Experimental Evaluation of Publish/Subscribe-based Spatio-Temporal Contents Management on Geo-Centric Information Platform

著者	Nagashima Kaoru, Taenaka Yuzo, Nagata Akira, Nakamura Katsuichi, Tamura Hitomi, Tsukamoto Kazuya
journal or publication title	Advances in Intelligent Systems and Computing
volume	1036
page range	396-405
year	2019-08-15
URL	http://hdl.handle.net/10228/00007846

doi: https://doi.org/10.1007/978-3-030-29029-0_37

Experimental Evaluation of Publish/Subscribe-based Spatio-Temporal Contents Management on Geo-Centric Information Platform

Kaoru Nagashima, Yuzo Taenaka, Akira Nagata,
Katsuichi Nakamura, Hitomi Tamura, Kazuya Tsukamoto

Abstract Cross-domain data fusion is becoming a key driver to growth of the numerous and diverse applications in IoT era. Nevertheless, IoT data obtained by individual devices are blindly transmitted to cloud servers. We here focus on that the IoT data which are suitable for cross-domain data fusion, tend to be generated in the proximity, and thus propose a Geo-Centric Information Platform (GCIP) for the management of Spatio-Temporal Contents (STCs) generated through the cross-domain data fusion. GCIP enables to keep STCs near the users (at an edge server). In this paper, we practically examine the fundamental functions of the GCIP from two aspects: (1) Geo-location aware data collection and (2) Publish/Subscribe-based STC production. Furthermore, we implement a proof-of-concepts (PoC) of GCIP and conduct experiments on a real IPv6 network built on our campus network. In this experiment, we showed that multiple types of IoT data generated in the proximity can be collected on the edge server and then a STC can be produced by exploiting the collected IoT data. Moreover, we demonstrated that the Publish/Subscribe model has a potential to be effective for STC management.

Kaoru Nagashima
Kyushu Institute of Technology , Japan e-mail: nagashima@infonet.cse.kyutech.ac.jp

Yuzo Taenaka
Nara Institute of Science and Technology, Japan e-mail: yuzo@is.naist.jp

Akira Nagata
iD Corporation, Japan e-mail: a-nagata@intelligent-design.co.jp

Katsuichi Nakamura
iD Corporation, Japan e-mail: k-nakamura@intelligent-design.co.jp

Hitomi Tamura
Fukuoka Institute of Technology, Japan e-mail: h-tamura@fit.ac.jp

Kazuya Tsukamoto
Kyushu Institute of Technology, Japan e-mail: tsukamoto@cse.kyutech.ac.jp

1 Introduction

With the rapid growth of both sensor devices and wireless technologies, various type of things can be connected to the Internet, that is the IoT. In the IoT era, since the combination of various "things' data" could bring us new and undiscovered contents, the cross-domain (horizontal-domain) IoT data fusion attracts much attention. However, IoT data transmitted from the numerous IoT devices are generally enforced to be gathered in cloud servers and be used for a specific IoT service, which is widely referred to as the vertical-domain.

In this paper, we focus on that some type of IoT data suitable for the cross-domain data fusion are generated in the geographical proximity. Along this line, we propose a new information platform, Geo-Centric Information Platform (GCIP) in which IoT data are kept in physically close edge servers (i.e., in the physical proximity) and then new contents can be produced as a result of analysis and processing of the data. Note that we define the produced contents, which are worth for users for the limited duration and at the limited location, as the Spatio-Temporal Contents (STCs).

First, we introduce a conceptual model for GCIP and then implement a proof-of-concept (PoC) with the fundamental functions of geo-location aware data collection and Publish/Subscribe-based STC production into a real environment. Next, we conduct the experiments to show the effectiveness of the GCIP by using IPv6 campus network. More specifically, we examine the feasibility of not only geo-location aware data collection but also Publish/Subscribe based STC production.

Rest of this paper is organized as follows, We first review the existing studies in Section 2. Then, the conceptual design of our proposed method is described in Section 3 and the demonstration environment is described in Section 4. We show and discuss the experimental results in Section 5, and finally Section 6 concludes this paper.

2 Related Work

In this section, we review the existing studies focusing on geo-location based network control and IoT data processing. In reference [1], MQTT has been extended to handle location information. Although all IoT devices are assumed to have its geo-location by using GPS, it is significantly difficult for IoT devices to load GPS because they are based on cheap, small, and low-powered design. As another example, the reference [2] discusses data processing on the premise of location information. Hence, if cloud server produces the contents from the collected IoT data, the server needs to be aware of location. As one of location-aware method, GCIP adopts network level approach. Thus, GCIP cloud collect IoT data in the geographical proximity by changing both of routing tables and identifier of the IoT devices.

