Detecting Network Intrusions Using Hierarchical Temporal Memory

Gift Khangamwa

Lecturer, Computing & Information Technology Department,
University of Malawi, The Polytechnic
giftkhangamwa@yahoo.com, gkhangamwa@poly.ac.mw

Abstract. Intrusion Detection Systems (IDS) are a very popular network security tool. These tools can allow network administrators, to identify and react to hostile traffic aimed at, or generated from their own network. In general there are two common Intrusion Detection approaches which are behavior or traffic anomaly based and knowledge or signature based. As a result of the increased sophistication of intrusion attacks, one very desirable feature of advanced IDS is to be capable of learning and generalizing from known traffic patterns of a system, process or a user's behavior. In this project we investigated the use of a novel Artificial Intelligence (AI) approach to intrusion detection based on network traffic anomaly detection. The AI technique used is based on the Hierarchical Temporal Memory (HTM) paradigm developed by Numenta, which is a relatively new AI concept that mimics the operation of the neocortex area of the human brain[11,14]. The developed AI scheme was evaluated using the corpus of data from Massachusetts Institute of Technology, Lincoln Laboratories in USA [20]. Our results show that HTM based intrusion detection can achieve relatively high success rates in identifying anomalous traffic in computer networks, furthermore our research also shows that HTM based schemes can achieve very fast detection rates making them a very good alternative for real time intrusion detection engine.

In this paper we present the results of our study as well as a discussion on our future work.

Keywords: Intrusion detection, Artificial Intelligence, Hierarchical Temporal Memory^{TM} , Network anomaly detection.

1 Introduction

The research presented in this paper is motivated by the need for intrusion detection mechanisms that are capable of dealing with the increasing security challenges faced by modern computer and data networks. Authors in [10] speak of the challenges that computers that are hooked to the internet face due to the security challenges in open environments like the internet. This has made intrusion detection systems an absolute necessity in this information age, more so in the developing world where ICT infrastructure developments are still in early stages. This means that as dependence on the internet grows so will the prevalence of intrusions and hence also the importance of intrusion detection systems to deter such attacks.

R. Popescu-Zeletin et al. (Eds.): AFRICOMM 2010, LNICST 64, pp. 41–48, 2011.

According to [10], the sophistication of attacks and tools used by attackers has been steadily advancing. This implies that new attacks are ever being devised and even old attacks are being modified to be made stealthier, and more undetectable.

The aim of this research was to investigate the usage of NuPIC (Hierarchical Temporal Memory technology) from Numenta Inc. for use in Network Intrusion detection. NuPIC is an artificial intelligence based platform that is built from a theory of the neocortex of the human brain. The platform used for this research was obtained freely under a research license. This technology has the capacity to detect both modified attack signatures as well as novel attacks, because it uses machine learning to learn and generalize from attack patterns that it is trained with. Consequently such a scheme will help to ensure that any known or unknown attacks are identified and reported.

The scheme was tested using the Massachusetts Institute of Technology Lincoln laboratories, 1999 DARPA Intrusion Detection Systems data corpora. This data is freely available online from www.ll.mit.edu and contains labeled anomalies; this data has also been used by authors in [1, 8 and 9].

2 Literature Review

Intrusion detection and Intrusion Detection systems have been defined in so many different ways by different authors in the literature. Some of the definitions that capture the important principles involved are as follows: Systems aimed at detecting attacks against computer systems and networks [5]. In his book on Intrusion Detection, Amoroso in [7] defined it as a process of identifying and responding to malicious activity targeted at computing and networking resources. Bace in [6] stated it as, a system that monitors computer networks and systems for violations of security policy. While Mukherjee et al, in [1] stated that, it's an approach of providing a sense of security in existing computers and data networks while allowing them to operate in their open Environment where threats are ever present. All these definitions capture the major aspects and principles of this discipline and provide an understanding of what IDS are; our work in this research is hence built upon this understanding. Authors in [1] used a technique based on neural networks, which also have the capacity for machine learning. The major weakness for neural networks is that it is not possible to analyze how the neural network learned and come to the conclusions that it makes and uses to make detections [3]. The major difference between their method and the strategy used in this study is the technology used, at the time of conducting this research there was no available literature on usage of this technology for Intrusion detection. HTM technology offers a better alternative to neural networks because once the network has been trained it can be analysed to access how the learning was done. Furthermore network training using HTMs is transparent and controllable through some parameters that can be changed. The authors in [3] also used neural networks for intrusion detection. The only difference in their method from that used in [1] is that they used misuse detection where detection is based on training a detection system using a known set of anomalies and the system then uses this knowledge to make detections, while the other authors in[1] used both anomaly and misuse detection using neural networks for intrusion detection.

