Study on Traffic Analysis in Darknet and Livenet for Network Security Management

Napaphat VICHAIDIS

Monitoring and understanding the activities of networked devices is essential for network security management. Understanding the activities of attackers' devices on the Internet provides important information for defensive preparation against attacks on the managed network. Understanding the activities of intranet devices is crucial for grasping ongoing fault and security incidents in the intranet. Network traffic monitoring and analysis play an important role in understanding the activity of a device since such activity is reflected in the traffic sent and received. In this thesis, traffic monitoring and analysis issues in darknet and livenet (i.e., a real operational network) are discussed. Monitoring and analysis of darknet and livenet traffic aid the understanding of attackers' devices and the activities of intranet devices, respectively. Darknet is a set of routed but unused IP address spaces on the Internet in which no active service resides. All traffic present in darknet is due to misconfiguration, malicious activities, attacks, experiments, or errors. Livenet is a real operational network, with diverse device activities on the network. Livenet traffic analysis is a good tool for understanding the diverse activities of network devices in the intranet. The contribution of this thesis involves the analysis of darknet traffic by focusing on its stability to detect and classify the cause. A simple traffic monitoring and analysis technique is proposed for covering the diverse network activity of intranet devices since it has low deployment cost and provides reasonable coverage. Darknet traffic is analyzed by focusing on traffic stability. The concept of traffic stability is that the relative volume of dominant traffic components does not drastically change. Two datasets of darknet traffic are analyzed, revealing several significant instabilities. Furthermore, the traffic instability is analyzed using different timescales to detect small-scale events and applying normalized entropy techniques to compare the differences between stable and unstable timeslots. Livenet traffic is analyzed using a simple model to cover intranet device activity. Traffic is analyzed to extract useful information and provide appropriate protection using a small number of monitors. The analysis results of monitored traffic are then presented for each traffic category and the monitoring network activity discussed to provide insight into potential cyber-attacks in the context of the cyber kill chain model.

1 Introduction

In this thesis, traffic monitoring and analysis issues in darknet and livenet (i.e., a real operational network) are discussed. Monitoring and analysis of darknet and livenet traffic are key for understanding the activities of attackers and intranet devices, respectively. Darknet is a set of routed but unused IP address spaces on the Internet in which no active service resides. Darknet traffic analysis is a powerful tool for monitoring and analyzing the activities of attackers' devices such as worms, distributed denial of service attacks, or scans. Livenet is a real operational network, with a diversity of device activities on the network. Livenet traffic analysis is a good tool for understanding the activities of devices in the intranet.

The objective of this thesis is to understand the activities of attackers' devices on the Internet and intranet. Darknet traffic analysis is proposed, by focusing on the stability of traffic to detect and classify the cause of stability. The traffic stability technique is useful for detecting and classifying abnormal activity in darknet traffic. For livenet traffic analysis, a simple traffic monitoring and analysis technique is proposed to cover diverse network activity since it has low deployment cost and provides reasonable coverage for detecting and classifying the cause of abnormal activities in livenet traffic.

The remainder of this paper is organized as follows. In the next section, an overview is provided of related work on traffic analysis techniques. Section 3 describes the proposed method for analyzing darknet traffic by focusing on its stability. Section 4 presents the researcher's proposal for simple traffic monitoring. It also describes a study on the proposed traffic analysis experiment technique for analyzing each traffic category in livenet traffic. Section 5 summarizes this research.

2 Related work

In this section, the existing work is explored, by focusing on the analysis of traffic on darknet and livenet.

2.1 Darknet analysis

Darknet traffic can be classified as attempts at infection by worm, botnet, misconfigured application, backscatter from spoofed denial of service attacks, and network scanning problems. The authors in [1] focus on Transmission Control Protocol (TCP) sessions, initiated by activities from botnets or worms on the Internet, and propose a multidimensional malicious packet monitoring architecture for Internet threat detection.

The authors utilized the traffic of different darknet monitoring systems. Shomura et al. [2] analyzed three different darknet datasets to study the nature and behavior of attacks by measuring packet distributions, protocol distributions, and protocol types.

