest of the nodes not takes to the consideration of alert of sinking behavior in network. This intrusion detection and results from these simulations are practically unreliable when IDS contained system becomes failed.

C. Proposed System

The proposed detection system uses a cross-layer approach to maximize detection accuracy. Node routing behavior is defined by collecting statistics from the protocol communication at network, MAC, and physical layers.

A machine learning algorithm in non linear manner named as Support Vector Machine (SVM) involved for training the detection system and used together with Fisher Discriminant Analysis (FDA).The proposed cross-layer approach aided by a combination of SVM and FDA reduces the feature set of MAC layer without reducing information content.

In this, detection methodology will be distributed to the entire participant in corresponding network hence all the nodes in network will be known about the attacking behavior node rather than source node to improve the security of the network. It's efficient when all the nodes know about the warning and aware message. so if the algorithm presents as distributed system then all the nodes will involved for further process after getting aware message.

III. Methodology

A. Architecture of the Proposed System

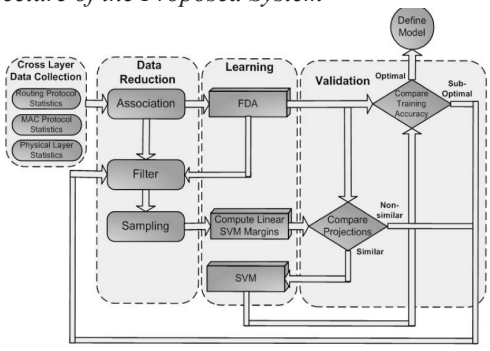

Fig.1 Architecture of the proposed system

The principal components of the training process are data collection module, data reduction module, and learning and validation modules. Data reduction is necessary to reduce the processing overhead incurred by machine learning methods, especially SVM algorithm. Data reduction module decreases the number of features and events in the training data. Validation is used to check the optimality of defined classification model. Since an event is mathematically represented as vector, hereinafter, event and vector will be used interchangeably.

B. Modules

- Cross Layer Data Collection
- Data Reduction
- Learning
- Justification,
- Detection

1) Cross Layer Data Collection:

In this, module depend upon two modules are physical layer and Mac layer. Physical layer and Mac layer contains events, network traffic and information of the packet to be transfer. The data collection module collects data from network, MAC, and physical layers. At each layer, the collection module monitors the events and computes time, traffic, and topology statistics and records the feature values. Any rare occurrences of out-of-bound events or outliers are removed from the collected data.

2) Data Reduction Techniques:

Three techniques are used for data reduction in this proposed IDS, namely association, filtering, and sampling.

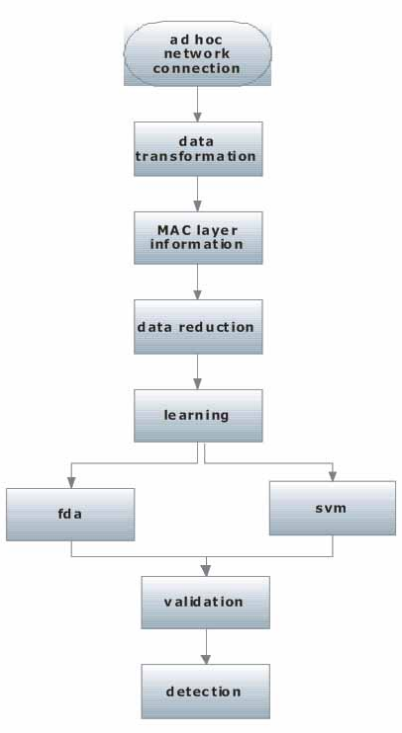

Fig. 2 Process flow of IDS

Association

Association reduces the feature set so that the overhead of learning is minimized. One or more features from different layers to a specific MAC layer feature. The features are classified based on dependency on time, traffic, and topology.

Filtering

In this, module is to remove the uninformative and redundant cases. Events which are inessential in defining or classifying a routing behavior are referred to as uninformative events. Removal of uninformative events consists of two stages. The linear decision boundary in the FDA's projection space is used.

Sampling

Sample is the process of selecting a subset of the original training data. Even after association and filtering, the volume of training data remains large enough to impose a huge computational overhead for nonlinear machine learning techniques.

3) Learning

Nonlinear SVM model is trained by reduced training data set, data reduction and validation of training data set adequacy is essential. Machine learning process consider set of threshold parameters to classify an unknown behavior. So consider a simple classification method, referred to as "linear machine". The vectors falling in the margin areas of the feature space are referred to as the support vectors. These support vectors govern the nonlinear boundaries of SVM classification model. Fisher discriminant analysis (FDA) is a variant of linear machine, which uses the ratio of between-class and within-class variance as a measure.

4) Justification

The prime purpose of validation is to check the adequacy of the sampled training data in SVM. SVM and FDA is compared to evaluate the adequacy of the SVM's training data and the SVM classification model. Fisher's linear discriminate are methods used in statistics, pattern recognition and machine learning to find a linear combination of features which characterize or separate two or more classes of objects or events. The resulting combination may be used as a linear classifier or, more commonly, for dimensionality reduction before later classification. FDA works when the measurements made on independent variables for each observation are continuous quantities. When dealing with categorical independent variables, the equivalent technique is discriminate correspondence analysis.

