

COMPARATIVE ANALYSIS of DARKNET TRAFFIC CHARACTERISTICS BETWEEN DARKNET SENSORS

Falguni Gadhia^{*,**}, Jangwon Choi^{**}, Buseung Cho^{**}, Jongsuk Song^{*,**}

**Korea University of Science & Technology, Daejeon, South Korea*

***Department of Advanced KREONET Security Service, Korea Institute of Science and Technology Information, Daejeon, South Korea*

gadhiafalguni@kisti.re.kr, jwchoi@kisti.re.kr, bscho@kisti.re.kr, song@kisti.re.kr

Abstract --- Today, Internet is incessantly attacked by wide variety of network-based threats. One of the ways to monitor or identify such prevailing threats is to monitor incoming traffic to unused network addresses popularly known as darknet and often also referred with various other names like network telescope or black hole. As, all the traffic arriving at darknet is mainly the result from malicious probing or misconfiguration in the network. It is expected that to have similar incoming traffic behaviour across different darknet sensors, however, various studies found it different. Various reason cited behind it is misconfiguration, certain kind of attack, difference in filtering parameter or system configuration itself. However, concrete reason beside this is still missing. In this regard, to get further understanding, in this study, we performed deeper comparative analysis between two darknet sensors (KISTI Darknet network) that are differently located but have similar filtering and system configuration. Comparative analysis considering total incoming packet, number of source host, targeting destination port and protocol revealed that there exists wide difference in incoming traffic characteristics between the darknet sensors. Moreover, for TCP and UDP comparison, UDP traffic showed more targeting behaviour to particular darknet block (difference in traffic characteristics between darknet sensors), in contrast to it, TCP traffic showed more scanning behaviour (similarity in traffic characteristics between darknet sensor).

Keywords—Darknet, network security, network monitoring, TCP, UDP

I. INTRODUCTION

From at least a decade, researchers are trying to identify and classify the network-based threats. But with the time as Internet usage is changing in same way new threats are also emerging. As security is becoming of prime importance, threats hindering smooth running of institution as well as Internet need to be detected at an early stage. Monitoring of unused address block to identify these malicious activities is an effective method to understand threats. Such system is

referred with different names like darknet, network telescope, blackhole etc. [1]-[3]. The unused address blocks means there are no active host present for these addresses i.e. no active services. So the activities or the packets coming to these addresses must be either due to misconfiguration (in which the network leads the packet to direct to darknet), scanning (for vulnerability or worm propagation), backscatter (from the spoofed source-ip addresses), DDoS activity or any other unwanted activity [1], [4]-[6]. Moreover, no active host in darknet helps to cover more address blocks for analysis without facing any privacy issue (majority faced for real network analysis). Large address block monitoring helps to quickly identify threats as it tries to affect other hosts in other blocks on Internet. As a result, Darknet helps to understand minute characteristic of anomalous traffic.

Many efforts have been made to identify abnormal traffic behavior one of which is by comparing different sized blocks. In [4], author used the Internet Motion Sensor (IMS), collection of blackhole sensor, deployed at different locations focusing on different sample of IPv4 address space. They observed that different blocks depicted different traffic and patterns. The possible reason behind such different patterns were given as (i) filtering policy at core and edge, (ii) propagation strategy, (iii) sensor address visibility, (iv) resource constraints and (v) statistical variation. But today, many Internet threats are internationally scoped like many worms, so one expects to see same behavior between the blocks. In [1], authors gave tempted reasons for these traffic differences related to targeting behavior (a) targeting those range of addresses that are identified to contain vulnerable hosts, so the attackers can easily target those systems, (b) placement of darknet before or after the firewall gives different visibility of threats of network. In their study, packets were normalized to size /24 to study different sized blocks located at different places, and such normalization of the packets helped to compare the results from different sized blocks. In [2], [8], Internet Background Radiation segments were studied by breaking the segments in terms of protocol,

application and port number. Both active and passive categorization was done in the study, which concluded that significant diversity was found in terms of blocks monitored as well as in terms of time. In this work, since the blocks that we are using to analyze traffic are of equal size, normalization of traffic packet is not required. Moreover, as the darknet is set of unused ip address, the pre-filtering of traffic is already done and the false positive and scaling issue is eliminated at large extent.

