October 2011

Role of Feature Selection in Intrusion Detection Systems for 802.11 Networks

N. Pratik Neelakantan

Intell Engineering College ,Akkampalli Cross Road, Anantapur, A.P., India, pratikcse08@gmail.com

C. Nagesh

Dept. of Computer Science & Engineering , Intell Engineering College, Anantapur,A.P., India, nagesh.pandu@gmail.com

Follow this and additional works at: <https://www.interscience.in/ijssan>

Part of the [Digital Communications and Networking Commons](#), and the [Electrical and Computer Engineering Commons](#)

Recommended Citation

Neelakantan, N. Pratik and Nagesh, C. (2011) "Role of Feature Selection in Intrusion Detection Systems for 802.11 Networks," *International Journal of Smart Sensor and Adhoc Network*: Vol. 1 : Iss. 2 , Article 14.
DOI: 10.47893/IJSSAN.2011.1030

Available at: <https://www.interscience.in/ijssan/vol1/iss2/14>

This Article is brought to you for free and open access by the Interscience Journals at Interscience Research Network. It has been accepted for inclusion in International Journal of Smart Sensor and Adhoc Network by an authorized editor of Interscience Research Network. For more information, please contact sritampatnaik@gmail.com.

Role of Feature Selection in Intrusion Detection Systems for 802.11 Networks

N. Pratik Neelakantan¹, C. Nagesh² M.Tech.,
Intell Engineering College¹, Dept. of Computer Science & Engineering²,
Akkampalli Cross Road, Anantapur, A.P., India¹, Intell Engineering College, Anantapur, A.P., India²
pratikcse08@gmail.com¹, nagesh.pandu@gmail.com²

Abstract—Intrusion Detection Systems are important for protecting network and its resources from illegal penetration. For 802.11 network, the features used for training and testing the intrusion detection systems consist of basic information related to the TCP/IP header, with no considerable attention to the features associated with lower level protocol frames. The resulting detectors were efficient and accurate in detecting network attacks at the network and transport layers, but unfortunately, not capable of detecting 802.11-specific attacks such as deauthentication attacks or MAC layer DoS attack. IDS systems can also identify and alert to the presence of unauthorized MAC addresses on the networks. The IDS is based a novel hybrid model that efficiently selects the optimal set of features in order to detect 802.11-specific intrusions. This model for feature selection uses the information gain ratio measure as a means to compute the relevance of each feature and the k-means classifier to select the optimal set of MAC layer features that can improve the accuracy of intrusion detection systems while reducing the learning time of their learning algorithm.

Index Terms—Intrusion Detection System, Information Gain Ratio (IGR), Dos attack, Feature Selection.

1. Introduction

IDS is not a firewall! Firewalls are designed to be outward looking and to limit access between networks in order to prevent an intrusion happening. IDS watch the wired and wireless network from the inside and report or alarm depending on how they evaluate the network traffic they see. They continually monitor for access points to the network and are able, in some cases, to do comparisons of the security controls defined on the access point with pre-defined company security standards and either reset or closedown any non conforming AP's they find. The distinction between placing IDS sensors on both wired and wireless networks is an important one as large corporate networks can be worldwide.

Traditional IDSs are able to protect the application and software components of TCP/IP networks against intrusion attempts, the physical and data link layers are vulnerable to intrusions specific to these communication layers. In addition to the vulnerabilities of wired networks, wireless networks are the subject of new types of attacks which range from the passive eavesdropping to more devastating attacks such as denial of service [22]. These vulnerabilities are a result of the nature of the transmission media [26]. Indeed, the absence of physical boundaries in the network to

monitor, meaning that an attack can be perpetrated from anywhere, is a major threat that can be exploited to undermine the integrity and security of the network. It is, therefore, essential to take into account these considerations when designing and deploying an intrusion detection system.

