EasyChair Preprint

№ 7115

Malware Detection Using Network Traffic Analysis and Predicting Accuracy Using Deep Learning Algorithms

Siddhant Khurana, Satwik Dash, Navneet Kumar Shiva and Rajat Vashist

November 28, 2021

"Malware Detection using Network Traffic Analysis and Predicting Accuracy using Deep Learning Algorithms."

Siddhant Khurana, Satwik Dash, Navneet Kumar Shiva, Rajat Vashist, *Member, IEEE*

Abstract— In this project, We will be analyzing malicious activity in our network like botnets, DDOS attack, SQL injection & erroneous packets in our Network traffic generated & analyze it using invaluable tools that allow for applied experimentation to find & calculate the working & performance of our networks, the infrastructure of our networks and the security preventive measures, by simulating and modelling the data packets and the payloads of those packets that would be generated by machines & devices on the network infrastructure like packets capturing & analysis using Wireshark. mainly for the secure & private applications, these networking tools shall be used to fluently simulate any kind of malicious or fraudulent activity on the network devices and testing the components that are designed & structured to mitigate ad detect the malicious activities, in a highly customizable and reliable way. The prediction and accuracy of performed results particularly depends on the reliability and performance of the used network traffic generator. So, here we will simulate & investigate the accuracy and performance of different network traffic tools which are most reviewed network traffic generators, namely Wireshark, Ostinato, Genesids and Cisco Trex. Most importantly, this analysis helps to examine & test the limitations and strengths of these networking tools, for any kind of bogus and malicious traffic. After the Analysis of this traffic, we will visualize the data and work with data sets trace files using to generate graphs using deep learning ANN algorithm to predict the accuracy of our analysis done for people to choose the best way of avoiding any kind of malicious activity in our network.

I. Introduction

Data is floating all over the internet. This information is transferred in the form of data packets and these routed packets are communicated through various protocols. The Intruders always try to perform attack in the network which leads to data breach. therefore, analysing of network traffic is of utmost importance. Analysing this traffic involves monitoring availability of network and activities to identify the anomalies & remove rogue attackers from the network to promote a safe & secure environment in the environment.

Some Common use case of network traffic analysis includes fetching the real-time and historical record of what's happening on the network, detecting malware such as ransomware activity, Detecting the use of vulnerable protocols and ciphers, troubleshooting a slow network, improving internal visibility, and eliminating blind spots. With such great importance different tools are used for Analysing the network traffic like Wireshark, Cisco Trex, Genesids etc.

These tools help in determining any such kind of attacks like botnets, DDOS or for detecting any erroneous packets in the network. These tools generate network logs which could be used to test for any known attacks so that it could be detected and removed from the network. This analysis of network traffic ensures the individual working in a safe and secure environment.

This analysis is done by setting up a command-and-control server. Ping blasts are performed to generate logs. This comparative study of tools also includes predicting its accuracy using working on different datasets that are generated from the tracer log file and its graphical representation for true and easy analysis of data packets.

Virtual machines are to be initialized for setting up the environment of client and server on different

machines using NAT address. These series of events are followed by ping the server from client and checking for incoming packets. Hence IP packets are analysed for results and further data visualization is performed.

Initial configuration and flow of project

1. Research and analyze the requirement for setting up the cloud environment for network attack.
2. Setting up of virtual machines and downloading the necessary tools.
3. Setting up of the command and control (C&C) server.
4. Attaching the victim Machines with our C&C server.
5. Testing the connection by a small ping & launching of ping blasts on the local server.
6. Extracting and gathering of logs & analyzing the fetched data using deep learning algorithm for predicting accuracy of fetched result.

