A Study of Packet Sampling Methods for Protecting Sensors Deployed on Darknet

Masaki Narita, Keisuke Kamada, Kanayo Ogura, Bhed Bahadur Bista, Toyoo Takata
Iwate Prefectural University
Graduate School of Software and Information Science
narita_m@iwate-pu.ac.jp, g231m010@s.iwate-pu.ac.jp, {ogura_k, bbb, takata}@iwate-pu.ac.jp

***Abstract*—A darknet monitoring system is developed to grasp malicious activities on the Internet in an early stage and to cope with them. The darknet monitoring system consists of network sensors deployed widely on the Internet. The sensors capture incoming unsolicited packets. A goal of this system analyzes captured malicious packets and provides effective information for protecting good Internet users from malicious activities. To provide effective and reliable information, sensors must be deployed in secret and hidden from outside. On the other hand, attackers intend to detect sensors for evading them. This attempt is known as localization attacks to darknet monitoring systems. If actual location of sensors is revealed to attackers, it is almost impossible to grasp the latest tactics used by attackers. Thus in our previous work, we proposed a packet sampling method, which samples incoming packets based on an attribute of packets sender, to increase a tolerance to a localization attack and to keep a high quality of information publicized by the system. As a result, we almost succeeded to counter from a localization attack, which generates spike on the publicized graph to detect a sensor. However in some cases, proposed sampling method works to attacker's advantage and spikes appear clearly on the graph. Therefore, we propose advanced sampling methods, which sample incoming packets based on multiple attributes of packets sender. In this paper, we present our improved methods and show a promising evaluation result obtained from the simulation.**

***Keywords*-darknet monitoring system; Internet threat monitoring; localization attack; packet sampling; network security;**

I. Introduction

Using the Internet is indispensable in our daily life. In the meantime, wide variety of malwares cause cyber attack that becomes serious threat to safe and reliable use of the Internet by stealing confidential personal data, launching Denial-of-Service (DoS) attacks to specific corporate enterprises to hinder their providing service, etc. Symantec security report published in 2015 reported that previously unknown 6,549 software vulnerabilities were found in 2014 [1]. This number is comparable to that of in 2013 and it is up by approximately 24% if we compare with the case of in 2012. Cyber attacks mainly exploit software vulnerability. Thus, it is essential to find such vulnerabilities in an early stage to prevent severe financial damage to corporate enterprises and/or general users and cope with them.

To find software vulnerabilities and keep up on the latest trend of attacks spreading on the Internet, darknet monitoring systems are developed. Fig. 1 depicts an overview of a darknet monitoring system.

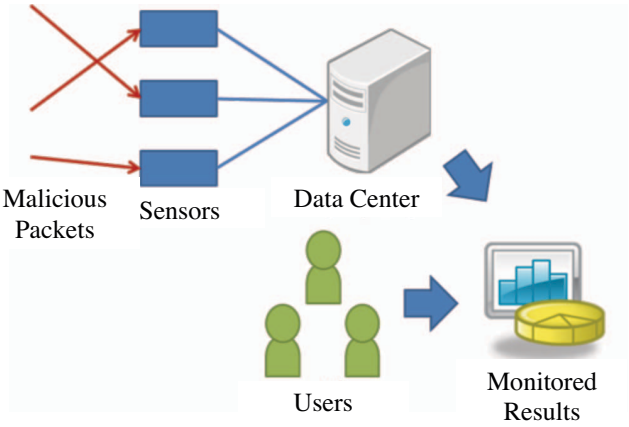

Figure 1. Overview of darknet monitoring system

A darknet monitoring system consists of multiple network devices, i.e., sensors, deployed in an unused IP address space on the Internet. Note that these sensors give no service outside and they are configured for capturing all incoming packets. In fact, sensors receive many unsolicited packets as long as they connect to the Internet directly. Arriving packets in an unused address space, i.e., darknet, is not general under normal circumstances. Thus, we can infer these packets are more likely to be sent by malicious purpose. Organizations operating a darknet monitoring system collect these malicious packets and analyze them. The analyzed results are opened to the public as effective information for protecting good Internet users from malicious activities.

Meanwhile, sensors must be deployed in secret and hidden from outside to provide effective and reliable information because attackers intend to detect sensors' IP addresses for evading them. This attempt is known as localization attack to darknet monitoring systems. If sensors' IP addresses are revealed to attackers, it is almost impossible to grasp the latest malicious activity as they bypass sensors artfully. Additionally, there is a possibility that a sensor is to become a target that is exposed to severe DoS attacks launched by attackers.