5) Detection

Intrusion detection system is distributed detection system. Detection of an attack is also performed locally by the nodes. This is commonly referred to as the kernel function. We address channel estimation, interference correlation estimation, and data detection for MIMO systems under both spatially and temporally colored interference.

IV. Performance Analysis

For brevity, only three factors, mobility, traffic density and drop ratio, are considered. The other three factors which are not considered in the analysis are traffic type, channel and fading conditions and attack duration. These omitted factors are strongly related to factors that are considered for analysis. Their effects highly resemble those of the three factors chosen. Both traffic type and attack duration are highly related to traffic density. This is because changing traffic types will actually cause the traffic density to change. Attack duration, though controlled by the attacker, also depends on the duration of inflow traffic density toward the sinking node. Therefore, similar to traffic type, attack duration has strong correlation with traffic density. Also, attack duration and selective drop ratio have similar characteristics.

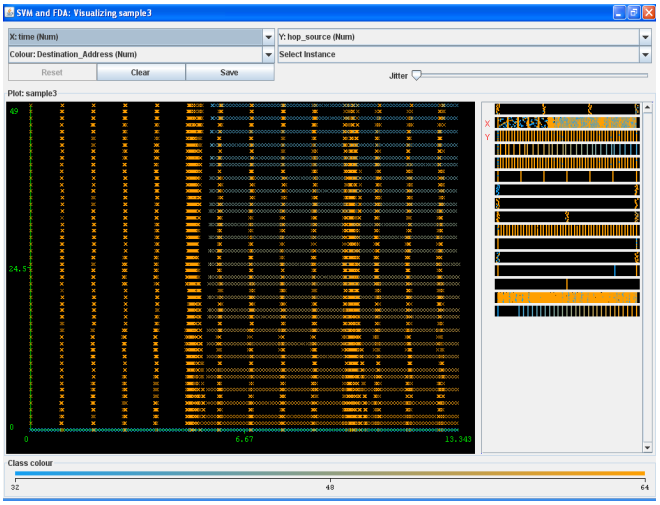

Fig.3 Detection accuracy with changing drop ratio.

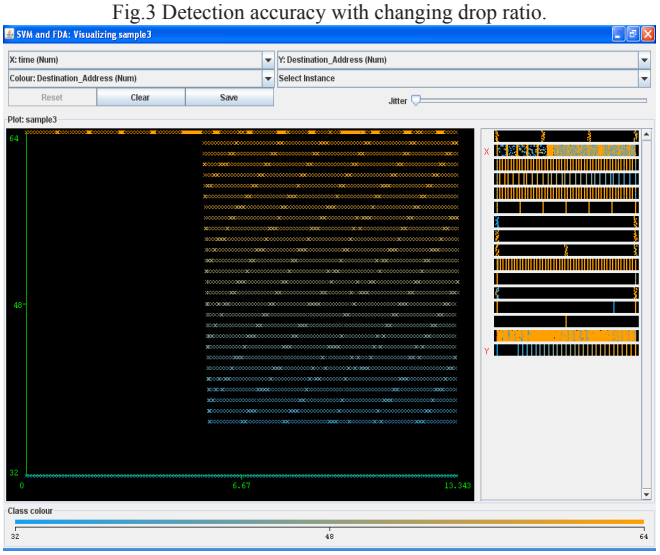

Fig.4 Detection accuracy with changing mobility

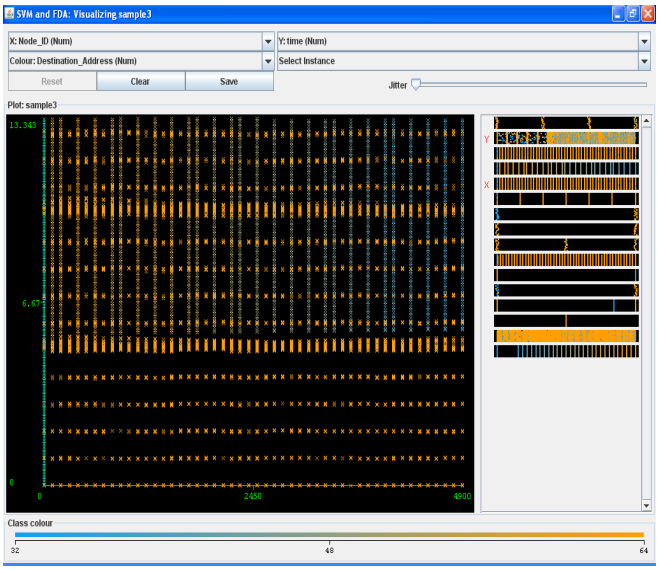

Fig.5 Detection accuracy with changing traffic density

A. Effect of Drop Selectivity Ratio

Selectivity of packet dropping affects the efficiency of single-layer methods more than cross-layer methods. Cross-layer methods experience a small drop in the efficiency, as the sinker drops less percentage of packets.

