（別紙）

専攻分野及び研究計画

Field of Study and Study Program

Full name in native language	PHAN	THIEN	XUAN
（姓名（自国語））	(Family name)	(First name)	(Middle name)

Nationality Vietnamese
（国　籍）

Proposed study program in Japan (State the outline of your major field of study on this side and the details of your study program on the backside of this sheet in concreteness. This section will be used as one of the most important references for selection. Statement must be typewritten or written in block letters. Additional sheets of paper may be attached if necessary.)
（日本での研究計画；この研究計画は，選考の重要な参考となるので，表面に専攻分野の概要を，裏面に研究計画の詳細を具体に記入すること。記入はタイプ又は楷書によるものとし，必要な場合は別紙を追加してもよい。）

If you have Japanese language ability, write in Japanese.
（相当の日本語能力を有する者は，日本語により記入すること。）

1　Field of study（専攻分野）

Internet traffic anomalies may cause a serious problem for the users and Internet service operators since they affect directly the availability of network services and prevent legitimate users from accessing the networks resources. Detecting and preventing the Internet traffic anomalies are very important tasks to maintain the network services. Existing anomaly detection approaches are typically based on the conventional network architecture that demand heavy processing in order to extract feature information needed for traffic analysis. This disadvantage results in the rather large amount of delay time to be able to detect abnormal traffics in case of network attacks, and thus the reaction against the attacks is rather inefficient. The abnormal traffic detection mitigation in large-scale networks is even more challenged since a large number of switches/routers need to be investigated for such tasks in the networks.

This research topic aims to provide an efficient architectural solution for early-stage anomaly detection and mitigation of Internet traffic on large-scale networks.

2　Study program in Japan in detail and concreteness（研究計画：詳細かつ具体に記入すること。）

Research Topic: **Early-stage Anomaly Detection and Mitigation in Large-scale Networks.**

The Research Proposal for the research topic is as follow.

I. Motivation

Internet services become more and more popular in human society nowadays. The growth of Internet services result in the instant growth of network traffic along with an increasing number of anomalies such as misconfiguration and remote attacks. These Internet traffic anomalies cause a serious problem for the users and Internet service operators since they affect directly the availability of network services and prevent legitimate users from accessing the networks resources. Therefore, detecting and preventing the Internet traffic anomalies are very important tasks to maintain the network services.

Existing anomaly detection approaches [1] are typically based on the conventional network architecture that demand heavy processing in order to extract feature information needed for traffic analysis. This disadvantage results in the rather large amount of delay time to be able to detect abnormal traffics in case of network attacks, and thus the reaction against the attacks is rather inefficient. The abnormal traffic detection mitigation in large-scale networks is even more challenged since a large number of switches/routers need to be investigated for such tasks in the networks.

This research topic aims to provide an efficient architectural solution for early-stage anomaly detection and mitigation of Internet traffic on large-scale networks. We exploit the flexibility and programmability of software-defined networking [2, 3, 4] for abnormal traffic measurement and mitigation in order to achieve an efficient anomaly protection solution.

II. Problem Statement

This research focuses on the problem of anomaly detection and protection of Internet traffic in production networks. The objective is to provide an architectural solution for anomaly detection on large-scale inter-networks in an early state and flexible reaction in case of network attacks. The DDoS attack [6, 9], which is the most widely used type of attack will be concentrated as a case study in our solution.

III. Related Works

This section introduces an overview about SDN concept, some existing approaches for anomaly measurement and detection based on SDN, as well as traditional proposals for anomaly detection in conventional networks.

Software-Defined Networks (SDN) [2, 3, 4] has been proposed and become more and more widely popular recently. This innovative network architecture provides the programmability to control the network by software in a centralized controller. It allows network operators to easily retrieve networking data from switches or routers in the network, as well as adding instructions to control the behaviors of these switches or routers. The flexibility and the programmability of SDN benefits in network management and a variety of solution has been proposed based on SDN to deal with current networking issues such as network traffic measurement [5] and anomaly detection [6].

The authors in [6] propose a lightweight method for detecting the DDoS flooding attack. This approach uses NOX/OpenFlow controller [7] to quickly extract traffic flows features and use the intelligent module based on Self Organizing Map [8] to analyze the traffic. By taking advantage of software-defined networking approach, the overhead of extracting switch information is low resulting in high rate detection and low rate of false alarms. Although the detection this research seems to be rather good, it can just be applicable for single network since it does not allow multiple interconnecting controllers for detecting in multiple networks. Although SDN approach can bring a lot of benefits network measurement and management, to the best of our knowledge, [6] is almost the only research that exploit the SDN approach for anomaly detection currently.