Tao Ban [3] analyzed the behavior of attackers to monitor the global trends of cyber-attacks to better predict the future status of attacking hosts on the Internet. Song et al. [4] provided detailed, in-depth analysis results, showing that a correlation analysis between darknet traffic and Intrusion Detection System (IDS) alerts is very useful for discovering potential attack hosts in organizations and installed malware. In [5], statistical estimation techniques were used to propose the use of moments, maximum likelihood, and linear regression estimators to quantify worm infections. Another example is the study by [6], who presented a novel method based on the usual behavior mode of real traffic data to realize rapid detection of distributed scan attacks in the darknet.

The above mentioned and other conventional methods mainly detect malicious and suspicious traffic in darknet by comparing it against legitimate traffic over networks. Some methods detect only specific malicious activity like worms, distributed denial-of-service (DDoS) attacks, or backscatter. In this research, darknet traffic analysis is proposed by focusing on traffic stability to detect and analyze network events on the Internet, as addressed in the next section.

2.2 Livenet analysis

One major line of research involves anomaly-based detection algorithms. Binkley et al. [7] presented an anomaly-based technique to explain and detect IRC botnets. While Kwon et al. [8] carried out a statistical analysis of network information to forecast network intrusion and predict the volume of a potential DDoS attack. Nychis et al. [9] proposed an entropy-based method using two types of distribution based on flow-header features e.g. address (source and destination), port (source and destination), and behavioral features for anomaly detection. The works basically target specific attacks and focus on improving detection accuracy. Generic attacks are not addressed.

The deterministic characterization of a network seems to be a major constraint in the context of a continuously evolving network usage scenario. Beukema et al. [10] proposed host clustering techniques and anomaly detection systems based on intranet traffic analysis to improve the efficiency of existing algorithms. Several papers focus on the Internet traffic flowing across the boundary of the intranet. Kumar et al. [11] examined packet flows between an ingress router, and a corresponding egress router for anomaly detection. While Zeidanloo et al. [12] proposed a generic botnet framework, monitoring traffic for similar communication patterns and behaviors among a group of hosts carrying out malicious activities. Several authors focus on intranet broadcast and multicast traffic, particularly Address Resolution Protocol (ARP) and Neighbor Discovery Protocol (NDP) packets, to obtain information on terminals connected to the intranet, man-in-the-middle attacks and DoS attacks. Whereas Nenovski et al. [13] monitored ARP packets to detect and prevent man-in-the-middle attacks. Barbhuiya et al. [14] presented a technique for detecting spoofed neighbor solicitations and advertisements in IPv6 NDP.

This thesis proposes that to detect malicious activities in the intranet it is necessary to carry out an audit of network activities. To perform a thorough audit, one must monitor all network activities. None of the works above attempt to examine and/or analyze all the network activities in the intranet. Thus far, the focus has generally been on specific attacks or information from a particular type of network traffic. It appears to be difficult, if not impractical, to monitor all network activities. In this paper, a simple traffic monitoring technique is proposed to cover device activities in the intranet. Understanding the intranet device activity is important for detecting anomalies and security incidents.

3 Darknet Traffic Analysis, by Focusing on Stability

3.1 Traffic stability

The concept of traffic stability is that the relative volume of dominant traffic components does not change drastically. The usefulness of traffic stability was initially discussed by Kanai et al. [15], who proposed a technique for traffic stability to detect several significant instabilities in darknet. The concept of traffic stability is applied in this study to understand the activities of an attackers' device on the Internet. Traffic is categorized into groups based on the values of packet header fields and the volume of the groups is calculated to evaluate stability.

Figure 1 illustrates the packets in three timeslots with the use of dotted lines. Each cube represents a packet and the numbers in the cube indicate the destination port number of the packet. Packets are grouped according to the destination port. There are two groups in the first and second timeslots, while the third slot has five. The third timeslot is considered unstable because the number of packet groups changes dramatically compared to the two previous periods.