B. Effect of Mobility

The node mobility is varied from no mobility to vehicular mobility. It can be observed in Fig. 4 that in single-layer methods, detection efficiency drops as mobility increases. However, the detection efficiency of cross-layer methods is not affected by mobility.

C. Effect of Traffic Density

Traffic density can be observed in Fig. 5 that single layer lies poor. Though the increase in traffic density actually creates higher background traffic, the incoming traffic toward the sinker also increases. Therefore, the number of packets dropped by the sink also increases and thereby increasing the detection rate.

V. Conclusion and Future Enhancement

In ad hoc networks, nonlinear machine learning techniques such as SVM were infeasible due to the computational complexity induced by the large size of training data. Furthermore, as the cross-layer approach forms a bigger feature set, the high complexity incurred has made combining cross-layer schemes and nonlinear machine learning techniques infeasible.

In this work, the number of features and the training data size are reduced by the process of association and filtering, respectively. To further reduce the overhead, only a sample of the original associated and filtered training data is used for training. A linear machine learning method, FDA, is used to check whether the chosen training data are always optimal. Experiments conducted by varying different network conditions and sink behavior were analyzed. Effects of network conditions such as mobility, traffic density have little or no effect on the proposed IDS. Similarly, sink behavior of selective dropping has no effect on the proposed IDS. This methodology was analyzed only for sinking behavior.

To extend this work, in future, co operative architecture for the proposed detection methodology will be considered. In this paper, the feature set is defined for Optimized Link State Routing (OLSR) routing protocol and MAC802.11b protocol. Analyzing the efficiency of the proposed IDS on other protocols is laid out for future work.

References

[1] P. Brutch and C. Ko, "Challenges in Intrusion Detection for Wireless Ad- Hoc Networks," *Proc. 2003 Symp. Applications and the Internet Workshops*, 2003.

[2] C.J.C. Burges, "A Tutorial on Support Vector Machines for Pattern Recognition," *Data Mining and Knowledge Discovery,* vol. 2, no. 2, pp. 121-167, 1998.

[3] A. Mishra, K. Nadkarni, and A. Patcha, "Intrusion Detection in Wireless Ad Hoc Networks," *IEEE Wireless Comm.,* vol. 11, no. 1, pp. 48-60, Feb. 2004.

[4] Y.A. Huang and W. Lee, "Attack Analysis and Detection for Ad Hoc Routing Protocols," *Proc. Symp. Recent Advances in Intrusion Detection,* pp. 125-145, 2004.

[5] M. Little, "TEALab: A Testbed for Ad Hoc Networking Security Research," Proc. *IEEE Military Comm. Conf. 2005* (MILCOM '05), 2005.

[6] P. Papadimitratos and Z. Haas, "Secure Routing for Mobile Ad hoc Networks," *Proc. SCS Comm. Networks and Distributed Systems Modeling and Simulation Conf.* (CNDS '02), 2002.

[7] G. Thamilarasu et al., "A Cross-Layer Based Intrusion Detection Approach for Wireless Ad Hoc Networks," *Proc. IEEE Int'l Conf. Mobile Adhoc and Sensor Systems* 2005, 2005.

[8] F. Anjum and P. Mouchtaris, Security for Wireless Ad Hoc Networks. Wiley, 2007.

[9] Y. Liu, Y. Li, and H. Man, "Short Paper: A Distributed Cross-Layer Intrusion Detection System for Ad Hoc Networks," *Proc. First Int'l Conf. Security and Privacy for Emerging Areas in Comm. Networks* 2005 (SecureComm '05), 2005.

[10] H. Deng, Q.-A. Zeng, and D.P. Agrawal, "SVM-Based Intrusion Detection System for Wireless Ad Hoc Networks," *Proc. IEEE 58th Vehicular Technology Conf. 2003* (VTC '03-Fall), vol. 3, pp. 2147-2151, 2003.

[11] Y. Liu, Y. Li, and H. Man, "MAC Layer Anomaly Detection in Ad Hoc Networks," Proc. Sixth Ann. *IEEE Systems, Man and Cybernetics (SMC) Information Assurance Workshop,* 2005.

[12] P. Ning and K. Sun, "How to Misuse AODV: A Case Study of Insider Attacks Against Mobile Ad-Hoc Routing Protocols," *Ad Hoc Networks,* vol. 3, no. 6, pp. 795-819, 2005.

[13] R.O. Duda, P.E. Hart, and D.G. Stork, Pattern Classification, second ed. Wiley Inter-Science Publication, 2000.

[14] V.N. Vapnik, Statistical Learning Theory. Wiley, 1998.

[15] P.-W. Yau and C.J. Mitchell, "Security Vulnerabilities in Ad Hoc Networks," *Proc. Seventh Int'l Symp. Comm. Theory and Applications* (ISCTA '03), 2003.

[16] M. Bykova, S. Ostermann, and B. Tjaden, "Detecting Network Inrusions via a Statistical Analysis of Network Packet Characteristics," *Proc. 33rd Southeastern Symp. System Theory,* 2001.