Anomaly Detection Model for Covert Intrusion Detection

N. Vadivelan
Research Scholar,
Department of CSE, St.Peter's University,
Chennai. India

Dr. S. Anbu
Professor, Department of CSE,
St.Peter's College of Engg. & Tech, Chennai.
India

Abstract:- **Covert communication is a hidden data transfer mechanism that is capable to communicate between non legal entities of the premises. Two instances of covert channels are defined as timing based and storage based. Both instance acts to be the modern covert data transfer system within the access to the backdoor. Hence a desirable anomaly model is required to detect various covert communications. Here we proposed a robust model which is likely to be unsupervised behavioural model in order to analyse the deviation range in the data communication. The proposed model is clearly tested in the test bed with two various flavour of Black linux. The behaviour and learned pattern of the system is allowed to run the packet sniffer to capture the general behaviour of the system. Experimental results reveal the promising results in identifying the covert communication.**

Keywords: Covert channels, covert communication, black linux, self-organizing map, K-means clustering

INTRODUCTION

Today networking and intercommunication of the hosts are increasing day by day. Every instance the adoptability towards a new communication space is also in increasing order. At the same time various attacks and successive intrusions are performed in all aspects. To defend this various pre built are available in the market in terms of security solutions both as hardware and software. These solutions provide security mechanism to the possible and known pattern. If the attack and its sequence is still in hidden format, no security solutions have not yet devised to identify these kinds of covert attacks.

An IDS is an one of the responsive intelligent systems which is used to generate alert when an successive attack is performed. Many IDS are available in the market as open source and commercial terms. Since no IDS is proposed so far to defend the guaranteed successive intrusion prediction for covert communication stated in figure 1 and figure 2. To overcome the above stated problem, we attempted our generic and unique work in identifying covert channels/communication. The first attempt was made successful and clear demonstration is also given in the next section.

According to the resource level, the IDS system is based upon two levels. The initial level is securing the host and alerting the sink i.e., host based intrusion detection system and secondly, network based intrusion detection system. Covert channels are successive terms of compromised host which infiltrates and exfiltrate data to outside source from the on premises. Hence in this paper our attempt is towards developing an optimal IDS and anomaly model for detecting covert anomalies in each hosts of the network.

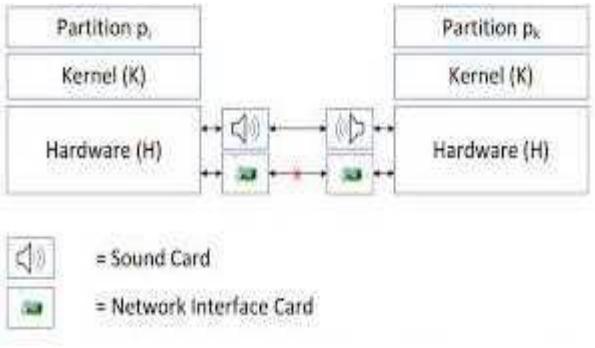

Figure 1: Two computers as the part of covert communication

PREVIOUS ATTEMPT

The covert channels can be easily established in the real time systems which are active in the networks [1]. Once the host establishes with the attacker (Server) the communication pattern will not be known and it still remains covert. The host and server only know the

transmission pattern, two types of active covert channels are timing channels and storage channels, and here the both channels are in active perspection. The key idea is the actual model of data's (Host and network) are learnt by the IDS in the network level and in detection phase these learning phase data's are applied and correlated to detect the attack sequence. If the data exhibits the correlation level below than that of the threshold, then the IDS detect the covert channels [1].

The main motto of the work is to detect covert communication in the active networks by analysing the network behaviour. Basically covert communication takes place between the compromised hosts to the server. The data transfer might be within the on premises (possible backdoor). General pattern is analysed for every host in the learning phase at periodical interval. During data transfer if the host which is likely to be compromised is sensitive in sending the vital information at a periodical pattern is analysed and its deviation is monitored from the learning phase. If the deviation ratio is intended to variation beyond the threshold level and the correlated parameters is again analysed to check the covert channels, if the threshold falls beyond the 1 then the pattern is recorded as the covert communication. Data correlation is achieved by comparing the data values with the learning phase and detection phase [1].