In [5], the authors proposed OpenSketch, a software-defined traffic measurement architecture that can be deployed in practice for a variety of measurement tasks in software-defined networks. OpenSketch separates measurement data plane from control plane. The data plane of OpenSketch provides a simple three-stage pipeline including hashing, filtering, and counting, which can be implemented with commodity switch components and support many measurement process. The control plane provides a measurement library for automatically configuring the pipeline and allocates resources for different measurement tasks. With its effectiveness showed by real measurement scenarios and the easily programmability, OpenSketch is one of the typical approach for measurement tasks in software-defined networks although currently it is just applicable for a single network.

Typical approaches for anomaly detection may include [10, 12, 14]. In [10], the authors present a signal analysis of network traffic anomalies in IP flow and SNMP data collected at the border router of a university network. This proposal analyzes four classes of network traffic anomalies: outages, flash crowds, attacks and measurement failures. It uses the Wavelet Analysis method [11] with the input for the analysis platform is a string of Internet traffic measurements. The results show that wavelet filters are quite effective at exposing the details of both ambient and anomalous traffic. The authors in [12] propose an approach called the subspace method to diagnose anomalies. The method can detect, identify and quantify traffic anomalies. The subspace method uses Principal Component Analysis [13] to separate network traffic into a normal component (predictable traffic) and an anomalous component (more noisy traffic). The method uses simple traffic measurements from links and can accurately detect when a volume anomaly is occurring, identify the underlying origin-destination flow which is the source of the anomaly, and accurately estimate the amount of traffic involved in the anomalous origin-destination flow. Evaluations show that the method consistently diagnoses the largest volume anomalies with a very low false alarm rate. In [14], the authors propose a reliable graph-based methodology that compares and combines the results from four anomaly detectors outputs. The approach consists of two main ingredients: a graph-based similarity estimator systematically uncovers the relations between the alarms

reported by the detectors, and a combiner classifies the similar alarms using a combination strategy. Experiments show that synergy between anomaly detectors permits to effectively detect twice more anomalous traffic than the most accurate combined detector, and to reject numerous false positive alarms reported by the detectors.

The above SDN – based measurement and DDoS detection proposal, as well as the conventional anomaly detection approaches such as [13, 15, 17] can be exploited to achieve an efficient solution for anomaly protection in large-scale networks. We explain these issues in the following sections.

III. Requirements of the expected solution and challenges.

In this section, we address the requirements and challenges of our expected solution. The expected solution include following requirements:

- Detect the abnormal traffic in the Internet in an early-stage.
- Quickly react to mitigate the attacks in order to ensure the quality of network services in case of a network attack.
- Be applicable for large-scale networks including a number of distributed networks.

In order to achieve a efficient for anomaly detection and mitigation on large-scalse networks, our research topic, some challenges need to be addressed as follow:

- The similarity between abnormal traffic raised by compromised machines in the attack and the normal traffic in the network.
- Detection techniques to detect the attack in an early stage is difficult because almost every detection method is based on the analysis of receiving traffic, it means we can just realize the anomaly after the attacks were raised for an amount of time.
- Experiments for the proposed solution need a big effort because the architecture need to be deployed in large-scale with a lot of computers, as well as system with a lot of compromised computers to simulate the attacks for the experiments.

IV. Proposed Method.

The idea of software-define networking [2, 3, 4] has been demonstrated to be more and more beneficial and applicable to our current networks. For anomaly detection and reaction, this approach is beneficial since it can support to retrieve networking data for measurement tasks and quickly process instructions for mitigation tasks. Realizing the benefits of SDN, we take advantage of software-defined networking approach in our architectural solution. We propose a software-defined networking architecture for early-stage anomaly detection and mitigation on large-scale production networks.

In the proposed architecture, each network contains a controller and a number of switches. Each switch contains an Interaction Channel that is responsible for communication with the controller, and a Measurement Module that can take on a variety of measurement tasks (e.g. counting a set of flows, measuring various traffic statistics, identifying specific flows,…), and a Mitigation Module that is responsible for protection tasks in the network. The Measurement Module and Mitigation Module are embedded based on the Data Plane of the switch.

Each controller contains a Measurement Library which provide APIs (Application Programming Interfaces) for traffic measurement tasks, a Operation Library which provide APIs for protection (prevention, detection and reaction) tasks, and Anomaly Protection Software which is programmed based on the APIs provided by the libraries to process the measurement and protection tasks for the network. The controller configures/ operates switches in its network or gets the measurement reports form switches by the interaction channels between it and the switches. Detective data can be shared between controllers through their specific channels for more efficient networks protection.

