Volume 4, Issue 10, October 2014 ISSN: 2277 128X
International Journal of Advanced Research in Computer Science and Software Engineering
Research Paper
Available online at: www.ijarcsse.com

Intrusion Detection System and Hidden Markov Models

Preeti Saini, Sunila Godara
Department of Computer Science and Engineering, Guru Jambheshwar University,
Haryana, India

Abstract— Intrusion detection is the process of identifying and responding to malicious activity targeted at computing and networking sources. We need an Intrusion Detection System to detect intrusion. HMM is a random, probabilistic and stochastic process with an underlying probabilistic process that is not observable, but can be observed through another set of stochastic or random process that produces the sequences of observed symbols. HMM can be used to model an IDS. In this paper, we have discussed how different number of HMMs trained affects the performance of IDS modelled using HMMs.

Keywords—Intrusion Detection System, Hidden Markov Model, system security, normal behaviour modelling, accuracy

I. INTRODUCTION

An intrusion detection system (IDS) is used to monitor network traffic, check for suspicious activities and notifies the network administrator or the system. In some instances, the IDS might also react to malicious or anomalous traffic and will take action such as barring the user or perhaps the IP address source from accessing the system. IDS are available in many different types and will approach the mission of uncovering shady traffic in various ways [8]. We have modelled IDS using alternative number of HMMs and compared the accuracy obtained.

II. HIDDEN MARKOV MODEL AND IDS

This section discusses HMM and IDS.

A. Hidden Markov Model

Hidden Markov model (HMM) is a partially hidden Markov Chain, named after Andrey Markov, 1856–1922 [11]. The Hidden Markov Model (HMM) is a finite set of states, each of which is associated with a probability distribution. Transitions among these states are governed by a set of probabilities called state transition probabilities. In a particular state an outcome or observation can be generated, according to the associated probability distribution. It is only the outcome, not the state visible to an external observer and therefore states are "hidden" to the outside; hence the name Hidden Markov Model [8].

Elements of HMM

1. There are a finite number, say N, of states in the model.
2. At each clock time, t, a new state is entered based upon a transition probability distribution which depends on the previous state (the Markovian property). (The transition may be such that the process remains in the previous state.)
3. After each transition is made, an observation output symbol is produced according to a probability distribution which depends on the current state. This probability distribution is held fixed for the state regardless of when and how the state is entered. There are thus N such observation probability distributions which, of course, represent random variables or stochastic processes [1].

Due to the nondeterministic nature of user behavior generally the probabilistic approach is the appropriate technique to be incorporated with anomaly detection technique for modeling IDS to model the user profile [7].

B. Intrusion Detection System

Intrusion detection is the process of identifying and responding to malicious activity targeted at computing and networking sources. an intrusion detection system (IDS) is used to monitor network traffic, check for suspicious activities and notifies the network administrator or the system. In some instances, the IDS might also react to malicious or anomalous traffic and will take action such as barring the user or perhaps the IP address source from accessing the system. IDS are available in many different types and will approach the mission of uncovering shady traffic in various ways [8].

There are a number of methods for constructing IDS models. Also it is possible to have IDS's deployment at different points in a working environment; like firewalls and application servers [7].

III. IMPEMENTATION OF IDS

In normal behavior modeling HMMs are trained using only normal TCP sessions. So, any deviation from normal behavior is taken as anomaly. Thus in proposed approach normal TCP sessions are used to train the model. The intrusion detection system is modeled using alternative number of hidden Markov models. Each HMM is trained by one of the TCP session sequence. First the system is modeled using ten HMMs. 10 sequences are used to train ten HMMs. The probability of sequence to be tested is calculated using these ten HMMs.

The maximum probability is taken as the probability of the sequence. Both normal TCP session sequence and attack sequence are used to test the model. 250 normal and 250 attack sequences are taken to test the model. TN, TP, FN, FP are calculated which are then used to calculate accuracy.

The accuracy obtained with 10 HMMs is 82.8%. The same process is repeated with 20, 30, 40, 50, 60, 70, 80, 90, 100 HMMs. The accuracy obtained with these HMMs is 86, 88, 88, 89.6, 90, 90.4, 90.4, 90.6, 90.6 respectively.

We have used NSL KDD dataset to train and test the proposed model.

IV. RESULTS

The results obtained are shown in the following table .

No. of HMMs trained	Accuracy Obtained with increased HMMs
10	82.8
20	86
30	88
40	88
50	89.6
60	90
70	90.4
80	90.4
90	90.6
100	90.6

V. CONCLUSION

We have modelled IDS using alternative number of HMMs and compared the accuracy obtained. The accuracy tends to remain at 90.6% after 70 HMMs. Thus it becomes stable after 70 HMMs. Also it takes larger time to train increasing number of HMMs.

REFERENCES

[1] Rabiner, Lawrence, and Biing-Hwang Juang. "An introduction to hidden Markov models." *ASSPMagazine, IEEE* 3.2 (1986): 4-16.

[2] Gideon Creech, Jiankun Hu, "A Semantic Approach to Host-based Intrusion Detection Systems Using Contiguous and Discontiguous System Call Patterns," IEEE Transactions on Computers, 28 Jan. 2013. IEEE computer Society Digital Library. IEEE Computer Society.