When attackers initiate a localization attack, they disguise general Internet users. They misuse information provided by organizations operating a darknet monitoring system for detecting a sensor. Hence, if organizations release a time series graphical information alone without taking coun-

2157-0426/16 $31.00 © 2016 IEEE
DOI 10.1109/NBiS.2016.37

termeasures, attackers would misuse the publicized graph and detect a sensor easily. Thus in our previous work, we proposed a packet sampling method, which samples capturing packets based on an attribute of packets sender, to increase a tolerance to a localization attack and to keep a high quality of information provided by a darknet monitoring system [2]. Consequently, we almost succeeded to counter from a localization attack, which generates spike on the publicized graph to detect a sensor. However in some cases, unexpected spike appear in the graph that can be a guide to location of a sensor.

Therefore, we propose advanced sampling methods, which sample incoming packets based on multiple attributes of packets sender. We designed proposed method by fully using previous knowledge from our previous work. Performance evaluation was conducted by simulating attackers' tactics and applying proposed methods. We used real captured packets provided by nicter darknet operated in Japan. In addition, we also discuss degradation of publicized information by sampling captured packets compared with no sampling case. In this paper, we present our improved methods and show a promising evaluation result.

II. Related Works

A wide variety of darknet monitoring systems are operated all over the world to grasp the latest trend of attacks on the Internet. In Japan, National Institute of Information and Communications Technology (NICT) operates nicter [3]. Nicter has many user interfaces. For example, *Cube* draws a cubical object on the center of a window and maps captured packets on the object based on source and destination of arrived packets. *Atlas* maps captured packets on the world map. *Atlas* indicates that which country sends malicious packets to Japan in real time. In addition, nicter counts the number of all captured packets from the aspect of country where those originated, and classifies TCP/UDP packets. All these information is accessible by anybody from the web site. Japanese National Police Agency also operates @police [4]. @police gathers Firewall's log and intrusion detection system's log at the gateway of institutions affiliated with the police to provide security-related information. @police publicizes their result in the form of time series graph and table on their web site.

From a global perspective, DShield [5] recruits many volunteers as informants from across the world to grasp the trend of malicious activities and established the world's largest darknet community. CAIDA (Cooperative Association for Internet Data Analysis) [6] manages UCSD (University of California, San Diego) Network Telescope system for capturing Internet traffic. This system dominates one whole /8 network. It captures all the packets of roughly 1/256 of the total Internet address space.

On the other hand, an attacker intends to detect sensors to bypass them. This attempt is known as localization attacks to a darknet monitoring system. Shinoda et al. [7] and Bethencourt et al. [8] indicate that an attacker can send a great number of probing packets to suspicious network that includes a sensor preliminary in the short term for detecting sensors. Subsequently, attackers affirm the presence of a sensor if spikes appear sharply on the time series graph publicized by the targeted system. Yu et al. introduces another localization attack inspired by spread spectrum technology [9], [10]. This method increases the stealthiness of attacks and is able to detect sensors with low probing packets, by sending probing packets in synchronization with value of a PN code sequence.

If IP addresses deploying sensors are unveiled by a localization attack, attackers can bypass sensors intentionally and perform malicious activities on the Internet. It may follow that grasping the latest trend of attacks becomes difficult. In addition, a sensor may become a target that is exposed to obstinate DoS attacks launched by attackers.

Basically, attackers infer IP addresses of sensors by investigating an open publicized data provided by a darknet monitoring system. In short, attackers put a mark on upcoming publicized data and verify the presence of the mark to detect a sensor. Hence, taking no countermeasure produces that attackers could exploit a publicized data and detect a sensor easily. As described above, to establish a countermeasure for localization attack becomes imperative.

Viecco and Camp indicate and emphasize the usefulness of packets sampling techniques for captured packets as a countermeasure to a localization attack [11]. However, they only discuss theoretical framework; any concrete sampling method has not been proposed and any numerical validation has not been done. Thus, we proposed a packet sampling method, which sampled capturing packets based on an attribute of packets sender. Sampling is essentially processing that decimates a part from entire and complete information. We assure that packets sampling alleviate the influence of probing packets sent by attackers in open publicized data and obtain a tolerance to a localization attack.

III. Localization Attack by Generating Spikes

In this section, we describe a concept of a localization attack, which generates spike on the publicized graph. This is based on the work by Shinoda et al. [7] and Bethencourt et al. [8]. Shinoda et al. [7] define a direct and intentional activity of detecting sensor as marking. They also define spikes used for marking as a marker. In this paper from this point forward, we use the same terms: marking and marker. An overview of a localization attack is shown in Fig. 2 and we explain procedures below.

