International Journal of Applied Engineering Research ISSN 0973-4562 Volume 10, Number 21 (2015) pp 42183-42188

Time Orient Multi Model Traffic Analysis for Efficient Botnet Detection in Internet Communication

P. Panimalar

Research Scholar, Research and Development Centre, Bharathiar University, Coimbatore-641046, India.

Dr. K. Rameshkumar

Research Supervisor, Research and Development Centre, Bharathiar University, Coimbatore-641046, India.

ABSTRACT:

The growth of internet communication and technology development paves the way for malicious user to perform variety of threats to spoil the communication efficiency. Among them, the botnet is the major threat performed by group of malicious and compromised nodes intended to spoil the data transfer happening in any internet network. There are many approaches has been presented either based on traffic or by behavior of nodes present in the network, but suffers with the accuracy and efficiency of identifying the botnet. To solve the problem of botnet detection, an efficient time orient multi model traffic analysis has been presented in this paper. The method performs botnet detection in many ways, by time orient traffic analysis and time orient behavioral analysis and Stream based approach. In traffic analysis approach, the node performs time orient traffic analysis on each of the routes available and with the list of routes being used to perform transmission of packets. At the behavioral based approach, the method uses the time variant behavior of each node in sending the data packets. The stream based classification approach, counts the stream type and the route being followed to perform the botnet detection. The method produces more efficient results in botnet detection by not only identifying the presence of botnet but identifies the members of the botnet also.

Key Terms: Botnet Detection, Internet Threat, Time orient Multi Model Traffic Analysis, Internet Communication

1. INTRODUCTION:

The internet communication has involved data transfer between variety of computers like routers, switches and servers. As of the data has to be transfer through different computers the data packet traverse through number of nodes in order to deliver the packets to the destination. There is a huge chance for the packets to be traverse through malicious nodes which can perform variety of attacks over the data transfer. The intermediate malicious node can learn many things from the data transfer like the destination where the service is running and the amount of data being sent by the source node and the kind of data, payload of data and TTL value of the packet and many more. By learning such features of the internet packets, the malicious node can perform many attacks like distributed denial of service attacks by producing enormous amount of packets or by generating packets with more pay load and so on. These kinds of internet attacks can spoil the quality of service of the service provided by the internet server and reduce the throughput of the network.

How these malicious nodes are located and how they perform such malicious activities is the question in real time scenarios. Sometimes a single or few nodes may present in the network and perform such malicious activities. But in other case, there will be group of nodes where there exist a controller and organize the network threat using the list of nodes. Such scenario can be named as botnet where there exist a controller and set of compromised nodes which acts based on the regulations of botnet controller. The internet traffic with the presence of botnet gets affected in more range where the generation of denial of service attack by the botnet will affect the internet traffic heavily. The presence of botnet can be identified in different ways like the flow based approaches, which uses the traffic arises in any specific path or in overall network. If the flow in the particular route is heavier than previous times then it can be concluded as there exist a botnet. On the other case, the botnet detection can be performed based on routing information.

The detection of botnet based on traffic flow or routing information will not be effective because the botnet controller or the compromised nodes may use different identity to get selected in the forwarding route. So to perform the botnet detection in efficient manner then the approach has to consider more features and some efficient measures has to be computed. To improve the efficiency of botnet detection, the time orient multimodel approach can be used. The time orient multi model approach, clubs different approaches like flow based technique which uses the internet traffic and the trace about the internet traffic, secondly the method considers the behavioral approach which monitors the traffic behavior of the nodes of network. Finally the method uses, the stream features to classify the internet traffic against different classes.

By considering more features and models in the classification of internet traffic the accuracy of botnet detection can be improved. In any network, the nodes selects the forwarding route based on the traffic, distance, number of hops or the time to live values. So whatever the routing strategy being used, the methods considers these features in group or single. Otherwise the nodes chooses shortest path or longer path if at all the packet has more TTL value. In case of

stream feature based approaches, the methods consider the packet features, for example, the method may choose a longer path for the more TTL valued packet or it may choose shortest path if the packet has only less TTL value.

