International Journal of Innovative Technology and Exploring Engineering (IJITEE)
ISSN: 2278-3075 (Online), Volume-8 Issue-12, October 2019

A Network Forensic Framework for Port Scan Attack based on Efficient Packet Capturing

Rajni Ranjan Singh Makwana, Deepak Singh Tomar

Abstract*: In the last two decades the networks become larger in scale, more complex in structure and more diversified in traffic. Which generate huge amount of network packets such as TCP, UDP and HTTP etc. Log files are repository for captured packets and play a vital role in investigation. However, a significant and obvious limitation of current packet logging is that, data storage volume increases rapidly depending on factors such as network bandwidth and the number of points in the network that need to be tapped. Forensic investigator needs to back up these recorded data to free up recording media and to preserve the data for future analysis. The objective of the proposed work is to build a network forensics framework that precisely scrutinizes only the relevant packets .In this work, a network forensic framework is developed subjected to port scanning attack to mitigate evidence gathering challenges faced by forensic investigator. It captures and processes only fine-grained evidences present in the network traffic stream. Moreover, in the captured relevant log, attack specific discovery is carried out to mine the exact packets utilized to execute the network attack. Hypotheses being developed to validate each machine against attack specific criteria's. Only those machine who full fill the criteria will be scrutinizes for further analysis. To test and validate the effectiveness of the proposed framework two scenario have been developed, It is observed that developed system preciously securitizes the attack patterns and improved decrement in the log size is observed for both scenario developed that is about 93.12% and 95.65% respectively. It is also observed that only 6.09% and 1.93% of total traffic being scrutinized for NULL, FIN and XMAS attack in scenario 1 and 2 respectively. Similarly 19.88% and 13.42% packets of total packets are scrutinized for TCP Connect and SYN (Half open) scanning variant in scenario 1 and 2 respectively.

Keywords: Network Forensics, Intrusion Detection System, Port Scanning Attack.

I. INTRODUCTION

Port scan techniques are utilized to detect live ports of systems which are remotely located. Generally port scanning is useful to troubleshoot the networking errors. However on the negative side attacker may utilize port scanning techniques to get information such as operating system details, open ports, firewall rules etc.

Revised Manuscript Received on October 30, 2019.
***** Correspondence Author
Rajni Ranjan Singh Makwana*, Ph.D Research Scholar, Department of CSE, Maulana Azad National Institute of Technology, Bhopal, India. ranjansingh06@gmail.com
Dr. Deepak Singh Tomar, Associate Professor, Department of CSE, Maulana Azad National Institute of Technology, Bhopal, India. deepaktomarmanit@gmail.com, deepaktomar@manit.ac.in

Once active ports are identified, further attacker may indentify the vulnerability of associated applications, which can be exploits. For example Suppose an attacker is able to determine that one of the IP address is live and runs a flavor of UNIX and has a ToolTalker database server running at 32775 has a known vulnerability called buffer overflow attack. It is observed that approximately half of the cyber attacks preceded by port scanning attack. [1]
TCP and UDP are transport layer protocols. They are used to provide service point addressing using port numbers. The classification of port scanning techniques based on UDP and TCP has been shown in the figure 1.

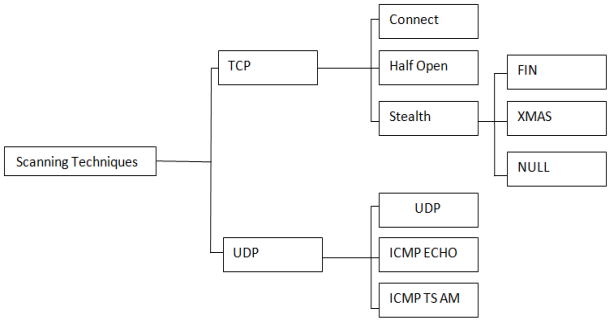

Fig. 1.Classification of port scanning techniques in context of Transport layer protocol

TCP port scanning techniques are most commonly used by the attacker because it is fast, reliable and can work efficiently in mostly platform (Linux, windows etc). It is reported that a firewall that utilize TCP filtering can detect or prevent about 68% of scanning activities [2]. Many systems which are infected by the worms remotely perform scanning activity in the internet searching for vulnerable systems to exploit. Forensic investigation of TCP port scanning is a challenges task. Details of TCP port scanning techniques are discussed in this section.

1. **TCP Connect scan –** In this approach scanner machine establish a three way handshaking connection by sending SYN packet to the target system. If the target port is open then SYN+ACK packet is returned otherwise reset packet returned.
2. **SYN Scan (Half open scanning)** – this method is very much similar to the connect scan however in this scan scanner machine does not complete the full handshaking procedure. If target port is active than target computer returned SYN /ACK packets. If port is closed then Reset packet will be returned.

3. **Stealth scanning:** - Stealth is a kind of scan that usually done not follows the normal characteristics of network connection. It looks like casual traffic. therefore these
4. scans usually bypasses firewall, network filters and routers etc. these method usually conduct scanning by setting single flags such as FIN, RESET, ACK, ALL or NULL. Most widely used approaches are presented here.