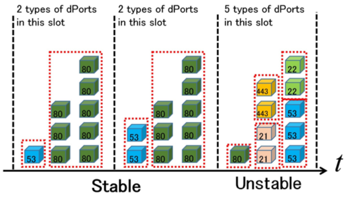
2 types of dPorts in this slot
2 types of dPorts in this slot
5 types of dPorts in this slot
53
80
53
80
80
21
443
22
53
t
Stable
Unstable

Figure 1: Event detection, by focusing traffic stability

3.2 Calculating TopN

In this study, $TopN$, is defined as the number of groups with occupancy exceeding a certain threshold (Th). Figure 2 illustrates the method used to calculate $TopN$. D_i represents the $i-th$ value of a certain attribute and $n(D_i)$ the number of elements related to D_i. The D_i values are sorted in descending order according to $n(D_i)$ (i.e., $n(D_i) \geq n(D_{i+1})$). The occupancy of D_i, $r(D_i)$, is calculated by Eq. 1 where f is the total number of attribute values,

$$r(D_i) = \frac{n(D_i)}{\sum_{x=1}^{f} n(D_x)} \times 100\% \tag{1}$$

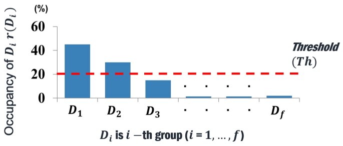

Figure 2: Event detection, by focusing on traffic stability

3.3 Evaluating stability

The proposed method evaluates the stability of traffic by observing $TopN$ groups. The $TopN$ value is denoted at timeslot t as $TopN_t$. Timeslot t is considered to be stable when the current $TopN_t$ does not show a notable difference to that of the previous consecutive T timeslots. Specifically, timeslot t is stable if $TopN_t$ satisfies Eq. 2.

$$\mu - 2\sigma \leq TopN_t \leq \mu + 2\sigma \tag{2}$$

In this formula, μ and σ denote the mean and standard deviation of $TopN$ for the last T consecutive timeslots from the timeslot t, respectively. If $TopN_t$ satisfies Eq. 2 condition, traffic in timeslot t is considered stable.

3.4 Results of darknet traffic analysis, by focusing on traffic stability

In this research, two types of datasets are analyzed: darknet traffic monitored by Japan Computer Emergency Response Team Coordination Center (JPCERT/CC) that operated an Internet threat monitoring system (TSUBAME) and darknet traffic monitored at LAB_DARKNET. LAB_DARKNET is unused IP address space in the researcher's laboratory network. The packets are organized according to the destination port in the TCP/UDP header, while $TopN$ is calculated using Th values of 1.0, 0.75, and 0.5%.

The findings show 76 unstable days in traffic monitored by JPCERT/CC TSUBAME system and six unstable days in traffic monitored in LAB_DARKNET. The 76 instances of instability in dataset 1 is due to scan activity, misconfiguration, DDoS attacks, DoS attack and backscatter. While six unstable days in dataset 2 is due to backscatter from an attack on a Domain Name System (DNS) server and a scan. It can therefore be concluded that the traffic stability is a potential candidate for understanding activities on the Internet.

3.5 Analyzing darknet TCP traffic stability at different timescales

The $TopN$ values are analyzed at two different timescales: daily and hourly. For the daily timescale analysis, the width of a timeslot is set to 24 hours, and the traffic stability in a given timeslot t is assessed on the basis of $TopN$ values in the previous seven days (i.e., $T = 7$). The term unstable-day is used to refer to a 24-hour timeslot, which is considered to be unstable in the daily-scale analysis. In the hourly-scale analysis, the width of a timeslot is set to one hour, and the traffic stability in a given timeslot t is assessed on the basis of $TopN$ values in the previous 24 hours (i.e., $T = 24$). The term unstable-hour is used to refer to a one-hour timeslot, which is considered to be unstable when detected in the hourly-scale analysis.

A dataset containing TCP packets observed for one year and nine months (2015/12/18 to 2017/08/31) is analyzed in LAB_DARKNET. These results show several anomalous events, detected by investigating $TopN$ values at different timescales. During daily-scale analysis, 17 unstable days were found, while in hourly-scale analysis, there were 140 unstable hours. The results of unstable days and hours were checked and 134 unstable hours were found to belong to stable-days and 24-hour timeslots did not show any instability. Therefore, the cause of instability was evaluated and classified and the number of senders and attack duration investigated at different timescales. It can therefore be concluded that the stability of TCP traffic, by focusing on $TopN$ values at different timescales (daily and hourly), is a useful method of analyzing events in darknet traffic.