2.1 Intrusion Detection

In this research our focus is mainly on approach to intrusion detection that falls under the classification that is based on detection method. We consider Intrusion Detection based on what data is used to establish a baseline for detection of anomalous or intrusive situations. Here two primary methods are identified:

2.1.1 Misuse Detection

Misuse detection [1, 2, 3, 6, 8, 9] or otherwise referred to as, signature based by [8, 9] or Knowledge based [5]. In this approach an Intrusion detection system is trained with patterns or signatures of all known anomalies. This implies that the system will have complete knowledge of all the anomalies for whose signatures or patterns it is equipped with. Therefore using this information, the system is able to identify these patterns and signatures in any current or ongoing network traffic [4].

2.1.1.1 Benefits. This approach has the benefit that it identifies the known anomalies with very high accurate rates, while at the same time yielding very low false positive rates [1].

2.1.1.2 Weaknesses. The major weaknesses of this approach are that it cannot identify novel (never seen before) attacks [1, 2, 8, and 9]; it also cannot identify known anomalies that leave different attack signatures in a system after every attack [1]. Hence even known attacks that evolve will not be identified by this methodology.

2.1.2 Anomaly Detection

Anomaly Detection [1, 2, 3, 8, 9] otherwise also referred to as Behavior Based [5] which models the normal anomaly free state of a network or system [7, 8, 18]. So in this scheme the system will mistrust any other pattern that deviates from what it knows as the normal state of the network or system.

2.1.2.1 Benefits. This scheme has the obvious benefit that anomalies that have never been seen before can be identified and detected [8, 9].

The AI method used in this study is suitable for both misuse detection based as well as anomaly detection based approaches to intrusion detection. Similar work was done by Ghosh and Schwartzbard in [1] where they used Artificial Neural Networks for both misuse and anomaly detection. Though we will discuss misuse detection in this paper our major focus in will be on anomaly detection.

2.1.2.2 Weaknesses. The problem with anomaly detection is that it is likely to raise many false alarms [2]. This is more true if the method used is incapable of recognizing novel legitimate behavior of the system as is the case where non machine learning approaches are used.

2.2 Hierarchical Temporal Memory

Numenta Platform for Intelligent Computing models the functioning of the neocortex of the human brain, without being a direct implementation of the same. This makes it potentially able to solve different kinds of problems that are easy for humans to solve but are extremely hard for computers to solve.

HTM uses a hierarchy of learning nodes that are capable of identifying spatial and temporal correlations to formulate a collection of beliefs of the existing causes in their world, as presented in the input data. According to [14], HTMs are organized as a tree-shaped hierarchy of nodes, where each learning node implements a common learning and memory function. HTMs store information throughout the hierarchy in a way that models the world. All objects in the world, be they cars, people, buildings, speech, or the flow of information across a computer network, have structure. This structure is hierarchical in both space and time. HTM memory is also hierarchical in both space and time, and therefore can efficiently capture and model the structure of the world.

Every node in an HTM network has 2 components a Spatial Pooler component and a Temporal Pooler component. On any given input the Spatial Pooler reduces a huge number of input occurrences to a smaller set of values called coincidences. These coincidences are fed to a Temporal Pooler which classifies them into groups basing on their similarities i.e. spatial and temporal correlations.

The HTM network therefore learns via belief propagation of knowledge from lower level nodes to nodes higher up in the hierarchy. This means what the HTM network knows finally is a summary of what all the nodes at different levels in the hierarchy have learnt during training.

At the time of the study the version of nupic used was 1.6.1 which was the latest at the time.

3 Experimental Setup

The experimental setup in this research involved setting up and configuring NuPIC platform and building NuPIC HTM networks in a computer lab. The HTM networks were then tested using corpora of data that was assembled by the Massachusetts Institute of Technology, Lincoln Laboratories as part of the 1999 DARPA Intrusion Detection Systems Evaluation project. According to authors in [20], the aim of the project was to produce corpora of data that were extensive covering a wide range of intrusions and making the data readily accessible to researchers and developers, to be used in the evaluation of intrusion detection systems worldwide.

In this research the tests were designed to assess if the NuPIC platform was suitable for the general problem of intrusion detection and more specifically anomaly detection. Furthermore more tests were conducted in order to ascertain other factors that might affect the performance of the developed HTM network based intrusion detection scheme under investigation.

4 Results

The experimental setup in this research involved setting up and configuring NuPIC platform and building NuPIC HTM networks in a computer lab. The HTM networks built were used to answer a number questions like is this technology appropriate for intrusion detection. Secondly how can detection of intrusions be improved when this scheme is being used. Below is a presentation of the results that were obtained during the experiments.

