Visualization of Invariant Bot Behavior for Effective Botnet Traffic Detection

Alireza Shahrestani, Maryam Feily*, Mona Masood, Balakrishnan Muniandy

*IEEE Graduate Student Member

Universiti Sains Malaysia (USM), 11800 Penang, Malaysia
alishasha@gmail.com, maryamfeily@gmail.com, msmona@usm.my, mbala@usm.my

Abstract— **Due to the sharp rise in computer network attacks through botnets, current security monitoring tools will be insufficient for effective botnet traffic detection. In fact, most of the existing tools are text-based and there is a lack of effective user friendly interface that can facilitate detection of botnet traffic in large datasets. Moreover, most of these tools are based on reactive approaches and will be triggered only after an attack is detected. Therefore, enhancement of botnet traffic detection is highly demanded. Knowledge discovery through information visualization is an avenue to solve these issues effectively. The aim of this research is to propose a proactive approach by adopting proper visualization techniques to increase the visibility of network traffic related to invariant bot behavior and botnet activities. The visualization techniques used in this research consist of graphs, scatter plots, and histograms. These visualization techniques are easy to interpret and good for visualizing large datasets. By adopting these techniques for invariant bot behavior visualization, it is possible to provide visual notification of bot existence in a network without distracting the user with huge volumes of data. In fact, the visual illustration of typical bot behavior improves the botnet traffic detection process by engaging human perception and intellectual capabilities. Overall, this visual approach can assist the security personnel to proactively detect invariant bot behaviors and botnet activities during the benign state of a botnet by providing a graphical user friendly interface. Exploiting the visual information, human analysts and security personnel will be able to gain more insights into their networks, leading to make correct decisions in critical situations and to prevent catastrophic botnet attacks.**

Keywords— Bot Behavior; Botnet Detection; Network Monitoring; Security; Visualization.

I. INTRODUCTION

Botnets are emerging as the most significant threat facing online ecosystems and computing assets due to their enormous volume and sheer power [1]. According to explanation in [2, 3] malicious botnet is a network of compromised computers called "Bots" under the remote control of a human operator called "Botmaster". Botnets are predominantly used for illegal activities. For instance, the botmaster can instruct his bot army to recruit new bots, launch coordinated Distributed Denial of Service (DDoS) attacks against critical targets, steal sensitive information from infected machines, malware dissemination and force distribution, phishing, click fraud, and so on [2, 4, 5]. However, the highlight value of botnets is the ability to provide anonymity through the use of a multi-tier command and control (C&C) architecture. Moreover, the individual bots are not physically owned by the botmaster, and may be located in several locations spanning the globe. Thus, differences in time zones, languages, and laws make it difficult to track malicious botnet activities across international boundaries [3, 4, 6]. These characteristics make botnet an attractive tool for cyber-criminals, and in fact pose a great threat against cyber-security [7].

Due to the emergence threat of botnets, botnet detection and mitigation has been a major research topic in recent years. Different solutions have been proposed in academia. However, most of the proposed techniques are reactive approaches that will be triggered only after an attack is detected. Nevertheless, it is possible to detect botnets proactively in the benign state of a botnet. Moreover, most of the existing solutions are text-based and require significant amount of time and effort to determine the malicious traffic related to botnet activities. The poor user interface of the existing tools leads to the insufficient utilization of the captured data, and do not consider utilization of human intellectual capability to ensure evidence of botnet activities. Therefore, enhancement of botnet traffic detection is highly demanded.

Knowledge discovery through information visualization is an avenue to solve these issues effectively. The visual presentation of information take advantage of visual and cognitive powers of humans to reduce effort required for processing complex information. The mapping of data parameters to locations, colors or shapes produces images that make it easier to detect pattern and relations. Therefore, visualization can be exploited as an ideal tool that allows active exploration of a knowledge space [8]. Specifically, in this research we have adopted information visualization techniques to facilitate knowledge discovery about invariant bot behavior to assist security personnel in botnet mitigation [7].

