Computer Network Monitoring Based on Iterative Dichotomiser 3 Approach

Prakai Nadee[1] and Chaiwat Dangchuntuk[2]

Received: October, 2015; Accepted: February, 2016

Abstract

The Network monitoring is an important part in ensuring computer network security for protecting network infrastructures. The problem of network system is very hard on generating the event occurring in a computer network and analyzing them for signs of anomalous traffic. This paper proposed an efficient technique of network monitoring system using the Nagios Application with Iterative Dichotomiser 3 (ID3) for monitoring and alerting of anomaly traffic detection. In addition, we improved the Multi - Router Traffic Grapher (MRTG) used to report traffic and manage devices based on the Simple Network Management Protocol (SNMP). As a result, it has shown that the proposed technique can improve performance of network security and network management. The Nagios application can produce information that defines signatures for anomaly detection.

Keywords: Nagios Application; Decision Tree; Multi-Router Traffic Grapher; Simple Network Management Protocol

[1] *Faculty of Engineering and Architecture, Rajamangala University of Technology Isan*

[2] *Office of Academic Resoureces and Information Technology, Rajamangala University of Technology Isan*

E - mail : chaiwat.d@rmuti.ac.th

Introduction

Nowadays, service computing has become an emerging science that is highly regarded as a necessary technology for the network communication system. The protective lines of attacks on computer networks are serious problems, because most deployed computer systems are vulnerable to those attacks. Computer network's security becomes a critical issue and it is important to develop mechanisms as a defense against the intrusions. Network monitoring is an important technology for the business sector in its effort to build systems for network security. The basic definition of monitoring a system is to see and check the process of operating system or application (Katsaros, G. et al., 2011).

There are many numbers of proposed network monitoring products in the market, both commercial and open – source. Commercial products usually offer comprehensive features, but they also cost a lot. Open – source products are with no cost, but they usually have some limitation, such as limited number of devices and services for monitoring, and most importantly no technical support. Examples of well-known open-source for network monitoring tool are Nagios, Zenoss, Groundwork, OpenNMS and Hyperic (Issariyapat, C. et al., 2012). This paper proposed the conceptual design of Nagios Application for traffic monitoring in a computer network. The newly designed hyper-map method shows host-group status and network/service links on web-based interface. These proposals offer an ID3 for traffic analysis, MRTG drawing tool used to graph all sorts of network devices and SNMP applied to managing devices on an IP network. It is an efficient technique of improved a computer management method.

The rest of this paper is organized as follows. In Section 2, we investigate background and related works of network monitoring. Section 3 explains the experimental design and setup of the framework. Section 4 presents the experimental results. Finally, discussions and some conclusions are Section 5.

Background and Related Work

Sun, H. (Sun, H., 2010) network Performance Monitoring has now become a central issue in network application and operation optimization. Recently, the application tool of network monitoring system has been of greatest significance for network security protection. Completely, detection and response are the two most important of the security processes. Detection is an extraction of specific information from a larger stream of information without specific participation with the sender. Generally, the monitoring system includes many main modules to evaluate the network performance. The performance function

is analyzing the sample data of network packets and evaluating the performance according to the rules as acquired by the flag of parameter changed by the network (Wei, X. et al., 2009). In our monitoring system, we thus propose a new technique that uses the Nagios Application to offer the network monitoring and analyze them for signs of nomalous traffic. The following section briefly states the Nagios application, Decision tree and evaluation outlined.

Nagios Application

The Nagios application has two major parts in the monitoring process, a passive and active monitoring. Active monitoring mechanism is a process the client takes to install the program on a computer being monitored. The client connects to a server that constantly controls the monitoring functions. These functions can range from system level information, disk space and memory usage to service information. Passive monitoring mechanism is a semi - timed query for the service to see, if it returns an expected response. It is processed to identify malicious packets and define the data on which routers travel through the computer network traffic (Murphy, J.W. et al., 2008). The Nagios is a program for monitoring and alerting again when the problem has been resolved. It is an open-source framework for monitoring the network host and service with the purpose of the failure detection. There are 3 stages of processing service performance scheduling. The first stage is the service state check, second stage is the host state check. Lastly, the parent state check of network status based on decision tree analysis (Imamagic, E. and Dobrenic, D., 2007), as shown in Figure 1.