4.1 Proof of Concept

In this test the objective was to test and find out if indeed using the NuPIC platform we could detect the presence of anomalies in the data. In order to do this a 3 layer network with 14 learning nodes was built and tested with 64 x 400 training data, the results obtained are:

- **88.25 %** Accuracy on Training data
- **88.25 %** Accuracy on Testing data

Accuracy is defined as the percentage of correctly matching the actual categories of anomalies learnt or discovered by the NuPIC network against the actual categories of anomalies passed to it in a category file. This is hence a good measure of how well a NuPIC network performs.

4.2 Factors Affecting Performance

In these set of tests different HTM networks were built in-order to discover the factors that affect their performance.

4.2.1 Varying Number of Anomalies in Training Data

In this test case, the number of anomalies in the training data is varied incrementally from 1 up to 4 different types of anomalies. This is done for scenarios where specific anomalies are used and also where categories of anomalies are used. The table below outlines some of the anomalies that were used both for the specific anomaly scenario and also for the categorized anomaly scenario. Below are the names of some of the anomalies used in the study;

Table 1. Specific and Categorized anomalies

Specific Anomalies	**Categorized Anomalies**
HTTPTUNNEL, PORTSCAN, XLOCK, XTERM, SECRET	DOS, PROBE, U2R , R2L , DATA

The graphs below show the outcome of plotting the percent accuracy of training and detection against different numbers of anomalies in the training data.

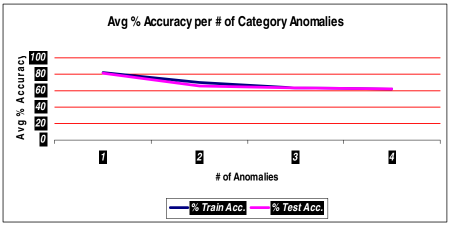

Fig. 1. Results for Categorized Anomalies

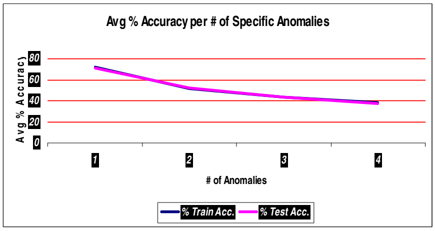

Fig. 2. Results for Specific Anomalies

The 2 lines in the graphs above refer to the percent accuracy obtained during training of the network (% train acc.) and the other line refers to the accuracy obtained with novel (unseen before) data during detection (% test acc.).

4.2.2 Varying Number of Learning Nodes in the HTM Network

In this test case the number of learning nodes per level in the HTM networks used as well as the number of levels of nodes was varied. Two general network designs were used in this case. The first design is a 3 layer network having 8 learning nodes at the first level, 4 learning nodes at the second level and 2 learning nodes at the third level. The second design is a 2 layer network having 4 learning nodes at the first level and 2 learning nodes at the second level. In all these designs the other factor that was also varied was the number of vectors or rows of values in each training file. HTM learning nodes accept integer vectors that it uses to build a model of its world [16]. Having done this each Network design was subjected to tests using a 32 vector input file as well as a 64 vector input file. This means the 32 vector training file had 32 columns while the 64 vector training file had 64 columns; the number of rows was kept constant per test.

Table 2. Test results for HTM networks

Anomaly	Test	set 1	set 2		anomaly	test	set 1	set 2
Secret	1	75.76142	72.26463		Secret	3	81.25	81.17048
Portsweep	1	75.892857	76.41026		portsweep	3	79.75	77.94872
Xlock	1	63.3587766	63.93862		Xlock	3	63.5	63.93862
Xterm	1	72.900763	74.93606		Xterm	3	81	78.77238
Mailbomb	1	73.7913486	73.91304		Mailbomb	3	75.5	75.19182
Httptunnel	1	68.298969	68.65285		Httptunnel	3	91.5	90.15544
Secret	2	76.14213	76.33588		Secret	4	80.75	79.13486
Portsweep	2	79.846939	80.51282		Portsweep	4	83.23	81.79487
Xlock	2	53.94402	54.4757		Xlock	4	68.25	68.28645
Xterm	2	62.595419	74.68031		Xterm	4	76.5	74.68031
Mailbomb	2	72.0101781	72.37852		Mailbomb	4	73	73.14578
Httptunnel	2	80.670103	82.38342		Httptunnel	4	88.75	87.56477

4.2.3 Measuring Time to Detect Anomaly in Test Data

The results presented here are for a series of tests to measure the amount of time taken to detect anomalies, using two network types as in the test above;

Table 3. Time Taken to Detect Anomaly

	2 Layer Network	3 Layer Network
# of Anomalies	Detection time (sec)	Detection time (sec)
1	0.85949996	1.023499932
2	1.7030002	2.358000135
3	2.24249998	3.625000057
4	3.18699984	2.898499315
Average	**1.997999995**	**2.47624986**

5 Conclusions

The conclusions arrived at in this study, after looking at all the experiments that were run and also, considering the results that were obtained are as follows:

NuPIC platform is suitable for solving the problem of Intrusion Detection; using either anomaly detection based approach or even misuse detection based approach can be accommodated. Nevertheless this method is most suitable for anomaly detection as it tends to perform best with a single anomaly training data scenario. This means that it is much easier to build an HTM based anomaly detection system that will be trained with a single data and be capable of making detections of any deviations from normal behavior as well as any modifications of some known anomalies. Furthermore the HTM networks tend to perform better if they are given the data representing a global view of a computer network.