Stealth scanning variant is capable to bypass conventional devices such as routers, filters and firewalls as it is appear as casual traffic. Most commonly used stealth scan approaches utilizes Reset, SYN, ACK, NULL and ALL flags for scanning. Types of steals port scanning methods are as follows.

a) SYN/ACK Scan: In this approach, scanner machine sends TCP packet by setting SYN/ACK flags to the victim computer. If destination port is active than no response returned otherwise Reset packet returned.

b) FIN Scan: This approach is mostly similar to SYN/ACK scan. If target system port is active, then target system does not return any packet. If target port is closed Reset packet returned.

c) NULL Scan: In this scan variant, scanner machine sends a TCP packet without setting any flags. If target port is active then none is returned otherwise Reset packet is returned.

d) XMAS Scan (Christmas tree): This approach is similar to NULL scan variant however in XMAS scan, PUSH, FIN and URGENT flags are set. Target does not reply, if target port is active. If target port is closed than it responds with Reset packet.

II. LITERATURE REVIEW

Literature survey has been carried out to perform deeper study and recognize the challenges associated with network forensics and TCP port scanning attack.

Atul Kant Kaushik et. al. [3] proposes network forensic system architecture to perform investigation of network attack, proposed architecture composed of three modules- Capturing module, Analysis Module and Presentation module. Capturing module is responsible to captures the packets from network interface card and mark the packets that are relevant for further analysis. They have tested the proposed architecture using TCP-SYN, TCP-ACK, TCP-FIN port scanning methods. Only the packets matched with predefined signature are marked as relevant. Further all relevant packets are stored in the hard disk of the host for further analysis. Further analysis module analyzes the relevant packet stored in the hard drive of the host and find out the origin of the attack. Finally presentation Module is responsible to marks the machines as suspicious or not based on the port count values. If port count is beyond the threshold it is treated as suspicious.

Ying Zhu [4] proposed an iterative algorithm to identify the suspicious attack patterns using feedback mechanism, proposed algorithm mine the network traffic data to discover the context of the attack. Degree of belief is calculated for each attack instances and passed on to the next iteration to refine the search. To measure the usefulness and effectiveness of the proposed work an experiment has been carried out using well known attacks such as HTTP DoS, password crack, ping sweep, port scan and buffer overflow attack

Tie-Shan ZHAO et al. [5] presented a computer immunology based adaptive intrusion detection system. Proposed system detects port scanning attack if it encounters the connection attempt to closed port. Active port marked as self body, anti body is a set of all active ports. Self body has been generated time by time to update the anti body set. Whenever a packet targeted to IP + port which is not in antibody set is marked as suspicious.

Mehiar Dabbagh et al. [6] presented an approach to detect FIN scanning attack. It works in two phase, first phase classifies an IP as malicious or not, depends on the no. of connection attempt to closed port. Second phase classifies the IP into 3 groups suspicious IPs , normal IPs and scanner IPs. Work is limited to FIN scanning variant. Author specifies that proposed work is suitable for slow port scanning detection, it would be better if they measure the storage requirement/speed of proposed approach.

W. Ren et al. [7] presented an approach to detect port scanning attack by identifying failed connection attempts. Each scanner machine is represented by a scan vector consists of target ports and scan rate. Work is limited to TCP SYN scanning variants. It should be better if author explain the approach to detect failed connection and to update consistence port status.

Yousra Chabchoub et al. [8] presented an algorithm to identify vertical port scanning attack in the IP traffic. Only destination ports information and destination IP addresses is stored using 2D bloom filter. However proposed method is suitable for only vertical port scanning variants.

Evgeny V. Ananin et al. [9] proposed a mathematics model to detect port scanning attack. It calculates the “index of anomaly” based on the types of flags utilized in the network packets and source IP address. Authors observed that scan detection is faster whenever unusual flag combinations are utilized. However proposed method may suffer by the many false positives because TCP SYN(connect & half open) packets may appears in the normal traffic.

Silvia Anandita et al. [10] presented anomaly based detection system which utilizes dendritic cell method. Author determines a threshold coefficient 0.4759933. ICMP destination host unreachable value and number of packets generated by the attacker is utilized to calculate the threshold. It is observed that the threshold of port scanning is about 0.6164136 which is an anomaly. Nevertheless, ICMP error is alone not sufficient to detect failure attempts. Port scanning may be slow i.e one packet per min/hour, it would be better to sample the traffic.

Keto et. al. [11] proposed an approach to detect port scanning attack by identifying connection attempts resulting RST+ACK packets, 15 minute time window is utilized to capture traffic. However work is limited to TCP port scanning attack techniques due to the fact that closed systems/UDP scan generate ICMP error instead of RST + ACK Packets.

Robertson et al. [12] presented an algorithmic approach to identify suspicious scanner by calculating anomaly score of each IP address.

Anomaly score is the count of no. destination contacted where response is not returned, however no response alone is not to be an indicator for scan.