2. RELATED WORKS:

There are number of approaches has been discussed for the improvement of botnet detection and we discuss some of the methods here in this section around the problem.

A Mulitiprocess Mechanism of Evading Behavior-Based Bot Detection Approaches [1], propose a novel evasion device of bot, multiprocess device. We first classify binary specific topographies of multiprocess bot: unravelling C&C connection after malicious performances, and assigning hateful behaviors to numerous processes. Then we hypothetically analyze why behavior-based bot discovery approaches remain less effective through multiprocess bot. After that, we current two critical contests of applying multiprocess bot. Then we instrument a single procedure and multiprocess bot, then use signature besides conduct detection methods to assess them. The results designate that multiprocess bot container effectively reduction the discoverylikelihood compared through single process bot. Finally we propose the conceivable multiprocess bot buildings and extension rules, besides expect they can cover most circumstances.

Unified P2P Botnet Detection by means of Behavioural Analysis besides Graph Analysis [3], propose a novel method for noticing P2P botnets. Discovery is based on uniting behavioural analysis with organized graph examination. First, our implication method exploits a important property of botnet design. Modern botnets use peer-to-peer message topologies which are fundamental to botnet pliability. Second, our method spreads conventional graph-based detection by joining behavioural examination into organized graph analysis, therefore uniting graph-theoretic discovery with behavioural discovery under a solitary algorithmic outline. We carried out assessment over real-world P2P botnet traffic and demonstration that the subsequent algorithm can localise the majority of bots with little false-positive rate.

On the Effectiveness of Different Botnet Detection Approaches [4], examine four different botnet uncovering approaches grounded on the method used and type of data working. Two of them are community rule based schemes (BotHunter and Snort) then the additional two are data mining based methods with dissimilar feature extraction approaches (packet payload based and traffic flow based). The presentation of these schemes range after 0% to 100% on the five publicly obtainable botnet data sets working in this work. We deliberate the assessment results for these dissimilar systems, their topographies and the models erudite by the data removal founded techniques.

A Technique for Detection of Bots which are using Polymorphic Code [5], The new-fangled technique of botnet uncovering which bots use polymorphic cipher was proposed. Performed uncovering is based on the multi-agent system by earnings of antiviral negotiators that cover sensors. For discovery of botnet, which bots use polymorphic cypher, the heights of polymorphism were examined and its replicas were constructed. A new device for polymorphic code discovery within antivirus agent of multi-agent system remained industrialized. Developed sensor does provocative movements against perhaps infected file, restarts of the doubtful file for probably modified code discovery, behavior examination for adapted code discovery, based on the values of documented levels of polymorphism.

PeerShark: Flow-clustering then conversation-generation aimed at malicious peer-to-peer traffic documentation [6], present a practice to detect P2P botnet traffic and distinguish it from benign P2P circulation in a net. Our method neither shoulders the obtainability of some ‘seed’ material of bots nor trusts on deep package inspection. It aims to perceive the surreptitious behavior of P2P botnets. That is, we intention to detect P2P botnets after they lie sleeping (to evade detection by intrusion detection systems) or though they complete malicious doings (spamming, password stealing, etc.) in a way which is not noticeable to a net administrator. Our approach PeerShark syndicates the welfares of flow-based then conversation-based tactics with a two-tier building, and speeches the limits of these methods. By removing statistical topographies from the net traces of P2P requests and botnets, we build oversaw mechanism knowledge models which can accurately distinguish between kind P2P requests and P2P botnets. PeerShark could also detect unidentified P2P botnet circulation with high accuracy.

Improved Detection of P2P Botnets through Network Behavior Analysis [7], suggests a model to discriminate P2P botnet knowledge and regulator network road traffic from standard traffic at higher proportion of both the programmes using collaborative of conclusion trees classifier named Random Forests. Further to augment the performance, this prototypical also addresses the problematic of unfair nature of dataset using methods like downsampling then cost sensitive knowledge. Performance examination has remained done on the proposed perfect and evaluation consequences demonstration that true optimistic rate for both botnet and genuine classes are additional than 0. 99 whereas untrue positive rate is 0. 008.