The aim of this research is to create a *"Visual Threat Monitor" (VTM)* which can detect botnet activities effectively prior to the attack. In order to achieve this goal, we have proposed a visual approach to enhance botnet traffic detection [7, 9]. The proposed visual network monitoring tool will assist the security personnel to recognize the threats more easily and effectively as the visualized information is easier to comprehend for human beings to gain useful knowledge. The visualization techniques adopted in this system are easy to interpret and good for visualizing large datasets. This paper emphasizes exclusively on the visualization features in *"Visual Threat Monitor"* development to presents the effective role of visualization to increase the visibility of network traffic related to the attributes of typical bot behaviors.

978-1-4673-4786-0/12/$31.00 ©2012 IEEE

The remainder of this paper is organized as follows: Section II provides a brief overview on typical bot behaviors and characteristics. Section III explains the appropriate visualization techniques adopted to detect invariant bot behaviors, supported with the screen shots of the implemented system. Section IV evaluates the prototype and finally, the paper will be concluded in Section V.

II. INVARIANT BOT BEHAVIORS AND CHARACTERISTICS

According to our survey on botnet phenomenon, a typical botnet can be created and maintained in five phases including: initial infection, secondary injection, connection, malicious command and control, update and maintenance [4]. Due to the evolving aspects of the botnet's life-cycle, botnet activities and bot behavior have become very dynamic. However, there are some invariant bot behaviors that can be detected during a benign state of a botnet. Since botmasters change their strategy of infecting new machines frequently, botnet activities and bot behavior in initial infection and secondary injection phases of botnet life-cycle are very dynamic, and considering bot behavior in these phases will not be practical for botnet detection. Therefore, in order to identify invariant bot behaviors, we focus on three phases namely: connection, malicious command and control, update and maintenance phases. These phases are depicted in Fig. 1.

Previous researchers have identified some of the primary bot behaviors. However, each detection algorithm or technique makes different set of assumptions about botnet behavior based on their own goals [7]. In this research, we are focused on a limited set of the invariant bot behaviors as follows:

Fast Response Time: Bots reply to the botmaster's command very fast. In other words, human response to a command or request is always much slower than a bot [10].

Small Size Commands: The lengths of command packets are typically very small. Despite, the normal packets that have unbounded size, a typical command packet from botmaster has a small size of *1KB* or even less [10].

Instant Execution of Commands: Bots may launch an executable application on the infected host machine immediately after receiving botmaster's command [10].

Figure 1. Three phases of botnet life-cycle considered for Invariant bot behavior identification.

The aim of the proposed *"Visual Threat Monitor"* or *VTM* is to detect evidence of bot existence by identifying such invariant bot behaviors in a monitored network. The proposed visualization techniques will be used to visualize the aforementioned bot behaviors. Thus, we intentionally restrict our data source to the level of session data. Here are the characteristics that the data should posses to reflect the specific bot behaviors [7]:

1. The *"Response Time"* should be fast enough to be considered as a bot response to some commands from botmaster. The threshold of *Response Time* for incoming packets is *100ms*, whereas the threshold for outgoing packets is *3sec*.

2. The size of sessions should be small enough to reflect the small size of command packets. Since the length of typical command packets are often *1KB* or less, the threshold for "*Session Size*" is set to *1KB*. If a Session Size is less than *1KB*, it indicates that the packet might be a command line from an infected host.

3. The "*Time Interval*" between receiving a packet and launching an executable application on the infected host should be small to emulate bot behavior in launching some specific executable applications according to the command received from the botmaster. Similar to the *Response Time*, the threshold of the Time Interval is *100ms*. The goal is to find out if a specific application has been triggered immediately after receiving an incoming packet. This case might indicate that automated bot software is in charge of executing the command coming through from a botmaster.

4. At the same instance of time, if both "*Session Count*" and "*Destination Count*" are high, but the "Average Size" is low, the machine is propagating the bot software.