Ertoz at al.[13] developed a system to detect slow and fast port scanning attack based on four information, i) No. of connection from source ii)No. of connection to destination iii) No. of connection to same port iv) No of connection from destination to source port. An anomaly score has been calculated to judge the system.

Streilein et al. [14] proposed a method to detect distributed port scanning attack, in the proposed method author maintain multiples tables to store connection information such as Source IP address, Time , Duration, Probe type, probe duration and alert generated by probe , clustering approach utilized to analyze table data to identify distributed attack or not.

Loai Zomlot et al. [15] presented an approach of intrusion detection. Presented approach utilizes dempster shafer theory (D-S theory) to calculate combined belief using proposed Hypothesis. D-S theory will prioritize the result. Experiments are carried out using SNORT alerts and DARPA dataset.

TIAN Zhihong, et al. [16] presented an approach to fuse evidence collected from various hosts and sub networks using dempster shafer evidence theory. Proposed work is validated using DRAPA 99 intrusion dataset.

III. PROPOSED WORK

The proposed framework for efficient capturing of TCP port scanning attack packets is composed of four phase- Evidence collection, Attack pattern Discovery, Analysis and Prioritization and Statistic & Visualization as shown in figure 2. Evidence collection phase performs relevant packet capturing in real time and store the captured packets in the local hard drive of the acquisition system. Attack pattern discovery has been carried out to scrutinize the context of attack in the captured relevant log. Analysis and prioritization phase analyses the traffic behavior of each machine in context of developed hypotheses. Visualization phase provides the details of findings.

3.1 Evidence Collection of TCP Port scanning attack Variants

An approach has been proposed which is based on efficient packet capturing to capture significant port scanning packets despite of capturing entire network traffic. Discarding bogus traffic improves the investigation efficiency. In order to capture only relevant packets, traffic behavior generated from the port scanners is analyzed. The identified characteristics of TCP port scanning attack methods are shown in the Table 1.

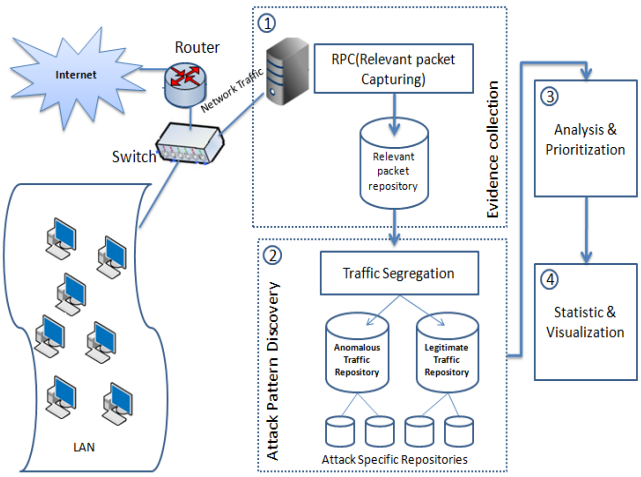

Figure 2: Proposed Network Forensic Framework

Table 1:Calssification of Port scanning Techniques [17,18]

Port scan Types	Protocol used	TCP flag set	Victims relay, if port open	Victims replay, if port closed
TCP Connect	TCP	SYN	ACK	RST
SYN scan(Half Open)		SYN	ACK	RST
FIN scan		FIN	No	RST
NULL Scan		No	No	RST
XMAS Scan		FIN/PSH/ URG	No	RST

As per the observation give in the Table 1 the scrutinize common features of TCP connect, SYN Scan, FIN Scan, NULL Scan and XMAS scan attacks are given in the Table 2

Table 2: Scrutinize Common Features

Protocol	*TCP*
Flags	SYN, RST,FIN, FIN+PUSH+URG and NULL.

Based on the scrutinize common features a Pseudo code has been developed to perform relevant packet capturing (RPC) shown in figure 3.

Developed pseudo code captures the packets having NULL, SYN, RST, and FIN flag set

The resultant relevant packet vector contains the all relevant packets utilized in the TCP port scanning attack. However in order to isolate the individual port scanning variant traffic attack pattern discovery will be carried out.

Relevant Packet Capturing (RPC)

Variables/functions used:

Packet $_i$ (flag): return the flag value of TCP packet

Packet $_i$ (Protocol): return the protocol of network packet

Input: Incoming packets

Output: Relevent_Packet _Vector

1.Initialize: Relevent_Packet_Vector [p_1,p_2……p_n] -> [0, 0….0];

2. Process incoming packet

3. if (Packet $_i$ (Protocol)= =TCP) then go to step 4 else go to step2.

4. if (Packet $_i$ (flag) == (NULL||SYN || RST ||FIN) then go to step 5 else go to step 2

5. Relevent_Packet _Vector. ->Packet $_i$ /* Add packet to vector*/ go to step 2.

Figure 3 Relevant Packet Capturing Approach

3.2 Attack Pattern Discovery

Attack pattern discovery is a systematic approach for finding context of port scanning attack in the traffic log generated by relevant packet capturing (RPC) approach. XMAS Scan, FIN scan and NULL Scan variant are similar in operation and they are belongs to the category of stealth port scanning attack. Similarly TCP Connect and SYN Scan (Half open) are alike in operation.

