ISBN: 978-989-8533-72-2 © 2017

CONCEPT ORIENTED DATA AGGREGATION IN EDGE COMPUTING

Ryuichiro Hiranuma, Noriko Matsumoto and Norihiko Yoshida
Department of Information and Computer Sciences, Saitama University
255 Shimo-Ohkubo, Saitama 338-8570, Japan

ABSTRACT

Edge Computing has been proposed as a novel architecture that changes traditional communication between Cloud and Internet of things (IoT). In Edge Computing, intermediate processing infrastructure called Cloudlet is interposed on the network between the Cloud and IoT. Cloudlets have arithmetic processing functions and storage, and bear part of the Cloud data processing. Edge Computing has made it possible to process substantial amounts of data from countless IoT devices. However, Edge Computing cannot cope with heterogeneity of the data. In this paper, we propose a method corresponding to heterogeneity of data by adding idea of concept newly to Edge Computing.

KEYWORDS

Edge Computing, Cloud Computing, Internet of Things, Big Data

1. INTRODUCTION

A new network model called Edge Computing [Shi, 2016] has been proposed. One usage of Edge Computing is to help analysis of Big Data [Jagadish, 2014] collected from IoT [Whitmore, 2015]. The model has a three-layer structure, including various networks formed by IoT devices in the lower layer, intermediate processing base of Big Data constituted by Cloudlets in the middle layer, final analysis base of Big Data composed of Clouds in the upper layer. The Cloudlet has arithmetic processing functions and storage, and it exists closer to end devices than the Cloud. With this model, Big Data generated by IoT can be locally distributed without concentrating on the Cloud.

However, there are problems that cannot be solved by merely applying Edge Computing. One of the problems is heterogeneity of data. Currently, diversity of systems and data sources is an issue in Big Data analysis. Simply arranging Cloudlets do not change the complexity of processing and operation. Therefore, in this paper, we propose a new Edge Computing architecture that can cope with heterogeneity of data. In the proposed method, we introduce the idea of concept newly to Edge Computing, and solve the problem of heterogeneity of data. Concept is a code representing various things such as the type, form, possession of data. The new architecture classifies/manages data based on the concept. We can control a wide variety of data collection and realize Edge Computing which is more suitable for Big Data analysis.

This paper is organized as follows: we explain previous research introducing the idea of concept into the sensor network in Section 2. In Section 3, we describe our proposed method. Subsequently, the proposed method is simulated assuming a simple situation in Section 4. Finally, Section 5 contains some concluding remarks and future work.

2. PREVIOUS RESEARCH

In this section, we introduce previous research [Ganz, 2011] incorporating the idea of concept into the sensor network.

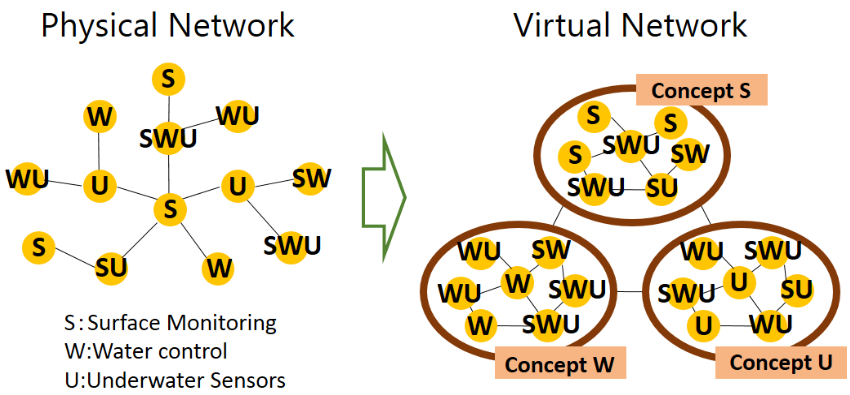

Figure 1. Virtualization of Sensor Network Using Concept

In the previous research, as shown in Figure 1, a network is virtually split into semantic similar groups based on the concept information of the network. The overlay networks are constructed according to the structure of concepts in a model and form a logical network. Concept is a code representing various things such as the type, form, possession of data.

In Figure 1, measurement of Surface Monitoring, Water Control, Underwater Sensors is assumed in environmental monitoring. Sensors that can handle multiple concepts are grouped into their respective concepts, so the number of nodes of the virtual network is larger than that of the physical network.

3. PROPOSED METHOD

As described in Section 1, this research aims to solve the problem of heterogeneity of data. Therefore, we incorporate the idea of the previous study introduced in Section 2 into the model of data aggregation explained in Section 1, Proposed method makes it possible to efficiently process a wide variety of data on Edge Computing.

