Procedia Computer Science

Volume 53, 2015, Pages 175–182

2015 INNS Conference on Big Data

Large-Scale Monitoring for Cyber Attacks by Using Cluster Information on Darknet Traffic Features

Hironori Nishikaze[1], Seiichi Ozawa[1], Jun Kitazono[1], Tao Ban[2], Junji Nakazato[2], and Jumpei Shimamura[3]

[1] Kobe University, Kobe, Japan
[2] National Institute of Information and Communications Technology, Koganei, Tokyo, Japan
[3] clwit, Inc., Mitaka, Tokyo, Japan

Abstract

This paper presents a machine learning approach to large-scale monitoring for malicious activities on Internet. In the proposed system, network packets sent from a subnet to a darknet (i.e., a set of unused IPs) are collected, and they are transformed into 27-dimensional TAP (Traffic Analysis Profile) feature vectors. Then, a hierarchical clustering is performed to obtain clusters for typical malicious behaviors. In the monitoring phase, the malicious activities in a subnet are estimated from the closest TAP feature cluster. Then, such TAP feature clusters for all subnets are visualized on the proposed monitoring system in real time. In the experiment, we use a big data set of 303,733,994 darknet packs collected from February 1st to February 28th, 2014 (28 days) for monitoring. As a result, we can successfully detect an indication of the pandemic of a new malware, which attacked to the vulnerability of Synology NAS (port 5,000/TCP).

Keywords: cybersecurity, visualization, darknet, packet traffic data, clustering

1 Introduction

Recently, cyber-attacks are one of the most serious problems in our life. There exist various kinds of cyber threads such as malware infection, DDoS attacks, probing to find security vulnerability, drive-by download attacks, phishing, spam mails to lure malicious web site, and targeted e-mail attack, which intend to steal money/important information and to stop/disturb public services, etc. To protect users form such cyber threads, it is important to grasp malicious activities not only within a local network domain but also on the Internet as a whole.

One way to observe large-scale events taking place on the Internet is to use a network telescope called *darknet*. The darknet is defined as an unused address-space of a computer network, which should not have any normal communications from other computers. Therefore, almost all traffic to the darknet is suspicious, and we can grasp the information either on cyber-attacks or insignificant communications due to misconfigurations. Observable cyber-attacks are mainly activities of random scanning worms and DDoS backscatter. Therefore, only a part of

Selection and peer-review under responsibility of the Scientific Programme Committee of INNS-BigData2015

doi:10.1016/j.procs.2015.07.292

cyber-attacks can be monitored by observing the darknet traffic. However, since the darknet can receive packets from the whole Internet space, it allows us to monitor the large-scale malicious activities on the Internet.

In this paper, we develop a darknet monitoring system that can grasp large-scale malicious activities by analyzing darknet traffic features. To actualize the real-time monitoring, the source IP space is divided into subnets, and the network traffic in each subnet is transformed into a feature vector based on the TAP (Traffic Analysis Profile) analysis that classifies short-time network traffic into 27 categories based on the numbers of packets, destination IPs/ports, and source IP/ports. In addition to TAP features, we use malware signatures to classify darknet packets. In the proposed monitoring system, the TAP type information of all active hosts in a subnet is accumulated and defined as a feature vector. Then, a hierarchical clustering algorithm based on the furthest neighbor method is applied to such feature vectors in order to find some useful information on malicious activities.

The rest of this paper is organized as follows. In Section 2, we first give a brief survey on machine learning approaches to grasping malicious activities using darknet. Section 3 presents a subnet-based monitoring system to capture the dynamical changes in malicious activities in real-time. In Section 4, we conduct a monitoring experiment using one-month darknet packet data that are sent from a subnet including 65,536 hosts to /16 darknet sensors including 65,536 IP addresses. Finally, we address the conclusions of this work and future work in Section 5.

2 Related Work

There have been many approaches to analyzing darknet traffic using machine learning methods.

Yamauchi et al. [10] proposed a botnet detection method that can extract the pure botnet transmission from darknet traffic data by nonnegative matrix factorization (NMF) [8, 9]. In this method, NMF is used to decompose signals related to the botnet transmission from darknet traffic, which are often buried in other frequent malware transmission. Ko et al, [7] proposed a classification method to identify malicious darknet packets that are caused by real cyber-attacks. In the proposed method, the three main modules dealing with ICMP, Torrent (a peer-to-peer file transfer protocol), and TCP protocols filter out unsuspicious darknet packets based on their communication characteristics. In the above approaches, specific information on cyber-attacks is intended to extract from the darknet traffic.