3.6 Using normalized entropy to compare traffic differences in stable and unstable timeslots

In this study, the traffic data for stable and unstable timeslots are compared from the perspective of randomness in IP addresses and ports by evaluating the normalized entropy. The normalized entropy is used to measure uncertainty or randomness of traffic in each timeslot.

The entropy values of IP addresses and ports are evaluated by focusing on the *TopN* destination ports found in unstable timeslots to extract unusual changes induced by traffic anomalies. $H(X)$ is defined as the entropy of histogram X. Histogram X={n_i} (i = 1, 2 . . ., Z), which means that feature i occurs n_i times in the sample, and the total number of observations in the histogram is $S = \sum_{i=1}^{Z} n_i$. The sample entropy is then defined as shown in Eq. 3

$$H(X) = -\sum_{i=1}^{Z} (\frac{n_i}{S}) \log_2 (\frac{n_i}{S}). \tag{3}$$

The entropy value lies in the range [0, $\log_2 Z$]. The entropy shows its minimum value 0 when all of the features are the same and its maximum value $\log_2 Z$ when all of the items are different. The normalized entropy of X, $NE(X)$, is calculated by Eq. 4 where n_0 is the number of distinct X values in the given timeslot.

$$NE(X) = \frac{H(X)}{\log_2 n_0} \tag{4}$$

The randomness of the source IP addresses, source ports, and destination IP address is investigated to identify a specific destination port on the basis of *TopN* destination port groups in unstable timeslots.

The researcher monitored and analyzed the traffic data in LAB_DARKNET for approximately two years (2015/12/18 to 2017/12/31). Packets were categorized according to the destination ports in TCP and UDP headers, and *TopN* was calculated for Th values of 1.0, 0.75, and 0.5%. Twenty-one unstable timeslots were found and the cause of instability investigated by focusing on the *TopN* values in each timeslot compared to the *TopN* values in the previous seven timeslots.

The stable and unstable timeslots of traffic data were compared from the perspective of randomness, in relation to the IP addresses and ports by evaluating the normalized entropy. If the randomness of the IP addresses and ports in unstable timeslots changes drastically, it can be assumed that an event has occurred in the unstable timeslot. In this study, 49 destination port groups were investigated by focusing on the sequence of packets in the *TopN* values to identify unstable timeslots. The analysis of the traffic in unstable timeslots identified backscatter, misconfiguration, and hostscan activities. The experimental results showed the potential usefulness of TCP traffic stability and normalized entropy as measures for understanding the characteristics of cyber threats on the Internet.

4 Cost and Effectiveness of Simple Traffic Monitoring for Network Security

4.1 A simple traffic monitoring model

This section discusses the cost and effectiveness of traffic monitoring in a livenet. The effectiveness of traffic monitoring is evaluated on the basis of coverage, which is the range of communication the monitor can access.

Network activities can be broadly categorized into three groups. Activities where the source and destination are both in the intranet are referred to as intranet activities or *I-activities*. Activities where either the source or destination is outside the intranet are referred to as external activities or *E-activities*. An I-activity or E-activity is usually preceded by an initial activity, which is generally of a broadcast nature and reveals the presence of the initiator in the intranet. Such activity is referred to in this study as a *P-activity*.

Network device communication can be classified into the following four categories, and the members of each category are referred to as communication elements in the following discussion.

- $h2h_u$: unicast communication with another host in the intranet
- $h2h_b$: unicast communication with another host in the intranet
- $h2I_u$: unicast communication with another host on the Internet
- $h2I_b$: broadcast/multicast communication with a group of hosts on the Internet

The I-activities of a host are manifested in the $h2h_u$ and $h2h_b$ communication elements. The E-activities of a host are manifested in the $h2I_u$ and $h2I_b$ communication elements of a host. The P-activities can be monitored in the $h2h_b$ communication elements.

The goal of traffic monitoring for security is to monitor all network activities. For this purpose, it is sufficient to monitor the communication elements of all devices in an intranet. This goal is referred to as "total coverage."

Traffic monitors can be broadly classified into the following three types:

1. A host traffic monitor is deployed between a host and its network attachment point. It has access to all traffic sent from or received by the host.

2. A border traffic monitor is deployed at the border between the target intranet and an upstream network. It has access to all traffic generated by communication between hosts in the intranet and on the Internet through the border.

