ISSN:0975-9646

DDoS Attack Detection based on Chaos Theory and Artificial Neural Network

Anjali. M, B.Padmavathi

Department of Computer Engineering, Pune University,
G. H. Raisoni College of Engineering and Management,
Wagholi, Pune, Maharashtra, India.

Abstract- DDoS attacks temporarily make the target system unavailable to the legitimate users. They don't steal anything. But they cause big headache for targeted companies and network engineers. Application layer DDoS attacks are difficult to detect because they mimic normal traffic. This paper proposes a novel method of detection of DDoS attacks based on Chaos theory and Artificial neural networks.

***Keywords*— AR model, Botnets, Chaos, DDoS attacks, Neural network.**

I. INTRODUCTION

Because of the rapid development in internet, number of online attacks are also getting increased. One of the strong and dangerous attack among them is DoS attack. . It is more harmful if DoS attack is distributed. A DDoS attack happens with the help of a mechanism called Botnets. A Bot is a harmful program which controls botnets. Botmasters control botnets remotely through C&C infrastructure. Specific, automated functions are performed by bots in small scripts. Bots are mainly used for negative purposes. It is used to create tools for the activities such as the widespread delivery of SPAM email, spyware installation, click-fraud, virus and worm dissemination, and DDoS attacks. There are three types of DDoS attacks. Resource attack, Bandwidth attack and Application layer attack. First two attacks concentrate on network layer or transport layer. Application-layer DDoS attacks are a bit more complex. They are some of the most difficult attacks to mitigate against because they mimic normal traffic as they interact with the user interface.

When suddenly the demand for a website increases rapidly, it can lead to a flooding attack. An example is a popular news posted on a website. It can lead to a bursty legitimate traffic. A Web-DDoS attack traffic is exactly like a legitimate burst traffic [1]. Distinguishing between a Web-DDoS attack and a bursty legitimate traffic is a tedious task.

II. RELATED WORK

Here a background on Botnets is provided and how they launch DDoS attacks. By using a software program named bots , botnets make the computers compromised with the help of command and control server. A series of systems got affected through numerous tools and through the installation of a bot that uses Internet relay chat (IRC) to remotely control the victim. Nowadays DDoS attacks are majorly caused by Botnets. Moreover, It is possible for attackers to change their communication approach during the creation of the bots. An example of Botnet attack on application layer is the HTTP flooding attack. HTTP server creates the bots launched by this attack. These bots are Web-based bots.

Flooding attacks can be mainly classified into three. They are Direct Attack, Distributed Attacks, Spoofing-Based Attacks [1].

A. *Direct Attack*

DoS attack performed with-out spoofing the attack packets is called direct attack. TCP SYN attack utilizes the vulnerability of 3 way hand-shake does not involve any spoofing, and is very easy to perform. Attacker sends number of SYN packets to the victim. They will not respond to the SYN-ACK packets which are sent by the victim. It is very easy to prevent these types of attacks when detected by using simple firewall rules.

B. *Spoofing-Based Attacks*

Another way to perform DDoS attack is spoofing the IP source address. In SYN flood attack, vulnerability of TCP three-way handshake is utilized. In TCP 3 way handshake client sends SYN packets to the server for requesting new connection. Server acknowledges by sending SYN-ACK packet back. Client responds with ACK packet as third step, thus a connection is established. During SYN flood attacks, SYN packets are sent by the attacker with spoofed source IP addresses. Because of spoofed SYN packets are sent to the server, connection is never established. ACK packets from the client are not received in SYN flood attack. The memory stack becomes full with each half open connection. Consequently, no further requests can be processed. All services from the system are denied and it becomes offline.

C. Distributed Attacks

Distributed DoS attacks are DDoS attacks. This attacks are launched by Botnets. If there is only one source it is easy to detect the attacker. This draw-back is remedied by distributed attacks. Another type of distributed attacks is distributed reflector denial of service attacks [11]. Here source of the attack traffic is concealed by third parties.

Botnet based DDoS attacks are divided into three. They are agent-handler, IRC-based, and Web-based models [2].

D. *Agent-Handler Model*

The agent-handler model of a DDoS attack contains clients, agents and handlers. Attacker communicates with clients in the DDoS attack system. Software packages located throughout the Internet are handlers. The client uses handler packages to communicate with the agents.