Phoenix: DGA-Based Botnet Tracking and Intelligence [8], propose Phoenix, a device that, in addition to effective DGA- and non-DGA-generated areas apart using a mixture of string and IP-based topographies, typifies the DGAs behind them, then, most importantly, discoveries groups of DGA-generated areas that are illustrative of the individual botnets. As a result, Phoenix can subordinate previously indefinite DGA-generated fields to these clusters, and produce novel information about the growing behavior of both tracked botnet. We appraised Phoenix on 1, 153, 516 provinces, including DGA-generated provinces from modern, well-known botnets: deprived of management, it properly illustrious DGA- vs. non-DGA-generated domains in 94. 8 out

of a hundred of the cases, branded relations of domains that fit to distinct DGAs, and helped investigators "on the field" in meeting intelligence on doubtful domains to classify the correct botnet.

Hybrid Botnet Detection Mechanism [9], analyzes plans of different detection methods. The newspaper tries to discovery features that, when joint together, accompaniment each other's strengths and remove the faintness and suggests a framework containing of a mixture of those features which, hypothetically, should overwhelmed most of the shared problems confronted by detection methods.

Detecting HTTP Botnet using Artificial Immune System [11], future a new general HTTP Botnet discovery outline for real time network by Artificial Immune System (AIS). Generally AIS is a novel bio-inspired model which applies to resolving various glitches in information safety; we used this idea in our future framework to make it additional well-organized in discovery of HTTP Botnet. Hence lastly in this paper, we rummage-sale AIS to detect effectively malevolent activities such as spam in addition port perusing in Bot diseased hosts to perceive these malicious adventures kit from a processor system. Our new evaluations show that our approach can perceives HTTP Botnet happenings successfully through high efficiency and low false optimistic rate.

All the above discussed approaches has the problem of detecting the presence of botnet and produces more missing cases due to the usage of minimum features. To overcome the problem of botnet detection, an more efficient comprehensive solution has to be developed.

3. TIME ORIENT MULTI MODEL TRAFFIC ANALYSIS:

The approach performs botnet detection using different methods namely flow orient detection, behavioral detection, and stream feature analysis approach. The flow orient approach counts the time orient traffic occur in the route being considered, the behavioral scheme uses the time orient style of sending data packets by each node, whereas the stream feature technique uses the features of the packet to perform botnet detection. The approach has the following stages namely: Feature Extraction, Flow Filtering Technique, Behavioral Monitoring Approach, Stream Analysis, Log Generation. Each will be discussed in detail in this section.

The Figure 1, shows the architecture of botnet detection approach proposed and shows the functional components in detail.