Here, our aim is to analyze traffic behavior between the differently located darknet sensors considering other parameters such as source host frequency, protocol, targeted port (destination port) etc. This may further help in to understand topological behaviors of darknet traffic and related health of network. Furthermore, the darknet traffic data used in analysis do not have much topological diversity (two different darknet sensor located in same country) and have similar filtering or other configuration parameter, which allow direct comparison between the sensors. In addition, freshness of this dataset (new dataset) could provide some additional important information or reconfirmation of previous findings.

II. DATASET

This section describes the information of the dataset we use and explain type of packets we use for the traffic analysis. Our dataset is comprised of darknet ip packets from two sensors captured over a span of 91 days i.e. Sept. 01, 2013 ~ Nov. 30, 2013 from Science and Technology Security Center (KISTI), South Korea. The sensors used for temporal study have same set of policies for the incoming darknet traffic. These sensors are of block size 16 each of Class C network with /24 netmask. From now onwards we will refer these sensors as sensor 1 (s1) and sensor 2 (s2).

III. COMPARATIVE ANALYSIS

This section is divided in three parts. §1 describes the health of network considering number of incoming packets. §2 explain proportionality of unique and common source-ip communicating with sensors. While §3 give details about way of communication by host for particular port in both the sensors.

A. Comparative Analysis of number of packets

Day by day trend of incoming traffic for 91 days in term of number of packet, with only TCP and with only UDP (packets per day) were compared between the sensor 1 and sensor 2 and depicted in figure 1. In certain time window number of incoming packet show similar trend but for certain time window it is dissociated, that means show different behaviors in packet inflow.

Similarities in inflow traffic, which maintain quite number of days endorse expectation of darknet traffic behavior. Moreover, it also suggests that there were no difference in filtering parameter or configuration of darknet sensor. Therefore, difference in pattern of darknet traffic between the

darknet sensors is not attribute to such difference in parameter or configuration, which is, one of the causes of difference in patterns between the darknet sensors [1], [4]. In our study, the observed differences in pattern for inflow traffic between the sensors suggest abnormal behavior particularly for sensor 1 (which is more targeted). Similar to previous study [2], [8], we also observed higher proportion of TCP packets than UDP packets coming to sensors.

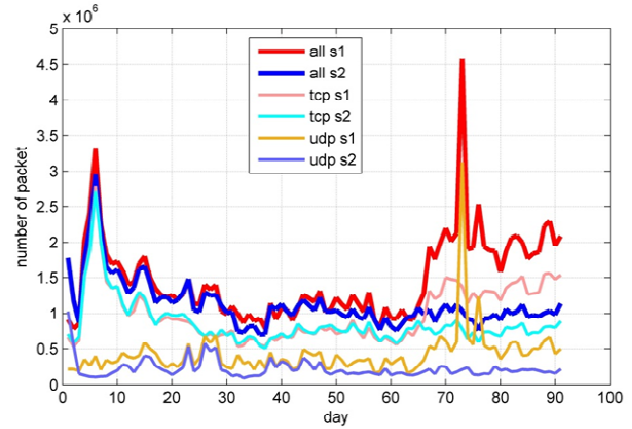


Figure 1: Temporal trend of total number of packets, total number of UDP packets and total of number of TCP packets coming to sensor 1 and sensor 2