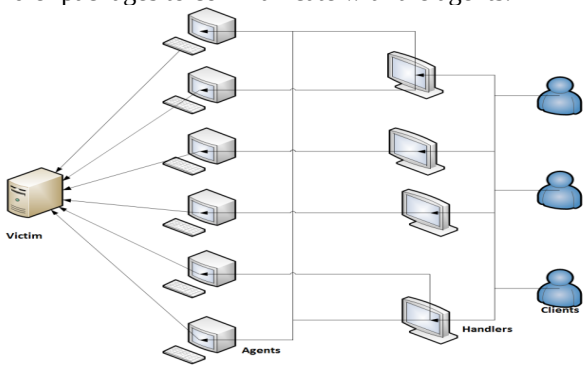

Fig. 1 Agent-handler model

E. *Internet Relay Chat (IRC) Model*

This model has similarities with Agent-Handler model. Here an IRC communication channel is used for the connection between clients and agents. Handler software's are not used. An Internet Relay Chat (IRC) channel contains IRC ports for sending commands mainly to agents. These ports are legitimate ports thus DDoS command packets are not getting tracked

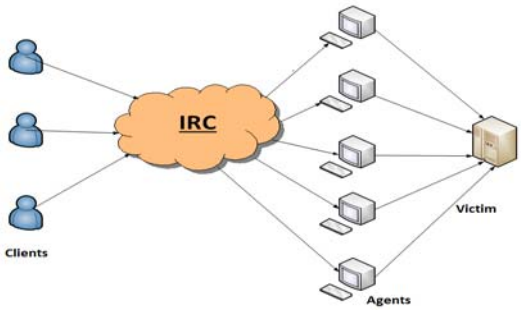

Fig. 2 Internet relay chat model

F. *Web-based Model*

For Botnet command and control (C&C), most easy way for attacker is the IRC-based model. But for last few years web-based reporting and command has emerged .Bots has to report statistics to website in Web-based model. A number of bots in the Web-based model simply report statistics to a Web site, whereas in other models bots are fully configured and controlled through encrypted communications and PHP scripts and over the 80/443 port and the HTTP/HTTPS protocol.

G. *The DDoS threat*

A DDoS attack directs a large number of "zombie" hosts, against a single target. Set of zombies are created quickly. Availability of attacking tools is another reason to make DDoS attacks wide spread. TFN, Trinoo, TFN2K are some examples of tools to launch DDoS attacks. A successful DDoS attack gives widespread impact. Violated SLA's, compromised site performance, diminished company reputations, productivity loss etc are some impacts of DDoS attacks. DDoS attackers are using complex spoofing techniques and legitimate protocols. Thus it is very difficult to detect and defeat.

H. *DDoS detection methods*

Entropy based DDoS detection is an information theoretic concept. It measures randomness. Entropy on each system is calculated based on up-stream traffic flow and down-stream traffic flow [1]. If this entropy value is greater than a pre-determined threshold value, then the traffic is attack traffic. After detecting it as attack traffic, attack source is found by using trace-back analysis. If the legitimate traffic is greater than seven times of the attack traffic, this detection is not efficient.

Detecting DDoS attack source is an essential step in defeating DDoS attacks [3]. Packet marking methods include the PPM and the DPM are two mechanisms to detect zombies and thus attackers. According to the probability on the local router PPM mechanism mark packets with the router's IP address information, victim can reconstruct the path according to this information. The PPM method has so many drawbacks. Spoofed marking information can be sent by the attackers to the victim. The accuracy of PPM is very less. Downstream routers can overwrite the packets. There occurs a storage space problem in PPM because it requires many number of packets for reconstruction.

The deterministic packet marking mechanism marks the packets with initial routers information example: IP address. Therefore, the victim can identify the starting location of the attack packets once the required information of the marks is obtained. For reconstructing attack path and for identifying the attack source a large number of marked packets are required. Pollution from attackers is another problem of DPM.