Tools Used

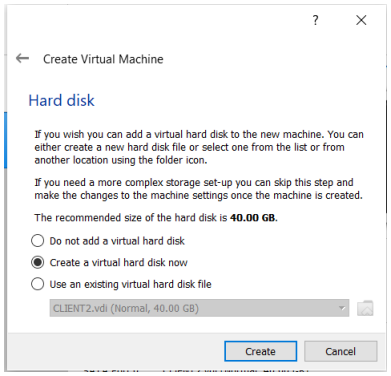
1. Windows 10/Mac/Linux
2. Wireshark, Cisco Trex, Ostinato & Genesids
3. Virtual Cloud machines (Oracle Virtual Box)
5. Deep Learning Visualization Tools

II. SETTING UP VIRTUAL MACHINES (ORACLE VIRTUAL BOX)

For Creating a newly virtual machine, first step is to start VirtualBox. Onto the host we installed and VirtualBox and Oracle VDI, click on **Applications** menu on the VirtualBox desktop, thereafter, click on **System Tools** menu option, and then on **Oracle VM VirtualBox**. Parallelly, we dry run out the **VirtualBox** command in a terminal.

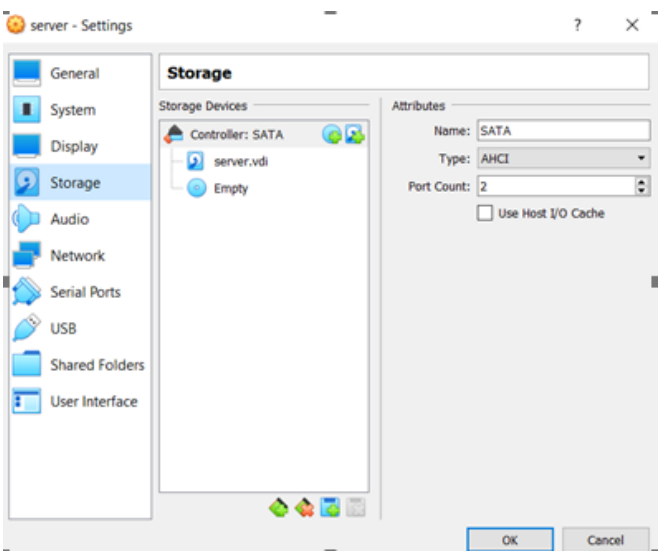
Click the Next button to traverse though the diverse steps of the wizard. The wizard permits you to configure the simple information of the digital system. On the VM Name and OS Type step, input a descriptive call for the digital system withinside the Name subject and pick out the working gadget and model that you're going to put in from the drop-down lists. It is vital to pick out the precise working gadget and model as this determines the default settings for VirtualBox makes use of for the digital system. We can further change to the desired settings later after you've got created the digital virtual system.

On the Memory step, we can merely settle for the default. this is often the quantity of host memory (RAM) that VirtualBox assigns to the virtual machine once it runs. you'll be able to amendment the settings of the virtual machine later once you import the model into Oracle VDI. On the Virtual disc step, guarantee Start-up Disk is chosen, select produce new hard disk and click on Next. The Virtual Disk (VD) Creation Wizard is showed **in a very** new window **thus** you can create and install the new virtual disk.



On the subsequent steps, pick out VDI (VirtualBox Disk Image) just as it is like the document type, dynamically allotted because the garage details, and take delivery of the defaults for the digital disk document region and size, after which click on Create to create the digital disk. When the digital disk is created, the Virtual Disk Creation Wizard is closed and you're back to the Summary step of the New Virtual Machine Wizard. Click Create to create the digital gadget. The wizard is closed and the newly created digital gadget is indexed in Oracle VM VirtualBox Manager as proven in image.

Since you would like to put in an software package within the virtual machine, you wish to form certain the virtual machine can access the installation media. to try to to this, you'll be able to edit the virtual machine settings. In Oracle VM VirtualBox Manager, choose the virtual machine so in the toolbar click the **Settings** button. The Settings window is displayed. within the navigation on the left, select **Storage.**



Click on **OK to** apply the storage settings. The Settings window is closed. If you connected the virtual machine' CD/DVD drive to the host' physical CD/DVD drive, insert the installation media **within the** host' CD/DVD drive **currently**. we tend to are now able to begin the virtual machine and install the operational system.

ATTACKS IN CYBERSECURITY

"Attack is Any **quite** malicious activity that **makes an attempt** to collect, disrupt, deny, degrade, or destroy **data system** resources or **the data** itself." Cybercriminals use **totally different strategies** to launch a cyberattack **that features** malware, phishing, ransomware, man-in-the-middle attack, or **alternative** methods.