In this section relevant traffic captured by RPC system is segregated into two parts, one is anomalous TCP traffic and another is legitimate TCP traffic. Here the attack packets of FIN, XMAS and NULL scanning attack have been scrutinizing form anomalous TCP traffic. The evidences of TCP connect and half open have been scrutinize from legitimate TCP traffic. Pseudo code for Traffic Segregation is shown in the figure 4.

Input: Relevent_Packet _Vector

Output: Anomalous_Traffic _Vector, Ligitimate_Traffic_Vector

1. Initialize:

Anomalous_Traffic _Vector [p1,p2……pn] - > [0, 0….0];

Ligitimate_Traffic_Vector [p1,p2……pn] - > [0, 0….0];

2. Process incoming packets in sequence i= 1, 2, 3…N, until all packet processed.

3. if (Packet i (flag) ! = SYN OR RST OR ACK) then go to step 4 else go to step 5

4. Anomalous_Traffic _Vector -> Packet i/* Add packet to vector*/ goto step 2.

5. Ligitimate_Traffic_Vector -> Packet i/* Add packet to vector*/ goto step 2

Figure 4: Pseudo code for Traffic Segregation

i) Discovery of TCP NULL, XMAS and FIN scanning attack patterns

Here Anomalous_Traffic_Vector has been utilized as an input to discover the attack patterns of individual NULL, XMAS and FIN scanning variants. An approach has been developed; pseudo of developed approach is shown in figure 5.

ii) **Discovery of TCP Connect and Half open Attack pattern discovery**

TCP Connect and SYN scan are legitimate TCP operations and does not contain any specific signature therefore it is a tedious task to mine the attack patterns directly. Hence it is needed to perform traffic behavior analysis to mine the attack patterns. A normal TCP connection always terminates with FIN packets. But connect scan and half open scan terminates with RST packets.

In TCP connect scan, first two steps are exactly similar as half open scan and instead of sending a RST packet, TCP connect scan sends an ACK packet. And after this it reset the connection. To scrutinize port scanner machines a traffic behavior analysis is required, details are discussed in the following section.

Input: Anomalous_Traffic _Vector

Output: XMAS_Scanning_Vector,FIN_Scanning_Vector and NULL_Scanning_Vector

1. Initialize:

XMAS_Scanning_Vector [p1,p2……pn] -> [0, 0….0];

FIN_Scanning_Vector [p1,p2……pn] -> [0, 0….0];

NULL_Scanning_Vector [p1,p2……pn] -> [0, 0….0];

2. Process incoming packets in sequence i= 1, 2, 3…N, until all packet processed.

3. if (Packet i (flag) = =NULL) then go to step 4 else go to step 5

4. NULL_Scanning_Vector->Packet i/* Add packet to vector*/

5. if (Packet i (flag) = =FIN && URG && PUSH) then go to step 6 else go to step 7

6. XMAS_Scanning_Vector -> packet i /* Add packet to vector*/

7. if ((Packet i (flag) = =FIN) && (Packet i (flag)!=URG OR PUSH) then go to step 8 else go to step 2

8. FIN_Scanning_Vector -> Packet i /* Add packet to vector*/ go to step 2.

Figure 5: Pseudo code for Attack pattern discovery of NULL, XMAS & FIN scanning variants

3.3 Analysis and Prioritization

Attack patterns discovery phase collects the attack specific packets of port scanning attack variants. However usually attacker sends huge number of packet to target system. Therefore it is needed to analyze the traffic behavior of each machine.

Analysis module validates the each source IP addresses through developed hypotheses to check whether they are scanner or not. The formation of Hypotheses is being presented in this section. Whenever a host scans a network, it doesn't matter which approach they are utilizing, data packets are usually targeted to the inactive systems and ports therefore these connections can be judged as abnormal connection.

Suppose if there are H hosts in a Network and the probability of one host being active is P_1.the probability of finding an active host after trying only once is p_2.

$$P_2 = H.P1/H……..(1)$$

Every host has 65535 ports. Usually very less no. of ports are active. If there are Hp open ports in the host and the probability that an attacker finds an active port after trying only once is P_3,

$$P_3=Hp / 65535..... (2)$$

It is observed that Hp is commonly less than 15 i.e Hp < 15, so

$$P_3<0.000228…….(3)$$

If the probability of finding an open port of an active host after trying only once is p4,

$$P_4= P_3.P_2……..(4)$$

$$P_4 < 0.000228 P2$$

$$P_4 < 0.000228$$

if scanner unaware about the target network/system then the probability of finding active port of an live host after trying only once is very small, and it is usually less than 1%. Following two hypotheses have been constructed to judge the source machine.

Hypothesis 1(h1): System which scans larger number of different destinations IP addresses (machines) and ports (services) is probably a port scanner.

Hypothesis 2(h2): System which attempt many failed connections is probably the port scanner.

A system can be judged as port scanner if an only if it satisfies both hypotheses. Suppose a node, which perform many connection attempts to many destinations however it does not attempt any failed connection therefore it cannot be a port scanner.