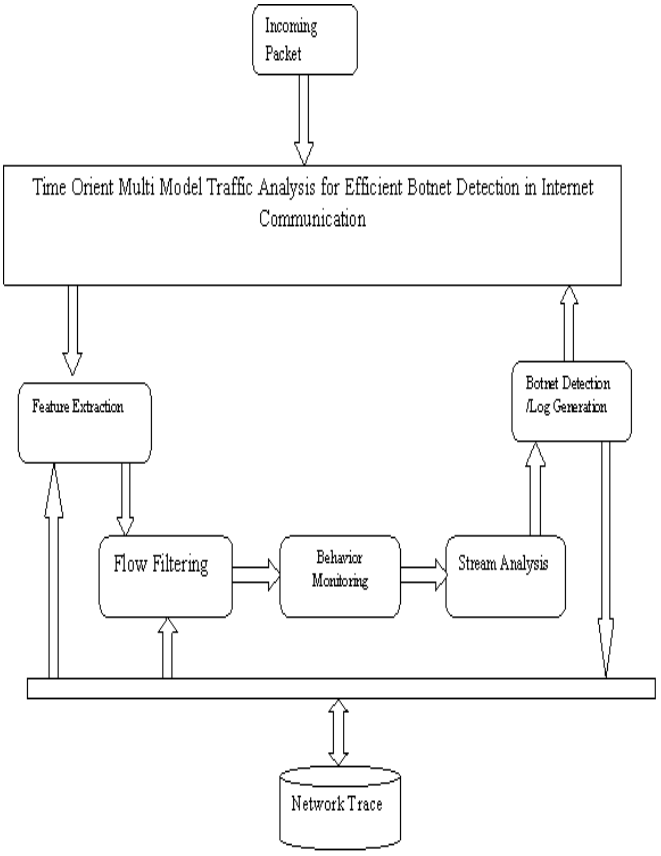


Figure 1: Architecture of proposed botnet detection approach

3. 1 Feature Extraction:

The method maintains different traces of packets being received from different source nodes which pass through number of hops. Whatever the packet being received for example, from the received packet P, the features like source node address Saddr, time being sent Ts, TTL value, Number of hops follows Nh, List of hop addresses Hlist, Payload of the packet Pl will be extracted and stored in the trace. So that the trace form the following pattern:

Trace instance Ti = {Saddr, Daddr, Ts, Ttl, Nh, Hlist{N1, N2, …, Nn}, Pl}. (1).

The above generated trace instance will be added to the set of traces TRS which will be used to perform botnet detection by different schemes.

The addition of the generated trace can be represented by the equation (2).

$$TRS = \sum_{i=1}^{size\ (TRS)} \sum TRS(i) \cup Ti \quad (2).$$

The above mentioned process can be named as the feature extraction phase and can be presented in form of algorithm as follows:

Pseudo Code of Feature Extraction: Input: Raw Network Packet Np. Output: Traffic Instance Ti

Read Packet Np.
Extract Features Saddr, Daddr, Ts, Ttl, Nh, Hlist, Pl.
Construct feature vector Fv_i according to equation (1).
Add feature vector to trace TRS based on equation (2).

The above discussed algorithm receives the incoming packets from the network and extracts the packet features to generate the feature vector. The generated feature vector has been added to the trace which will be used to perform botnet detection later.

3. 2 Time Orient Flow Filtering Technique:

The methodology reads the generated feature vector from the trace and extracts the traces generated for the source being identified. Using the traces belongs to the source S_i, the method splits them according to different time domain Tm_i. For each time domain the method identifies the possible routes in the network and computes the frequency of packets being followed by the route through which the current packet travelled. Now the method computes the traffic rate for the current time window and performs analysis. If the current traffic rate is higher than the previous one then it is concluded as the presence of botnet. Once the presence of botnet is identified then the method extracts the list of hops participated in previous time window and list of hops present in current time window. Based on both the hops list, the method identifies the varying hops from the packet traversal routes. The varying hops represents the probable hops participated in botnet may be of controller or compromised nodes.

Pseudo Code of Time Orient Traffic Analysis:
Input: Traffic Trace Ts.
Output: List of Hops Hl.
Read Traffic Set Ts.
Read packet P.
Extract Features Fv.
Extract traces belongs to the source Identified Ss = $\int_{i=1}^{size\,(Ts)} \sum Ts(i).Saddr = Fv.Saddr$
Split traces into different time domain.
Tss = $\int_{i=1}^{size\,(tw)} \sum Ts(i).Time == Twi$
For each time window Twi
Compute traffic introduced TI = $\frac{\sum Traces eTss(i)}{size\,(SS)}$
End
Compute standard deviation Tstd = $\sqrt{\sum Ti} \in Twi$
If Dist(CTI-CTI_{-1})>Tstd Then
Collect set of all routes followed.
Rf = $\int_{i=1}^{size\,(SS)} \sum Hlist(SS(i)) \nexists Rf$
Extract Uncommon nodes Hl = $\int_{i=1}^{size\,(Rf)} \sum Hlist(Rf(i)) \nexists UN$
End

The above discussed algorithm performs botnet detection and extract some uncommon hosts or newly used hosts to perform identification of botnet members.