CATEGORY OF ATTACKS

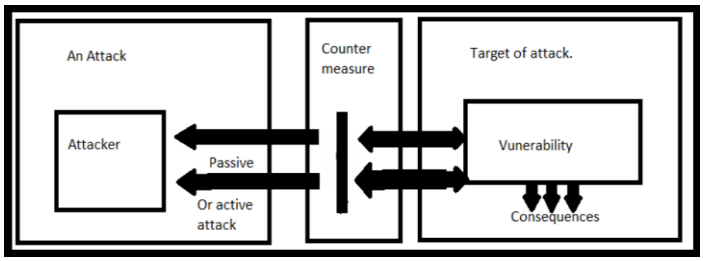
• **Active Attack**: - associate degree "active attack" tries to change system resources or influence their operation

• **Passive Attack**: - A "Passive Attack" attempts to find out or build use of knowledge from the system however doesn't affect system resources.

Associate degree attack may be perpetrated by a business executive or from outside the Organization:

• associate degree "**inside attack**" is an attack initiated by an entity within the protection perimeter (an "insider"), i.e., an entity in a system that's qualified to grant access to system resources but they uses them in a particular way which is not approved by those that have granted the authorization.

• associate degree "**outside attack**" is particularly initiated from generally outside of the scope or organization, by an (unauthorized/ illegitimate) user in a pc (an "outsider"). On the Internet, the potential outside attackers vary greatly from amateur pranksters to being the organized and expert criminals, international terrorists & hostile governments agencies**.**



A resource (both physical or logical), referred to as an asset, can have one or additional vulnerabilities that {may} be exploited by a threat agent in a very threat action. As a result, the confidentiality, integrity or convenience of resources is also compromised. Potentially, the injury may be resources additionally to the one at the start known as vulnerable, as well as any resources of the organization, and therefore the resources of alternative concerned parties (customers, suppliers). Central Intelligence Agency triad is that the basis of data security.

TYPES OF CYBER ATTACK

- **Denial-of-service attack:**

A denial-of-service attack fills systems, servers, or networks with traffic that exhaust resources and bandwidth. that creates the system incapable to fulfil legitimate requests. Attackers additionally use multiple compromised devices to launch this attack. this can be referred to as a distributed-denial-of-service (DDoS) attack.

- **SQL injection**:

  A Structured command language (SQL) injection happens once an wrongdoer inserts malicious code into a server that uses SQL and forces the server to reveal info it unremarkably would not. AN attacker may perform a SQL injection just by pushing a infectious code into a potentially infected web site search box.

- **Phishing:**

  Phishing is that the technique of causation dishonest communications that appears to return from a respected source, sometimes through email. The goal is to steal or get sensitive information like MasterCard and login info or to put in malware on the victim's machine. Phishing is an progressively common cyberthreat.

- **Man-in-the-middle attack:**

  Man-in-the-middle (MitM) attack, also popularly known as eavesdropping attack. This attack occur once attackers/intruder insert themselves into a two party gateway transaction. Once the attackers is successful in interrupting the traffic, they will filter out and breach data.

Two common points of entry for MitM attacks:

On unsecure public Wi-Fi, offenders will insert themselves between a traveller's device and also the network.

INTRODUCTION TO KALI LINUX

Mati Aharoni and **Deavon Kearns** are the core developers of Kali UNIX system. it had been a rewrite of return Linux, that was another penetration testing centrical Linux distribution. it's a superbly crafted OS that especially serves to especially the wants of penetration testers & also to the community analysts. Here, presence of a inordinateness of substances that come back pre-mounted with OS Kali Linux changes it into ethical hacker's tool alike of a Swiss knife. Kali Linux is especially used for and Security Auditing & generally used for advanced Penetration Testing. Kali contains many hundred tools which are double-geared towards varied info security tasks, admire Penetration Testing, Security research, pc Forensics and Reverse Engineering Some tools that are trending involves Wireshark, Air crack-ng, Nmap, Nessus etc…