2. Features

Feature selection is the important step in building intrusion detection models [1], [2], [3]. During this step, the set of attributes or features deemed to be the most effective attributes is extracted in order to construct suitable detection algorithms (detectors). A key problem that many researchers face is how to choose the optimal set of features, as not all features are relevant to the learning algorithm, and in some cases, irrelevant and redundant features can introduce noisy data that distract the learning algorithm, severely degrading the accuracy of the detector and causing slow training and testing processes. Feature selection was proven to have a significant impact on the performance of the classifiers. Experiments in [4] show that feature selection can reduce the building and testing time of a classifier by 50 percent. There are currently two models in the literature for feature selection: the filter model and the wrapper model [5]. The wrapper model uses the predictive accuracy of a classifier as a means to evaluate the “goodness” of a feature set, while the filter model uses a measure such as information, consistency, or distance measures to compute the relevance of a set of features. These approaches suffer from many drawbacks: the first major drawback is that feeding the classifier with arbitrary features may lead to biased results, and hence, we cannot rely on the classifier's predictive accuracy as a measure to select features. A second drawback is that for a set of N features, trying all possible combinations of features (2^N combinations) to find the best combination to feed the classifier is not a feasible approach. For example, the DARPA data set contains 41 features [6], and the data set would be larger if we add to it the OSI Layer 2 (MAC layer) features, resulting in thousands of billions of different feature combinations. Different techniques have been used to tackle the problem of feature selection. In [7],

Sung and Mukkamala used feature ranking algorithms to reduce the feature space of the DARPA data set from 41 features to the six most important features. They used three ranking algorithms based on Support Vector Machines (SVMs), Multivariate Adaptive Regression Splines (MARSs), and Linear Genetic Programs (LGPs) to assign a weight to each feature. Experimental results showed that the classifier's accuracy degraded by less than 1 percent when the classifier was fed with the reduced set of features. Sequential backward search was used in [8], [9] to identify the important set of features: starting with the set of all features, one feature was removed at a time until the accuracy of the classifier was below a certain threshold. Different types of classifiers were used with this approach including Genetic Algorithms in [9], Neural Networks in [8],[10], and Support Vector Machines in [8].

3 WIRELESS NETWORKS INTRUSIONS

A deauthentication attack is an example of an easy to mount attack on all types of 802.11 networks. Likewise, a duration attack is another simple attack that exploits the vulnerability of the virtual carrier sensing protocol CSMA/CA and it was proven in [11] to deny access to the network. Many open source tools are available on the Internet which allow hackers to exploit these protocol weaknesses to deny access to a network, as can be seen in [12], where a collection of tools to attack 802.11-based networks is available for download. These tools operate on WEP- and WPA-protected networks.

Deauthentication Attack

The attacker fakes a deauthentication frame as if it had originated from the base station (Access Point). Upon reception, the station disconnects and tries to reconnect to the base station again. This process is repeated indefinitely to keep the station disconnected from the base station. The Attacker can also set the receiving address to the broadcast address to target all stations associated with the victim base station. However, we noticed that some wireless network cards ignore this type of deauthentication frame.

4 APPROACH

Select the best set of MAC layer features that efficiently characterize normal traffic and distinguish it from abnormal traffic containing intrusions specific to wireless networks. Our framework uses a hybrid approach for feature selection that combines the filter and wrapper models. In this approach, we rank the features using an independent measure: the information

gain ratio. The k-means classifier's predictive accuracy is used to reach an optimal set of features which maximize the detection accuracy of the wireless attacks. To train the classifier, we first collect network traffic containing four known wireless intrusions, namely, the deauthentication, duration, fragmentation, and

```

Input:
F - Full set of features
IGR: Information Gain Ratio Measure
C: K-means classifier
T: Gained Accuracy Threshold

For each feature f compute IGR(f)
Rank features in F according to IGR(f)

//Optimal Set Selection Algorithm
Initialize: S={}, ac=0
Repeat
(1) ap=ac
(2) f=getNext(F)
(3) S=S U {f}
(4) F=F- {f}
(5) ac= accuracy(C,S)
Until (ac-ap)<T Or ac<ap
    
```

The selection algorithm (Fig. 1) starts with an empty set S of the best features, and then, proceeds to add features from the ranked set of features F into S sequentially. After each iteration, the "goodness" of the resulting set of features S is measured by the accuracy of the k-means classifier. The selection process stops when the gained classifier's accuracy is below a certain selected threshold value or in some cases when the accuracy drops, which means that the accuracy of the current subset is below the accuracy of the previous subset.