Secondly the results indicated that training a network with a single pattern led to more effective HTM networks. This means that HTMs are more suitable for anomaly detection based systems as opposed to misuse detection. The HTM network can be trained with normal state data and be used to detect any deviation from this norm as an anomaly.

Thirdly the design of a NuPIC network is very important as it affects how best a network will be trained and hence also its ability to make predictions and inferences. In addition to this, the input training file needs to have more vectors to ensure that each learning node receives sufficient data for it to learn the patterns in data easily and pass its belief distribution to the nodes above it.

Finally the trained NuPIC networks were capable of making detections in less than 3 seconds. This implies that HTM based IDS can be used for real-time detections. This is very important for real-time online IDS since some attacks are rather short and require very fast anomaly detection in order to secure data and other computing resources.

References

1. Ghosh, A.K., Schwartzbard, A.: A Study in Using Neural Networks for Anomaly and Misuse Detection. In: Proceedings of the 3rd USENIX Windows NT Symposium, Seattle, Washington, July 12-15 (1999)
2. Ryan, J., Lin, M.-J., Miikkulainen, R.: Intrusion Detection with Neural Networks. AAAI Technical Report, Vol. WS-97-07, pp. 72–77 (1997)
3. Cannady, J.: Artificial Neural Networks for Misuse Detection. In: Proceedings of the 1998 National Information Systems Security Conference (NISSC 1998), October 5-8, pp. 443–456 (1998)
4. Helmer, G.G., Wong, J.S.K., Honavar, V., Miller, L.: Intelligent Agents for Intrusion Detection, Iowa State University
5. Debar, H., Dacier, M., Wespi, A.: Towards a Taxonomy of Intrusion Detection Systems. Computer Networks 31, 805–822 (1999)
6. Bace, R.G.: Intrusion Detection. Macmillan Technical Publishing, Indianapolis (2000)
7. Amoroso, E.G.: Intrusion Detection: An Introduction to Internet Surveillance, Correlation, Traps, Trace Back, and Response. AT&T Laboratories (1999)

8. Paschalidis, I.C., Smaragdakis, G.: Spatio-Temporal Network Anomaly Detection by Assessing Deviations of Empirical Measures
9. Paschalidis, I.C., Smaragdakis, G.: A large Deviations Approach to Statistical Traffic Anomaly Detection
10. Allen, J., Christie, A., Fithen, W., McHugh, J., Pickel, J., Stoner, E.: State of the Practice of Intrusion Detection Technologies, Carnegie Mellon, Software Engineering Institute, Technical Report CMU/SEI-99-TR-028, Networked Systems Survivability Program (January 2000)
11. Numenta Inc., Getting Started with NuPIC, Document Version 1.2.1 (2008)
12. Numenta Inc., Advanced NuPIC programming, Document Version 1.8.1 (2008)
13. George, D., Jaros, B.: Numenta Inc., The HTM Learning Algorithms (2007)
14. Hawkins, J., George, D.: Numenta Inc., Hierachical Temporal Memory: Concepts, Theory and Terminology (2006)
15. Numenta Inc., Hierarchical Temporal Memory: Comparison with Existing Models, version 1.01 (2007)
16. Numenta Inc., Problems that Fit HTM, version 1.0 (2007)
17. Koutsoutos, S., Christou, I.T., Efremidis, S.: A Classifier Ensemble Approach to Intrusion Detection for Network Initiated Attacks. In: Emerging Artificial Intelligence Applications in Computer Engineering. IOS Press, Amsterdam (2007)
18. Mukherjee, B., Heberlein, L.T., Levitt, K.N.: Network Intrusion Detection. IEEE Network (1994)
19. Lee, W., Stolfo, S.J., Chan, P.K., Eskin, E., Fan, W., Miller, M., Hershkop, S., Zhang, J.: Real Time Data Mining-based Intrusion Detection
20. Haines, J.W., Lippmann, R.P., Fried, D.J., Zissman, M.A., Tran, E., Boswell, S.B.: DARPA Intrusion Detection and Procedures, February 2001, Technical Report 1062 (1999)