Similarly a node which performs many failed connection attempts. However it does not send packets to many destinations. Thus it cannot be a port scanner.

Prioritization is beneficial to handle uncertainly using Dempster-Shafer theory that uses probabilities called beliefs score for a given hypotheses. DS theory provides unique ability to combine beliefs from multiples sources.

In this work sources are sensing devices deployed at various strategic positions in the network. Strategic position might be in-line to router or inside the subnet etc. each position may have different capturing/analysis capabilities, for example it might be possible that sensor which is deployed inside the subnet may capture very less traffic instead of sensor at router position due to timeouts. However sensor at router position can capture everything. This should be a wise decision to deploy sensor at appropriate position by the network engineer.

Here numerical values of hypothesis 1 and 2 have been calculated by the various sensor devices deployed at various position in the network. For combining the scores D-S theory is utilized to calculate combined belief for each source/evidence. Higher the belief shows high probability to be a port scanner.

Here let Θ be a frame of discernment is a set of host machines represented by their IP addresses.

$$\Theta = \{ IP_1, IP_2 \ldots . IP_n \}$$

Two BPA function h_1 and h_2 distributes the belief over the set of the frame of discernment after normalization. Here $m_1(h_1)$ and $m_2(h_2)$ are the numerical values observed by the hypothesis 1 and 2 for each IP addresses. Dempster-Shafer method calculates the overall combined belief of both hypotheses for each element of frame of discernment (for each IP address). Following combination rules have been utilized to calculate combined belief.

$$m_{1,2}(h) = \frac{1}{1-K} \sum_{h_1 \cap h_2 = h} m_1(h_1).m_2(h_2)$$

$$K = \sum_{h_1 \cap h_2 = \{\}} m_1(h_1).m_2(h_2)$$

Combined belief helps the investigator to prioritize the further analysis.

5.1.4 Statistic and Visualization

To visualize the result a data structure *portscannerlist* is developed with the following fields: IP address, Port Observed, data, Start time, End Time and Scanning variant.

- IP address: contains IP address of scrutinize machine
- Port Observed: contain the list of destination ports observed by the scanner
- Date: Date of scanning attack(according to first connection attempt)
- Start Time: Time when scanner machine sent first packet to target system
- End Time: Time when scanner machine send last packet to target system
- Scanning Variant: Scanning techniques such as XMAS, NULL, FIN, Connect & SYN.

IV. RESULT AND DISCUSSION

Proposed framework based on efficient packet capturing has been implemented and tested on a small local area network. Two different scenarios have been developed with varying size of port scanning traffic.

A testbed is developed comprised of a victim machine. A switch (layer 2) and seven other machines generate scanning and normal traffic as shown in figure 6. Scanners are equipped with NMAP performing TCP Connect, Half Open, FIN, XMAS and NULL scanning targeted to the victim machine. The summary of all the machines involves in the testbed are shown in Table 3.

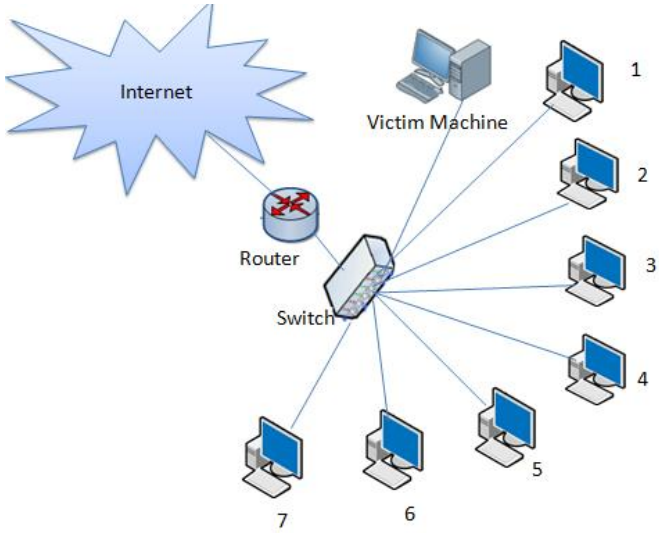

Figure 6: Testbed Topology

Table 3: Testbed System Configuration

	Scenario 1	Scenario 2	
Machie No.	**IP address of machines**		**Developed Attack**
1	172.16.51.253	172.16.51.154	Half Open
2	172.16.50.157	172.16.51.121	Connect
3	172.16.51.174	172.16.49.216	NULL
4	172.16.50.201	172.16.49.199	XMAS
5	172.16.49.37	172.16.51.147	FIN
6	172.16.51.249	172.16.51.40	Normal
7	172.16.49.132	172.16.49.221	Normal

Java socket application has been developed and deployed on Machine 6 and 7 to generate normal traffic (Attack free). To capture the network traffic as a evidence an open source Intrusion detection system SNORT[19] is configured in NIDS(Network Intrusion Detection System) mode to carry out traffic collection. SNORT is a signature based intrusion detection system able to capture selected packets based on the rules. Snort captures the packet in tcpdump format and generates alerts in text format. Snort alerts are converted into port scanning simultaneously to the victim machine (in the presence of normal Internet traffic) CSV format and uploaded to open source MYSQL[20] database for further Query driven analysis. Scanners performing port scanning simultaneously to the victim machine (in the presence of normal Internet traffic)

Traffic statistics generated during scenario 1 and scenario 2 is analyzed and presented as shown in table 4.