For a monitoring purpose, Ban et al. [1] presented a monitoring system that characterizes the behavior of long term cyber-attacks by mining the darknet traffic data collected by the nicter project [5]. In this system, machine learning techniques such as clustering, classification, function regression are applied to the analysis of darknet traffic. Bou-Harb et al. [2] proposed a multidimensional monitoring method for source port 0 probing attacks by analyzing darknet traffic. This method is aiming for extracting and fingerprinting malicious darknet traffic from received packets. By performing unsupervised machine learning techniques on the extracted traffic, the activities by similar types of hosts are grouped by employing a set of statistical-based behavioral analytics. However, this approach is targeted only for source port 0 probing attacks.

3 Subnet-based Darknet Monitoring System

3.1 Traffic Analysis Profile (TAP) and Malware Signature

As mentioned in Section 1, packets delivered to the darknet are basically generated as the results of malicious acts such as network scan and DDoS attack. Therefore, to infer the types

		#dest. ports = 1		
		#dest. addr. = 1	#dest. addr. > 1	
			sequential dest. addr.	random dest. addr.
#src ports = 1	#packets = #src ports	Pong(1.1)	---	---
	#packets > #src ports	HammerN(2.N)	SequencedScan1(4.1)	RandomScan1(5.2)
#src ports > 1	#packets = #src ports	MultiPong(3.1)	SequencedScan2(4.2)	RandomScan2(5.4)
	#packets > #src ports	MultiHammer(3.2)	SequencedScan3(4.3)	RandomScan3(5.6)

		#dest. ports > 1		
		#dest. addr. = 1	#dest. addr. > 1	
			sequential dest. addr.	random dest. addr.
#src ports = 1	#packets = #src ports	---	---	---
	#packets > #src ports	FunSpan(6.1)	SequencedFunSpan(7.1)	RandomFunSpan(8.2)
#src ports > 1	#packets = #src ports	LinearSpan(6.2)	SequencedLinearSpan(7.3)	RandomLinearSpan(8.4)
	#packets > #src ports	MultiSpan(6.3)	SequencedMultiSpan(7.5)	RandomMultiSpan(8.6)

Figure 1: Traffic Analysis Profile (TAP) to classify darknet traffic features.

of malicious activities, temporal traffic features are analyzed for darknet packets. For known malwares, their activities are featured by some typical temporal patterns of packet traffic to specific ports called *malware signatures*. Therefore, if the packet traffic from a host is matched with a certain malware signature, the infection type is easily identified. However, there are many suspicious traffic patterns which are not exactly matched with any known malware signatures.

To classify such suspicious malicious activities, Suzuki et al. [4, 6] have proposed Traffic Analysis Profile (TAP) analysis (see Table 1) for the darknet analysis. As seen in Table 1, there are 19 TAP types to categorize traffic features of darknet packets. A TAP type is determined based on the number of packets, the number of destination/source ports, the number of destination IP addresses, and scan types (sequential or random) for every 30-second darknet traffic. For example, if darknet packets are sent from a source host with a single port to a specific port on multiple random destination IPs, the TAP type is classified as 'RandomFunSpan(5.2)' which is possibly a network scan to find hosts with security vulnerability (see Fig. 1). For 'Hammer$N$(2.N)', multiple types of attacks can be considered depending on the number of packets N. Hammer is a type of malicious behaviors which intermittently sends multiple packets to a specific port of a specific host. Our preliminary observation tells that 'Hammer$N$(2.N)' with more than $N = 10$ does not happen frequently. Therefore, we consider the 9 types of 'Hammer$N$(2.N)'($N = 2 \cdots 10$) as TAP features. Here, we assume that 'Hammer10(2.10)' cover all the attacks with more than $N = 10$.

From the above discussion, we consider 27 TAP types in total as darknet traffic features to discriminate cyber-attacks.