B. Comparative Analysis of source ip

There is possibility that some of the source-ip is common in both while some of them are unique in each sensor. The set of source-ip are said to be common if they send proportionally same number of packet to both sensors and they are said to be unique if they send proportionally high amount of packet to either of sensor than other sensor. Here, we define proportionality in term of relative proportionality as R.P (equation 1). The value R.P=0, suggests same number of packet send to both sensor, R.P=1 suggests number of packets are send only to sensor 1 while R.P=-1 suggests number of packet send only to sensor 2. In this study, source-ip with $R.P. \leq |0.2|$ is defined as a common source-ip while source-ip with $R.P. > |0.2|$ is defined as a unique source-ip.

$$R.P = \frac{N1-N2}{N1+N2} \quad \dots (1)$$

where,

N1= number of packet from single source-ip in sensor 1 per day

N2= number of packet from single source-ip in sensor 2 per day

Relative proportion of source-ip with the higher number of packet either to one or both darknet sensors provide more confidence to assert similarity or difference of behavior between the sensors. So, four packets frequency (PF) ranges are studied i.e. $PF > 10, 100, 1000$ and 10000 . Higher the

frequency of packet, more confidence is developed for similarity or uniqueness behavior between darknet sensors. Temporal trend of total number of source-ip, common source-ip and unique source-ip for TCP and UDP protocol is given in figure 2.

At around day 70, dissociation is observed in terms of increase in number of source-ip for sensor 1 and sensor 2 for TCP. But, the proportion escalation is more for sensor 1 than sensor 2. So, for TCP traffic, majority of source-ip are

common source-ip that means they sent traffic to both darknet sensors equally. However, UDP traffic shows just opposite behavior where unique source-ip outnumber common source-ip that means they sent traffic to either of darknet sensor in more targeted manner than other sensor. Moreover, TCP traffic across the darknet sensor closely follows each other at certain time window (1-72) compared to UDP traffic, which deflects in majority of time window (more targeted behavior to sensor 1).