Traffic analyser module

State of the art of the proposed research is to develop two modules namely traffic analyser module which as a promiscuous mode in nature and secondly anomaly module using K-means and SOM. The typical architecture of the proposed module is defined in figure 3. Here the module is used to capture all the network traffic and redirects it into the single communication path where all the traffic packets are inspected carefully. The module is specifically designed to apt for TCP, UDP, ICMP and SNMP for managing the state of promiscuous mode.

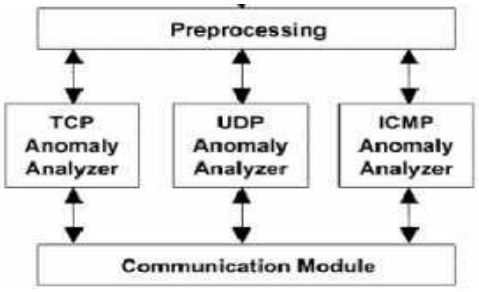

Figure 2: Analyser module for traffic management

The three protocols were the key term of data theft. TCP/UDP is the main protocols used to transfer the streams from one host to another. Hence all the traffic from this two source protocol is analysed mandatory. ICMP protocol is used to revert back the communication in terms of port scanning, probing, and pinging. Hence it is mandatory to analyse this particular protocol too.

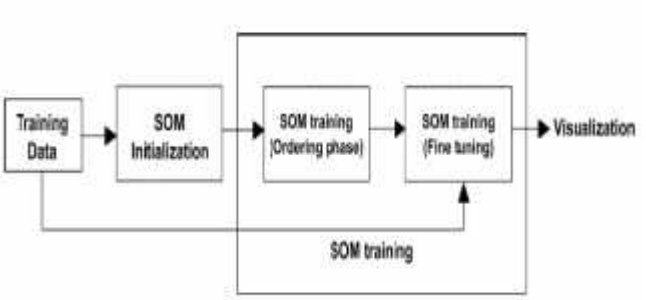

Figure 3: Typical architecture of training phase

Anomaly detection module

Anomaly is defined to be the major term of intrusion detection and response system in which the system detects the anomalies by estimating the deviation range from the normal behaviour of the system to the abnormal behaviour. Anomaly detection model has an advantage of detecting new order of attack patterns and intrusions as the estimation of normal behaviour deviation and its range.

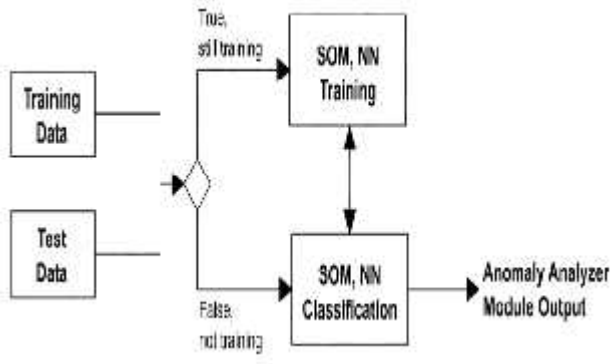

Figure 4: Architecture of Test and Training bed of SOM

In this research, the network traffic is the basic parameter used and it is examined using traffic analyser module and traffic features are extracted. Then those features are given as the input to the system and trained periodically. Then one of inter connected host is compromised and installed with backdoor client, here we used Meterpreter tool. The backdoor is installed through timing channel in the periodical interval. Each anomaly module uses the algorithm which is defined in the section algorithm for classifying the anomalies. Extracted features are

clustered using K-means clustering algorithm and clustered features are given as input to the kohenen's self- organising map which is a neural based approach and SOM classifies the anomalies based on the defined threshold conditions or values.