3.2 Subnet-based Darknet Traffic Features

The above-mentioned TAP analysis is applied to individual source hosts to feature the type of cyber-attacks. However, monitoring all of the individual hosts in the internet is hardly carried out because a huge number of hosts could give security alerts simultaneously. Therefore, we propose a large-scale monitoring system that observe the infection states of subnets: a group of hosts with a certain range of IP space such as ISP, company, institution, and so on. Obviously,

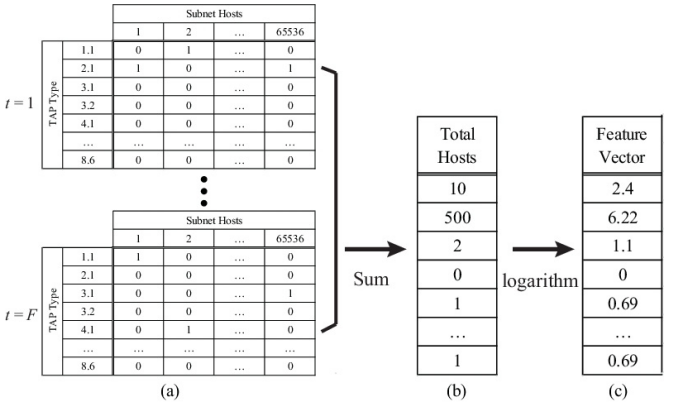

Figure 2: An example of feature vector transformation: (a) TAP vectors for subnets from $t = 1$ to F, (b) summation of TAP vectors, and (c) subnet feature vector representing infection status.

such a subnet-based monitoring system cannot detect security incidents of individual hosts. However, it can be considered as a first warning system to grasp the collective trends of malicious activities within a subnet.

To represent the attack type of a source host, let us define a 27-dimentional TAP vector $\boldsymbol{v}_i^{(t)} = \{v_{i1}^{(t)}, \cdots, v_{i27}^{(t)}\}$ for the ith host at the tth time flame (30 sec.), where

$$v_{ij}^{(t)} = \begin{cases} 1 & \text{if an attack by the } i\text{th host is categorized into the } j\text{th TAP type,} \\ 0 & \text{otherwise.} \end{cases} \tag{1}$$

Since we construct a subnet-based monitoring system, a feature vector should include the collective information on TAP types of all active hosts within a subnet. Then, let us define the following TAP feature vector $\boldsymbol{V}_k$ by accumulating TAP vectors of all hosts in a subnet for F time flames:

$$\boldsymbol{V}_k^F = \ln\left(\sum_{t\in F}\sum_{i\in H_k} \boldsymbol{v}_i^{(t)} + 1\right) \tag{2}$$

where H_k is the set of hosts in the kth subnet. The reason to take the logarithm in Eq. (2) is that the difference in a small number of hosts is important compared with that in a large number. Let us take an example in Fig. 2 to explain how a TAP feature vector $\boldsymbol{V}_k^F$ is calculated.

An example case in Fig. 2(a) illustrates a table of TAP types of 65,536 hosts (i.e., a /16 subnet is assumed). As seen in Fig. 2(a), each column corresponds to a TAP vector of a host at the time flame t ($t = 1, cdots, F$). For example, the TAP type of host #2 is 'Pong(1.1)' and that of host #65,535 is 'Hammer1(2.1)' at the time flame $t = 1$, Note that if a host is non-active (i.e., only a few or no packet is transmitted), no TAP type is provided; therefore, a TAP vector for a non-active host is a zero vector. Figure 2(b) shows a TAP vector obtained by summing all TAP vectors in Fig. 2(a). Then, a TAP feature vector in Eq. (2) is obtained as shown in Fig. 2(c).

As mentioned in 3.1, a known malware can be identified by their signatures. For a host infected by known malwares, they should not be considered into a TAP feature vector because the proposed monitoring system is developed to detect unknown cyber-attacks. Therefore, the signature matching is first performed to every hosts in a subnet, and the TAP type is provided

only for the hosts whose packet traffic pattern is not matched with any malware signatures. Then, a subnet feature vector is created by accumulating TAP vectors for such unmatched hosts. If the majority of hosts in a subnet have known malware signatures, such a subnet should be immediately identified as an infection state by the malware without transforming into a TAP feature vector.