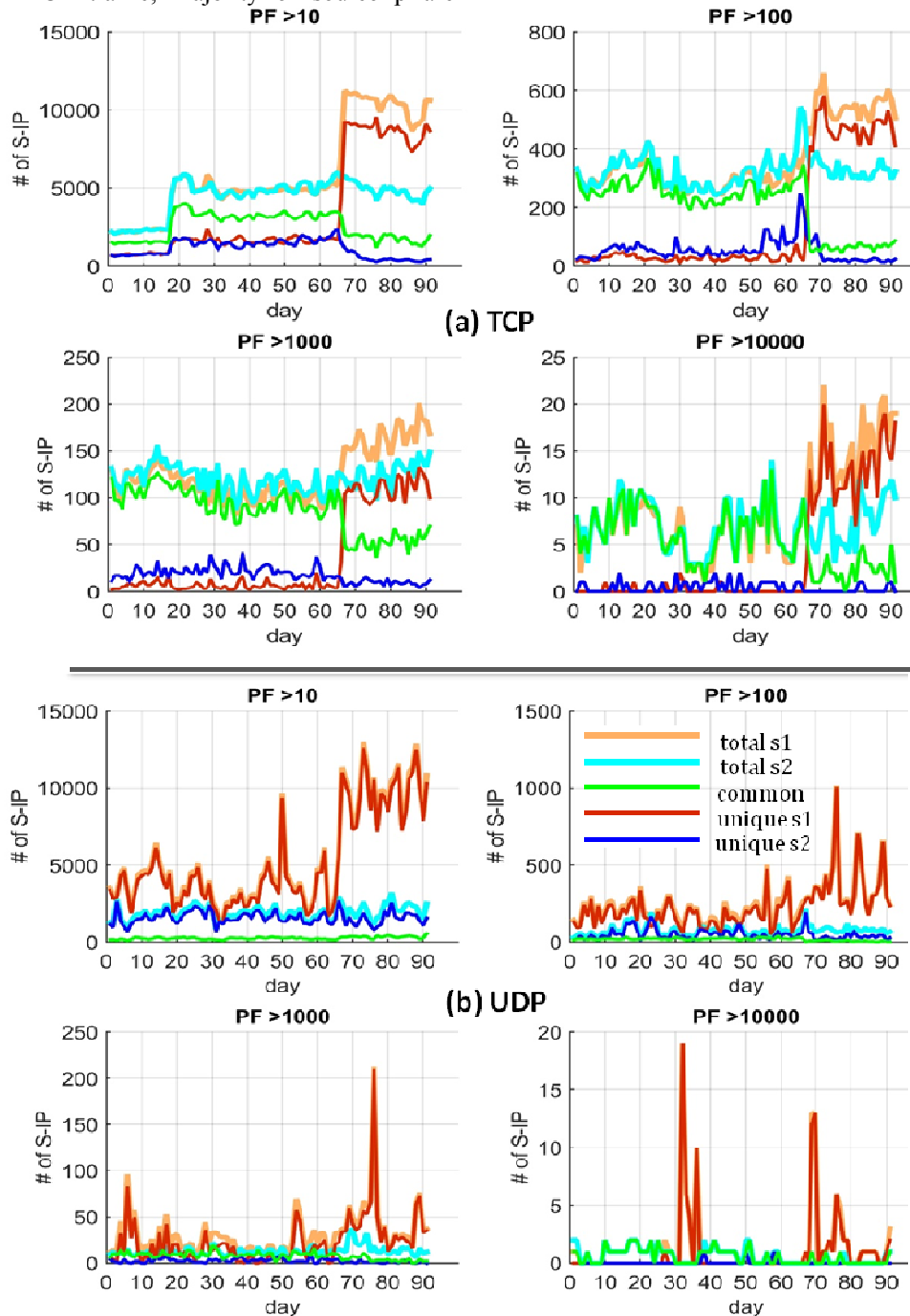


Figure 2: Temporal Trend of total source-ip, unique source-ip and common source-ip. Common source-ip is $R.P < 0.2$ and unique source-ip is $R.P > 0.2$. Packet frequency (PF) is the number of packet sent to either of sensor (a) for TCP and (b) for UDP

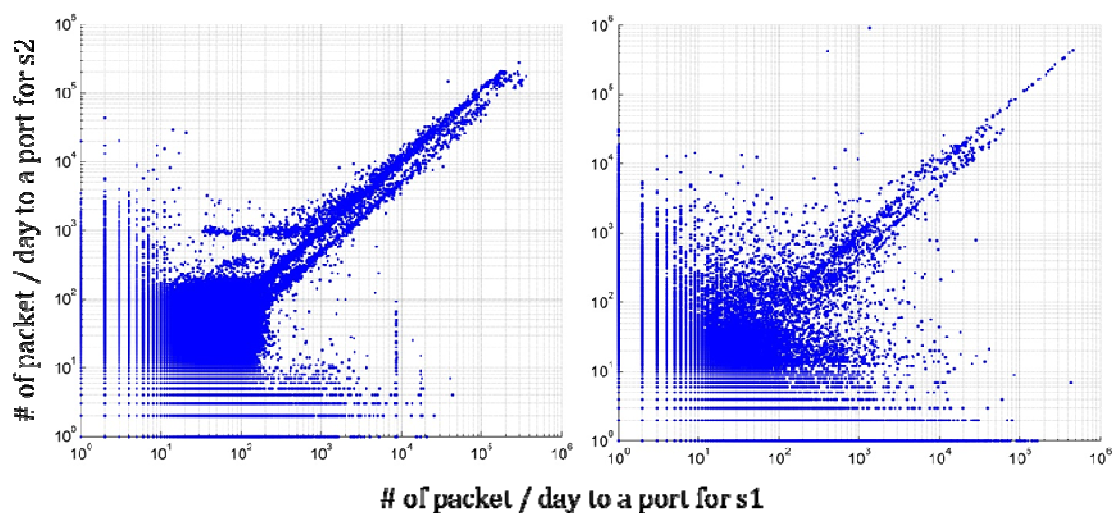


Figure 2: Comparison between sensor 1 and sensor 2 for Total number of packet to a port per day: a) using TCP protocol, b) using UDP protocol. Each dot represents single port per day