Table 4: Traffic Statistics generated by SNORT NIDS

	Packets	Time Span	Avg. pps	Avg. Pkt, Size	Avg. Bytes/s	Avg. Bits/s	Rate(ms)
Scenario 1	42019	123.152	341.2	259	88k	706k	0.3412
Scenario 2	62453	198.528	314.6	306	96k	769k	0.3146

i) Relevant Packet Capturing (RPC)

In order to select only relevant packets from incoming traffic, SNORT rules are developed and configured. Captured traffic static is shown in table 5.

It is observed that, out of 42019 packets, 10920 are scrutinized by the RPC approach during scenario 1. Similarly out of 62453 packets only 9591 have been scrutinized. Traffic behavior is shown in the figure 7 and 8

Table 5: Relevant packet capturing statistics generated by SNORT NIDS

	Packets	Time Span	Avg. pps	Avg. Pkt. Size	Avg. Bytes/s	Avg. Bits/s	Rate(ms)
Scenario 1	10920	117.640	92.8	59	5460	43k	0.0928
Scenario 2	9591	189.896	50.5	60	3012	24k	0.0505

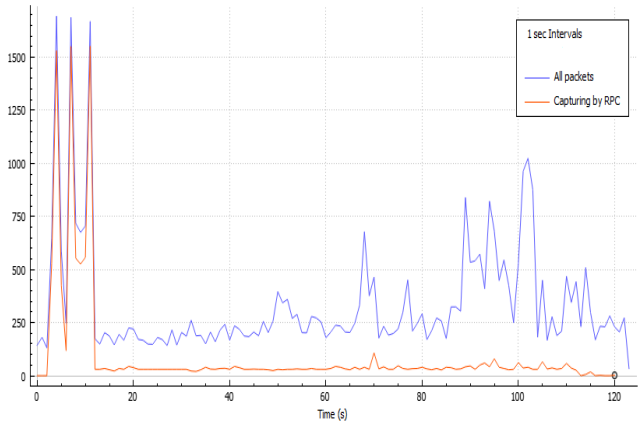

Figure 7: Comparison of Normal capturing with capture using RPC Approach, X-Axis; Time, Y- Axis: Packets (Scenario 1)

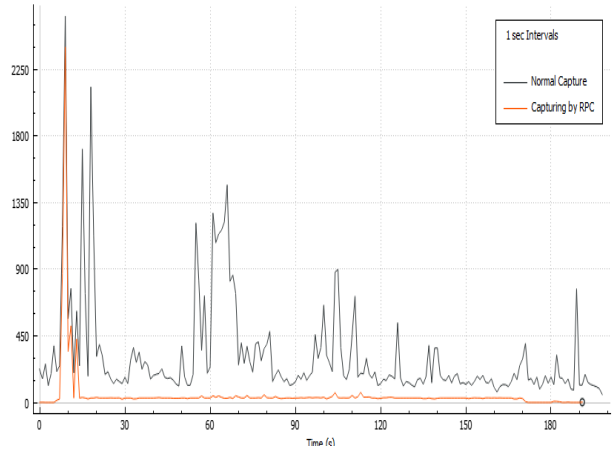

Figure 8: Comparison of Normal capturing with capture using RPC Approach, X-Axis; Time, Y- Axis: Packets (Scenario 2)

ii) Attack pattern Discovery

Traffic captured by relevant packet capturing system (RPC) is now segregated into two parts, one for anomalous traffic and another for legitimate traffic. The attack patterns of XMAS, NULL and FIN scanning are extracted from anomalous traffic archive. Similarly the attack patterns of TCP connect and half open scanning will be extracted from legitimate traffic archive.

After segregation two network traffic repositories have been created, the Statistics of both are shown in table 6.

Table 6: Traffic repositories statistics

	Scenario 1		Scenario 2	
	No. of packets	**Size**	**No. of Packets**	**Size**
Anomalous Traffic	2563	190KB	1206	108KB
Legitimate Traffic	8357	607 KB	8385	764 KB

iii) Analysis and Prioritization

After successful packet scrutinization, next is to analyze the behavior of the each machine against developed hypotheses. The network traffic behavior is shown in the table 7.

System 6 and 7 does not generate any failed connection therefore they are eliminated as per the hypotheses h1 and h2. Combined belief of each IP addresses has been calculated by D-S theory that shows the belief on hypotheses h1 and h2 for each host machines. Here basic belief is calculated by taking average of data given in the table 7. The basic belief and combined belief value of each scrutinize machine are shown in the table 8 and 9 respectively. Combination rules are utilized to calculate combined belief of each machine; this value will help to prioritize the further research.

It is observed that in scenario 1, machine 1 ,3 and 5 having highest combined belief values, similarly in scenario 2, machine 1 having highest combined belief value.