Gil and Poletto introduced a detection method which contain a heuristic and a data structure called MULTOPS (MUlti-Level Tree for Online Packet Statistics), that monitor certain traffic characteristics to detect and eliminate DDoS attacks. MULTOPS [9] is a tree of nodes that tracks packet rate statistics for subnet prefixes at different aggregation levels. According to the pre-specified memory size, expansion and contraction of the tree occurs. A network device using MULTOPS detects bandwidth DDoS attacks by the presence of a disproportional difference between packet rates going to the victim and coming from the attacker.

Wang et al. presented modelling of DDoS attacks using Augmented Attack Tree (AAT) [12]. He also introduced an attack detection algorithm based on AAT. Subtle incidents triggered by a DDoS attack and the corresponding state transitions are captured by this model from the view of the network traffic transmission on the primary victim server.

All these proposed approaches depend on monitoring the traffic volume on victim. They are not able to differentiate attack traffic and flash crowd (legitimate burst).

III. PROPOSED WORK

A. DDoS detection based on Chaos theory

For analysing and forecasting network traffic, time series models, such as AR, ARMA, ARIMA, ARFIMA [10] and FARIMA etc [7] are used. Predictability analysis on network traffic shows that low-pass filtering and multiplexing can provide better predictability. However, due to the bursty network traffic, there is a possibility of large prediction error. Therefore these time series models should be relatively stable. Network traffic prediction model predicts trends of sometimes things in the future under the guidance. Network traffic model is divided into two categories. Traditional traffic model and new traffic model. In this approach, after collecting network traffic packets and flow information, all network traffic is sampled. Let x_n denote the state of traffic, so the sequence of network traffic is

$\mathbf{x_1, x_2, \ldots\ldots x_k, \ldots\ldots x_n}$

Next step is the prediction of network traffic. To get an accurate result, network traffic should be suppressed. This is done by pre-processing the traffic by cumulatively averaging the sequence $\mathbf{x_n}$ with a time range.

$$\bar{X}_k = (x_1 + x_2 + \ldots\ldots\ldots + x_k)/t_k \quad (1)$$

After finding out cumulative average , prediction is done based on AR model. That is

$$\hat{X}_j = \sum_{k=1}^{m} ak\, \bar{X}_{j-k} \quad (2)$$

From the above equations x_k can be predicted

$$\hat{x}_k = t_k \hat{X}_k - t_{k-1} \hat{X}_{k-1} \quad (3)$$

Prediction of $\mathbf{x_k}$ is $\hat{x}_k$. t_k is the time of the kth sequence of network traffic. Prediction error can be found out from above formulas

Prediction error $\Delta x_k = x_k - \hat{x}_k$ (4)

Now assuming the behaviour of propagation error Δx_k is chaotic, Lyapunov constant is used to analyse it.

Lyapunov constant $\lambda_k \approx \{\ln(\Delta x_k / \Delta x_0)\}/t_k$ (5)

If $\lambda k > 0$, the Δx_k is chaotic. This means that the change is not caused by DDoS attack traffic but because of new legitimate traffic entering the system[4],[5].

If $\lambda k = 0$, Δx_k and Δx_0 do not differ in value .Propagation error is constant and there is no new traffic and thus no attack traffic.

If $\lambda k < 0$, the Δx_k is not chaotic. This says it as an attack traffic.

B. Algorithm

Step 1: Collect network traffic packets and flow information in real-time.

Step 2: Pre-process network traffic by cumulatively averaging it as in (2)

Step 3: By using AR model, predict the network traffic.

Step 4: Find out the prediction error by (4)

Step 5: Detect the abnormal traffic by analyzing prediction error based on chaos theory

Step 6: Detect DDoS by using trained neural network..

To improve the detection efficiency, trained neural networks are used. Artificial neural network is a type of network which considers nodes as artificial neurons [13]. In artificial intelligence 2 types of learning are there. Supervised and Unsupervised.

C. Supervised Learning

In supervised learning, one set of observations, is assumed to be the cause of another set of observations. They are termed as inputs and outputs respectively. Supervised learning builds an artificial system, and predicts he output of the system with new inputs, by learning the mapping between existing inputs and outputs. If the output takes a limited set of discrete values that indicate the class labels of the input, the learned mapping leads to the classification of the input data.The back-propagation algorithm uses supervised learning, Error is the difference between actual and expected results. The aim of the back-propagation algorithm is to reduce this error, until the ANN learns the training data.