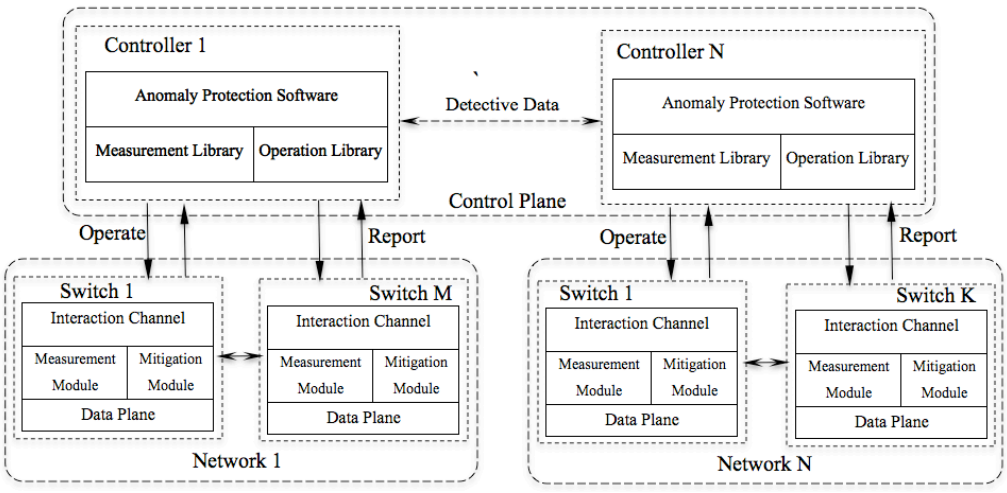

Fig. 1: The early-stage anomaly detection and mitigation architecture.

By the support of this architecture, the measurement results obtained by the Measurement Data Planes in switches will be sent to the controller frequently. The controller analyzes these results to determine if there is any abnormal traffic transmitting in its network. If yes, the controller instructs related switches to react against the abnormal traffic, and share the detective data to the other controllers to warn them about the abnormal traffic. These detective data help the other controllers to

prevent and react against the anomaly traffic more efficient in an early stage.

This architecture can serve a variety of anomaly detection as well as reaction to mitigate the possible attacks. Since the control plane in this architecture can quickly request the networking data from switches/routers and configure them easily by programming, it can analyze and detect the abnormal traffic in an early stage, as well as quickly react in case of network attacks. The collaboration between controllers benefits in more efficient protection for the networks. For instance, in case of an abnormal traffic is detected by a controller of a network, the controller can send the detective data as well as suggested instructions to the other controllers to help these controllers protect their networks better in a very early stage.

As a case study for our solution, an efficient DDoS detection and mitigation method will be build based on the proposed architecture to protect inter-networks from DDoS attack.

V. Implementation Plan

In order to implement the proposed architecture, existing software-defined networking (SDN) implementations will be exploited as follow. We extend the SDN architecture to enable interconnecting multiple controllers in the control plane in our architecture. We will embed a connection module in the controller that is responsible for opening and managing the communication channels to the other controllers for exchanging detective data and protective data. NOX [7], a widely used open-source SDN controller, will be exploited in this task. The instructions set defined by OpenFlow protocol [4] can be used to build the Operation Library in the controller. For the Measurement Library, we exploit the sketch library in OpenSketch controller [5] to support a variety of measurement tasks, including: couting a set of flows, measuring various traffic statistics, identifying specific flows, as well as user-programmed measurement tasks.

For the implementation of switches in our architecture, Open vSwitch [15], a production-quality open source implementation of a distributed virtual multilayer switch that supports multiple networking protocols and standards including OpenFlow protocol [4], will be exploited with extensions for our measurement functionality. Our Measurement Module will be inserted to the switch pipeline. It includes hashing function to reduce the measurement data, classification function to select flows, and counting function to accumulate traffic statistics as the measurement technique in OpenSketch [5]. The Mitigation Module will be build based on the action set defined by OpenFlow protocol [4] to redirect traffic flows for balancing the network traffic, or simply drop the abnormal traffic in case of network attacks.

VI. Evaluation Method

In order to evaluate the efficiency of the proposed method, a real system implemented from the method will be deployed in a large-scale testbed including a number of networks, each network will be controlled by a controller as in the proposed architecture. Another set of computers will be used to simulate network attacks in a variety of scenarios for experiments. The evaluation will be focused on how quickly the system can detect the anomaly traffics (defined by the amount of time the system issue the alert since the time first packet of the attack traffic arrive a network in the system), the accuracy of the detection, and the effectiveness of the applied reaction after an abnormal traffic is detected.