3.3 Monitoring Subnet Infection Status Using TAP Features

To monitor malicious activities on subnets, we adopt a vector quantization approach, in which a subnet state is associated with the closest cluster representing a typical type of malicious activities. For this purpose, we perform a clustering algorithm for collected darknet packets $\boldsymbol{X} = \{\boldsymbol{x}_i\}_{i=1}^{N}$ (i.e., training data set). In this work, we adopt a hierarchical clustering in which two clusters with the closest distance are merged until the cluster distance reaches a threshold θ. Here, we adopt the following farthest neighbor method as a distance measure $D(C_i, C_j)$ for two clusters C_i, C_j:

$$D(C_i, C_j) = \max_{k,l} \|\boldsymbol{x}_k - \boldsymbol{x}_l\| \quad \text{where } \boldsymbol{x}_k \in C_i \text{ and } \boldsymbol{x}_l \in C_j. \tag{3}$$

This clustering algorithm is also known as the complete-linkage clustering [3]. This clustering avoids a drawback of the alternative single linkage method - the so-called chaining phenomenon, where clusters formed via single linkage clustering may be forced together due to single elements being close to each other, even though many elements in each cluster may be very distant to each other. Complete linkage tends to find compact clusters of approximately equal diameters. The outputs of the clustering algorithm are cluster data sets $\{C_i = \{\hat{\boldsymbol{x}}_{ij}\}_{j=1}^{N_i}\}_{i=1}^{M}$ where $\hat{\boldsymbol{x}}_{ij}$ is the jth data of the ith cluster C_i and N_i is the number of data in C_i.

The monitoring phase is summarized as follows. As explained in 3.2, the transformation of darknet packets into a feature vector is conducted for every subnets at every time flames. At the time flame F, a TAP feature vector for the kth subnet $\boldsymbol{V}_k^F$ is calculated by Eq. (2). Then, the closest cluster C_{i^*} is obtained based on the farthest neighbor method where the distance is given by

$$d(\boldsymbol{V}_k^F, C_i) = \max_{j} \|\boldsymbol{V}_k^F - \widehat{\boldsymbol{V}}_{ij}\|. \quad \text{where } \widehat{\boldsymbol{V}}_{ij} \in C_i. \tag{4}$$

Finally, the cluster information C_{i^*} is displayed on the monitoring system.

4 Experiments

4.1 Experimental Setup

In the experiment, we use the data set of darknet packs collected from February 1st to February 28th, 2014 (28 days) by the National Institute of Information and Communications Technology (NICT), Tokyo, Japan. The used darknet sensor covers /16 IP space (i.e., the number of sensing destination IP addresses is 65,536), and the total number of collected darknet packets is 303,733,994. For the sake of convenience, the subnet mask is assumed to be /16 in the experiments; that is, 65,536 subnets are monitored simultaneously on the proposed system.

In the above data set, the number of effective non-zero TAP feature vectors is 503,148 out of 1,835,008 (65,536 subnets $\times$ 28 days). To reduce the computational costs, we randomly select 20,000 TAP feature vectors for training, and the hierarchical clustering is performed to obtain typical malicious activities. In this experiment, the threshold parameter in the clustering is empirically set as $\theta = 5$ so that we can recognize the transitions of malicious activities easily.

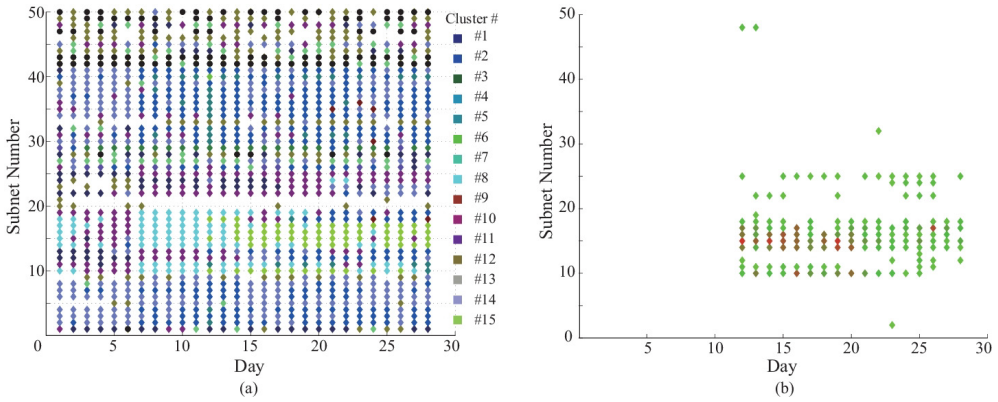

Figure 3: Monitoring results. (a) Transitions of TAP feature clusters for adjacent 50 subnets. (b)Transitions of attacks to port 5,000 for adjacent 50 subnets. The color of a dot shows the intensity of attacks (i.e., the number of darknet packets within a day). The red and green dots show high and low intensity, respectively.

4.2 Monitoring Results

Before deploying the developed monitoring system in practical environments, we conduct a preliminary monitoring test by using the above-mentioned data set collected from February 1st to February 28th, 2014 (28 days). Since only 20,000 TAP features are used to create 15 cluster prototypes, the other 483,148 TAP feature vectors are used for a testing purpose.