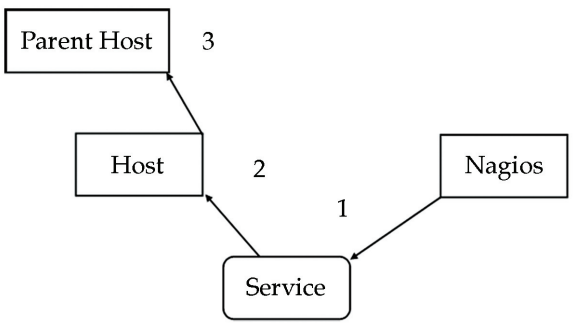

Figure 1 Nagios application architecture

Decision Tree

Tran, K.N. and Jin, H. (Tran, K.N. and Jin, H., 2010) describes the decision tree as a predictive model that classifies an item based on machine learning. It is built from a training data set that results in a mapping from the independent variable to a dependent variable. The decision tree "rules" are summarized into a tree structure based on the conditions

defined by the rule. The tree has three types of nodes: Root, internal and leaf nodes. The root node contains nodes and edges. Each internal node is a test node which corresponds to one of the input variables. There are edges, leaving a node for every possible values that input variable (Katz, G. et al., 2014). Finally, the leaf nodes of the tree contain the expected class value of transactions matching the path from the root to the leaf as shown in Figure 2.

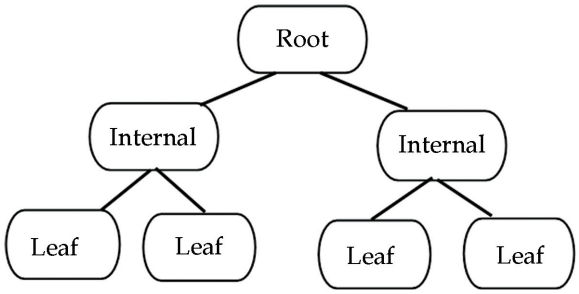

Figure 2 Decision tree architecture

Evaluation

Tubnakog, S. (Tubnakog, S., 2007) measured the performance of network monitoring and evaluated it with respect to detecting change in the patterns of behavior on the system. The effect of these flag's limits was evaluated based on the veritable behavior and predicted classifications done by a classification system. The performance of such a system is commonly evaluated using the data in the matrix, shown in Table 1.

Table 1 The matrix for performance evaluation

		Predicted Class	
		Yes	No
Actual Class	Yes	True Positive	False Negative
	No	False Negative	True Negative

TP. (True Positive): is the number of prospects predicted as being responses which are actually responses.

FN. (False Negative): is the number of prospects predicted as being non-responses, which are actually responses.

FP. (False Positive): is the number of prospects predicted as being responses, which are actually non-responses.

TN. (True Negative): is the number of prospects predicted as being non-responses, which are actually non-responses.

Most widely-used matrix is accurately defined by the following formula, defined as equation (1) (Tubnakog, S., 2007).

$$Accuracy = \frac{TP}{TP + TN + FP + FN} \quad (1)$$

Experimental Design and Setup

We developed a new monitoring framework by using the object - oriented Application Programming Interface (API). Given hybrid HTML, PHP and Java script for accessing data from Nagios application. It is installed on the target machine and executes commands that supported the server architecture. The overall architecture is shown in Figure 3.