C. Comparative Analysis of destination port

In network we have 65536 standard defined ports on a computer. These ports are divided in three large categories. (i) 0-1023 are Well Known Port (ii) 1024-49151 are Registered Ports and (iii) 49152-65535 are Dynamic Ports. Host uses port scanning to know the vulnerability of computer to launch an attack [9]. This does not damage the system directly but gives useful information to host about the ports of computer. So to detect interest of host for particular ports is significant for behavior analysis.

Direct comparison of number of packets sent to particular port (per day) was performed between darknet sensors (figure 3). Port that have similar or near similar number of packet to both darknet sensor appear along the diagonal of plot. On other hand, ports that sent different number of packet between darknet sensors appears deviated to either of side from diagonal. This difference suggests that particular port is more targeted in one of darknet sensor compare to another one. Doing figurative comparison between TCP and UDP traffic, we clearly seen that, TCP traffic show more similar behavior than UDP that suggest comparative more similarity of traffic behavior between darknet sensors. On other hand, UDP traffic shows more deviated behavior to either side, which suggests comparative more uniqueness of traffic behavior.

In deeper analysis, we made the direct comparison of number of packets to top 25 ports (which have high frequency of packets on particular day) between darknet sensors. Here, number of packets to top port in one darknet sensor is directly compared to number of packets to same port in another darknet sensor. The ratio of number of packets to top ports between sensor 1 and sensor 2 for all 91 days for TCP and UDP traffic is depicted in figure 4. The value of ratio=1 suggests that similar number of packets to both darknet

sensors and deviation from it suggests difference in behavior of number of packets between darknet sensors. Larger the ratio more is the targeting behavior i.e. large amount of packet sent to one sensor than other sensor. So, for TCP top ports in both sensors have same frequency of packets and even the top port numbers are majority seem to be same in both sensors. On other hand, for UDP top ports in both the sensors, frequency of packets for the ports are different when top ports of sensor is compared to other sensor. Since we observe almost same behavior in sensor 1 and sensor 2 for TCP, similar behavior of top ports are observed. Unlike TCP, UDP have different behavior for top ports in both darknet sensors.

As a whole, comparative analysis between two darknet sensors explains about behavior of network parameters in term of number of packets, source-ip and destination port. Interestingly, behavioral pattern of darknet sensors follows each other in certain time window but at other time window it deflects and behaves differently. Therefore, observation of difference of behavioral pattern for darknet sensors at different location could provide more information about dynamics of abnormal traffic particularly about targeting or scanning behavior. In contrast to previous studies, the difference found here is not likely due to filtering parameters because they are same in this study. This is directly reflected as similarity for certain network dynamics between two sensors. Therefore, reported difference is more likely due to some abnormal traffic. With this study, we are able to notify about targeting and scanning behavior prevail in sensors located at different locations, which may help to understand further probing activities more precisely.