3. An intranet broadcast traffic monitor can be deployed anywhere in the intranet. It has access to all broadcast packets sent from hosts in the intranet.

Figure 3 illustrates the relationship between monitor types, network activity types, and accessed communication elements. As indicated in Figure 3, a host traffic monitor covers three network activity types (I-activity, E-activity, and P-activity of a host) with the ability to access all four communication element types. A border traffic monitor covers the E-activity, which has access to two communication element types. An intranet broadcast traffic monitor covers the P-activity, which has access to a single communication element type.

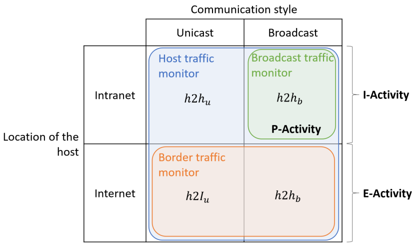

Figure 3: Communication in the intranet and monitor types

The simple traffic monitoring scheme proposed in this study effectively combines the three monitor types and provides reasonably good coverage at an acceptable cost.

Figure 4 depicts the proposed simple traffic monitoring concept and its coverage. The blue rectangle, orange cuboid, and green cuboid in the figure indicate the coverage of a host monitor, a border monitor, and an intranet broadcast monitor, respectively.

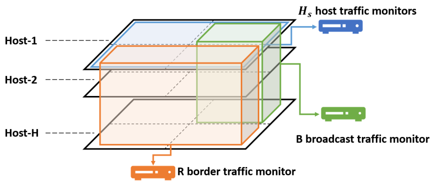

Figure 4: Simple traffic monitoring concept and coverage of intranet communication

The proposal in this study consists of host traffic monitors, deployed only for certain hosts that act as activity hubs, such as servers, in the intranet. Border traffic monitors are deployed at each of the R routers in the intranet, and these cover $h2I_u$ and $h2I_b$ communication elements for every host in the intranet, as indicated by the orange cuboid in Figure 4. Intranet broadcast monitors are deployed in all B broadcast segments in the intranet, and these cover $h2h_b$ communication elements of every host in the intranet, as indicated by the green cuboid in Figure 4.

As illustrated in Figure 4, the proposal in this study accepts that not all $h2h_u$ communication elements are covered. In this case, $h2h_u$ communication elements in hosts other than activity hubs are not covered. These hosts are usually client PCs. At least, the P-activity of such communication is covered by the broadcast traffic monitor and the administrator can determine the presence of such communications.

The effectiveness of the proposed scheme is demonstrated by evaluating the coverage and number of monitors required. For simplicity, the following discussion assumes the simplest case, with an intranet consisting of a single broadcast segment, connected to the Internet via a single router. There are H hosts in the intranet. For total coverage C_t, it is sufficient to monitor all $4H$ communication elements using H host monitors. In this case, the required number of monitors M_t is H. In the proposal in this study, the coverage C_p is achieved using M_p monitors, where

$$C_p = 4H_s + 3H_c \quad (5)$$
$$M_p = H_s + R + B, \quad (6)$$

in which H_s is the number of activity hubs, H_c is the number of hosts not counted as activity hubs, R is the number of routers, and B is the number of broadcast segments in the intranet. In the simple case of an intranet with a single broadcast segment connected to the Internet via a single router, $R = B = 1$. The H_s host traffic monitors cover all four communication elements of every server in the intranet. The border traffic monitors cover $h2I_u$ and $h2I_b$ elements of the H_c client PCs, while the intranet broadcast monitors cover the $h2h_b$ elements of the H_c client PCs. In terms of the number of communication elements, the coverage C_p is H_c less than the total coverage C_t. However, the proposed scheme uses a significantly smaller number of monitors, because $H_s + R + B$ is generally much less than H in an average intranet.

4.2 Traffic analysis experiment

4.2.1 Experimental environment

The traffic monitoring and analysis experiment was carried out on the university's laboratory network, with 50 devices in one broadcast segment, and connected to the Internet via a router. In this network, 15 devices with statically assigned IP addresses provided network services to other devices in the network. These devices qualified to be categorized as activity hubs. The remaining 35 devices were classified as clients. For simplicity and a leaner experimental setup, only the four most active servers were categorized as activity hubs with host traffic monitors being deployed for these. Among the four activity hubs, one was a Windows host, two were Linux hosts, and one was a printer. The host monitors were set with different OSs to demonstrate the feasibility of the host traffic monitor for the selected OSs. A border traffic monitor and broadcast traffic monitor were deployed at the router and inside the broadcast segment, respectively.