Due to the space limitation, let us show only a part of monitoring results for 65,536 subnets. Figure 3(a) shows a monitoring result for adjacent 50 subnets, in which a colored dot represents the closest TAP feature cluster showing a typical type of malicious activities, and each row corresponds to a temporal transition of closest TAP feature clusters for a subnet. As seen in Fig. 3(a), the subnets #10 and #14-#17 have a similar transition pattern: clusters #10 $\rightarrow$ #8 $\rightarrow$ #15. This transition pattern is commonly detected among some adjacent subnets. It may show the process that an unknown malware proliferates and dominates over these adjacent subnets from the beginning to the middle of February, 2014.

To see what happens in the subnets, we analyzed darknet packets within the subnets #10 and #14-#17. Then, it was found that a lot of darknet packets were sent to the destination port 5,000. Figure 3(b) illustrates the number of packets to the port 5,000 (i.e., intensity of port attacks) with colored dots. The red and green dots mean high and low intensity, respectively. As clearly seen in Fig. 3(b), the attack to the port 5,000 starts from Feb. 12th in the subnets #10 and #14-#17. Comparing the results in Figs. 3(a) and (b), the TAP feature cluster #8 seems to be related to the start of the attack to the port 5,000. Actually, it is known that this port attack was aiming for attacking the vulnerability of Synology NAS (port 5,000/TCP), and the surge of this attack was observed around Feb. 28th, 2014.

Since the cluster #8 starts on Feb. 7th, it is considered to be an indication of the attack to the port 5,000, and it was observed 20 days before the pandemic of the attack. This is an encouraging result because an obvious difference in TAP features were observed on the proposed monitoring system (see Fig. 3(a)). Therefore, we could be aware of the emergence of the unknown malware around Feb. 7th, which was quite earlier than the pandemic. Figure 4 shows the distributions of TAP features in the clusters #10, #8, and #15. From the transitions of

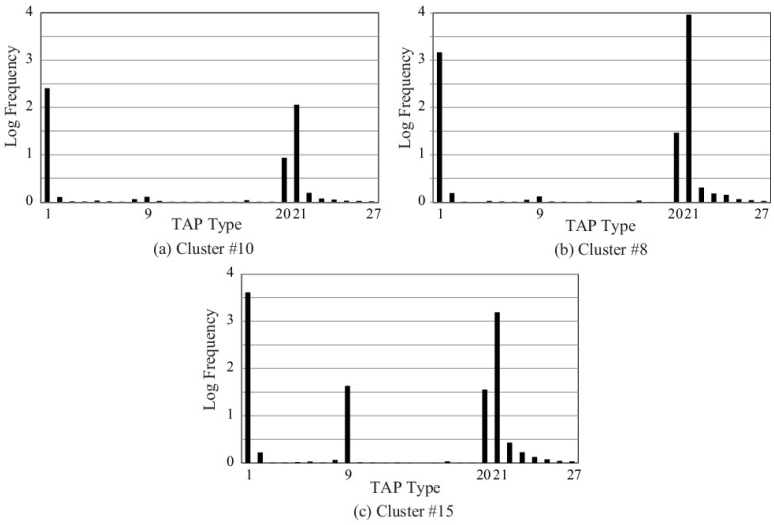

Figure 4: Transitions of TAP feature distributions: clusters #10 → #8 → #15.

the TAP feature distributions, we can recognize that the intensity of TAP type #21 was clearly increased along with the increase of the attack to the port 5,000. Then, the TAP type #9 was intensified during the sustained weak attack to the port 5,000, which led to the pandemic on Feb. 28th. Although we haven't completed the detailed analysis on such a transition of TAP features, we expect that it could be used as an signature of the malicious activity.

5 Conclusions

In this paper, we propose a large-scale monitoring system for cyber attacks using darknet traffic information. In the system, /16 darknet sensors (unused 65,536 destination IPs) are deployed to collect packets which are sent from a /16 subnet (65,536 source hosts). The darknet packets collected at every subnet in one day are transformed into a TAP feature vector representing the activities of unknown malwares. Then, a hierarchical clustering algorithm is applied to TAP feature vectors to obtain typical types of malicious activities (TAP feature clusters). In the monitoring phase, darknet packets in a subnet are transformed into a TAP feature vector, and the closest TAP feature cluster is obtained to estimate the malicious activities in the subnet.