In reference [3], various IoT data such as social data, media data, etc. are collected and analyzed to extract beneficial information with unique features in each region. However, there is a problem in scalability and performance because all of the IoT data is processed on a specific cloud server. The references [5] and [2] have mentioned several use cases in which various types of data are processed, i.e., cross-domain data fusion. However, since the processing are performed on the cloud server, the amount of traffic between the users and servers is significantly increased, thereby increasing not only the latency between the users and server but also packet losses. In contrast, GCIP produces STC on the edge server(s) , which is relatively geographically closer to the users.

Reference [4] widely summarizes the existing studies focusing on the contents search method. However, none of existing studies tries to find the realtime/on-demand contents dynamically generated based on the collected IoT data. On the other hand, GCIP dynamically generates dynamic contents by considering the temporal and spatial characteristics.

From these points, we can say that GCIP provides new aspects of (1) geolocation-aware communication to IoT devices and (2) on-demand Spatio-Temporal content production based on the geographical proximity.

3 Geo-Centric Information Platform (GCIP)

This section first describes the GCIP design concepts. After that, we describe PoC design in terms of data collection and STC production, respectively.

3.1 Conceptual design of GCIP

Majority of IoT devices are deployed and dedicated for specific services. That is, device and service are tightly associated. Moreover, each of network is individually managed by network operator without the consideration of the physical location. These situations make the cross-domain data fusion quite difficult. However, as stated in the introduction, cross-domain IoT data fusion is becoming a key driver to accelerate the production of new and beneficial services in the IoT era. To achieve cross-domain data fusion, we here focus on that the data suitable for the cross-domain data fusion tend to be generated in the geographical proximity. Thus, we introduce a concept of Geo-Centric Information Platform (GCIP), which allows us to collect, process the IoT data with consideration of geographic location where they generated. As a result, GCIP efficiently produces various types of STCs.

Fig. 1 shows the conceptual design of GCIP. Procedures of GCIP consist from the following 4 steps. First, the data transmitted from IoT devices within some area are replicated at the intermediate router(s), irrespective of the network type (Step 1). Then, the replicated data are forwarded to a proximity edge server(s) having the analysis/process functions (Step 2). The server(s) generates STCs as a result of processing with the collected data (Step 3). Finally, the server sends the produced

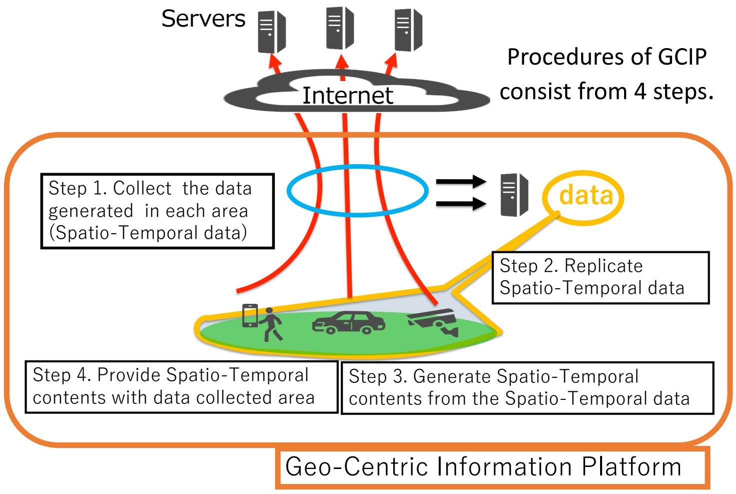


Fig. 1 Conceptual design of GCIP

STC in response to users (Step 4). In this way, the proposed GCIP has a potential to be a fundamental infrastructure for the IoT era.

3.2 PoC design for geo-location aware IoT data collection

There are two requirements to collect data based on geographical location.

(i) Geo-location area for collecting IoT data can be identified by the intermediate routers.
(ii) Modification for geo-location aware communication has a backward compatibility.