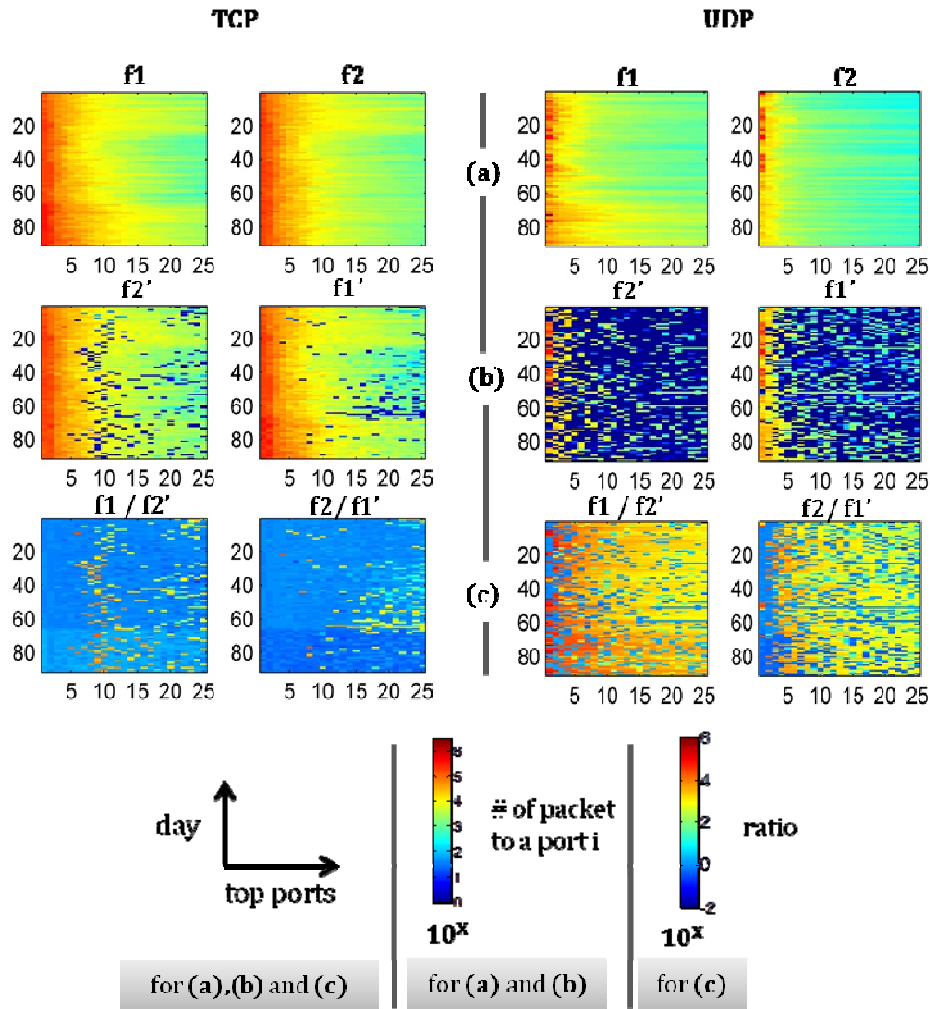


Figure 3: Comparison of total number of packets to top 25 ports between sensor 1 and sensor 2 Total number of packet (f1) to top 25 ports in sensor 1 and total number of packets (f2') at same port in sensor 2. Total number of packet (f2) to top 25 ports in sensor 2 and total number of packets (f1') at same port in sensor 1.

IV. CONCLUSION

In this study, various parameters such as total frequency of packet, number of source host, and targeted port were compared between two differently located darknet sensors. Deeper analysis revealed that even though total number of incoming packets are similar however, its composition differ between the sensors. TCP traffic coming to darknet sensors showed more similar behavior between differently located darknet sensors. In compared to this, UDP traffic showed much different traffic behavior that targeted either of the ip blocks of sensor. Such phenomenon is not reported in previous studies. However the exact reason behind dissociation of similar and unique behavior for TCP and UDP found in this study is yet to answer.

REFERENCES

- [1] M. Bailey, E. Cooke, F. Jahanian, A. Myrick and S. Sinha, 'Practical Darknet Measurement', in *40th Annual Conference on Information Sciences and Systems (CISS)*, Princeton, NJ, 2006, pp. 1496-1501.
- [2] E. Wustrow, M. Karir, M. Bailey, F. Jahanian and G. Huston, 'Internet Background Radiation Revisited', in *Proceedings of the 10th annual Conference on Internet Measurement (IMC'10)*, ACM, 2010.
- [3] J. Song, J. Choi and S. Choi, 'A Malware Collection and Analysis Framework Based on Darknet Traffic', in *19th International Conference on Neural Information Processing (ICONIP 2012)*, Doha, Qatar, 2012, pp. 624-631.
- [4] E. Cooke, M. Bailey, Z. Mao, D. Watson, F. Jahanian and D. McPherson, 'Toward understanding distributed blackhole placement', in *Proceedings of the ACM workshop on Rapid malware, WORM '04*, Washington DC, USA, 2004, pp. 54-64.
- [5] S. Choi, J. Song, S. Kim and S. Kim, 'A model of analyzing cyber threats trend and tracing potential attackers based on darknet