III. VISUALIZATION OF INVARIANT BOT BEHAVIORS

The proposed security monitoring tool offers various data visualizations to aid security personnel or human analyst in identifying invariant bot behaviors in a network. We have considered the most important principles of graphic interface design and interaction in the HCI (Human Computer Interaction) design model of this system, as described in [7]. Specifically, in VTM consistency has been carefully considered. In VTM consistent sequence of actions will be required in similar situations and identical terminology has been used in prompts, menus, and display screens. Moreover, consistent colors, layouts, and icons have been employed throughout the system. This makes the interface more familiar and predictable for the user. On the other hand, two interaction styles including direct manipulation and menu selection have been employed in this system. The visualization techniques used in this system consist of graphs, scatter plots, and histograms. These visualization techniques are easy to interpret and good for visualizing large datasets. Hence, these techniques will be used to visualize a limited set of invariant bot behaviors including Fast Response Time, Small Size Commands, and Instant Execution of Commands. The proposed visualization techniques will be described in details in this Section.

A. Graph Visualization of Traffic Overview

The graph visualization in this system can help the user to get an overview of the network traffic. The nodes of this graph represent In-house (gray boxes) and Out-bound (black boxes) active machines during the monitoring period specified in the system. The monitoring period depends on the volume of network traffic, and could be hourly or daily. The nodes are then connected using lines with three specific attributes, namely thickness, pattern, and color. The thickness of each line indicates the traffic load between the two nodes measured by the number of sessions between these nodes. Therefore, a thicker line indicates more connections, and accordingly more sessions between the two nodes. In addition, since different threshold are assigned to incoming and outgoing traffic, different patterns are used to visualize these connections. A dotted line is used to show the incoming traffic to an In-house PC, whereas a straight line is used to show the outgoing traffic from an In-house PC. On the other hand, different colors are used to illustrate the speed of responses. If the Response Time of any connection is fast enough to be considered as a response from automated bot software, the line will be shown in red color to illustrate fast response time as an invariant bot behavior. Another metric is the size of sessions. By clicking on each line a pop up bar chart will appear to show the number and the size of sessions. The threshold for the Session Size is 1KB. If a Session Size is less than 1KB, it indicates that the packet might be a command line from an infected node.

Finally, based on the above metrics (Response Time, Session Size and Session Count) the security personnel can determine whether a node is behaving as a bot in the network. A default set of thresholds for these connections is set for the system, but these thresholds can be changed from the setting menu of the system to meet certain requirements in different situations. This property of the proposed system provides flexibility for future security demands [7, 11].The default layout of graph visualization in VTM is shown in Fig. 2. However, as it can be seen in Fig. 2, the default visualization of traffic overview may not be clear to the user due to the dense nodes in the graph. Therefore, a special feature has been added to the system so that the user can drag different nodes and rearrange them in a way that all links and their attributes such as thickness, pattern and color become more visible. A sample snapshot of the rearranged nodes is shown in Fig. 3.

The user can get additional information about the incoming or outgoing flow by simply clicking on the desired line, and see the total number of session and the size of these sessions in a bar chart as shown in Fig. 3. This chart helps the security personnel to determine the status of each computer in the monitored network.

B. Scatter Plot Visualization of Time Intervals

Scatter plot is another visualization technique that is used to facilitate invariant bot behavior detection, so that the analyst can compare different metrics to decide on the status of certain machines in the network. The advantage of this technique is the capability to show an alternative metric on a chart that has its own mapping. This capability is not provided by conventional charts. For instance, in this system color is used to map the "Time Interval". The goal is to find out if a specific application has been triggered immediately after receiving an incoming packet. This case might indicate that automated bot software is in charge of executing the commands coming through from a botmaster [7, 11].