ADVANTAGES OF USING KALI LINUX

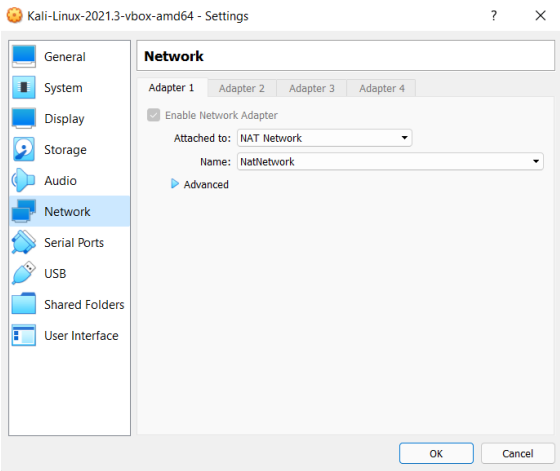
- **As free because it will get**: - Kali UNIX system has been and can perpetually be liberal to use.
- **A lot of tools than you'll suppose of**: - Kali Linux operating system generally has over 600 completely different security analytics connected tool. And penetration testing
- **Open-source**: - since Kali, belongs to a part of the Linux family, it chases the wide appreciated ASCII text file simulates Their growth and dev. tree is out in public seeable on so-and-so and everyone the code is out there for analyzing purposes.
- **Multi-language Support**: - though penetration tools tend to be written in English, it's been ensured that Kali includes true polyglot support, permitting a lot of users to work in their linguistic communication and find the tools they have for the job.
- **Fully Customizable**: - The developers at offensive security.

INTRODUCTION TO WIRESHARK

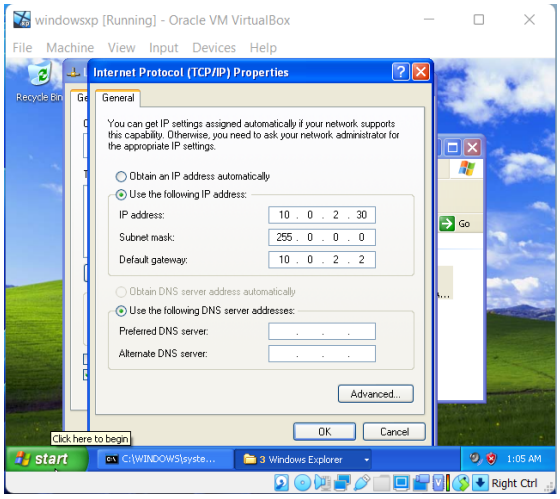
Wireshark could be a free and ASCII text file packet analyzer. it's used for network troubleshooting, analysis, software package and protocol development, and education. Discovered originally as Ethereal, this great project was renamed as Wireshark in the year 2006 because of trademark issues. Wireshark is cross-platform, victimisation the Qt contrivance toolkit in current releases to implement its user interface and using pcap to capture packets; it runs on Linux, macOS, BSD, Solaris, another Unix-like operative systems, and Microsoft Windows. there's conjointly a terminal/console dependent (non-GUI) version referred to as TShark. Wireshark, and therefore the other programs distributed with it comparable to TShark, are free software, free below the terms of the wildebeest General Public License version two or any later version.

SIMULATION OF DDOS ATTACK

1. Connect two virtual machines (Kali Linux and Host) by changing NAT to NAT Network.



2. Change the network setting in host machine. Set default gateway & IP-address of host computer. Also, Switch Off firewall on machines.