1) Attackers preliminary narrow down lists of suspicious target networks that include a sensor by analyzing materials on the web, obtaining handouts distributed in a workshop by an operating organization of the target darknet monitoring system. The investigated results

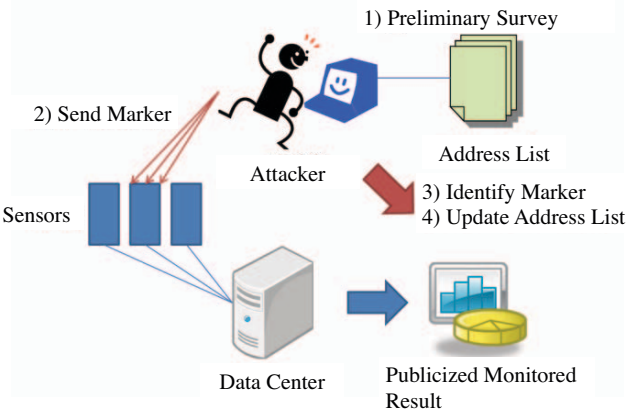

Figure 2. Overview of localization attack

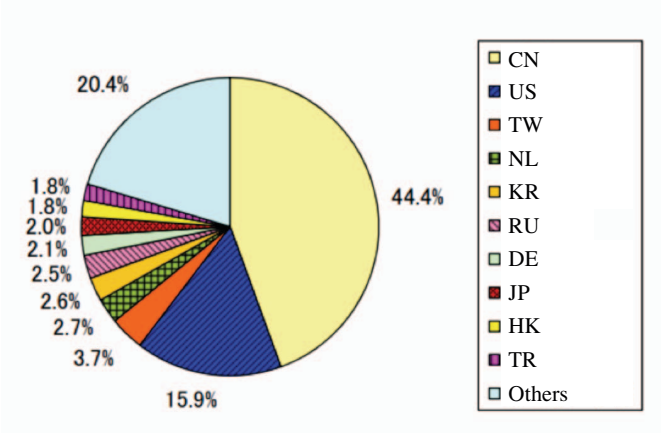

Figure 3. Ratio of source country of captured packets

are preserved by attackers and those become the lists of suspicious IP addresses that act as sensors on the Internet. Then, attackers select a range of IP addresses from the lists and initiate localization attack by means of sending markers.

2) Attackers send markers in the short term to the suspicious networks that include a sensor, i.e., marking. If the sensor receives packets in short period of time, a discriminating spike appears in the time series graph publicized by the target darknet monitoring system.
3) Attackers disguising general and good Internet users access to a web site that the organization of the target darknet monitoring system operates. They verify the presence of a sensor in target network by identifying trace of marker in publicized time series graph publicized by the system.
4) Following the result of procedure 3), attackers update the lists of suspicious IP addresses. They identify precise and concrete sensor's IP address by repeating from procedure 2) to procedure 4).

The above procedures are an overview of localization attack, which produces spike on a publicized graph. Since markers are sent as standard packets generated by port scan, it is not easy to distinguish between markers and other packets. Consequently, a darknet monitoring system is to give instructive feedback to attackers unintentionally by publicizing monitored results that is based on an all captured packets including markers.

IV. A Proposal of Packet Sampling Methods

We have proposed a packet sampling method, which samples incoming packets based on an attribute of packets sender, to increase a tolerance to a localization attack so far [2]. In this section, we propose two advanced sampling methods, which sample incoming packets based on multiple attributes of packets sender based on our previous work.

A. Method 1: Sampling Captured Packets by Focusing on Arrival Time and Source IP Address

As a first step, we decide the time of capturing packets by using random numbers in common with our previous work, which uses arrival time of packets. Consequently, using random numbers makes attackers difficult to infer when sensors capture packets and we also keep captured packets in chronological order. This is effective in contributing a tolerance to a localization attack.

However, whether success ratio of preventing a localization attack or not depends on a matter of probability. In the worst case, each sensor may capture all packets sent by attackers based on the time selected by random numbers. Thus, we additionally sample captured packets based on source IP address as a second step.

Generally, global IP address is administrated by a registration system. Users of global IP address must register their certain amount of identity to WHOIS. Because WHOIS information contains which IP address is assigned to which country, it is easy to access to information where captured packets came from [12]. As shown in Fig. 3 provided by @police [13], a ratio of source country of captured packets is greatly biased toward several countries. We assume that attackers manipulate multiple hosts in botnet for sending markers when they attempt a localization attack. If botnet is composed of exploited hosts whose source IP addresses are greatly biased toward several countries, sampling captured packets uniformly in terms of source countries reduces a spike produced by attackers on a publicized graph even with the containing markers in the time selected by random numbers.

The procedure of the sampling method 1 is described below. We use GeoIP [14] for mapping from a source IP address to a country that originate captured packets.