In this visualization, the horizontal axis shows the In-house IPs (Destination IP), whereas the vertical axis shows the IPs from which packets are coming to the monitored network (Source IP). Each point on the scatter plot indicates the Time Interval between the packet arrival time and the time that an application has been executed. This Time Interval is mapped by three different colors including red, yellow and green to illustrate the node behavior as bot-infected, suspicious, or normal, respectively. Moreover, the legend for this visualization is placed on top of the screen, indicating the color mapping of the points for the reference of the user [7, 11]. Fig. 3 displays a sample snapshot of the scatter plot visualization. Here, if more than one executable session exists in the traffic, the worst option that is "Infected" will be set for that point on the scatter plot. In addition, by clicking on each point on the scatter plot, user can find out the number of executed sessions in an additional bar chart. This information is also mapped in to three different colors in the bar chart, accordingly. A sample view of the bar chart is shown in Fig. 3.

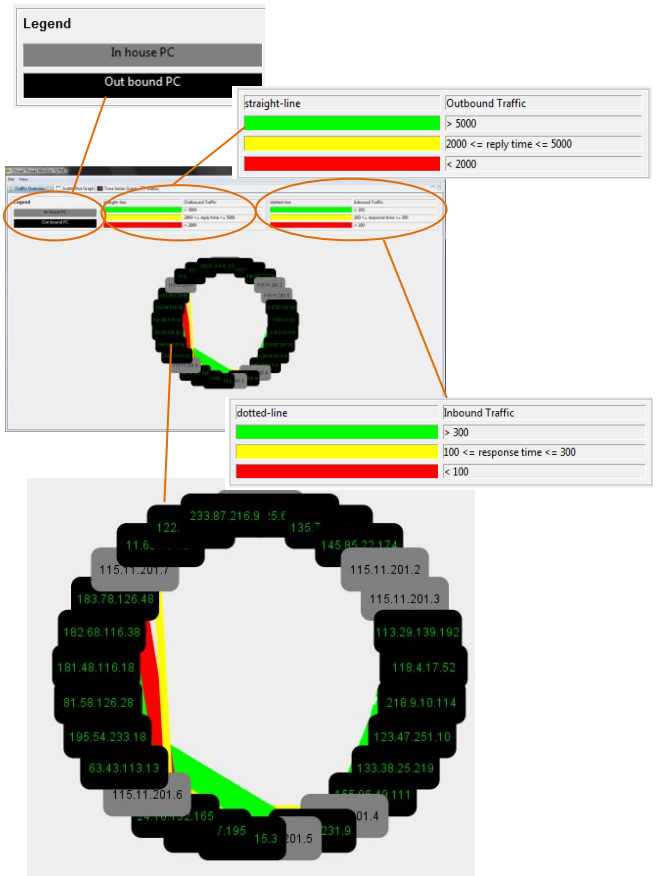

Figure 2. Traffic Overview Visualization.
(The magnified view of important components of the screen shot)

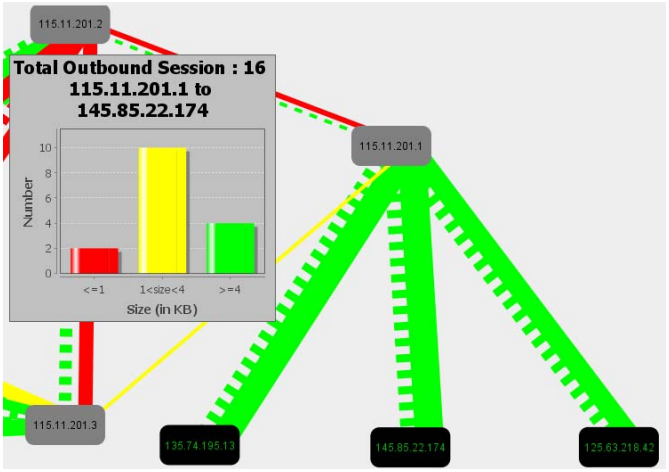

Figure 3. A sample snapshot of the rearranged nodes with the additional bar chart.