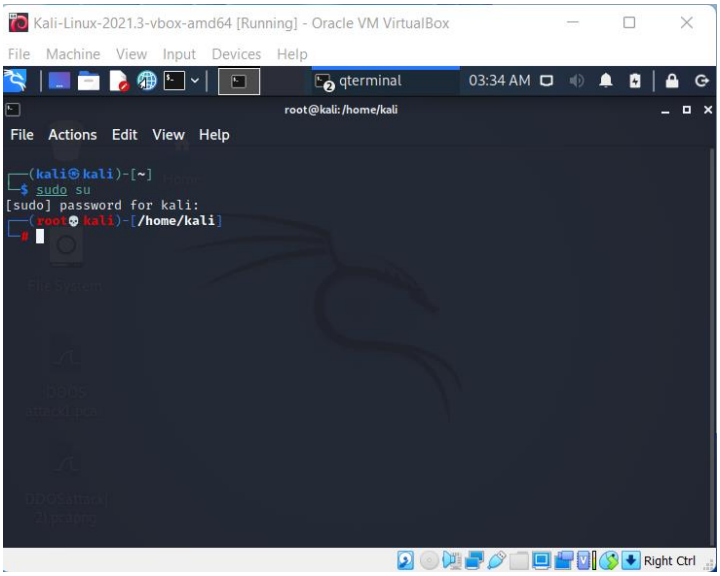


In DDOS attack ping flood (ping death) is performed on the host PC which leads to max CPU utilization of the resources and sometimes system may even go to halt state. To achieve this, we use Metasploit Framework which is an essential and most important tool for finding out any hidden vulnerabilities/bugs using a different tools and utilities.

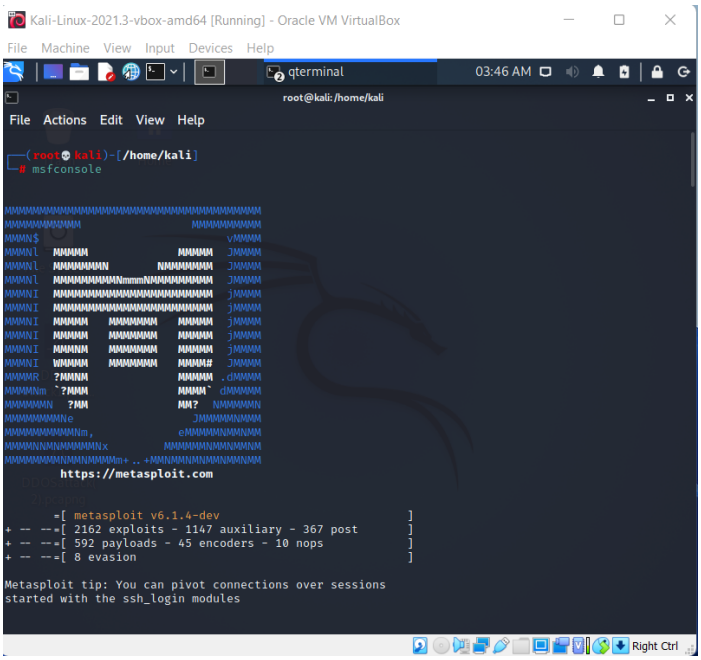
:

3. In Attacker (Kali Linux) terminal Emulator, we entered the root directory of Kali using #sudo su command and reconfigure the system to password less privilege escalation for adding "sudo" group to "Kali-Trusted" group.

Most of the attacks in Kali Linux begins with entering in its root directory,
Which can cause changes in the OS and PC may be at risk therefore Kali Linux is preferred by ethical hackers to simulate. such attacks.

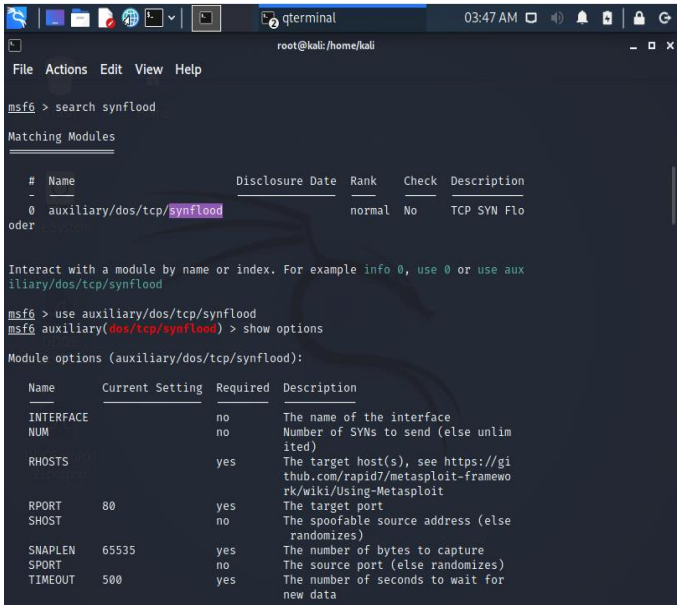


4. After configuration, we enter to Metasploit console using #msfconsole.
Metasploit comes with inbuilt features to find hidden vulnerabilities in a System.