We propose Physical Location-Aware Communication (PLAC) method that can satisfy the requirements described above. To achieve the first requirement, we design the hierarchical mesh-structured network topology, as shown in Fig. 2. The geographical space is divided into hierarchical meshes based on latitude and longitude lines, and each of which has a unique mesh code [1]. Size of the minimum mesh area is a square area of 39 m each side, by extending rule of Open-i area). As shown in Fig. 2, the length of the mesh code increases as the decrease in the size of geographical area decreases. Since a unique mesh code is allocated to each of intermediate router, the router can identify its own belonging geographical area and handle all of IoT data containing the same mesh code. That is, the first requirement can be satisfied.

[1] The mesh code is basically followed by the NTT Docomo open-i area[6].

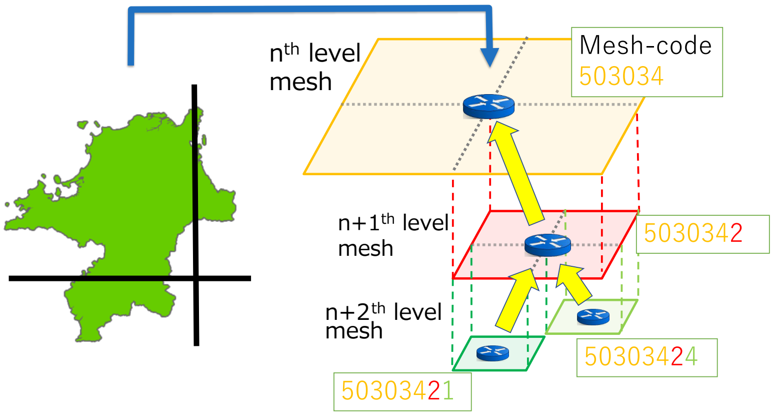


Fig. 2 Example of hierarchical mesh area and mesh code

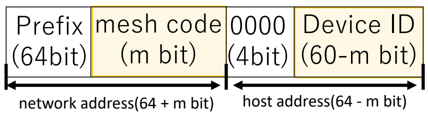


Fig. 3 Physical location-aware address

Next, to achieve PLAC with backward compatibility with the traditional Internet technology, we already proposed the geo-location aware address format [7]. As shown in Fig. 3, the mesh code with variable length depending on the mesh level is embedded with the IPv6 address. More specifically, it is located in the beginning of the lower 64 bits and used as a part of network address. Moreover, to maintain the backward compatibility with the Internet, we do not modify the part of network prefix (the upper 64 bits). From this, the extended network address (prefix + mesh code) can express the certain geographical area and the principle of the traditional IP routing process (i.e., longest match routing) can be used for PLAC, thanks to the backward compatibility. In this way, PLAC can be effective for IoT data collection at the routers on the mesh-structured network. Note that we hereinafter define intermediate routers on the mesh-structured network as ”mesh router”.

3.3 PoC design for Publish/Subscribe-based STC production

In this section, we consider how various STCs, which are worth for users in the IoT area, are produced on the proposed GCIP. First, we need to consider that the STC is produced as a result of analyzing/processing to the collected IoT data; Therefore, some dedicated server(s) is necessity for IoT data analysis and processing, in addition to the mesh router. Furthermore, to flexibly and efficiently produce STCs on the server, the following three requirements should be satisfied.

Requirement 1	Type of IoT data should be uniquely identified.
Requirement 2	Collected IoT data should be flexibly stored and processed.
Requirement 3	To produce the STC with accurate geo-location information, the edge server needs to know where the IoT data was generated.

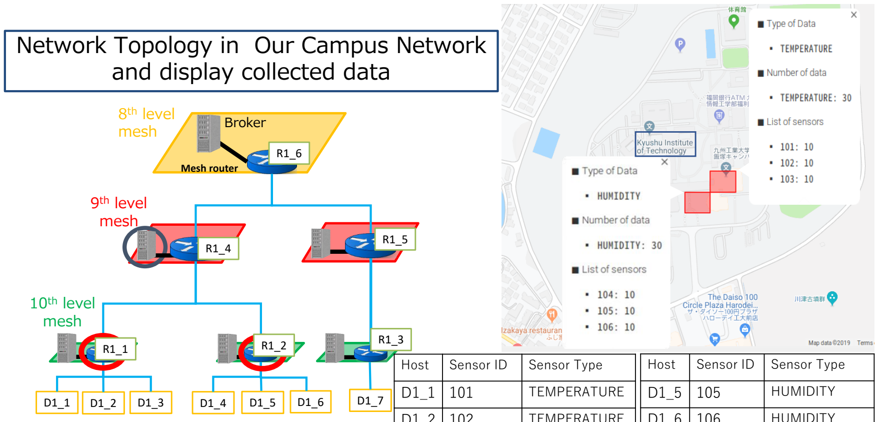
To satisfy these requirements, we focus on the Publish/subscribe communication model, which is a new concept for decoupling data collection and distribution (interactive communication is not mandatory). By employing the topic-based Publish/Subscribe model, requirement 1 can be achieved. Moreover, if the processing functions are implemented in the broker, requirement 2 also can be achieved. As for requirement 3, if the location information is included in the application data, it is possible. However, if the mesh router collecting IoT data from some geographical area serves as the publisher, we cannot touch the application data directly. Therefore, we need to achieve requirement 3, even without modification of the application data at all.