C. *Parallel Histogram Visualization for Time Series*

Parallel histograms are used for outgoing packets to find evidence of bot existence in the network. These histograms visualize three different metrics on the time line of 12 hours for each machine in the monitored network. The first histogram shows the "Session Count" that is the total number of sessions in every hour. The second histogram shows the "Destination Count" which is the total number of unique destination addresses in the same hour, and the third one shows the "Average Size" of the packets in that hour. If the total number of sessions and destination addresses are both high, and the average size of same instance is relatively small (less than 1KB), there is high probability that this machine is propagating the bot command or software to other machines [7, 11]. A sample snapshot of the parallel histograms is also shown in Fig. 5. The vertical red lines in this figure indicate the probability of the bot command propagation as an invariant bot behavior.

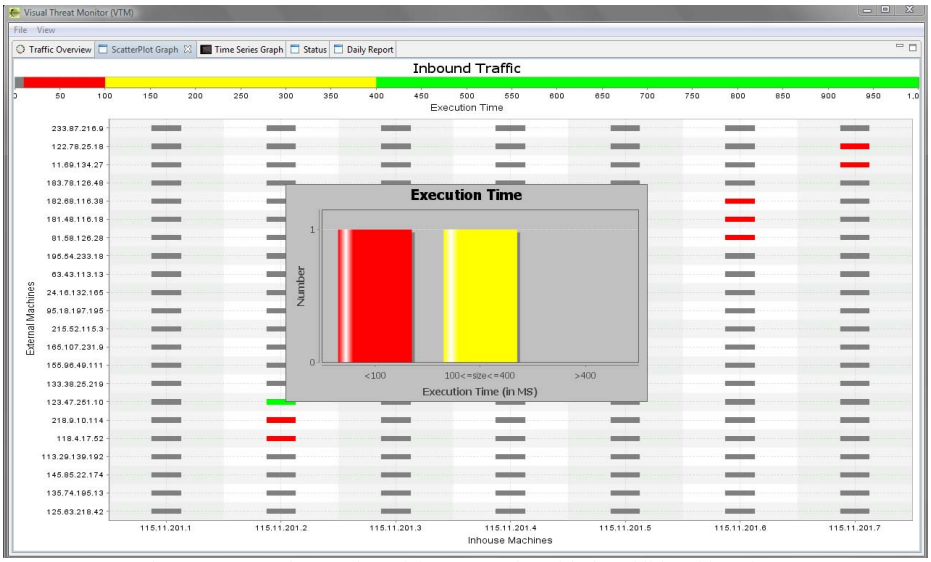

Figure 4. A sample snapshot of the scatter plot with the additional bar chart.