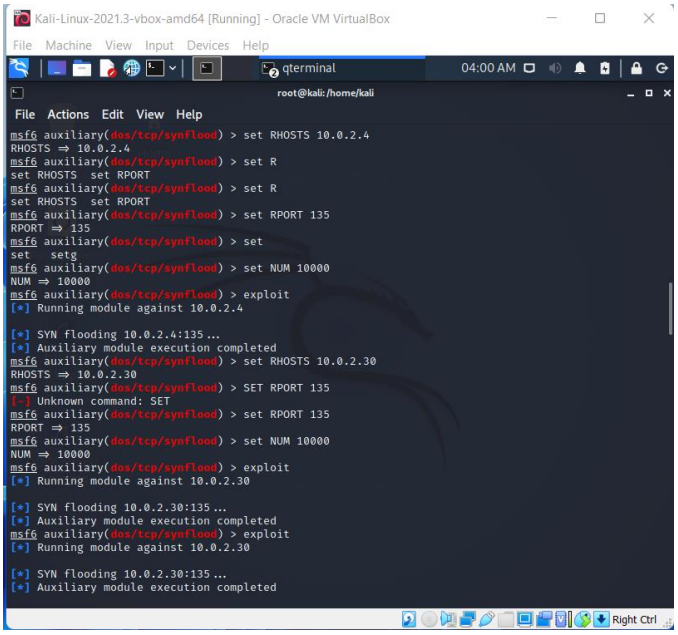
5. In Metasploit console, we determine appropriate option to search for sync flood and choose option auxiliary/doc/tcp sync flood.

SYN flood attacks work by exploiting the handshake process of a TCP connection. The server afterwards then responds to this initial packet with a (SYN / ACK) packet, to confirm the communication. Finally, the client then sends back an ACK packet to acknowledge receipt of the packet from the server.



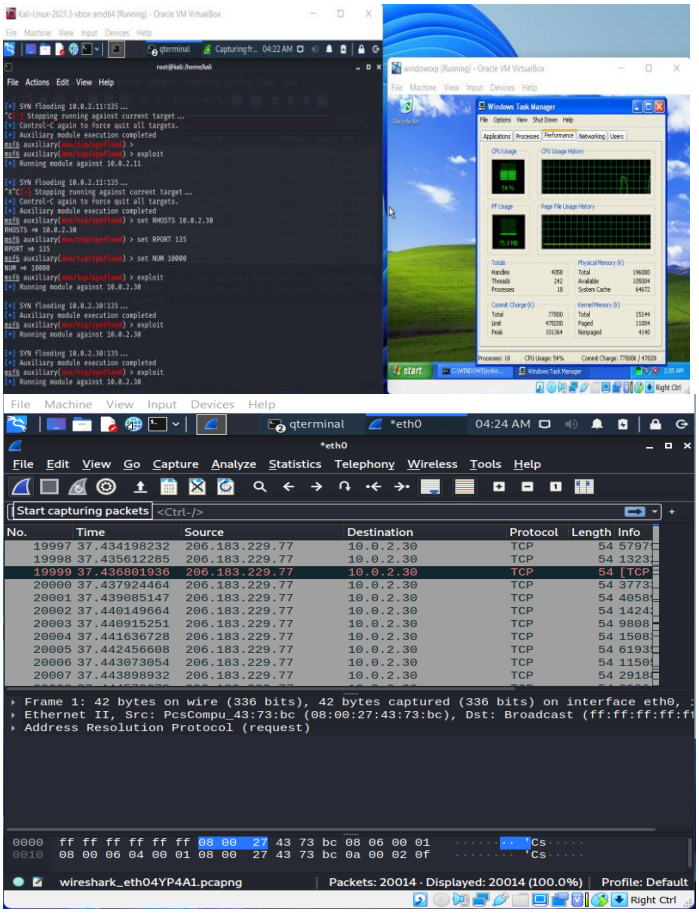
6. Now we look up for target host and set RHOST, RPORT, NUM to perform a DOS attack on the host computer.

- RHOST will look up for target Host.
- RPORT will refer to a Target Port.
- NUM will set the total number of SYN packets.
- Exploit will start performing the action desired on the target.



7. Before performing Exploit command on the host, Open Wireshark and start Capturing the packets for network traffic analysis.