Procedure 1, Decision of Capturing Time:

1) Configure two parameters, m and n as $0 \leq m < n \leq 60$. Hereafter, (m, n) is called as range of random

Table I
INITIAL TTL VALUE OF MAJOR OPERATING SYSTEMS

OS	Initial Value
UNIX	255
Windows	128
Linux	64

numbers.
2) Divide captured packets into separate data per one hour. The data is denoted as $..., D_{i-1}, D_i, D_{i+1}, ...$ Generate l random numbers for each D_i. Note that $l \in [m, n]$.
3) Generate distinct l random numbers r_j under following conditions, $1 \leq j \leq l$ and $0 \leq r_j < 60$.
4) Sample packets arrived at k minutes at each D_i, $k \in \{r_1, r_2, ..., r_l\}$.

Procedure 2, Sampling Packets whose Originated Countries is Distributed Equally:

1) Identify originated country by referring source IP address of a captured packet.
2) Sample or discard packets based on the result of 1) as described below.
   a. If a packet is not captured from the country yet, sample it.
   b. If a packet has been captured from the country, sample it as long as packets sent from the country are below predefined threshold θ_1. Otherwise, discard it.
3) Reset a list of the sampled country per one hour.

B. Method 2: Sampling Captured Packets by Focusing on Arrival Time and TTL

As a first step in method 2, we decide the time of capturing packets by using random numbers in common with the method 1. Then, we sample captured packets based on TTL (Time-To-Live) as a second step.

TTL is a configured value in an IP header to prevent endless loop of packets caused by misconfiguration on the network. The value of TTL is reduced one by one whenever a packet gets through a router. When the value of TTL is reduced to zero, the corresponding packet is discarded. Generally, initial value of TTL is defined by operating system as shown in Table I.

Thus, TTL becomes a clue to infer the number of routers passing packets on a communication path and we can use it as an index of distance on a network. As is the case with the method 1, if source distance of captured packets is greatly biased, we believe sampling captured packets uniformly in terms of the number of network hops reduces a spike produced by attackers on a publicized graph. The procedure of the sampling method 2 is described below.

Procedure 1, Decision of Capturing Time:

1) Configure two parameters, m and n as $0 \leq m < n \leq 60$. Hereafter, (m, n) is called as range of random numbers.
2) Divide captured packets into separate data per one hour. The data is denoted as $..., D_{i-1}, D_i, D_{i+1}, ...$ Generate l random numbers at each D_i. Note that $l \in [m, n]$.
3) Generate unique l random numbers r_j under following conditions, $1 \leq j \leq l$ and $0 \leq r_j < 60$.
4) Sample packets arrived at k minutes at each D_i, $k \in \{r_1, r_2, ..., r_l\}$.

Procedure 2, Sampling Packets Uniformly in Terms of the Number of Network Hops:

1) Infer the number of network hops by TTL value. If captured TTL value is t_1, we obtain t_2 as the t_2 is greater than the value of t_1 and lower than minimum value listed in Table I. We define the network hops as $t_2 - t_1$.
2) Sample or discard packets based on the result of 1) as described below.
   a. If the number of network hops is not captured yet, sample it.
   b. If the number of network hops has been captured, sample it as long as the number of network hops is below predefined threshold θ_2. Otherwise, discard it.
3) Reset a list of the number of network hops per one hour.

V. PERFORMANCE EVALUATION

A. Evaluation of a Tolerance to a Localization Attack

1) Assumption: In the first place, we simulate a localization attack described in section III and evaluate the effectiveness of our proposed methods to counteract the attack. Simulations were carried out by using a dataset [15], which included packets captured in November of 2015. The dataset is provided by nicter, i.e., an actual darknet monitoring system operated in Japan. An assumption of a target darknet monitoring system and a localization attack are shown below.

Assumption of a target darknet monitoring system:
In this simulation, the target darknet monitoring system captures packets at each port. Then, the system publicizes and updates the monitored results at the same time interval. As shown in Fig. 4, a transition of the number of packets captured by the system is plot on time series graph per one hour within the compass of one week. An X-axis indicates the number of packets and a Y-axis does elapsed time. Since any general Internet users can access the publicized result, malicious users also can access the same information.

Assumption of a localization attack attempted by attackers:
We assume that an attacker sends markers to 445/tcp port as a destination. In this simulation, because the number of packets arrived at 445/tcp port draws relatively smooth graph

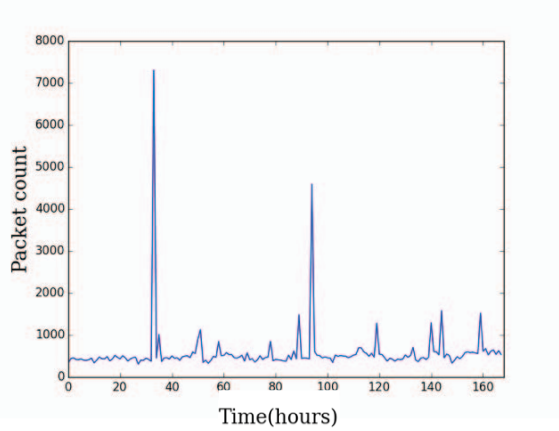

Figure 4. Publicized graph that the target darknet monitoring system publicizes

Table II
THE BREAKDOWN OF BOTNET A IN TERMS OF SOURCE COUNTRIES

Country (Code)	The Number of Hosts
CN	2
US	1
RU	1

compared with that of other ports as shown in Fig. 4, we think that attackers are easy to identify their spikes on the publicized graph. In other words, we believe that this port is suitable for a localization attack.