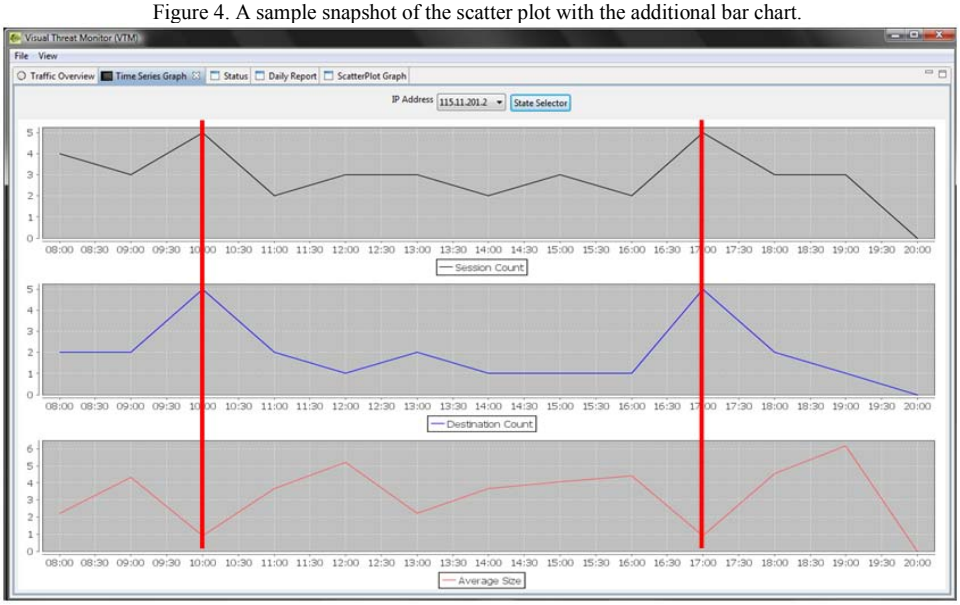

Figure 5. A sample snapshot of the parallel histograms.

IV. PROTOTYPE EVALUATION

Most of previous solutions to combat botnet threats aim to detect botnet attacks. Therefore, these techniques are mostly evaluated based on the detection rate of the system. However, VTM cannot be evaluated based on detection rate, as it is not expected to detect any attack, but to visualize some typical bot behavior in the benign state of a botnet prior to be used for an attack. In fact, since the main focus of this research is on the visualization of invariant bot behavior, VTM is evaluated based on the effectiveness of the proposed visualization techniques and the design of the visual interfaces of the system. The users have assessed the effectiveness of the proposed visualization techniques and the design of the visual interfaces of the system. Moreover, different characteristics of the system have been ranked as the overall assessment of the proposed system.

Overall, majority of users had a good knowledge about network security and network monitoring as they were mostly experts, researchers or developers in the area of network security and network monitoring. Nevertheless, they were less familiar with information visualization and botnet phenomenon. According to the user assessment results, about 64.28% of the polled users were aware of botnet threats emergence. However, all users believed that botnets are great threats against cyber-security, and organizations should protect their networks against botnet threats. Moreover, they all agreed that visualization can assist security personnel to combat and mitigate cyber-threats through botnets [11].

A. *Effectivness of the Proposed Visualization Techniques*

The effectiveness of the proposed visualization techniques and the design of the visual interfaces of VTM have been evaluated by user assessment. According to the evaluation results, 57.14% of experts agreed that the proposed visualization techniques enhance the visibility of network traffic related to typical bot behavior. Besides, all non experts expressed the same opinion. Altogether, 78.57% of the polled users confirmed the effectiveness of the proposed visualization techniques in enhancement of the visibility of network traffic related to invariant bot behavior [11]. Furthermore, the effectiveness of each visualization technique was evaluated separately, and the results are summarized in Fig. 6.

According to the results, all visualization techniques are effective enough to enhance the visibility of network traffic related to typical bot behavior. However, the graph visualization technique which is used to visualize the traffic overview of the network has been ranked as the most effective visualization technique used in VTM [11].

B. *Overall Assessment of VTM*

Finally, in order to perform the overall assessment of VTM users were asked to rank different characteristics of the system. The result of this ranking is summarized in a chart in Fig. 7.

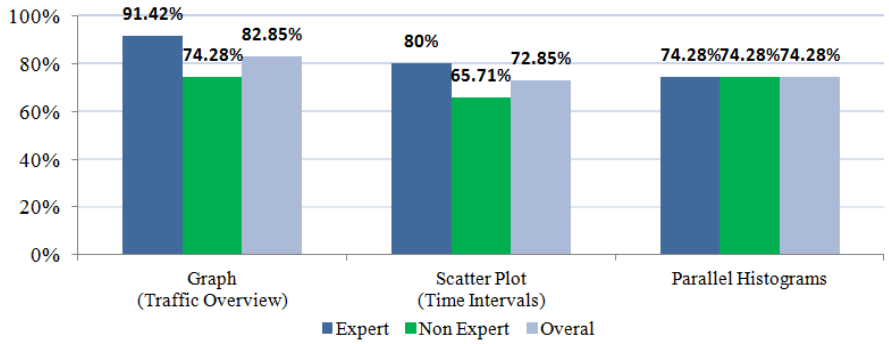

Figure 6. Effectiveness of the proposed visualization techniques.