VII. Research plan

	Name	Duration
1	⊟**Semester 1: Research and Refine the proposed preliminary method**	130 days?
2	Review more related works about software-denifed measurement and DDoS detection techniques	50 days?
3	Refine the preliminary proposal into an optimal solution	30 days?
4	Do preliminary experiment to evaluate the feasibility of the solution	30 days?
5	Write the first conference paper presenting the obtained result	20 days?
6	⊟**Semester 2: Implementation and preliminary experimentation**	130 days?
7	Implement the software-defined measurement architecture and the anomaly protection method	80 days?
8	Deploy the module for anomaly protection on a single network and experiment its effectiveness	30 days?
9	Write the second conference paper based on the expermental results	20 days?
10	⊟**Semester 3, 4: Experimentation and writing a jourrnal paper**	250 days?
11	Deploy the solution in a large-scale inter-networks testbed	30 days?
12	Experiment the promptness and accuracy of the detection and the effectiveness of the reaction	160 days?
13	Conclude all the obtained results and write a journal paper	60 days?
14	⊟**Semester 5: Futher experimentation and write the second journal paper**	130 days?
15	Do further experiments together with optimizing the solution if neccessary	70 days?
16	Write a the second journal paper	60 days?
17	⊟**Semester 6: PhD Thesis Defense**	90 days?
18	Conclude all theorical and experimental results to write the PhD thesis	60 days?
19	PhD thesis defense	30 days?

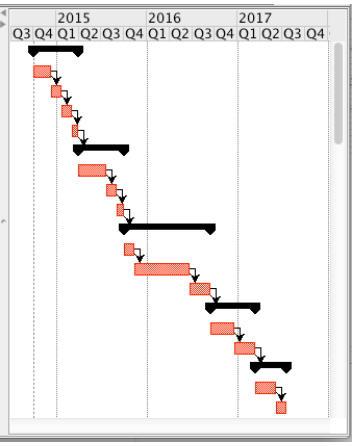

Fig. 2: Research plan for 3-year PhD program.

VIII. Impact

The proposed architecture can be applied to build early stage anomaly detection systems for large-scale networks. These systems can quickly detect the anomaly in the Internet traffic to warn the network administrators. They can also support to quickly react against the attacks by various mitigation techniques controlled by software in the operating center.

The system build on the proposed solution can be used to protect the important networks like the networks of government organizations, research institutes and universities. For instance, the system can be applied to protect government networks in Tokyo against a variety of anomaly traffic and networking attacks to maintain the availability and quality of these network services.

References:

[1] A. Patcha, J. M. Park, "An overview of anomaly detection techniques: Existing solutions and latest technological trends." Journal on Computer Networks, Volume 51, Issue 12, August 2007.

[2] N. McKeown, T. Anderson, H. Balakrishnan, G. Parulkar, L. Peterson, J. Rexford, S. Shenker, and J. Turner, "Openflow: enabling innovation in campus networks," SIGCOMM Comput. Commun. Rev., vol. 38, no. 2, pp. 69–74, 2008.

[3] "Software-Defined Network architecture overview", December 2013. Available: https://www.opennetworking.org/images/stories/downloads/sdn-resources/technical-reports/SDN-architecture-overview-1.0.pdf.

[4] "OpenFlow Switch Specification v1.4," Open Networking Foundation, October 2013. Available: https://www.opennetworking.org/images/stories/downloads/sdn-resources/onf-specifications/openflow/openflow-spec-v1.4.0.pdf.

[5] M. Yu, L. Jose, R. Miao, "Software-defined traffic measurement with OpenSketch", In NSDI 2013.

[6] BRAGA, R. S. ; MOTA, E ; PASSITO, A. "Lightweight DDoS Flooding Attack Detection Using NOX/OpenFlow", In: 35th Annual IEEE Conference on Local Computer Networks, 2010, Denver, Colorado – USA

[7] N. Gude, T. Koponen, J. Pettit, B. Plaff, M. Casado, and N. McKeown, "Nox: Towards an operating system for networks," In *ACM SIGCOMM CCR: Editorial note,* July 2008.

[8] T. Kohonen, "The self-organizing map," *Proceedings of the IEEE,* vol. 78, no. 9, pp. 1464–1480, 1990.

[9] Dhwani Garg, "DDoS mitigation techniques – a survey", In International Journal of Advances in Computer Networks and Its Security.

[10] P. Barford, J. Kline, D. Plonka, A. Ron, "A signal analysis of network traffic anomaly", In Proceedings of ACM SIGCOMM Internet Measurement Workshop, 2002.

[11] D. T. L. Lee, A. Yamamoto, "Wavelet analysis: theory and application", Hewlett Packard Journal, December 1994.

[12] A. Lakhina, M. Crovella, C. Diot, "Diagnosing network-wide traffic anomalies", In ACM SIGCOMM'04, Oregon, USA, 2004.

[13] Principle Component Analysis, http://en.wikipedia.org/wiki/Principal_component_analysis.

[14] R. Fontugne, P. Borgnat, P. Abry, K. Fukuda, "MAWILab: Combining diverse anomaly detectors for automated anomaly labeling and performance benchmarking", In ACM CoNEXT 2010, Philadelphia, USA, 2010.

[15] Open vSwitch. http://openvswitch.org