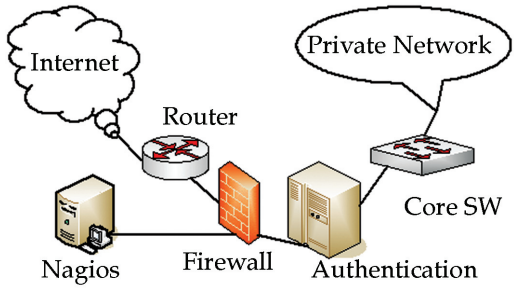

Figure 3 Overall architecture of system

Then the application is the process of collecting, identifying, verifying and escalating suspicious events of Figure 4. It has traditionally been the heart of the reasoning behind deploying to intrusion detection systems.

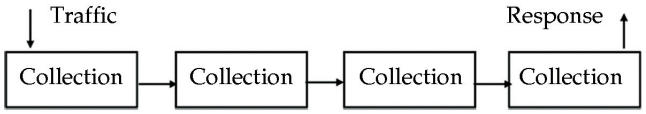

Figure 4 The detection process (Richard Bejtlich, 2005).

Collection: The process begins with all traffic. Once the sensor performs collection, it outputs observed traffic to the analysis. With respect to the full content collection, the data is a subset of all the traffic the sensor sees.

Identification: The analyst performs identification on the observed traffic, judging it to be normal, suspicious, or malicious. Overall event is sent to the next stage.

Verification: The analyst categorizes the event into one of several incident categories. Verification produces indications and warning.

Escalation: The analyst forwards incidents to decision makers. Incidents contain actionable intelligence that something malicious has been detected.

Traffic Analysis

Suthampan, A. (Suthampan, A., 2005) proposed the decision tree learning technique that applied the Iterative Dichotomiser 3 (ID3) to network traffic analysis. It is a predictive model that predicts the value of a target variable based on several input variables. Given R is a set of attributes that have not been associated with a node, C is the class attribute, and T is a set of transactions. ID3 method aims to classify transactions one of possible categories; assumes that each attribute is categorical, that contains discrete data only, in contrast to continuous data. The tree is simply constructed top - down in a recursive function. The best attribute is then chosen from feature values that are information gainers. The ID3 algorithm is iterative and stated as follows:

Step 1: If R is empty, return a leaf - node with the class value of the majority of the transaction in T.

Step 2: If T contains transactions with the same value C for the class attribute, return a leaf - node with the value C to finish the classification path.

Step 3: Otherwise,

Step 3.1: Find the attribute that best classified transaction in T, let it be A.

Step 3.2: Let $a_1,...,a_m$ be the values of attribute A and let $T_{a1},...,T_{am}$ be partitions of T such that every transaction in T_{ai} having the attribute a_i value.

Step 3.3: Return the tree whose root is labeled A of the test attribute and has edges labeled $a_1,...,a_m$ such that for every i, the edge a_i goes to the tree $ID3(R-\{A\},C,T_{ai})$

The exact test for determining the best attribute is defined as letting $c_1,...,c_n$ be the class attribute values. Let $T(c_i)$ be the set of transactions with c_i class. Then information needed to identify the class of a transaction in T is the entropy, defined as Equation (2).

$$Entropy(T)=\sum_{i=1}^{n}\frac{|T(c_i)|}{|T|}\log\frac{|T(c_i)|}{|T|} \tag{2}$$

ID3 can predict the class of a new transaction T as the attributes of a root node be A, where A obtains possible values $a_1,...,a_m$. Then the m edges, leaving root node are labeled $a_1,...,a_m$ respectively. Let A be a non - class attributes that improve to quantify the information needed to identify the class of a transaction in T given that

the value of A has been obtained. Let A obtain value $a_1,...,a_m$ and let T_{ai} be the transactions, obtaining value a_i for A. Then used for conditional information of T given A, defined as Equation (3).

$$Entropy(T|A) = \sum_{j=1}^{m} -\frac{|T_{aj}|}{|T|} Entropy(T_{aj}) \quad (3)$$

For attribute A is the information gain, as defined by Equation (4).

$$Gain(T, A) = Entropy(T) - Entropy(T|A) \quad (4)$$

The attribute A which has the maximum gain over all attributes in R is then chosen (Li, H. and Zhijian, L., 2010).