Thus, in this monitoring environment, H_s, H_c, R, and B B were 4, 46, 1, and 1, respectively. Based on Eq. 5, 154 out of 200 communication elements were monitored using six monitors.

4.3 Peer analysis

Peer analysis is performed to ascertain and validate the purpose of the network activity (that is, access) covering both I-activity and E-activity based on the peer address (IP address and/or domain name) and port number. When an activity satisfies both of the following conditions, the purpose of the activity is considered valid.

1. The peer host is not blacklisted
2. The endpoint (peer host and port number) can be validated.

For simplicity, this study focuses on the activities from host traffic that are not directed by a user, which is referred to as non-user-initiated (NUI) activities. To this end, the researcher extracted data from the logs of the activity hubs for the period during which no user was logged in. This does not imply that an activity is initiated when a user is logged in, or that an activity initiated by a user is legitimate. It only implies that the NUI activity is the first location to search for suspicious activities.

The corresponding domain name for each packet was checked by observing the packet payload or by examining the DNS transaction relating to the IP address. When a domain name corresponding to an IP address was not found, validation was carried out based on the IP address.

The URLVoid [16] was used to determine whether or not the peer was blacklisted. When a peer was not blacklisted, the endpoint was validated according to the following criteria.

- The endpoint is listed in the official documents released by the developers of the OS and applications running on the source host.
- The IP address and domain name of the endpoint is owned by the vendors of the OS and applications running on the source host.
- The endpoint is in the intranet and valid.

Through this experiment, it was observed that a surprisingly large volume of NUI activities were carried out by OSs and applications on a device. Such activities were very likely carried out under the software agreements “agreed to” when the software was installed. It is important for the researcher to gain awareness and understanding of NUI activities in order to efficiently detect malicious activities hidden in such activities.

4.4 Volume analysis

During the experiments, the researcher observed variations in the traffic volume patterns in the form of packet count, high-resolution monitoring, encrypted packets, and the amount of traffic in each category. The limitation of this section is that the discussion only concerns high-resolution traffic monitoring that can be exploited using the researcher’s simple traffic monitoring scheme.

4.4.1 High-resolution traffic monitoring

High-resolution traffic monitoring [17] is a technique for monitoring the traffic at millisecond intervals. This technique was tested on border traffic by counting the number of packets in 10 ms intervals and looking for significant variations.

In this experiment, the total number of 10 ms timeslots was 276,480,000. The average and median number of packets per timeslot were 0.58 and 0.0, respectively. Only 0.001% of the timeslots (2,729 timeslots) had 340 or more packets. Such a slot is referred to as a “peak slot,” and the highest number of packets in a peak slot is 1,152.

Figure 5 compares the number of packets every minute and every 10 ms over a period of 10 minutes. The variation in the number of packets per minute was almost stable. However, the number of packets per 10 ms exhibited frequent and significant changes, and 143 peak slots occurred during this period. These hidden peaks can provide important insight for network administrators, and indicate the necessity for traffic to be investigated in greater detail. In this case, it can be confirmed that most of the packets in the peak slots belonged to secure shell sessions between hosts in the intranet and hosts on the Internet. A secure shell session generates numerous peaks when a user uploads or downloads many files via the SSH tunnel.

4.5 Broadcast analysis

By monitoring the ARP/NDP broadcast packets, network administrators can detect newly connected ter-

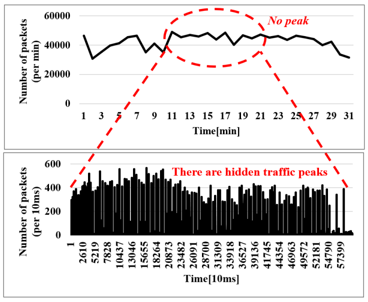

Figure 5: High-resolution monitoring

minals and obtain pairs of IP and media access control (MAC) addresses. The results may also be used to generate the usage logs of terminals. An ARP or NDP packet indicates that the packet source is present in the network and attempting to communicate with another device. Hence, the number of broadcast packets from a device can be used to estimate the terminal activity in the intranet.