D. Unsupervised Learning

In unsupervised learning, the learning can continue orderly from the observations into more abstract levels of representation. Clustering is an unsupervised learning approach. Clusters reflect the underlying structure of the data based on similarity groups within the data. Cluster analysis or clustering is the task of grouping a set of similar objects.

IV. IMPLEMENTAION AND RESULTS

Two scenarios are created by using Opnet modeler 1.5 simulator. Legitimate normal and burst traffic are created in first scenario. Second scenario contained DDoS attack traffic.

Both scenarios are executed in Opnet, and traffic values are collected. Fig 3 shows legitimate burst traffic, in which clear spikes occur in some intervals. Fig 4 shows DDoS attack traffic. When time goes on attack traffic is increased.

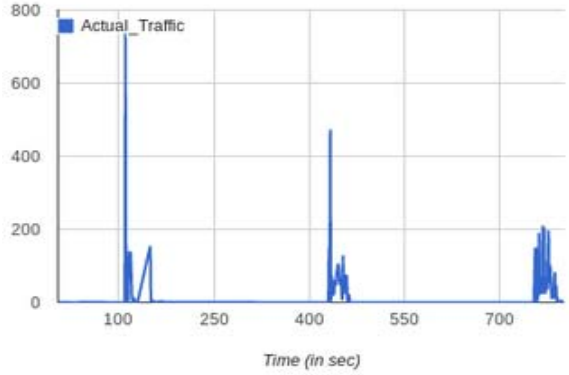

Fig. 3 Legitimate burst traffic

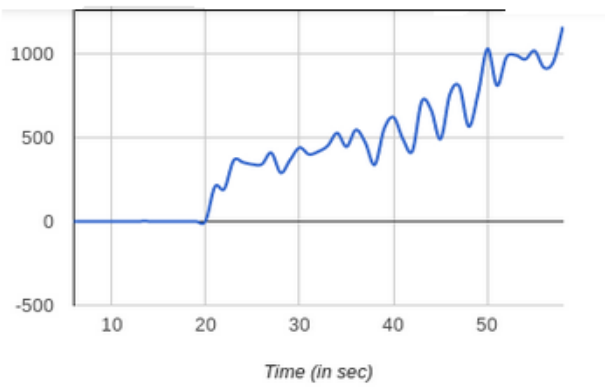

Fig. 4 DDoS attack traffic

By using Java SE 1.7, Lyapunov coefficient values are calculated and plotted

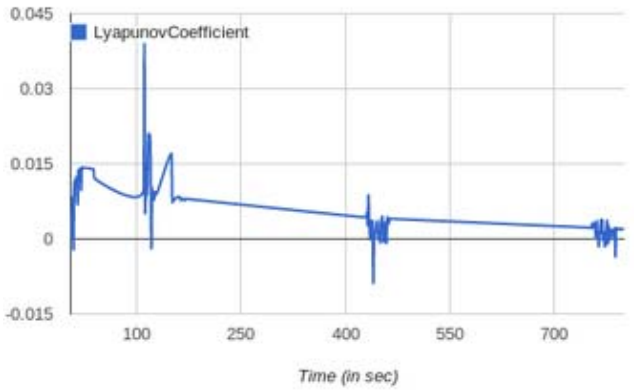

Fig. 5 Traffic chaos pattern- Legitimate burst

Above graph is a legitimate burst traffic pattern Values of Lyapunov coefficient is positive at maximum places.

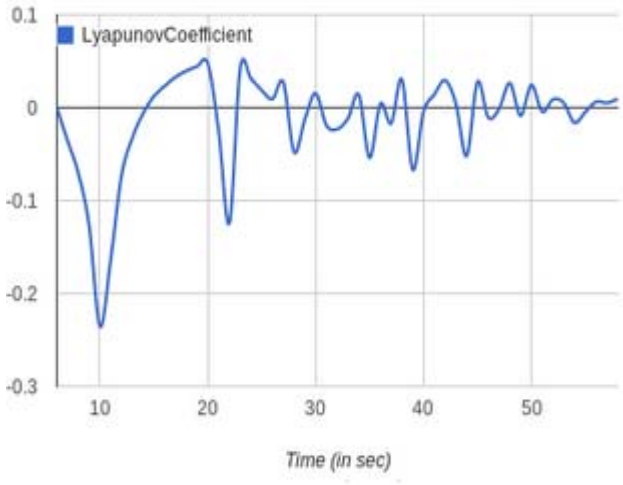

Fig. 6 Traffic chaos pattern- DDoS attack

Above graph is a DDoS attack pattern. In this Lyapunov coefficient values are negative at maximum places.