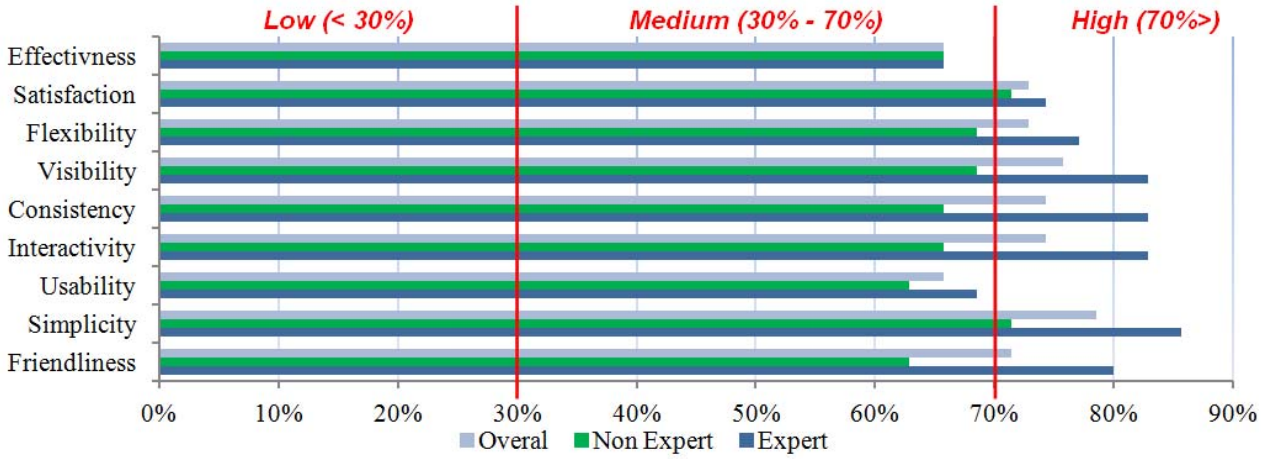

Figure 7. System Characteristics.

Comparing to previous approaches to combat botnet threats, VTM has several advantages. First of all, VTM is a proactive approach to combat botnet threats. As mentioned earlier, security personnel prefer a proactive approach that can detect threats prior to the attack. Unlike reactive approaches that will be triggered only after an attack is detected, VTM is capable to discover invariant bot behavior during the benign state of a botnet. Thus, it helps security personnel to protect their network against further malicious botnet activities to mitigate massive botnet attacks in future.

Moreover, although VTM is a monitoring tool based on passive network monitoring and analysis, it does not require great deal of configuration to provide accurate alerts for suspicious traffic. In fact, the proper visualization techniques used in VTM, can effectively enhance the visibility of suspicious network traffic related to invariant bot behavior and botnet activities. In other words, VTM provides visual notification of bot existence without distracting the user with huge volumes of data that can result in alarm fatigue. The visual illustration of typical bot behavior improves the botnet traffic detection process by engaging human perception and intellectual capabilities.

Last but not least important is the visual interfaces of VTM. Despite most of the existing tools are text based, provides several visual interfaces. The poor user interface of the existing tools leads to the insufficient utilization of the captured data. However, the user-friendly and interactive interfaces of VTM consider utilization of human intellectual capability to ensure evidence of botnet activities. The rich visual interfaces of VTM provide a quick view of the informational goal as it can show much more information in the same viewing space rather than text based information. In addition, details on demand allow the security personnel to access to a deeper level of information interactively.

More importantly, the visualized information is easier to be processed and comprehend for human beings to gain useful knowledge about bot existence in a network. In other words, by incorporating human perception into detection process through visualization, the security personnel can detect evidence of bot existence in a network more easily and effectively.