These machines attempts highest number of probing packet to closed & unique ports as shown in table 7.

Table 7: Observed Traffic behavior

	SCENARIO 1		SCENARIO 2	
Machine Number	**No of Unique connection attempt**	**No of Connection to closed ports**	**No of Unique connection attempt**	**No of Connection to closed ports**
1	1024	1019	1000	995
2	512	509	800	795
3	1024	1019	600	596
4	512	509	400	397
5	1024	1019	200	198
6	01	00	01	00
7	01	00	01	00

Table 8: Basic belief value of each scrutinize host

Scenario 1					
M/C No.	**1**	**2**	**3**	**4**	**5**
m1(h1)	0.250	0.125	0.250	0.125	0.250
m2(h2)	0.250	0.125	0.250	0.125	0.250
Scenario 2					
M/C No.	**1**	**2**	**3**	**4**	**5**
m1(h1)	0.333	0.267	0.200	0.133	0.067
m2(h2)	0.333	0.267	0.200	0.133	0.067

Table 9: Combined belief of each machine

Machine No.	**Combined belief**	
	Scenario 1	**Scenario 2**
1	0.2857	0.453
2	0.0714	0.293
3	0.2857	0.163
4	0.0714	0.072
5	0.2857	0.018

iv) Statistic and Visualization

Query driven analysis has been conducted and the findings are presented in this section as shown in table 10 and 11.

Table 10: Identified and scrutinize suspicious activity (Scenario 1)

	Scenario 1				
IP address	**Port requested**	**Date**	**Start Time**	**End Time**	**Scanning Variant**
1	1-1024	13/2/2018	15:34:57.394203	15:34:58.802146	Half open
2	1-512	13/2/2018	15:34:57.394203	15:36:55.034522	Connect
3	1-1024	13/2/2018	15:35:04. 492520	15:35:05. 892285	NULL
4	1-512	13/2/2018	15:35:02.355569	15:35:03.562295	XMAS
5	1-1024	13/2/2018	15:35:00.062478	15:35:01.462206	FIN

Table 11: Identified and scrutinize suspicious activity (Scenario 2)

	Scenario 2

Machine No.	Port requested	Date	Start Time	End Time	Scanning Variant
1	1-1000	3/12/2018	16:35:31.903737000	16:35:33.290378000	Half open
2	1-800	3/12/2018	16:35:29.715945000	16:38:14.663601000	Connect
3	1-600	3/12/2018	16:35:31.590386000	16:35:32. 864897000	NULL
4	1-400	3/12/2018	16:35:34.078283000	16:35:38.827403000	XMAS
5	1-200	3/12/2018	16:35:36.598464000	16:35:37.702439000	FIN

In the first Scenario, acquisition system processed 42019 packets and scrutinizes 10920 packets. Total reduction in number of packet captured by RPC is about 74.02%, similarly improved decrement in the log size is observed, that is about 93.12% of total traffic size.

In the second Scenario, acquisition system processed 62453 packets and scrutinizes 9591 packets. Total reduction in number of packet captured by RPC is about 84.64%, similarly improved decrement in the log size is observed, that is about 95.65% of total traffic size.

It is observed that only few relevant packets are scrutinized; only 19.88% and 13.42% Packets of total packets are scrutinized for TCP Connect and SYN (Half open) scanning variant in scenario 1 and 2 respectively. The total reduction in No. of Packets is 80.12% and 86.58% respectively.

It is also observed that only 6.09% and 1.93% of total traffic being scrutinized for NULL, FIN and XMAS attack in scenario 1 and 2 respectively. The reduction is about 93.91% and 98.07% of total traffic respectively.

In context of storage requirement, only 5.23% and 3.80% traffic (in KBytes) have been scrutinized for TCP connect and SYN(Half open) Scanning and total reduction is about 94.77% and 96.2.01% of total traffic size in scenario 1 and 2 respectively.

Similarly only 1.63% and 0.5% traffic (in K.Bytes) being scrutinized for TCP NULL, FIN and XMAS scanning variants for scenario1 and 2 respectively and total reduction is about 98.37% and 99.5% of total traffic size for scenario 1 and 2 respectively.

V. CONCLUSION

The work presented in this paper focuses the problem of effective capturing of attack evidences for the forensic investigation. In this work a framework for network forensics has been developed, which is based on the efficient packet capturing approach. The framework has been implemented using well known open source utility such as SNORT and MYSQL. NMAP is utilized to perform port scanning attack. This work helps the forensic investigator to discover the evidences of port scan attack variants presents in the network traffic log and finally provide all necessary finding such as IP address, date and time of scanning, scanning method and port observed.

The future work will be focused on, the fusion of evidences collected from the multiple sensor devices deployed at various strategic position in the network. Detection and forensic investigation of hybrid port scanning attack shall be explored in future with all necessary improvement in the exiting framework.