5 LIST OF FEATURES

The Features are Version, Type, SubType, ToDS, FromDS, More Fragment, Retry, Power Mgmt, More Data, WEP, Order, Duration, RA,TA,MA etc.,

Feature	Description
Version	Two bits indicate which version of the 802.11 MAC is contained in the rest of the frame.
Type	Indicate the type of the frame (Mgmt, Ctrl, Data).
SubType	Indicate the subtype of the frame.
ToDS	Indicate if a frame is destined to the Distribution System.
FromDS	Indicate if a frame is originated from Distribution System.
More Fragment	Indicate whether a frame is non final fragment or not.
Retry	Indicate if the frame is a retransmitted frame.
Power Mgmt	Indicate whether the station is active or in Power Saving Mode.
More Data	Indicate whether an access point has buffered frames for a dozing station.
WEP	Indicate if the frame is processed by the WEP protocol.
Order	Indicate if the "strict ordering" delivery is employed.
Duration	The number of microseconds the medium is expected to be busy
RA	The MAC address of the receiving station.
TA	The MAC address of the transmitting station.
MA	Depending on the values of ToDS and FromDS fields, this address can be the MAC address of the Sending, Destination or Base Station.
FCS	A Frame Check Sequence, which contains a 32 bit Cyclic Redundancy Code.

6 INFORMATION GAIN RATIO MEASURE

We used the Information Gain Ratio (IGR) as a measure to determine the relevance of each feature. Note that we chose the IGR measure and not the Information Gain because the latter is biased toward the features with a large number of distinct values [5]. IGR is defined in [18] as

$$\text{IGR}(Ex, f) = \text{Gain}(Ex, f) / \text{SplitInfo}(Ex, f)$$

where Ex is the set of vectors that contain the header information.

$$\begin{aligned} \text{Gain}(Ex, f) &= \text{Entropy}(Ex) \\ &- \sum_{v \in \text{Values}(f)} \frac{|Ex, v|}{|Ex|} * \text{Entropy}(Ex, v), \\ Ex, v &= \{x \in Ex / \text{value}(x, f) = v\}. \end{aligned}$$

The entropy function is the Shannon's entropy defined as

$$\text{Entropy}(Ex) = - \sum P_i \log_2(P_i),$$

where P_i is the probability of a class i .

$\text{SplitInfo}(Ex, f)$ is defined as

7 DATA SET

The data we used to train and test the classifiers were collected from a wireless local area network. The local network was composed of three wireless stations and one access point. One machine was used to generate normal traffic (HTTP, FTP). The second machine simultaneously transmitted data originating from four types of attacks. The last station was used to collect and record both types of traffic (normal and intrusive). The attacks we used to test our system were: deauthentication, duration, fragmentation, and chopchop. The source code for the attacks is available in [12]. The data collected were grouped in three sets (Table 4): learning, validation, and testing sets. The first set is used to reach the optimal weight of each synapse. The learning set contains the input with its desired output. By iterating on this data set, the neural network classifier dynamically adjusts the weights of the synapses to minimize the error rate between the output of the network and the desired output. The validation data set is necessary to avoid the effect of overfitting. Indeed, in some cases, the neural network classifier might produce an excellent performance on the learning data set, but still have a low performance on the testing data set. In general, the learning algorithm stops when the error between the output of the validation data set and the desired output is below a predefined threshold. Once the network is trained and validated, it should be able to predict the output of each entry in the testing data set.

Top Ten Features

Rank	Feature	IGR
1	IsWepValid	1.02
2	DurationRange	1.01
3	More_Frag	0.98
4	To_DS	0.89
5	WEP	0.85
6	Casting_Type	0.82
7	Type	0.73
8	SubType	0.65
9	Retry	0.46
10	From_DS	0.41
11-38	Remaining features	< 0.23

CONCLUSION

Presenting an approach to select the best features for detecting intrusions in 802.11-based networks. Approach is based on a model which combines the filter and wrapper models for selecting relevant

features. By reducing the number of features from 38 to 8. The impact of feature selection on the performance of different classifiers based on neural networks. Planning to do a comparative study of the impact of the reduced feature set on the performance of classifiers-based ANNs.