Traffic Report

The main goal of this method is to operate the computer network traffic and notification systems to configured device for alerting to the network administrator. Moreover, overall information sends email alert and SMS notification to the network administrator for taking further action to maintening the Quality of Service (QoS) in the network, shown in Figure 5 (Bin Mohd Shuhaimi, M.A.A. et al., 2011).

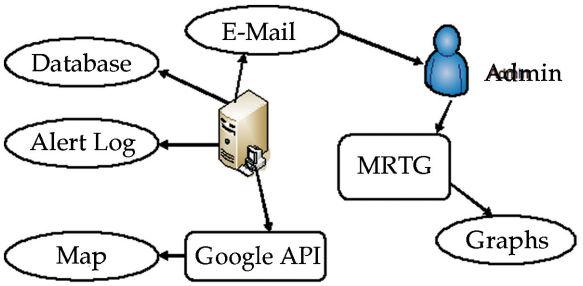

Figure 5 Monitoring and alerting process

Certainly, we apply a new report tool by using the Multi - Router Traffic Grapher (MRTG) to generate HTML suitable for display in a web browser. The MRTG provides long-term network usage statistics. Based on such observation is tight linked for estimating the available bandwidth that is specifically designed of a computer network environment (Xing, X. and Mishra, S., 2009). The Simple Network Management Protocol (SNMP) on the network device is for managing devices on IP networks, because, MRTG works by polling SNMP-enabled network device for interface statistics. Replace public and private in configurations with community string which is not easily guessed by an intruder.

SNMP is a method of managing, such as server computers, routers, and switch from a centrally located computer running network management software. This service is an optional feature that is installed after the TCP/IP protocol has been successfully configured (Salvador, E.M. and Granville, L.Z., 2008).

Experimental Results

New conceptual design, new hyper map, ID3 offer and summaries of the new design and test results of network monitoring were responsible for addressing content and processing of the map service on a web page. There were 3 active status of network connection, up, warning and down status. Up status is a normal connection or established connection warning status is error connection, network traffic is being collected from midstream and down status is a lost connection.

When a connection cannot establish, such as request time out, connection failed, system has created the action report from the monitoring system. Overall detail of the problem event is sent via email to the network administrator. Each traffic information report, date, month and year traffic, administrators take a Quality of Service (QoS) of bandwidth, as shown in Figure 6.

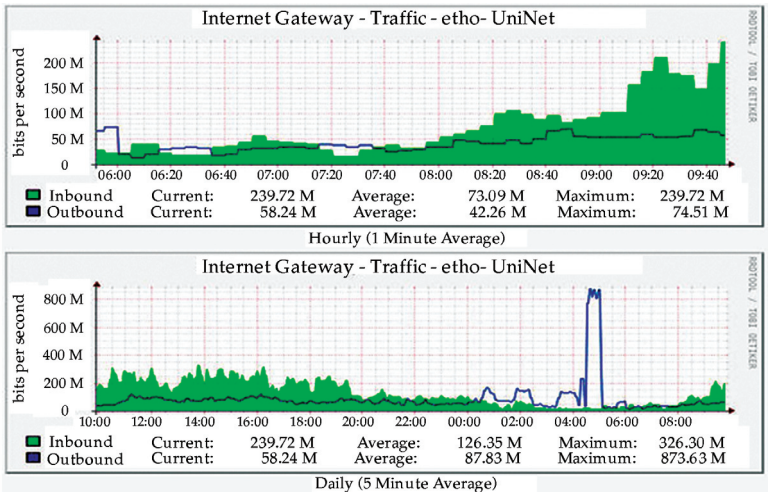

Figure 6 Information of traffic report

Each activity is a record of everything that goes in and out of a particular server. It essentially records everything that goes on within the server and in the event, usage statistics for log - analysis. Successfully implementing a security process requires maintaining a network capable of being defended. The administrator has a chance of detecting intruder who communicates with system compromise.