Figure 6 presents a breakdown of the ARP requests in the laboratory network. The ARP requests are categorized as those sent from one server to another, from a server to a client, from a client to a server, and from one client to another. These categories are indicated by the blue, orange, green, and red areas, respectively. As illustrated in Figure 6, as expected, ARP requests sent to a server (blue and green areas) were dominant. It should be noted that ARP requests from one client to another were rarely observed, and constituted only 0.41% of the entire ARP requests monitored for one month. These analysis results support the researcher's assumption that little communication takes place among client PCs, as described in Section 4.1.

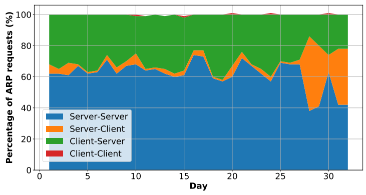

Figure 6: Breakdown of ARP request traffic

4.6 Discussion

This research also discusses the effectiveness of the simple monitoring technique in detecting a network attack in the context of a cyber kill chain model® [18]. According to the results of the analysis techniques presented in this work, the possibility of observing the traces of an attack in progress in each phase was examined. It can be concluded that the simple traffic monitoring model and traffic analysis techniques are useful for validating observed suspicious and/or malicious activities in the intranet within the context of the cyber kill chain.

5 Summary

In this thesis, traffic monitoring and analysis issues in darknet and livenet are discussed. Monitoring and analysis of darknet and livenet traffic assist in understanding the activities of attackers' devices and the activities of intranet devices, respectively.

Darknet traffic analysis is proposed in this study, by focusing on the stability of traffic. In the traffic monitored at JPCERT/CC TSUBAME, 76 unstable days were found due to scan activity, misconfiguration, DDoS attacks, DoS attack, and backscatter. While the six days of instability in LAB_DARKNET were due to backscatter from an attack on a DNS server and a scan. It can be concluded that the stability of traffic is a potential candidate for understanding activities on the Internet. Traffic analysis techniques are proposed in this study to detect small-scale events, where a small number of attackers send packets to the destination targets in a short duration. Several case studies on detected unstable timeslots indicate the presence of scanning activities in the traffic. The analysis results of traffic found in unstable timeslots are then reported in this study. The traffic data for stable and unstable timeslots is compared from the perspective of randomness in IP addresses and ports by evaluating the normalized entropy. The analysis of the traffic in unstable timeslots identified backscatter, misconfiguration, and hostscan activities. The experiment results demonstrate the potential usefulness of TCP traffic stability and normalized entropy as measures for understanding the characteristics of cyber threats on the Internet.

A simple traffic monitoring model proposed in this research is an abstraction of network activities in the intranet. It is complete in the sense that it covers the activities in terms of four communication elements ($h2h_u$, $h2I_u$, $h2h_b$, and $h2I_b$). It provides a mapping between the monitors and activities. It serves as a basis for evaluation of the coverage of monitoring and the cost of monitoring in terms of the number of monitors and the complexity of monitors. This study demonstrates the effectiveness of peer analysis techniques for intranet host traffic, traffic volume analysis techniques for border traffic, and broadcast analysis techniques for intranet broadcast traffic. Peer analysis is useful for validating network activity in the intranet. Volume analysis indicates E-activities with the use of high-resolution traffic monitoring, encrypted traffic monitoring, and advanced techniques such as category transformation. Broadcast analysis techniques can be effectively used to monitor P-activities e.g., network attachment. The simple traffic monitoring model and traffic analysis techniques are useful for validating suspicious and/or downright malicious activities in the intranet in the context of the cyber kill chain.

It can be concluded that the analysis of traffic in darknet and livenet can be used to retrieve useful information for network administrators and to understand the activities of attackers' devices on the Internet and the activities of intranet devices

References

[1] A. Shimoda, T. Mori, and S. Goto, "Extended darknet: Multidimensional internet threat monitoring system," *IEICE Transactions on Communications*, vol. E95-B, no. 6, pp. 1915–1923, 2012.