3. 3 Time Orient Behavioral Technique:

The behavior of any node represents not only how the data being set but also the way how it propagates the packet towards destination. There may be number of routes to reach a destination but the genuine user always use the protocol being enforced in the network. The protocol may be to follow a shortest route or traffic free route. Based on identifying this variance the presence of botnet can be identified. In this method, the approach first identifies the possible occurrence of network threat and then identifies the list of hops may be participated. The method collects the list of routes and computes the hop count and based on the route followed the method analyze that whether the node has followed appropriate routing strategies or not.

Pseudo Code of Time Orient Behavioral Technique:
Input: Traffic Trace Ts
Output: Hop List Hl.
Read Traffic Trace Ts.
Read Packet P.
Extract Features Fv from P.
Split Trace into different time window Tw.
Identify set of all routes available AR = $\int_{i=1}^{size\,(Ts)} \sum (Routes \in Ts(i)) \cap AR$
Generate Topology TP.
Compute list of all routes RV.
For each time window Ti
Split traces SS = $\sum Traces\,(Ts).Time == Ti$
Compute Traffic Rate Tr = $\int_{i=1}^{size\,(Tw)} \frac{\sum TraceseTs}{size\,(Ts)}$
End
Compute standard traffic deviation Tdev = $\sqrt{\sum Ti} \in Twi$
If (Tr. current-Time-window –Tr. CTW-1)>Tdev Then
Collect the routes followed.
Rf = $\int_{i=1}^{size\,(SS)} \sum Hlist(SS(i)) \nexists Rf$
Compute common hops Chops = $\sum hops \in \forall (Rf)$
Extract the routes with the Chops from the current trace.
Extract the hops Active-Hops Hl = $\sum Hops(Chops) \in T\,races\,(Ti).Routes$
End

The above discussed algorithm performs behavioral analysis to identify the presence of botnet and collects the list of common hops available in the current trace which may be participated in botnet attacks.

3. 4 Stream Based Botnet Detection:

The stream based approach consider the features of the packet like the payload, TTL value and the route being followed. Using the above mentioned features the method perform monitoring of the incoming packet and if the packet is identified with more payload then the method identifies the

botnet with the help of traffic analysis approach, and if the TTL value has crossed the limit then it performs the behavioral technique to get the common hops. The result will be returned to the botnet detection mechanism.

```
Pseudo Code of Stream based Approach:
Input: packet P
Output: Hop List Hl.
Read Packet p.
Extract Payload Pl.
Extract TTL.
If Pl> Payload-Threshold then
Hl = Time-Orient_Traffic_Analysis(P)
Else
Hl = Behavioral_approach(P).
End
```

The above discussed algorithm extract the stream features and the extracted features are used to identify the presence of botnet in the network.

3. 5 Botnet Detection:

At this stage, the method combines all the multiple models of botnet detection to identify the nodes which are compromised with the botnet controller. The method invokes all the models whenever a packet is being received. Based on the result from the all three models, the method selects the common hops which are present in the hop list returned from the all three models. The identified common nodes are the compromised nodes which are involved in botnet attack.

```
Pseudo Code of Botnet Detection:
Input: Trace Ts, Packet P.
Output: List of malicious nodes.
Start
Hop list Hl = Traffic_Analysis(P).
Hop List HlB = Behavioral_Analysis(P).
Hop List HLS = Stream_Analysis(P).
Identify common hops from HL, HLB, HLS.
Propage them as malicious in the network.
Stop.
```