In this paper, we assume two groups of botnets, botnet A and botnet B. Breakdown of two botnets are shown in Tables II–V. Botnet A is the group composed of exploited hosts whose source IP addresses and network hops are greatly biased. Botnet B is the group that is the same assumption in our previous work for the comparison. The number of sending packets, markers, duration of sending packets and time interval are shown in Table VI.

Table III
THE BREAKDOWN OF BOTNET A IN TERMS OF NETWORK HOPS

The Number of Network Hops	The Number of Hosts
18	2
15	1
12	1

Table IV
THE BREAKDOWN OF BOTNET B IN TERMS OF SOURCE COUNTRIES

Country (Code)	The Number of Hosts
CN	7
US	3
TW	2
NL	2
IN	1
KR	1
TR	1
RU	1
FR	1
MX	1

Table V
THE BREAKDOWN OF BOTNET B IN TERMS OF NETWORK HOPS

The Number of Network Hops	The Number of Hosts
18	4
19	4
15	2
17	2
20	1
21	1
16	1
14	1
22	1
12	1
28	1
27	1

Table VI
PARAMETERS OF MARKING

Destination Port	445/tcp
The Number of Sending Packets	3,000
The Number of Markers	4
Time Interval among Markings	20 hours
Duration of Marking	within 1 minutes

2) How to Determine Markers by Attackers: After attackers attempted marking, they disguise general good Internet users and access to a darknet monitoring system to identify their markers on a publicized graph for ascertaining presence of a sensor. On this occasion, attackers need to determine whether appeared markers, i.e., spikes, on the graph are generated by themselves or not. In this paper, we assume attackers determine their marker by an outlier detection method. If the spikes are identified as statistical outlier, attackers are able to regard the spikes as their markers. In our evaluation, we adopt the most general statistical method, 2σ outlier test. Fig. 5 represents a simulated graph where attackers insert four markers. If organizations operating a darknet monitoring system provide the graph as it is without taking countermeasures, attackers would acquire advantageous feedback and an IP address of a sensor is easily unveiled.

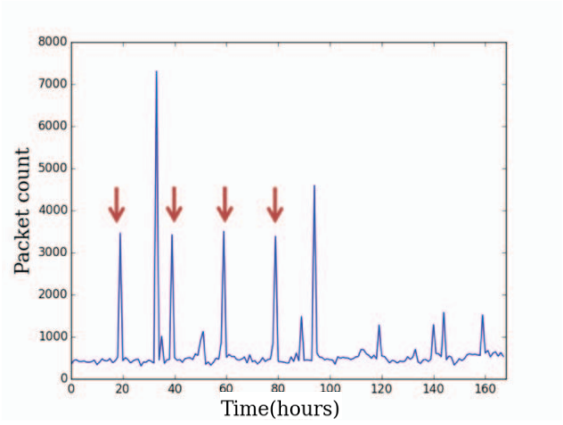

Figure 5. Markers generated by attackers

B. Evaluation of Information Quality

Sampling is essentially processing that decimates a part from entire and complete information. Thus, it is necessary to consider a degradation of information quality that a darknet monitoring system provides when we use a sampling method. In this paper, we also evaluate our proposed methods in terms of information quality of sampling case compared with that of no sampling case. We decide to evaluate a quality by comparing between before and after of sampling in the well-known publicized format, i.e., time series graph and table.

1) Rate of Concordance of the Most Accessed Port: A number of darknet monitoring systems generally publicize the most accessed top 10 port numbers on their web site in the form of table. The most accessed top 10 port numbers mean destination ports that attackers intend to exploit. We assumed two sets that included top 10 port numbers each. One is the set X that is obtained result by no sampling. The other is the set Y that is obtained result by sampling. We evaluated the rate of concordance of these two sets. To evaluate the rate of concordance of two sets, we computed the following Simpson's coefficient,

$$Sim = \frac{|X \cap Y|}{min(|X|, |Y|)}.$$

2) Similarity of Time Series Graph: A number of darknet monitoring systems also publicize the amount of arriving packets at each port as a time series graph on their web site. In this paper, we compare a time series graph with no sampling $\{X_u\}$ and a time series graph with sampling $\{Y_u\}$. We adopt the following Bhattacharyya coefficient L as an index to compute similarity of above two graphs,

$$L = \sum_{u=1}^{m} \sqrt{X_u Y_u},$$

$$\sum_{u=1}^{m} X_u = \sum_{u=1}^{m} Y_u = 1,$$

$$(0 \leq L \leq 1).$$

Bhattacharyya coefficient is essentially an index used for computing similarity of histograms. In this paper, we define the value of $1 - L$ as discrepancy of graphs. In short, two graphs are similar as the value gets close to zero.