Unsupervised learning is done on traffic data, by using clustering technique. Groups of similar data or clusters are created by competitive network algorithm. Assuming normal data, bursty legitimate traffic and DDoS attack traffic are three different clusters. Supervised learning is done on the traffic data by using clustered data. Back propagation algorithm is used to reduce the error. This Artificial neural network based detection gave more than 95% accuracy in detection of DDoS attacks.

V. CONCLUSION

Proposed detection method based on Chaos theory effectively detects DDoS attacks. It differentiates DDoS attacks and legitimate burst traffic. To improve he detection efficiency we used supervised and unsupervised learning techniques of artificial neural networks. More than 95% accuracy is obtained because of this.

ACKNOWLEDGMENT

With immense pleasure, we are presenting this paper as a part of the curriculum of M.E Computer Engineering. We are very thankful to our guide, for guidance, encouragement, co-operation and timely help during the preparation phase, because of which we could complete our work.

REFERENCES

[1] K Munivara Prasad , Dr A Rama Mohan Reddy , Dr K VenugopalRao ”An Efficient Detection of Flooding Attacks to Internet Threat Monitors (ITM) using Entropy Variations under Low Traffic” 10.1109/ICCCNT.2010.

[2] Esraa Alomari, Selvakumar Manickam, B. B. Gupta, Shankar Karuppayah, Rafeef Alfaris “Botnet-based Distributed Denial of Service (DDoS) Attacks on Web Servers: Classification and Art” *International Journal of Computer Applications (0975 – 8887)*, vol. 49, no.7, July 2012.

[3] Robin Doss, Shui Yu “Traceback of DDoS Attacks Using Entropy Variations,” *IEEE Transactions on parallel and distributed system*, vol. 22, no. 3, March 2011.

[4] Chonka, J. Singh, and W. Zhou, “Chaos theory based detection against network mimicking DDoS attacks,” *IEEE Communication Letters*., vol. 13, no. 9, pp. 717–719, 2009.

[5] Yonghong Chen, Xinlei Ma, Xinya Wu, “DDoS detection algorithm based on pre-processing network traffic predicted method and Chaos theory,” *IEEE Communications Letters*, vol. 17, no. 5, May 2013.

[6] Carlos Gershenson “Artificial neural networks for beginners”.

[7] Gerhard Munz, Georg Carle “Real-time analysis of flow data for network attack detection,” *IEEE international symposium*, 2007.

[8] Vidushi Sharma Sachin Rai Anurag Dev “A Comprehensive Study of Artificial Neural Networks” IJARCSSE, vol. 2, issue. 10, 2012.

[9] Monowar H. Bhuyan, H. J. Kashyap, D. K. Bhattacharyyaand J. K. Kalita “Detecting Distributed Denial of Service Attacks: Methods, Tools and Future Directions,” *The Computer Journal (2013)* doi: 10.1093/comjnl/bxt031.

[10] Dingding Zhou, Songling g, Shi Dong “Network traffic prediction based on ARFIMA model ” IJCSI arXiv: 1302.6324[cs.NI], February 2013.

[11] Vern Paxson “ An analysis of using reflectors for distributed denial of service attacks,” *ACG SIGCOMM Computer Communication Review*, vol. 31, issue. 3, July 2001.

[12] Jie Wang, Raphael C.-W. Phan, John N. Whitley and David J. Parish “Augmented Attack Tree Modeling of Distributed Denial of Service and Tree Based Attack Detection Method,” *10th IEEE International Conference on Computer and Information Technology*, 2010.

[13] Carlos Gershenson “Artificial Neural Networks for Beginners,” arXiv.org, 2003.