traffic', *Security Comm. Networks*, vol. 7, no. 10, pp. 1612–1621, 2014.

- [6] N. Furutani, 'Detection of DDoS Backscatter Based on Traffic Features of Darknet TCP Packets', in *9th Asia Joint Conference on Information Security (ASIA JCIS)*, Wuhan, 2014, pp. 39 - 43.
- [7] A. Lakhina, M. Crovella and C. Diot, 'Diagnosing network-wide traffic anomalies', *SIGCOMM Comput. Commun. Rev.*, vol. 34, no. 4, p. 219, 2004.
- [8] R. Pang, V. Yegneswaran, P. Barford, V. Paxson and L. Peterson, 'Characteristic of internet background radiation', in *Proceedings of the 4th ACM SIGCOMM conference on Internet measurement (IMC '04)*, Taormina, Sicily, Italy, 2004, pp. 27-40.
- [9] M. Bhuyan, D. Bhattacharyya and J. Kalita, 'Surveying Port Scans and Their Detection Methodologies', *The Computer Journal*, vol. 54, no. 10, pp. 1565-1581, 2011.

machine learning, security issues on IPv6, spam analysis, and cryptography theory.



from March 2013. Her research intersect include network monitoring, network and data management, network security.

Falguni Gadhia received her B.S degree in 2010 from Atmiya Institute of Technology and Science, India. She worked as an intern in Korea Institute of Science and Technology Information from Sept. 2011 to Aug. 2012. She is currently pursuing her master degree in grid and supercomputing from Korea University of Science and Technology. She is working as a master student in Department of Advanced KREONET Security Service, Korean Institute of Science and Technology Information



research networking, future internet and future network operation, knowledge-based network configuration and fault management, optical network and network management and data modeling and standardization for plant. He is member of Technology Program Committee, International Conference on Networks (ICN), GLIF and GOLE Partnership, GLORIAD Partnership and APAN.

Buseung Cho received his B.S, M.S and Ph.D. degree from Sungkyunkwan University, Korea in 2000, 2002, and 2012 respectively. He was a researcher in Institute for Advanced Engineering from Sept. 2002 to July 2005. He is working as a senior researcher in Korea Institute of Science and Technology Information(KISTI) from July 2005. Currently, he is a senior researcher and director of Dept. of Advanced KREONET Operation and Service. His research interest include international



security issues on IoT, and cryptography theory

Jangwon Choi received his B.S. and M.S degrees in Electronic Engineering from Hongik University, Korea in 1996 and 1998, respectively. He received his Ph.D. degree in the Department of Computer Science and Engineering, Korea University, Korea in 2009. He is a principal researcher and director of Dept. of Advanced KREONET Security Service at Korea Institute of Science and Technology Information. His research interests include network, cloud computing, grid computing, network security,



researcher at KISTI (Korea Institute of Science and Technology Information), Daejeon, Korea. His research interests include network security, data mining,

Jungsuk Song received his B.S. and M.S. degrees in Information and Telecommunication Engineering from Korea Aerospace University, Korea in 2003 and 2005, respectively. He received his Ph.D. degree in the Graduate School of Informatics, Kyoto University, Japan in 2009. He worked for NICT (National Institute of Information and Communications Technology), Tokyo, Japan, as an expert researcher from Apr. 2009 to Sep. 2010 and as a researcher from Oct. 2010 to Sep. 2011. He is currently a senior