Open Task manager to look up for spike in CPU Utilization and memory consumption when exploit operation is performed by attacker.



III. CONCLUSION

The spike in CPU utilization of the host depicts those resources have been consumed and DOS attack has been performed on the system and all the TCP packets have been traced in Wireshark and trace file of the attack simulation has been saved for future use.

IV REFERENCES

[1] S Gowrishankar, A Time Series Modeling and prediction
[2] of wireless Network Traffic (Georgian Electronic
Scientific Journal: Computer Science and
[3] Telecommunications,2008) |No.2(16).
[4] 19. Y.Yu, M. Song, Z. Ren, l. Song, Network Traffic
[5] Analysis and Prediction Based on APM (IEEE, 2011)
978-1-4577-0208-2/11.
[6] 20. N.Sadek, A. Khotanzad, Multi-scale High Speed
Network .
[7] Traffic Prediction Using K-Factor Gengendaue ARMA
Model (IEEE, 2004) 2148-2152.
[8] D. Zeng, J. Xu1, J. Gu , L.Liu , G. Xu, Short Term Traffic
[9] Flow Prediction Using Hybrid ARIMA and ANN model (
[10] IEEE, 2008) 978-0-7695-3342-1.
[11] W. Peng1 ,L.Yuan, Network Traffic Prediction Based
[12] on Improved BP Wavelet Neural Network (IEEE,2008)
[13] 978-1-4244-2107-7.
[14] L.J fei, S. Lei, T. Yongan, Prediction Of Network
[15] Flow Based On Wavelet Analysis And ARIMA Model (
[16] IEEE,2009) 978-0-7695-3901-0.
[17] 24. H.Zhao, Multiscale Analysis and Prediction of
Network.
[18] Traffic (IEEE, 2009) 978-1-4244-5737-3.

[19] Y.Yu, M. Song, Z. Ren, l. Song, Network Traffic Analysis and Prediction Based on APM (IEEE, 2011) 978-1-4577-0208-2/11.

[20N.Sadek, A. Khotanzad, Multi-scale High Speed Network Traffic Prediction Using K-Factor Gengendaue ARMA Model (IEEE, 2004) 2148-2152.

21. D. Zeng, J. Xu1, J. Gu , L.Liu , G. Xu, Short Term Traffic Flow Prediction Using Hybrid ARIMA and ANN model (IEEE, 2008) 978-0-7695-3342-1.

[22]W. Peng1 ,L.Yuan, Network Traffic Prediction Based on Improved BP Wavelet Neural Network (IEEE,2008) 978-1-4244-2107-7.

[23]L.J fei, S. Lei, T. Yongan, Prediction Of Network Flow Based On Wavelet Analysis And ARIMA Model (IEEE,2009) 978-0-7695-3901-0.

[24]H.Zhao, Multiscale Analysis and Prediction of Network Traffic (IEEE, 2009) 978-1-4244-5737-3.

[25]Y. Yu, J. Wang, M. Song, J. Song, Network Traffic prediction and result analysis based seasonal and ARIMA and Correlation Coefficient (IEEE, 2010) 978-1-4244-8333-4.

[26]Y.Zhou, Guangmin, H.W He, Using Graph to Detect Anomaly (IEEE, 2009) 978-1-4244-4886-9

[27]N.Gupta, N.Singh, V. Sharma, T. Sharama, A.S. Bhandra, Feature Selection and Classification of intrusion detection using rough set (International Journal of Communication Network Security, 2013) ISSN: 2231 – 1882, Volume-2, Issue-2.

[28]A.R Syed, A.S.M Burney, B. Sami, Traffic Forecasting Network Loading Using Wavelet Filter and seasonal Autoregressive Moving Average Model (International Journal of Computer and Electrical Engineering, 2010) Vol.2, No.6.

[29]Y. Shu, M.Yu, J. Liu, O.W.W. Yang, Wireless Traffic Modeling and Prediction Using Seasonal ARIMA Models (IEEE, 2003) 0-7803-7802-4.

[30]M. F. Iqbal, L. K. John, Power and Performance Analysis of Network Traffic Prediction Techniques (IEEE, 2012) 978-1-4673-1146-5/12.