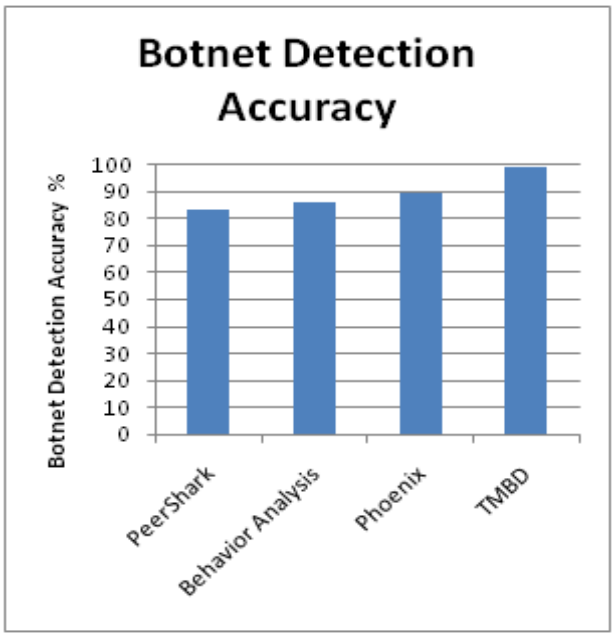
4. RESULTS AND DISCUSSION:

The time orient multi model botnet detection approach has been implemented and evaluated for its efficiency in botnet detection. The approach has been validated with different data sets and real time analysis. The details of simulation has been presented here:

Table 1: Details of simulation being used

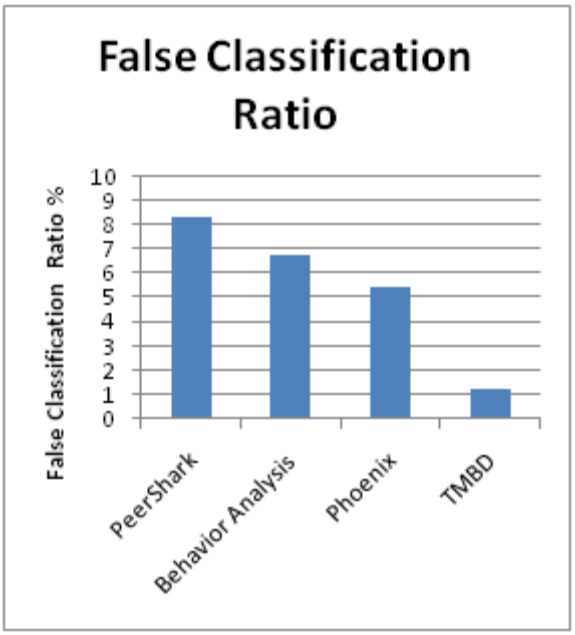
Parameter	Value
Simulator	Advanced Java
Number of Nodes	200
Simulation Time	5 Minutes
Protocol	TMBD
Simulation Area	1000×1000 meters
Node Range	100 meters

The Table 1, shows the details of simulation being used to evaluate the performance of the proposed multi model approach for botnet detection.



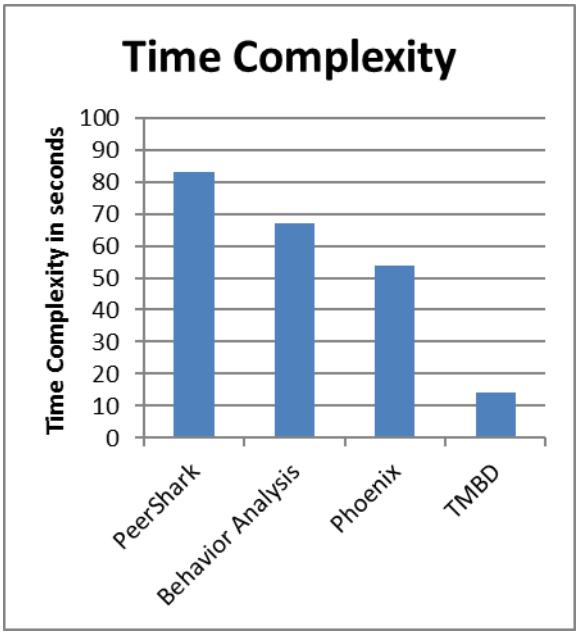
Graph 1: Comparison of botnet detection accuracy

The graph1, shows the comparative result on botnet detection produced by different methods and it shows clearly that the proposed method has produced more efficient accuracy than other methods.



Graph 2: Comparison of false classification ratio

The Graph 2, shows the comparison of false classification ratio produced by different methods and it shows clearly that the proposed method has produced less false classification ratio than other methods.