REFERENCES

1. Panjwani, S., Tan, S., Jarrin, K.M. and Cukier, M., 2005, June. An experimental evaluation to determine if port scans are precursors to an attack. In Dependable Systems and Networks, 2005. DSN 2005. Proceedings. International Conference on (pp. 602-611). IEEE.
2. Bou-Harb, E., Debbabi, M. and Assi, C., 2014. Cyber scanning: a comprehensive survey. IEEE Communications Surveys & Tutorials, 16(3), pp.1496-1519.
3. Kaushik, A.K., Pilli, E.S. and Joshi, R.C., 2010, February. Network forensic system for port scanning attack. In Advance Computing Conference (IACC), 2010 IEEE 2nd International (pp. 310-315). IEEE.
4. Reducing the Cost of Incident Response with Selective Packet Capture. https://www.network-visibility.com/downloads/Savvius_TechBrief_Selective-Packet-Capture.pdf [Last Accessed November 2018]
5. Zhao, T.S., Li, Z.Z., Wang, Z.M. and Lin, X.J., 2007, December. An adaptive LAN intrusion detection system based on computer immunology. In Robotics and Biomimetics, 2007. ROBIO 2007. IEEE International Conference on (pp. 2234-2238). IEEE.
6. Dabbagh, M., Ghandour, A.J., Fawaz, K., El Hajj, W. and Hajj, H., 2011, December. Slow port scanning detection. In Information Assurance and Security (IAS), 2011 7th International Conference on (pp. 228-233). IEEE.
7. W. Ren and H .Jin "Modeling the network Forensic behavior "Proc 1st Int'l conf. security and privacy for Emerging Areas in communication Networks (SecureComm 2005) pp 1-8.2005, IEEE
8. Chabchoub, Y., Fricker, C. and Robert, P., 2012, October. Improving the detection of on- line vertical port scan in IP traffic. In Risk and Security of Internet and Systems (CRiSIS), 2012 7th International Conference on (pp. 1-6). IEEE.
9. Ananin, E.V., Nikishova, A.V. and Kozhevnikova, I.S., 2017, November. Port scanning detection based on anomalies. In Dynamics of Systems, Mechanisms and Machines (Dynamics), 2017 (pp. 1-5). IEEE.
10. Anandita, S., Rosmansyah, Y., Dabarsyah, B. and Choi, J.U., 2015, November. Implementation of dendritic cell algorithm as an anomaly detection method for port scanning attack. In Information Technology Systems and Innovation (ICITSI), 2015 International Conference on (pp. 1-6). IEEE.
11. Kato, N., Nitou, H., Ohta, K., Mansfield, G. and Nemoto, Y., 1999. A real-time intrusion detection system (IDS) for large scale networks and its evaluations. IEICE transactions on communications, 82(11), pp.1817-1825.
12. Robertson, S., Siegel, E.V., Miller, M. and Stolfo, S.J., 2003, April. Surveillance detection in high bandwidth environments. In DARPA Information Survivability Conference and Exposition, 2003. Proceedings (Vol. 1, pp. 130-138). IEEE.
13. Ertoz, L., Eilertson, E., Lazarevic, A., Tan, P.N., Dokas, P., Kumar, V. and Srivastava, J., 2003, November. Detection of novel network attacks using data mining. In Proc. of Workshop on Data Mining for Computer Security, pp 29-39.
14. Streilein, W.W., Cunningham, R.K. and Webster, S.E., 2001, June. Improved detection of low-profile probe and denial-of-service attacks. In Proceedings of the 2001 Workshop on Statistical and Machine Learning Techniques in Computer Intrusion Detection, pp.11–13.
15. Zomlot, L., Sundaramurthy, S.C., Luo, K., Ou, X. and Rajagopalan, S.R., 2011, October. Prioritizing intrusion analysis using Dempster-Shafer theory. In Proceedings of the 4th ACM workshop on Security and artificial intelligence (pp. 59-70). ACM.
16. Zhihong, T., Wei, J., Yang, L. and Lan, D., 2014. A digital evidence fusion method in network forensics systems with Dempster-shafer theory. China Communications, 11(5), pp.91-97.

17. Bhuyan, M.H., Bhattacharyya, D.K. and Kalita, J.K., 2011. Surveying port scans and their detection methodologies. The Computer Journal, 54(10), pp.1565-1581.
18. whitepaper by dethy@synnergy.net "Examining port scan methods - Analyzing Audible Techniques" www.ouah.org/portscandethly.pdf [Last Accessed September 2019]
19. SNORT https://www.snort.org/ [Last Accessed September 2019]
20. MYSQL *https://www.mysql.com* [Last Accessed September 2019]

AUTHOR PROFILE

Mr. Rajni Ranjan Singh, Ph.D Research Scholar, Department of Computer Science & Engineering, Maulana Azad National Institute of Technology, Bhopal, M.P., India. His research interest includes Algorithm Design, Network security, Network Forensics and Computer Networks

Dr. Deepak Singh Tomar, Associate Professor, Department of Computer Science & Engineering, Maulana Azad National Institute of Technology, Bhopal, M.P. India. Dr. Tomar's research area of interest includes Data Mining, Internet Technology, Computer & Network Security, Digital Forensics and Machine Learning