In the experiment, we use a big data set of 303,733,994 darknet packs collected from February 1st to February 28th, 2014 (28 days) by the National Institute of Information and Communications Technology (NICT). The number of transformed TAP feature vectors is 503,148; due to our limited computational resources, randomly selected 20,000 TAP feature vectors are used for clustering and all the feature vectors are used for monitoring. As a result, we found an interesting transition of TAP feature clusters for some adjacent subnets, which have distinctive transitions from the other subnets. The detailed packet analysis implies that this distinctive transition seems to be related to the pandemic of the attack to the vulnerability of Synology NAS (port 5,000/TCP). Since an indication of this attack was detected on the proposed monitoring system about 20 days earlier than the pandemic, we can conclude that the proposed system is promising as a large-scale monitoring system for cyber-attacks.

There still remain open problems in the proposed system. In the system, malicious activities in a subnet are estimated from the closest TAP feature cluster, which are obtained by performing a hierarchical clustering for preliminary collected darknet packets. Therefore, if new malwares emerge, there is no guarantee if the obtained TAP feature clusters can represent the activities of the new malwares. To solve this, we should adopt an online clustering for a large amount TAP feature vectors that are generated continuously, and such a system should be implemented on a distributed storage and distributed processing platform such as Hadoop. In addition, it is unclear how we can select a suitable threshold for the hierarchical clustering. To detect a new trend on malicious activities accurately, it is important to have a good visualization property in the monitoring system. However, if a threshold is too low, redundant clusters are created and it makes us difficult to find a distinctive transition of TAP feature clusters. On the other hand, if a threshold is too large, only a few clusters are created and we may miss an important transition due to the coarse representation of states in a TAP feature space. Furthermore, we need to prove that the proposed system can capture indications of other pandemic of unknown malwares. The above open problems are left as our future work.

Acknowledgments

This work is partially supported by the Ministry of Education, Science, Sports and Culture, Grant-in-Aid for Scientific Research (C) 24500173.

References

[1] T. Ban, L. Zhu, J. Shimamura, S. Pang, D. Inoue, and K. Nakao. Behavior analysis of long-term cyber attacks in the darknet. In et al. Z. Zeng, editor, *Neural Information Processing*, LNCS 7667, pages 620–628, 2012.

[2] E. Bou-Harb, N.-E. Lakhdari, H. Binsalleeh, and M. Debbabi. Multidimensional investigation of source port 0 probing. *Digital Investigation*, 11:S114–S123, 2014.

[3] B. S. Everitt, S. Landau, and M. Leese. *Cluster Analysis (5th ed.)*. Wiley, 2010.

[4] D. Inoue, K. Yoshioka, M. Eto, M. Yamagata, E. Nishino, J. Takeuchi, K. Ohkouchi, and K. Nakao. nicter: An incident analysis system toward binding network monitoring with malware analysis. In *Proc. of WOMBAT Workshop on Information Security Threats Data Collection and Sharing 2008*, pages 58–66, 2008.

[5] D. Inoue, K. Yoshioka, M. Eto, M. Yamagata, E. Nishino, J. Takeuchi, K. Ohkouchi, and K. Nakao. An incident analysis system nicter and its analysis engines based on data mining techniques. In et al. M. Köppen, editor, *Advances in Neuro-Information Processing*, pages 579–586, 2009.

[6] H. Wada K. Nanano H. Takakura K. Suzuki, S. Baba and Y. Okabe. Development and evaluation of traffic behavior analysis system for understanding network condition. *IEICE Trans. on Communications*, (7):916–927, 2010.

[7] S. Ko, K. Kim, Y. Lee, and J. Song. A classification method of darknet traffic for advanced security monitoring and response. In et al. C. K. Loo, editor, *Neural Information Processing*, LNCS 8836, pages 357–364, 2014.

[8] D. D. Lee and H. S. Seung. Learning the parts of objects by non-negative matrix factorization. *Nature*, 401:788–791, 1999.

[9] P. Paatero and U. Tapper. Positive matrix factorization: A non-negative factor model with optimal utilization of error estimates of data values. *Environmetrics*, 5(2):111–126, 1994.

[10] S. Yamauchi, M. Kawakita, and J. Takeuchi. Botnet detection based on non-negative matrix factorization and the mdl principle. In et al. Z. Zeng, editor, *Neural Information Processing*, LNCS 7667, pages 400–409, 2012.