Graph 3: Comparison of time complexity

The Graph 3, shows the comparative result of time complexity produced by different methods and it shows clearly that the proposed method has produced less time complexity than other methods.

5. CONCLUSION:

The problem of botnet detection has been studied in different dimensions and the botnet detection approach has been presented in different models. The proposed multi model time orient analysis method utilizes various features and metrics of network traffic to perform botnet detection. The method clubs features and efficiency of three different models like traffic orient method, behavior analysis approach and stream based methods. The results from all these three methods has been used to perform botnet detection and the method has produced efficient results.

REFERENCES

1. Yuede Ji, Yukun He, Dewei Zhu, Qiang Li, Dong Guo, A Mulitiprocess Mechanism of Evading Behavior-Based Bot Detection Approaches, Springer, Information Security Practice and Experience, Volume 8434, 2014, pp 75-89.
2. Ahmad Karim, Rosli Bin Sa lleh, Muhammad Shiraz, Syed Adeel Ali Shah, Irfan Awan, Nor Badrul Anuar, Botnet detection techniques: review, future trends, and issues, Springer, Journal of Zhejiang University SCIENCE CNovember 2014, Volume 15, Issue 11, pp 943-983.
3. Shishir Nagaraja, Botyacc: Unified P2P Botnet Detection Using Behavioural Analysis and Graph Analysis, Springer, Computer Security - ESORICS 2014Lecture Notes in Computer Science Volume 8713, 2014, pp 439-456.
4. Fariba Haddadi, Duc Le Cong, Laura Porter, A. Nur Zincir-Heywood, On the Effectiveness of Different Botnet Detection Approaches, Springer, Information Security Practice and Experience Lecture Notes in Computer Science Volume 9065, 2015, pp 121-135.
5. Oksana Pomorova, Oleg Savenko, Sergii Lysenko, Andrii Kryshchuk, Andrii Nicheporuk, A Technique for Detection of Bots Which Are Using Polymorphic Code, Springer, Computer Networks Communications in Computer and Information Science Volume 431, 2014, pp 265-276.
6. Pratik Narang, Chittaranjan Hota, VN Venkatakrishnan, PeerShark: flow-clustering and conversation-generation for malicious peer-to-peer traffic identification, Springer, EURASIP Journal on Information SecurityOctober 2014, 2014:15,
7. Shree Garg, Anil K. Sarje, Sateesh Kumar Peddoju, Improved Detection of P2P Botnets through Network Behavior Analysis, Springer, Recent Trends in Computer Networks and Distributed Systems Security Communications in Computer and Information Science Volume 420, 2014, pp 334-345.
8. Stefano Schiavoni, Federico Maggi, Lorenzo Cavallaro, Stefano Zanero, Phoenix: DGA-Based Botnet Tracking and Intelligence, Springer, Detection of Intrusions and Malware, and Vulnerability AssessmentLecture Notes in Computer Science Volume 8550, 2014, pp 192-211.
9. Katha Chanda. Article: Hybrid Botnet Detection Mechanism. International Journal of Computer Applications 91(5):12-16, April 2014.
10. Hossein Rouhani Zeidanloo, Azizah Bt Abdul Manaf, "Botnet Detection by Monitoring Similar Communication Patterns", (IJCSIS) International Journal of Computer Science and Information Security, Vol. 7, No. 3, March 2010, pp. 36- 45.
11. Amit Kumar Tyagi and Sadique Nayeem. Article: Detecting HTTP Botnet using Artificial Immune System (AIS). International Journal of Applied Information Systems 2(6):34-37, May 2012.
12. Zeidanloo, H. R. ; BT Manaf, A. ; Vahdani, P. ; "Botnet Detection Based on Traffic Monitoring", International Conference on Networking and Information Technology, 2010.
13. Hossein Rouhani Zeidanloo, Azizah BT Abdul Manaf et. al "A proposed framework to detect P2P Bots", IACSIT International Journal of Engineering and Technology, Vol. 2, No. 2, April 2010