REFERENCES

- [1] T.M. Khoshgoftaar, S.V. Nath, S. Zhong, and N. Seliya, "Intrusion Detection in Wireless Networks Using Clustering Techniques with Expert Analysis," *Proc. Fourth Int'l Conf. Machine Learning and Applications*, Dec. 2005.
- [2] A. Boukerche, K.R.L. Juc, J.B. Sobral, and M.S.M.A. Notare, "An Artificial Immune Based Intrusion Detection Model for Computer and Telecommunication Systems," *Parallel Computing*, vol. 30, nos. 5/6, pp. 629-646, 2004.
- [3] Y.-H. Liu, D.-X. Tian, and D. Wei, "A Wireless Intrusion Detection Method Based on Neural Network," *Proc. Second IASTED Int'l Conf. Advances in Computer Science and Technology*, Jan. 2006.
- [4] Y. Chen, Y. Li, X. Cheng, and L. Guo, "Survey and Taxonomy of Feature Selection Algorithms in Intrusion Detection System," *Proc. Conf. Information Security and Cryptology (Inscrypt)*, 2006.
- [5] H. Liu and H. Motoda, *Feature Selection for Knowledge Discovery and Data Mining*. Kluwer Academic, 1998.
- [6] <http://kdd.ics.uci.edu/databases/kddcup99/task.html>, 2010.
- [7] A.H. Sung and S. Mukkamala, "The Feature Selection and Intrusion Detection Problems," *Proc. Ninth Asian Computing Science Conf.*, 2004.
- [8] A.H. Sung and S. Mukkamala, "Identifying Important Features for Intrusion Detection Using Support Vector Machines and Neural Networks," *Proc. Symp. Applications and the Internet (SAINT '03)*, Jan. 2003.
- [9] G. Stein, B. Chen, A.S. Wu, and K.A. Hua, "Decision Tree Classifier for Network Intrusion Detection with GA-Based Feature Selection," *Proc. 43rd ACM Southeast Regional Conf.—Volume 2*, Mar. 2005.
- [10] A. Hofmann, T. Horeis, and B. Sick, "Feature Selection for Intrusion Detection: An Evolutionary Wrapper Approach," *Proc. IEEE Int'l Joint Conf. Neural Networks*, July 2004.
- [11] J. Bellardo and S. Savage, "802.11 Denial-of-Service Attacks: Real Vulnerabilities and Practical Solutions," *Proc. USENIX Security Symp.*, pp. 15-28, 2003.
- [12] <http://www.aircrack-ng.org/>, 2010.
- [13] A. Boukerche and M.S.M.A. Notare, "Behavior-Based Intrusion Detection in Mobile Phone Systems," *J. Parallel and Distributed Computing*, vol. 62, no. 9, pp. 1476-1490, 2002.
- [14] A. Boukerche, R.B. Machado, K.R.L. Juca', J.B.M. Sobral, and M.S.M.A. Notare, "An Agent Based and Biological Inspired Real- Time Intrusion Detection and Security Model for Computer Network Operations," *Computer Comm.*, vol. 30, no. 13, pp. 2649- 2660, Sept. 2007.
- [15] S. Zhong, T.M. Khoshgoftaar, and S.V. Nath, "A Clustering Approach to Wireless Network Intrusion Detection," *Proc. 17th IEEE Int'l Conf. Tools with Artificial Intelligence (ICTAI '05)*, Nov 2005.
- [16] A. Bittau, M. Handley, and J. Lackey, "The Final Nail in WEP's Coffin," *Proc. 2006 IEEE Symp. Security and Privacy*, May 2006.
- [17] IEEE 802.11-1999, Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications: Higher Speed Physical Layer Extension in the 2.4 GHz Band.
- [18] J.R. Quinlan, "Induction of Decision Trees," *Machine Learning*, vol. 1, pp. 81-106, 1986.
- [19] Z. Zhang and C. Manikopoulos, "Investigation of Neural Network Classification of Computer Network Attacks," *Proc. Int'l Conf. Information Technology: Research and Education*, pp. 590-594, Aug. 2003.
- [20] NeuroSolutions, Inc., <http://www.neurosolutions.com/>, 2010.
- [21] CERT, <http://www.cert.org/stats/>, 2010.
- [22] J. Yeo, M. Youssef, and A. Agrawala, "A Framework for Wireless LAN Monitoring and Its Applications," *Proc. ACM Workshop Wireless Security*, pp. 70-79, 2004. 6 *IEEE TRANSACTIONS ON PARALLEL AND DISTRIBUTED SYSTEMS*, VOL. 21, NO. X, XXX 2010
Fig. 5. False Positives Rate (%) for the three types of neural networks using 8 and 38 features
Fig. 6. False Negatives Rate (%) for the three types of neural networks using 8 and 38 features. This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.
- [23] Y. Chen, Y. Li, X. Cheng, and L. Guo, "Survey and Taxonomy of Feature Selection Algorithms in Intrusion Detection System," *Proc. Conf. Information Security and Cryptology (Inscrypt)*, 2006.
- [24] A. Boukerche, *Algorithms and Protocols for Wireless and Mobile AdHoc Networks*. Wiley, 2008.
- [25] A. Boukerche, *Handbook of Algorithms for Wireless Networking and Mobile Computing*, CRC/Chapman and Hall 2005.