VI. Results and Discussion

The evaluation results are shown in Tables VII–XI. The number of markers at each table means the number of spikes that attackers successfully identified in the publicized graphs.

Table VII is the result obtained by our previous method. The other tables are the results obtained by our proposed methods. Our proposed sampling methods succeed to increase a tolerance to a localization attack if we focus the number of appearing markers compared with the previous method. However, if we refer to discrepancy of graphs and rate of concordance of the accessed port, procedure 2, i.e., sampling source countries or network hops uniformly, brings a degradation of information quality.

As for the rate of concordance of the accessed port, method 2 obtains better results as compared to method 1. We conclude that our proposed methods achieve higher tolerance to a localization attack by decreasing sampling packets as a range of random number defined in procedure 1 and threshold defined in procedure 2 become small. While on the other hand, degradation of information quality occurs.

In our previous work, prevention of a localization attack depends on the probability. As shown in Fig. 6, if each sensor may capture all packets sent by attackers based on the time selected by random numbers, spikes are further emphasized. This is a major problem to solve. On the other hand, method 1 and method 2 succeed to alleviate a problem by sampling source countries or network hops uniformly in procedure 2 if sensors capture all packets sent by attackers in procedure 1. As shown in Figs. 7 and 8, our proposed methods succeed to curb markers if the number of hosts managed by attackers are small and attributions are greatly biased. Fig. 9 shows that our method can prevent the markers as threshold defined in procedure 2 is low. If the number of hosts managed by attackers is large and attributions are not biased, we confirm that our proposed methods can alleviate an impact of markers.

VII. Conclusion and Future Works

In this paper, we proposed advanced packet sampling methods based on multiple attributes of packets sender to increase tolerance to a localization attack. As a result, our two kinds of proposed methods succeeded to overcome the

Table VII
Sampling Results of Our Previous Work

(m, n)	The Number of Markers	Discrepancy, $1 - L$	Sim
$(10, 20)$	1.02	0.0694	0.833
$(25, 35)$	2.08	0.0352	0.910
$(40, 50)$	3.05	0.0140	0.949

Table VIII
Sampling Results of Method 1 based on Arrival Time and Source Countries against Botnet A

(m, n)	θ_1	The Number of Markers	Discrepancy, $1 - L$	Sim
$(10, 20)$	10%	0.18	0.0621	0.742
$(10, 20)$	15%	0.62	0.0556	0.753
$(25, 35)$	10%	0.10	0.0509	0.744
$(25, 35)$	15%	0.76	0.0417	0.771
$(40, 50)$	10%	0.20	0.0466	0.743
$(40, 50)$	15%	1.82	0.0312	0.803

Table IX
SAMPLING RESULTS OF METHOD 1 BASED ON ARRIVAL TIME AND SOURCE COUNTRIES AGAINST BOTNET B

(m, n)	θ_1	The Number of Markers	Discrepancy, $1 - L$	Sim
$(10, 20)$	10%	1.00	0.0621	0.743
$(10, 20)$	15%	1.00	0.0617	0.750
$(25, 35)$	10%	1.92	0.0419	0.743
$(25, 35)$	15%	2.24	0.0372	0.773
$(40, 50)$	10%	3.06	0.0320	0.743
$(40, 50)$	15%	3.08	0.0226	0.796

Table X
SAMPLING RESULTS OF METHOD 2 BASED ON ARRIVAL TIME AND TTL AGAINST BOTNET A

(m, n)	θ_2	The Number of Markers	Discrepancy, $1 - L$	Sim
$(10, 20)$	10%	0.34	0.0601	0.786
$(10, 20)$	15%	0.60	0.0556	0.822
$(25, 35)$	10%	0.32	0.0482	0.839
$(25, 35)$	15%	1.18	0.0412	0.874
$(40, 50)$	10%	0.38	0.0434	0.883
$(40, 50)$	15%	1.86	0.0327	0.912

weakness of our previous method. In other words, we concluded our method prevented a localization attack without depending on the sampling time decided by random numbers along of sampling packets uniformly in terms of source countries or the number of network hops if a publicized graph might contain markers.

In our future works, we plan to investigate optimal trade-off between a tolerance to a localization attack and quality of publicized results by a darknet monitoring system. Additionally, we continue to make improvements of packet sampling algorithm further because attackers shall also devise more complicated localization attacks.