[2] Y. Shomura, K. Yoshida, A. Sato, S. Matsumoto, and K. Itano, "A traffic analysis using cardinalities and header information," *Proceedings - 2010 1st International Conference on Networking and Computing, ICNC 2010*, pp. 55–62, 2010.

[3] T. BAN, "3-3 data mining applied to Darknet Traffic Analysis," *Journal of the National Institute of Information and Communications Technology*, vol. 63, no. 2, pp. 45–54, 2016.

[4] J. Song, Y. Lee, J. W. Choi, J. M. Gil, J. Han, and S. S. Choi, "Practical In-Depth Analysis of IDS Alerts for Tracing and Identifying Potential Attackers on Darknet," *Sustainability (Switzerland)*, vol. 9, no. 2, pp. 1–18, 2017.

[5] Q. Wang, Z. Chen, and C. Chen, "Darknet-Based Inference of Internet Worm Temporal Characteristics," *IEEE Transactions on Information Forensics and Security*, vol. 6, no. 4, pp. 1382–1393, 2011.

[6] Y. Feng, Y. Hori, K. Sakurai, and J. Takeuchi, "A Behavior-based Method for Detecting Distributed Scan Attacks in Darknets," *Journal of Information Processing*, vol. 21, no. 3, pp. 527–538, 2013. [Online]. Available: http://www.scopus.com/inward/record.url?eid=2-s2.0-84880150397partnerID=tZOtx3y1

[7] J. R. Binkley and S. Singh, "An algorithm for anomaly-based botnet detection," *SRUTI'06: Proceedings of the 2nd conference on Steps to Reducing Unwanted Traffic on the Internet*, 2006.

[8] D. Kwon, H. Kim, D. An, and H. Ju, "DDoS Attack Volume Forecasting Using a Statistical Approach," *Integrated Network and Service Management (IM), 2017 IFIP/IEEE Symposium on*, pp. 2015–2018, 2017.

[9] G. Nychis, V. Sekar, D. G. Andersen, H. Kim, and H. Zhang, "An empirical evaluation of entropy-based traffic anomaly detection," *Proceedings of the 8th ACM SIGCOMM conference on Internet measurement conference - IMC '08*, p. 151, 2008.

[10] W. J. B. Beukema, T. Attema, and H. A. Schotanus, "Internal Network Monitoring and Anomaly Detection through Host Clustering," *Proceedings of the 3rd International Conference on Information Systems Security and Privacy (ICISSP 2017)*, pp. 694–703, 2017.

[11] H. Kumar, S. Kumar, S. Singh, A. Kumar, R. Joseph, and P. Kumar, "Network Traffic Monitoring , Analysis and Anomaly Detection," vol. 2, no. 2, pp. 489–496, 2013.

[12] H. R. Zeidanloo, A. B. Manaf, P. Vahdani, F. Tabatabaei, and M. Zamani, "Botnet detection based on traffic monitoring," *International Conference on Networking and Information Technology*, pp. 97–101, 2010.

[13] B. Nenovski and P. Mitrevski, "Real-World ARP Attacks and Packet Sniffing , Detection and Prevention on Windows and Android Devices," in *12th International Conference on Informatics and Information Technologies (CiiT 2015)*, no. Ciit, 2015, pp. 186–191.

[14] F. A. Barbhuiya, G. Bansal, N. Kumar, S. Biswas, and S. Nandi, "Detection of neighbor discovery protocol based attacks in IPv6 network," *Networking Science*, vol. 2, no. 3-4, pp. 91–113, 2013.

[15] T. Kanai, H. Tsunoda, and G. Mansfield Keeni, "Darknet Traffic Analysis by Focusing on Variations in Dominant Traffic," Tech. Rep. 3, 10 2015.

[16] "Check if a Website is Malicious/Scam or Safe/Legit — URLVoid," https://www.urlvoid.com/. [Online]. Available: https://www.urlvoid.com/

[17] G. Mansfield, S. Karakala, and T. Saitoh, "High Resolution Traffic Measurement," *Workshop on Passive and Active Measurements on the Internet (PAM2001)*, 2001.

[18] "GAININGTHEADVANTAGE Applying Cyber Kill Chain® Methodology to Network Defense," Tech. Rep. [Online]. Available: https://docplayer.net/23006368-Gaining-the-advantage-applying-cyber-kill-chain-methodology-to-network-defense.html