Overall, VTM is a visual security monitoring tool that can discover existence of botnet threat proactively prior to the attack. The visual interface design of VTM addresses the most important requirement of usability and principles of interface design. By exploiting visual information, human analysts and security personnel will be able to gain more insights into their networks which lead to make correct decision in critical situations.

V. CONCLUSION

In this research, we have developed an effective visual network monitoring tool which is called *"Visual Threat Monitor"* or *VTM*. VTM is a RCP (*Rich Client Platform*) application designed to enhance knowledge discovery about invariant bot behavior in small to medium size networks. The HCI design model of this system is user-centered and maximizes the engagement of users. The visual illustration of typical bot behavior improves the botnet traffic detection process by engaging human perception and intellectual capabilities. Exploiting the visual information, human analysts and security personnel will be able to gain more insights into their networks which lead to make correct decision in critical situations.

Overall, the proposed visual security monitoring system will be highly useful to enhance the capability to combat and mitigate cyber-threats through botnets in small to medium size networks. It can effectively address the usability requirements of security personnel by providing a graphical and interactive user friendly interface. Furthermore, VTM is a flexible and scalable visual approach that can be adjusted according to future security demands. However, it is important to mention that the current system is only designed to visualize some invariant bot behaviors to aid human analyst in botnet traffic detection, and therefore, at this stage the system is not expected to detect any botnet attack. Yet, this system is under further development to become capable for automatic detection of botnet attacks by appending a data mining layer in the system architecture.

ACKNOWLEDGMENT

The authors graciously acknowledge the great support from the UNIVERSITI SAINS MALAYSIA (USM) through the USM Fellowship awarded to Miss Maryam Feily.

REFERENCES

[1]. B. Saha and A. Gairola. "Botnet: An Overview." CERT-In White Paper, CIWP-2005-05, 2005.

[2]. N. Ianelli and A. Hackworth. "Botnets as a vehicle for online crime." CERT Coordination Center, 2005.

[3]. R. Moheeb Abu, Z. Jay, M. Fabian and T. Andreas, "A multifaceted approach to understanding the botnet phenomenon," in *Proc.* 6th ACM SIGCOMM conference on Internet measurement, 2006.

[4]. M. Feily, A. Shahrestani and S. Ramadass, "A Survey of Botnet and Botnet Detection," in *Proc.* SECURWARE 2009, 2009, pp. 268-273.

[5]. T. H. Paul Bächer, Markus Kötter and Georg Wicherski. "Know your Enemy: Tracking Botnets," Retrieved 7th March, 2009, from http://old.honeynet.org/papers/bots/.

[6]. G. P. Schaffer, A. Commun and L. Rock, "Worms and viruses and botnets, oh my!," Rational responses to emerging Internet threats. IEEE security & privacy, 4(3), pp. 52-58, 2006.

[7]. A. Shahrestani, M. Feily, R. Ahmad and S. Ramadass, "Discovery of Invariant Bot Behaviour through Visual Network Monitoring System," in *Proc.* SECURWARE 2010, 2010, pp. 182-188.

[8]. G. B. Judelman, "Knowledge Visualization." Problems and Principles for Mapping the Knowledge Space, 2004.

[9]. A. Shahrestani, M. Feily, R. Ahmad and S. Ramadass, "Architecture for Applying Data Mining and Visualization on Network Flow for Botnet Traffic Detection," in *Proc.* ICCTD 09, 2009, pp. 33-37.

[10]. M. M. Masud, T. Al-khateeb, L. Khan, B. Thuraisingham and K. W. Hamlen, "Flow-based identification of botnet traffic by mining multiple log files," in *Proc.* DFmA 2008, 2008.

[11]. A. Shahrestani, "Discovery of Invariant Bot Behaviour through Visual Network Monitoring System" Dissertation Submitted for the Master of Software Engineering, Faculty of Computer Science and Information Technology, University of Malaya (2011).