ACKNOWLEDGMENT

The authors thank MWS 2015 committee for their provision of valuable Internet monitoring dataset. This work was partially supported by JSPS KAKENHI Grant Number 26330159.

REFERENCES

[1] Symantec, 2015 Internet Security Threat Report, vol.20, 2015.

[2] K. Kamada, M. Narita, K. Ogura, B. B. Bista, and T. Takata, "Evaluation of Packet Sampling Methods for Protecting Sensors on the Dark Net," *Proc. 32nd Symposium on Cryptography and Information Security*, January 2015. (in Japanese)

[3] M. Eto, D. Inoue, J. Song, J. Nakazato, K. Ohtaka, and K. Nakao, "nicter: A Large-Scale Network Incident Analysis System: Case Studies for Understanding Threat Landscape," *Proc. 1st Workshop on Building Analysis Datasets and Gathering Experience Returns for Security*, pp.37–45, April 2011.

[4] @police. http://www.npa.go.jp/cyberpolice/english/

[5] DShield. http://www.dshield.org/

[6] CAIDA. http://www.caida.org/home/

Table XI
SAMPLING RESULTS OF METHOD 2 BASED ON ARRIVAL TIME AND TTL AGAINST BOTNET B

(m, n)	θ_2	The Number of Markers	Discrepancy, $1 - L$	Sim
$(10, 20)$	10%	0.76	0.0656	0.789
$(10, 20)$	15%	1.16	0.0637	0.820
$(25, 35)$	10%	2.00	0.0422	0.835
$(25, 35)$	15%	2.08	0.0392	0.873
$(40, 50)$	10%	3.00	0.0302	0.885
$(40, 50)$	15%	3.04	0.0245	0.914

[7] Y. Shinoda, K. Ikai, and M. Itoh, "Vulnerabilities of Passive Internet Threat Monitors," *Proc. 14th USENIX Security Symposium*, pp.209–224, July 2005.

[8] J. Bethencourt, J. Franklin, and M. Vernon, "Mapping Internet Sensors with Probe Response Attacks," *Proc. 14th USENIX Security Symposium*, pp.193–208, July 2005.

[9] W. Yu, X. Wang, X. Fu, D. Xuan, and W. Zhao, "An Invisible Localization Attack to Internet Threat Monitors," *IEEE Trans. Parallel and Distributed Systems*, vol.20, no.11, pp.1611–1625, November 2009.

[10] W. Yu, N. Zhang, X. Fu, R. Bettati, and W. Zhao, "Localization Attacks to Internet Threat Monitors: Modeling and Countermeasures," *IEEE Trans. Computers*, vol.59, no.12, pp.1655–1668, December 2010.

[11] C. H. Viecco and L. J. Camp, "A Risk Based Approach to Limit the Effects of Covert Channels for Internet Sensor Data Aggregators for Sensor Privacy," *Proc. 3rd IFIP International Conf. on Trust Management*, pp.234–251, June 2009.

[12] Spammers & Hackers: Using the APNIC Whois Database to Find Their Network. http://www.apnic.net/apnic-info/whois_search/using-whois/abuse-and-spamming

[13] Japanese National Police Agency, Internet Threat Report (The First Half of the Year in 2015). http://www.npa.go.jp/cyberpolice/detect/pdf/20150917.pdf (in Japanese)

[14] MaxMind GeoIP. http://dev.maxmind.com/geoip/

[15] Anti Malware Engineering WorkShop (MWS) 2015 Dataset. http://www.iwsec.org/mws/2015/en.html

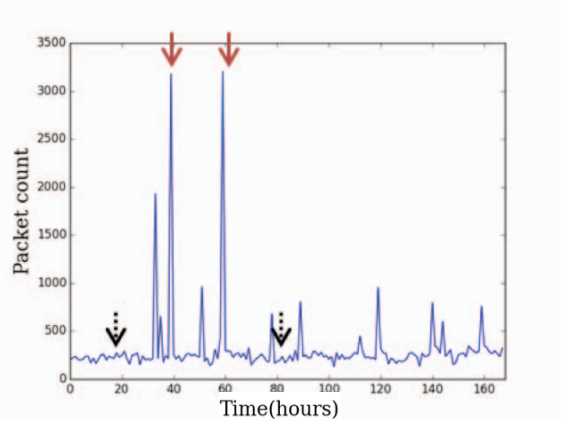

Figure 6. Publicized graph produced by our previous method under the conditions of $(m, n) = (25, 35)$