4 Demonstration Environment

In this section, we conduct experiments to show the effectiveness of the proposed GCIP from the view point of the following aspects.

- Can a mesh router on the mesh network topology collect various type of IoT data by using PLAC method?
- Can a mesh router collect various types of IoT data generated in the geographical proximity but transmitted from physical different network ?
- Can a mesh router (broker) produce the STCs by using the collected IoT data?

For the first experimentation, we used kyutech campus network. Fig. 4 shows the mesh network topology. We placed 6 mesh routers, which are denoted as R1_1 , R1_2, and so on, for each mesh level. In contrast, 7 elements like D1_1, D1_2, and so on are IoT devices with various types of sensor, which are developed by Raspberry Pi. D1_1, D1_2 and D1_3 are located in the 10th level mesh network whose mesh router is R1_1. Similarly, D1_4, D1_5 and D1_6 are located in 10th level mesh network whose mesh router is R1_2. And, D1_7 are located in one of 10th level mesh network belonging to different 9th level mesh network, which belong to R1_3.



Host	Sensor ID	Sensor Type	Host	Sensor ID	Sensor Type
D1_1	101	TEMPERATURE	D1_5	105	HUMIDITY
D1_2	102	TEMPERATURE	D1_6	106	HUMIDITY
D1_3	103	TEMPERATURE	D1_7	107	ILLUMINATION
D1_4	104	HUMIDITY			

Fig. 4 Demonstration Environment for examining heterogeneous IoT data collection

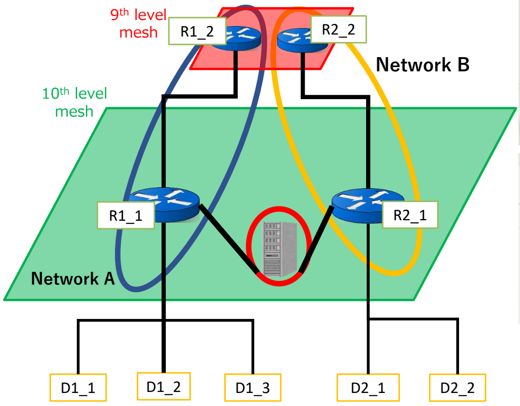
These IoT devices periodically transmit sensor data obtained from sensor to the cloud server on the Internet every 5 minutes. Since the transmission path is determined by the PLAC method, intermediate mesh routers including 9th level and 10th level mesh routers can collect the IoT data while associating with the geographical location. Then, the router replicates the collected data and transmits them to the broker. In this experiment, we confirm whether the R1_8 and R1_9 collect different types of IoT data by using PLAC method.

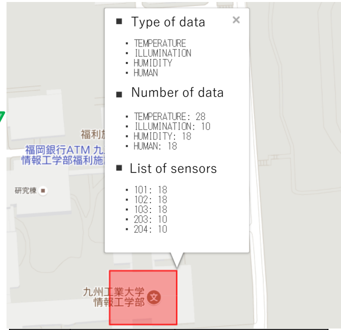
Next, we examine that PLAC method can collect the IoT data even when the data are transmitted through the different networks (those subnets/prefixes are different). Fig. 5 shows the topology used for experimentation. We particularly focus on 10th level mesh network, which settings are the same on Network A, but we additionally assume Network B has a coverage in the same 10th level mesh area. R1_1 and R2_1 denote mesh routers for each of mesh network. In contrast, D1_1 to D1_3 and D2_1 to D2_2 denote the IoT devices with various types of sensor, which are developed by Raspberry Pi. IoT devices periodically (every 5 minutes) transmit data as in the same experimentation. In this experiment, we investigate that GCIP collects the IoT data generated in the same proximity but transmitted over different networks.