Algorithm

For each host in the network

Calculate the value pair using feature vector (fv)

Classify the incoming data packets using

protocol
analyser module

Classify the data packets as

attacks, If fv> Thresholds

Classify the data packets as
normal

If fv<
Thresholds

Generate classified results

Alert the system

End

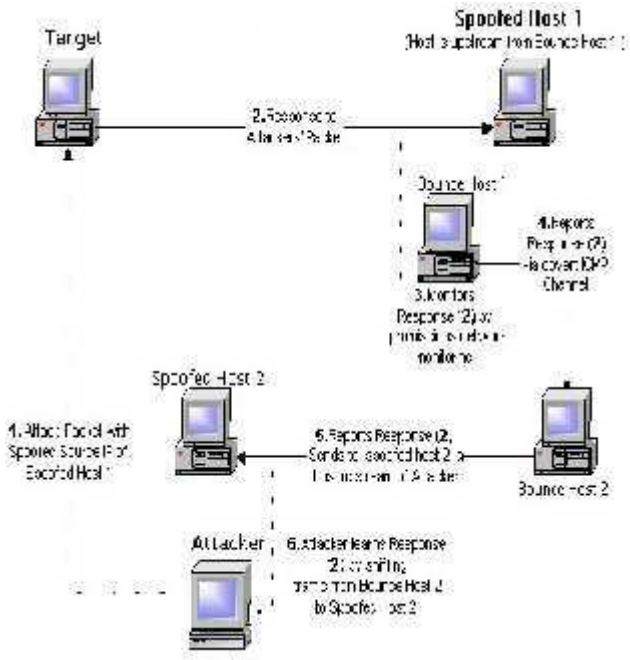

Figure 5: State of the art – Defined problem

Experimental setup and result analysis

The test bed of the proposed model consists of 5 hosts which are connected in an network. Out of five hosts, Machine 2 and machine 5 runs on linux and rest of all the machines is window based. Promiscuous mode is enabled on all the machine and achieved through redirecting all the traffic using PROC command in linux. The traffic and its pattern is given as the learning pattern to the system and data's are collected for 4 days. The observed pattern is trained as the input to the system. Initially all the network parameters were analysed and features are selected in the learning phase, then the live data is given as input to the anomaly module, here the live data is captured and then features are extracted and given as input to the detection module and finally the proposed anomaly detection module classifies the anomalies if the feature vector exceeds the threshold value. Figure 6 denotes the storage channel identification by our proposed methodology and figure 7 denotes the timing channel identification by our proposed methodology. We collected all the feature parameters and analysed both the host learnt data along with the attacking data which is simulated in MATLAB R2013b. Our parametric analysis is considered for all the network parameters which were clearly discussed in the above sections.