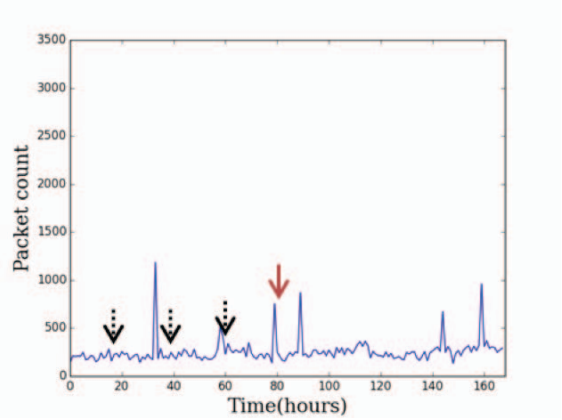

Figure 7. Publicized graph produced by method 1 to counteract botnet A under the conditions of $(m, n) = (25, 35)$, threshold $\theta_1 = 15\%$

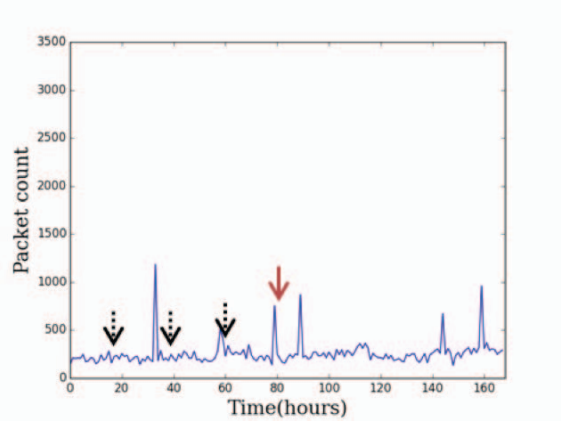

Figure 8. Publicized graph produced by method 2 to counteract botnet A under the conditions of $(m, n) = (25, 35)$, threshold $\theta_2 = 15\%$

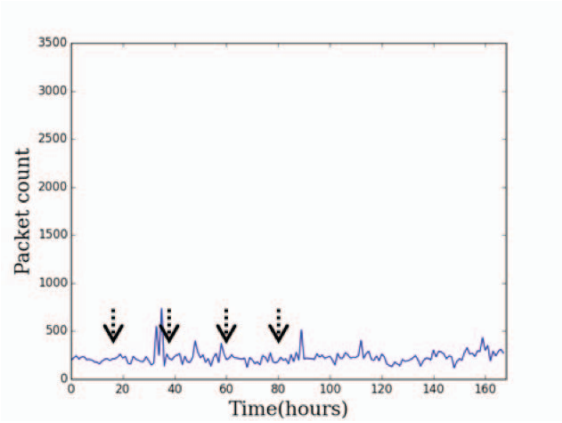

Figure 9. Publicized graph produced by method 1 to counteract botnet A under the conditions of $(m, n) = (25, 35)$, threshold $\theta_1 = 10\%$

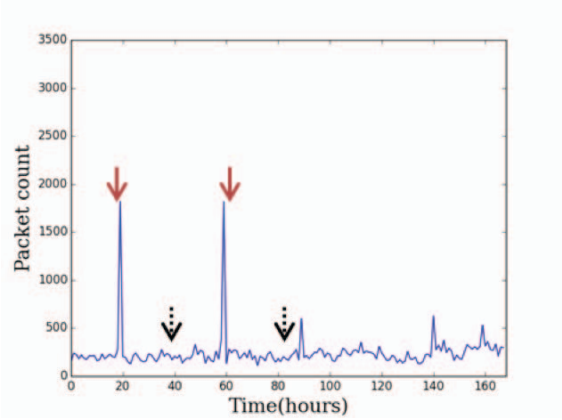

Figure 10. Publicized graph produced by method 1 to counteract botnet B under the conditions of $(m, n) = (25, 35)$, threshold $\theta_1 = 10\%$

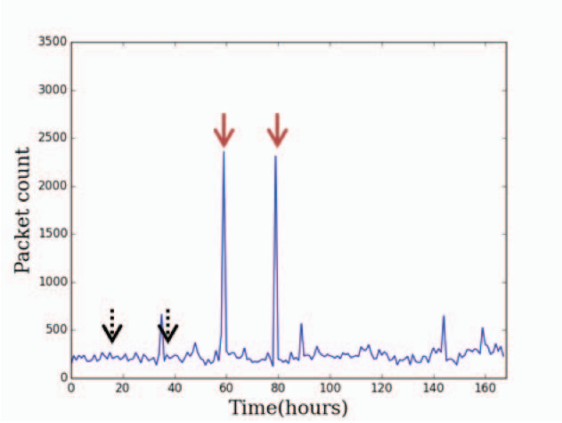

Figure 11. Publicized graph produced by method 2 to counteract botnet B under the conditions of $(m, n) = (25, 35)$, threshold $\theta_2 = 10\%$