For the final experiment, we examine the feasibility on producing a STC using heterogeneous IoT data in a specific area. We here use the network topology of first experimentation again and employ the discomfort index [9] as the STC. The discomfort index (DI) is calculated as follows:

$$DI = T - (0.55 - 0.0055RH)(T - 14.5) \quad (1)$$

where T indicates the mean value of air temperature in C and RH indicates the 5 minutes average relative humidity (Since the IoT data are received by the broker at





Host	Sensor ID	Sensor Type
D1_1	101	TEMPERATURE
D1_2	102	HUMIDITY
D1_3	103	HUMAN
D2_1	203	TEMPERATURE
D2_2	204	ILLUMINATION

Fig. 5 Demonstration Environment for examining IoT data collection over multiple networks

the intervals of 5 minutes, the broker calculates the DI every 5 minutes based on the information of temperature and humidity. Note that the distribution method of the produced STC is out of scope in this paper. So, we use typical Web API (application) to demonstrate the detailed information of STC generated in the broker.

5 Result and Discussion

5.1 Experimental results

The data collected by the routers R1_1 and R2_1 are shown on the right side of Fig.4. We also illustrate the detailed information of the sensor number and the sensed values, which are collected by mesh routers, in the right map. Fig. 4 shows that the mesh routers can collect the IoT data, irrespective of the data type because R1_1 can collected temperature data from D1_1, D1_2 and D1_3 and R1_2 can collected humidity data from D1_4, D1_5 and D1_6. Therefore, this result shows that various types of IoT data could be collected thanks to both the by PLAC address and its routing function.

Fig. 5 shows the results of second experiment. From Fig. 5, we can see that broker can collect data over different networks because R1_1 transmits data from D1_1, D1_2 and D1_3 , while R2_1 transmit data from D2_1 and D2_2. As we can find from the sensor ID, we can see that the data passing through R1_1 and R2_1 is can be successfully collected. From this result, we can say that IoT data can be collected based on geographical proximity over different networks.

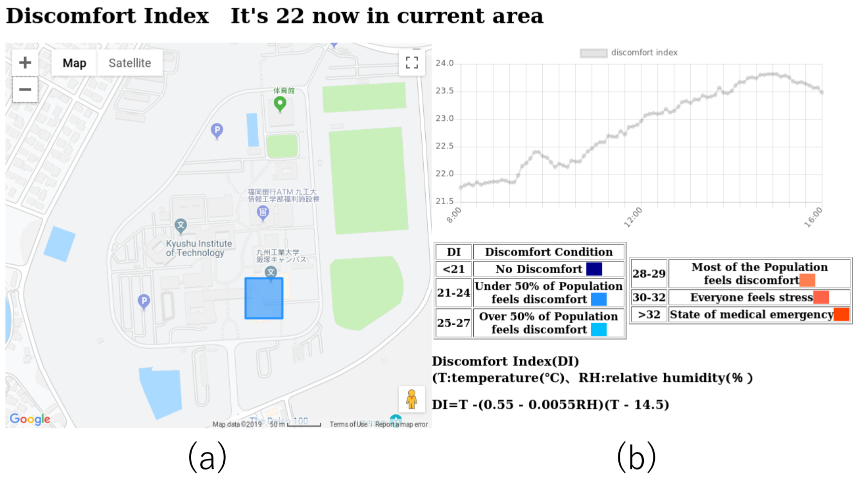


Fig. 6 Produced STC (left: spatial view, right: temporal view)

We describes the third experiment. Fig. 6 shows content generated by the 9th mesh broker. This result can be viewed in the web application. Fig. 6 (a) shows a spatial area in where the produced STC is beneficial for the users. The broker (mesh server) inherently identifies the geo-location area by its own IPv6 address including mesh code. Moreover, Fig. 6 (b) shows a time series variation of STC value. Since the collected data are stored in the broker (mesh server), we can naturally produce the time series variance of the STC.

In this way, the broker can identify the type of IoT data by the topic ID defined by introducing the topic-based Publish/Subscribe model (i.e., Apache Kafka) (requirement 1). Furthermore, we found that the broker can store and process the IoT data collected by using PLAC method (requirement 2). Finally, as for requirement 3, the current implementation indicates the STC for 9th level mesh only. As a result, the broker naturally produces the STC with time series variance.

5.2 Discussion

First, we discuss the data collection performance of the proposed GCIP. The first and second experimentation show that the proposed PLAC mechanism can handle the IoT data based on the geographical location, thereby achieving the geo-location aware data collection by the router on the mesh network. More specifically, we also show that the IoT data can be naturally collected at mesh routers, independent of the access network.