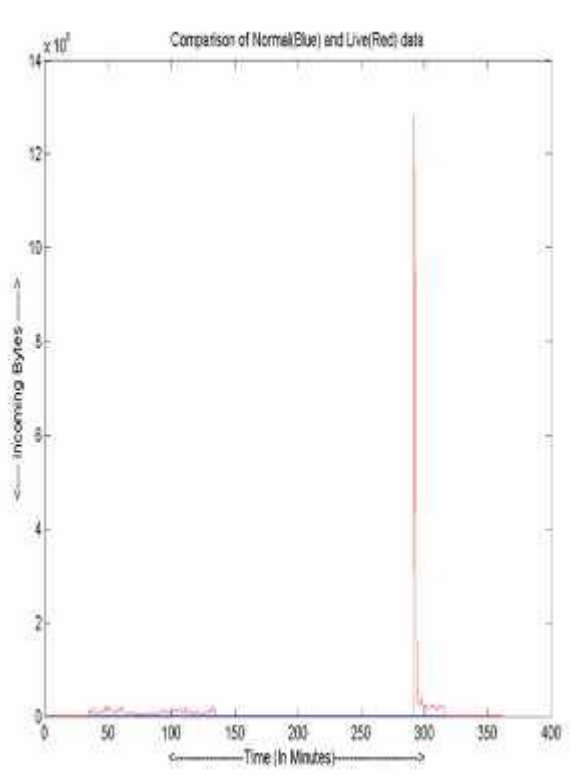

Figure 6: SOM classified estimation for feature: incoming bytes – Timing Channel

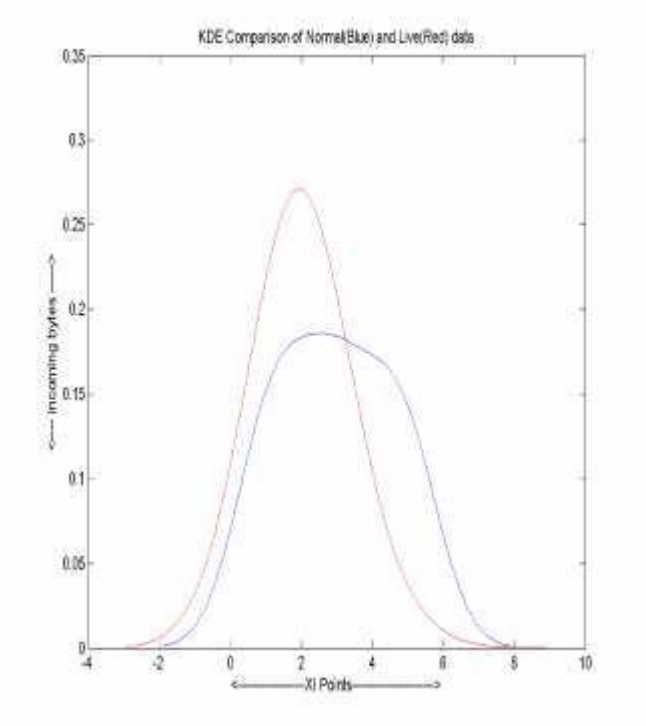

Figure 7: SOM classified estimation for feature: incoming bytes – Point deviation

CONCLUSION

Hence we conclude the paper by proposing a novel methodology for anomaly detection and simulation results shows the accuracy rate of nearly about 78 % and
1% of false positive rate which an aggregate of manual estimation.

```
wrong_fragment <= 0
|  num_compromised <= 0
|  |  count <= 236
|  |  |  dst_host_srv_diff_host_rate <= 0.24
|  |  |  |  dst_host_same_srv_rate <= 0.01
|  |  |  |  |  src_bytes <= 1
|  |  |  |  |  |  rerror_rate <= 0.98
|  |  |  |  |  |  |  serror_rate <= 0.32
|  |  |  |  |  |  |  |  count <= 2
|  |  |  |  |  |  |  |  |  protocol_type = tcp: portsweep
|  |  |  |  |  |  |  |  |  protocol_type = udp: satan
```

Figure 8: Results of feature vectors – Anomaly module

Figure 8 denotes the feature vector along with its detection ratio. The proposed model has an advantage of detecting the near real time feed patterns which can be a live feed from any packet sniffers or analysers like Wireshark etc. In future the work would be carried to the live streaming of data along with deploying any queuing model for better accuracy.

REFERENCES

[1] Hong Zhao," Covert Channels in 802.11e Wireless Networks",IEEE,2014
[2] Kamalanaban Ethala, R. Sheshadri, S. Sibi Chakkaravarthy," WIDS Real-Time Intrusion Detection System Using Entrophical Approach" Artificial Intelligence and Evolutionary Algorithms in Engineering Systems, SPRINGER, Volume 324, 2015,pp 73- 79
[3] R. Goncalves, M. Tummala and J. McEachen "A MAC Layer Covert Channel in 802.11 Networks", The Third International Conference on
[4] Emerging Network Intelligence (EMERGING 2011), pp. 88-99.
[5] Asrar, "Android Threat Tackles Piracy Using Austere Justice Measures," Irfan Asrar's blog, 2011.
[6] B. Morrow, "BYOD security challenges: control and protect your most sensitive data," Network Security, vol. 2012, no. 12, pp. 5-8, Dec. 2012.
[7] M. Graa, N. Cuppens-boulahia, and A. Cavalli, "Detecting Control Flow in Smarphones: Combining Static and Dynamic Analyses," in The
4th International Symposium on Cyberspace Safety and Security. Melbourne, Australia: Springer, Dec. 2012, pp. 33-47.
[8] R. Browne "Mode Security: An Infrastructure for Covert Channel Suppression", *Proceedings of the IEEE Computer Society Symposium on Research in Security and Privacy*, pp.39 -55 1994
[9] Vishal Bharti, Practical Development and Deployment of Covert Communication in IPv4, Journal on Theoretical and Applied Information Technology, Apr 2007.
[10] Gustavus J Simmons, The Prisoner's Problem and the Subliminal Channel,
Springer-Verlag, 1996.
[11]"Common methodology for information technology security evaluation,"July 2009, http://www.commoncriteriaportal.org.