[3] Khosronejad, Mahsa, et al. "Developing a hybrid method of Hidden Markov Models and C5.0 as a Intrusion Detection System." International Journal of Database Theory & Application 6.5 (2013).

[4] Wang, Wei, Xiao-Hong Guan, and Xiang-Liang Zhang. "Modeling program behaviors by hidden markov models for intrusion detection." Machine Learning and Cybernetics, 2004. Proceedings of 2004 International Conference on. Vol. 5. IEEE, 2004.

[5] Ourston, Dirk, et al. "Applications of hidden markov models to detecting multi-stage network attacks." System Sciences, 2003. Proceedings of the 36th Annual Hawaii International Conference on. IEEE, 2003.

[6] Joshi, Shrijit S., and Vir V. Phoha. "Investigating hidden Markov models capabilities in anomaly detection." Proceedings of the 43rd annual Southeast regional conference-Volume 1. ACM, 2005.

[7] Badajena, J. Chandrakanta, and Chinmayee Rout. "Incorporating Hidden Markov Model into Anomaly Detection Technique for Network Intrusion Detection."International Journal of Computer Applications 53 (2012).

[8] Devarakonda, Nagaraju, et al. "Intrusion Detection System using Bayesian Network and Hidden Markov Model." Procedia Technology 4 (2012): 506-514.

[9] Tavallaee, Mahbod, et al. "A detailed analysis of the KDD CUP 99 data set."Proceedings of the Second IEEE Symposium on Computational Intelligence for Security and Defence Applications 2009.

[10] Farhadi, Hamid, Maryam AmirHaeri, and Mohammad Khansari. "Alert Correlation and Prediction Using Data Mining and HMM." ISeCure 3.2 (2011).

[11] Langin, Chet, and Shahram Rahimi. "Soft computing in intrusion detection: the state of the art."Journal of Ambient Intelligence and Humanized Computing 1.2 (2010): 133-145.

[12] Hu, Jiankun, et al. "A simple and efficient hidden Markov model scheme for host-based anomaly intrusion detection."Network, IEEE 23.2 (2009): 42-47.

[13] Ariu, Davide, Roberto Tronci, and Giorgio Giacinto. "< i> HMMPayl</i>: An intrusion detection system based on Hidden Markov Models."computers & security 30.4 (2011): 221-241.

[14] Jain, Ruchi, and Nasser S. Abouzakhar. "A Comparative Study of Hidden Markov Model and Support Vector Machine in Anomaly Intrusion Detection." Journal of Internet Tecnology and Secure Transactions Vol.2 (2013).

[15] Shahboddin Shamshirband et al, "An appraisal and design of a multi-agent system based cooperative wireless intrusion detection computational intelligence technique",Engineering Applications of Artificial Intelligence , ScienceDirect , Volume 26,Issue 9, October 2013, Pages 2105– 2127

[16] Sendi, Alireza Shameli, et al. "Real Time Intrusion Prediction based on Optimized Alerts with Hidden Markov Model."Journal of Networks 7.2 (2012).

[17] Gao, Bo, Hui-Ye Ma, and Yu-Hang Yang. "Hmms (hidden markov models) based on anomaly intrusion detection method." Machine Learning and Cybernetics, 2002. Proceedings. 2002 International Conference on. Vol. 1. IEEE, 2002.

[18] Rabiner, Lawrence. "A tutorial on hidden Markov models and selected applications in speech recognition." Proceedings of the IEEE 77.2 (1989): 257-286.

[19] Warrender, Christina, Stephanie Forrest, and Barak Pearlmutter. "Detecting intrusions using system calls: Alternative data models." Security and Privacy, 1999. Proceedings of the 1999 IEEE Symposium on. IEEE, 1999.

[20] Yeung, Dit-Yan, and Yuxin Ding. "Host-based intrusion detection using dynamic and static behavioral models." Pattern recognition 36.1 (2003): 229-243.

[21] Hoang, Xuan Dau, Jiankun Hu, and Peter Bertok. "A multi-layer model for anomaly intrusion detection using program sequences of system calls." Proc. 11th IEEE Int'l Conf. Networks. 2003.

[22] Khanna, Rahul, and Huaping Liu. "System approach to intrusion detection using hidden Markov model." Proceedings of the 2006 international conference on Wireless communications and mobile computing. ACM, 2006.

[23] Hoang, X. A., and Jiankun Hu. "An efficient hidden Markov model training scheme for anomaly intrusion detection of server applications based on system calls." Networks, 2004.(ICON 2004). Proceedings. 12th IEEE International Conference on. Vol. 2. IEEE, 2004.

[24] Srivastava, Abhinav, et al. "Credit card fraud detection using hidden Markov model." Dependable and Secure Computing, IEEE Transactions on 5.1 (2008): 37-48.

[25] Cho, Sung-Bae. "Incorporating soft computing techniques into a probabilistic intrusion detection system." Systems, Man, and Cybernetics, Part C: Applications and Reviews, IEEE Transactions on 32.2 (2002): 154-160.

[26] Hu, Jiankun, et al. "A simple and efficient hidden Markov model scheme for host-based anomaly intrusion detection." Network, IEEE 23.2 (2009): 42-47.

[27] http://www.mathworks.in