Conclusion

This paper has presented a new design method and approach network monitoring. There are implementations of anomaly traffic that improved hybrid method, Nagios, MRTG, and SNMP for Managing device. It is a process of maintaining an acceptable level of perceived risk. The aim is to minimize the risk, defenders must be vigilant by implementing assessment, protection, detection and response procedure. As a result, the proposed method showed a high performance of network security and network management system. Countermeasures can also be applied against computer network threat. They can act against the offending party capabilities and intentions. It is the collection, analysis and escalation of indications and warnings to detect and respond to intrusion. In the future, we plan to apply other theories and techniques of computer network security in our work.

References

Bin Mohd Shuhaimi, M.A.A., Binti Roslan, I., Binti Zainal Abidin, Z. and Binti Anawar, S. (2011). The new services in Nagios: Network bandwidth utility, email notification and sms alert in improving the network performance. In **Proceeding of 7th International Conference on Information Assurance and Security**. pp. 86-91

Imamagic, E. and Dobrenic, D. (2007). Grid Infrastructure Monitoring System Based on Nagios. In **Proceeding of the 2007 workshop on Grid monitoring**. pp. 23-28

Issariyapat, C. Pongpaibool, P., Mongkolluksame, S. and Meesublak, K. (2012). Using Nagios as a groundwork for developing a better network monitoring system. In **Proceeding of Technology Management for Emerging Technologies**. pp. 2771-2777

Katsaros, G., Kü bert, R. and Gallizo, G. (2011). Building a Service-Oriented Monitoring Framework with REST and Nagios. In **Proceeding of IEEE International Conference on Services Computing**. pp. 426-431

Katz, G., Shabtai, A., Rokachm, L. and Ofek, N. (2014). A statistical method for improving decision trees. **Journal of Computer Science and Technology**. Vol. 29. No. 3. pp. 392-407

Li, H. and Zhijian, L. (2010). The Study and Implementation of Mobile GPS Navigation System Based on Google Maps. In **Proceeding of the International Conference on Computer and Information Application**. pp. 87-90

Murphy, J.W. (2008). **SnoScan: An iterative functionality service scanner for large scale networks**. Master Thesis, In Graduate school of sciences, Iowa State University: Ames, Iowa, p. 45

Richard Bejtlich. (2005.). **The Tao of Network Security Monitoring Beyond Intrusion Detection**. Addison Wesley. Pearson Education Inc.

Salvador, E.M. and Granville, L.Z. (2008). Using Visualization Techniques for SNMP Traffic Analyses. In **Proceeding of IEEE Symposium on Computers and Communications**. pp. 806-811

Sun, H. (2010). An Integrated Network Performance Monitor System. In **Proceeding of the Third International Symposium on Intelligent Information Technology and Security Informatics**. pp. 88-91

Suthampan, A., (2005). **Privacy Preserving Decision Tree in Multi Party Environment**. Master Thesis. In Faculty of Engineering. King Mongkut's University of Technology Thonburi, County: Thailand - Bangkok

Tran, K.N. and Jin, H. (2010). Detecting Network Anomalies in Mixed-Attribute Data Sets. In **Proceeding of the Thied International Conference on Knowledge Discovery and Data Mining**. pp. 383-386.

Tubnakog, S., (2007). **Risk modeling using decision tree algorithm for voluntar motor insurance**. Master Thesis. In Graduate School of sciences, Mae Fah Luang University, County: Thailand-Chiang Rai.

Wei, X., Wu, W. and Liu, Y. (2009). A Network Monitor System Model with Performance Feedback Function. In **Proceeding. of International Conference on E-Business and Information System Security**. pp. 1-5

Xing, X. and Mishra, S. (2009). Where is the Tight Link in a Home Wireless Broadband Environment. In **Proceeding of IEEE International Symposium on Modeling**. Analysis & Simulation of Computer and Telecommunication Systems. pp. 1-10