Next, we discuss the IoT data processing performance of the proposed GCIP. We have already explained that requirement 1 and 2 are completely satisfied, but the requirement 3 is not completely satisfied. The reason is that the spatial information (mesh code) of the IoT data is included in the IP header as the part of the source IP address. The application layer of the broker cannot obtain the geo-location information of narrow area (such as 10the level mesh). However, since the Apache Kafka handles data at the application layer, the IP address is discarded and thus the mesh code of the IoT devices is not passed to the Apache Kafka. In such case, the broker illustrates the spatial area by using its own mesh code. Therefore, if a broker is always located in all of the 10th level mesh, the problem can be solved. However, since the assumption is not realistic, some kind of cross-layer mechanism for informing the mesh code in network layer to the Apache Kafka in application layer is needed.

Through experiments, the usefulness of GCIP was actually confirmed. First, GCIP can collect IoT data based on geography location by exploiting PLAC mechanism. When the cloud server tries to achieve the same thing, the geo-location information should be added to all of application packets (payload). That is, modification of application is mandatory. Second, GCIP naturally supports the store of the received IoT data on the mesh server. Therefore, the mesh server flexibly (e.g, temporally) process the stored data, thereby producing the STC effectively.

6 Conclusion

In this paper, we proposed the Geo-Centric Information Platform (GCIP) that can manage the Spatio-Temporal Contents (STCs) produced by the cross-domain data fusion. We practically examined the effectiveness of the GCIP in terms of (1) geo-location aware data collection and (2) Publish/Subscribe-based STC production through the experiments on campus IPv6 network. As a result, we confirmed the following outcomes.

- By using PLAC mechanism, IoT data transmitted from diverse devices can be collected at the edge router based on the geographical area independent of their access networks..
- STC can be successfully produced by exploiting the existing implementation (Apache Kafka) employing Publish/Subscribe model. Note that processing function should be developed on the implementation.

These results demonstrated the feasibility of the fundamental functions in the proposed GCIP. However, although the current implementation assumes that each of mesh network has one mesh router, it is difficult in the real environment. Therefore, we will extend the GCIP to adapt to the practical environment in the near future.

Acknowledgements This work was supported by JSPS KAKENHI Grant Number JP18H03234 and the Commissioned Research of National Institute of Information and Communications Technology （NICT).

References

1. Robert Bryce and Gautam Srivastava, “The Addition of Geolocation to Sensor Networks, ” 13th International Conference on Software Technologies, pp.762-768, 2018.
2. Keiichi Yasumoto, Hirozumi Yamaguchi, Hiroshi Shigeno, “Survey of Real-time Processing Technologies of IoT Data Streams, ” Journal of Information Processing, Vol.24, No.2, March 2016.
3. Gonzales E, Ong B T, and Zettsu, K, “Searching inter-disciplinary scientic big data based on latent correlation analysis, ” in Proc. of Workshop on Big Data and Society (in conjunction with IEEE BigData 2013), pp. 9-12, October 2013.
4. Pattar, S.Buyya, R.Venugopal, K.R.Iyengar, S.S.Patnaik, L.M. “Searching for the IoT Resources:Fundamentals, Requirements, Comprehensive Review, and Future Directions, ” IEEE Communications Surveys & Tutorials, Vol.20, pp. 2101-2132, April 2018.
5. Ala Al-Fuqaha, Mohsen Guizani, Mehdi Mohammadi, Mohammed Aledhari, Moussa Ayyash, “Internet of Things: A Survey on Enabling Technologies, Protocols and Applications, ” IEEE Communications Surveys & Tutorials, Vol. 17, no. 4, pp. 2347-2376, June 2015.
6. NTT docomo : Open iArea guideline, https://www.nttdocomo.co.jp/binary/pdf/service/developer/make/content/iarea/domestic/open-iarea.pdf (in Japanese)
7. Hitomi Tamura, “Program for determining ip address on the basis of positional information, device and method, ” JP Patent 6074829, 2017-01-20.
8. Apache Kafka, https://kafka.apache.org/
9. Md Sarwar Jahan Talukdar, Md. Shahadat Hossen and Abdul Baten, “Trends of Outdoor Thermal Discomfort in Mymensingh: an Application of Thoms’ Discomfort index, ” Journal of Environmental Science and Natural Resources, Vol 10, No 2 , pp.151-156, November 2018.