In the proposed method, as in the conventional Edge Computing, it forms a three-layer structure consisting of IoT devices, Cloudlets, and a Cloud. In this section, we explain the connection between each layer, and advantages of the proposed method.

3.1 Connection between Iot Device Layer and Cloudlet Layer

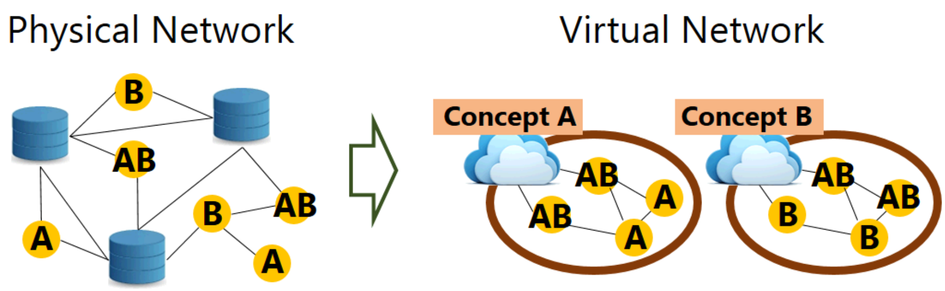

Figure 2. Virtualization of Edge Computing Using Concept

In the proposed method, data collection based on concept is realized based on the idea of the previous research. In the physical network of Figure 2, the IoT device is connected to any one of physical servers. At this time, the IoT device may be directly connected to a physical server or indirectly being relayed by another IoT device. In this example, since two types of data are collected, the concept also include two types: A and B respectively. Three types of IoT devices can be considered: a device capable of measuring only the concept of A, a device capable of measuring only the concept of B, and a device capable of measuring the concepts of both A and B. In the proposed method, this physical network is virtualized based on the concept.

In the virtual network of Figure 2, IoT devices are logically grouped based on the concept. An IoT device capable of measuring the concepts of both A and B participates in both groups. Each group is managed by a dedicated Cloudlet which performs specific processing for that concept. A logical network structure of each group is constructed and managed by a Cloudlet which grasps the whole. IoT devices transmit the data to the Cloudlet based on the overlay network. Whichever IoT device sends data to any server on the physical network, it logically reaches the target Cloudlet. Therefore, IoT devices do not have to care about the physical server and relay nodes.

Cloudlets specialized for each concept are logically created on multiple physical servers. Cloudlets specialized for each concept is created. In addition, Cloudlets are added/deleted dynamically. Therefore, the proposed method can flexibly reconfigure the virtual network even if the type of data to be collected dynamically changes. So, the proposed method can cope with the mobility of IoT devices.

It is likely that IoT devices such as vehicles and wearable devices move. Therefore, the physical server to which the IoT device is connected in the real may change constantly. However, applying the proposed method does not change the logical structure even if the physical structure changes.

In IoT, various kinds of communication protocols conscious of power saving and narrowband are developed, and IoT gateway converts these communication protocols to IP (Internet Protocol) and connects to the Internet. In the proposed method, the Cloudlet also serves as the IoT gateway, and the IoT device can transmit data to the Cloudlet no matter what communication protocol is used.

3.2 Connection between Cloudlet Layer and Cloud Layer

In the proposed method, the Cloud grasping the whole manages the concept. In short, the Cloud can send feedback to the Cloudlet on the edge side of the network.

For example, when a new concept occurs in a certain area, the Cloud orders the creation of a corresponding new Cloudlet. Also, when the demand for a specific concept rises sharply in a certain area, the Cloud orders the Cloudlet to strengthen the process concerning that concept. In this way, feedback control is performed according to the area where the Cloudlet is installed in the proposed method. Since the configuration of the Cloudlet is optimized according to the processing suited to the area, more efficient data collection can be performed. Also, when the demand in the area changes, the Cloud can grasp it and return the feedback, so that the Cloudlet can be readjusted.

3.3 Advantages of Proposed Method

The following advantages can be obtained by using the proposed method.

1) Dealing with data heterogeneity
Since the data is structured by the concept, it is possible for the Cloud to control a wide variety of data collection and to manage it efficiently.

2) Free of IoT devices from physical network
As the network is virtualized by the Cloudlet, the IoT device can transmit data without being conscious of the physical network. Furthermore, mobility of IoT devices can be accommodated.

3) Freedom of communication protocol
As the Cloudlet also serves as an IoT gateway, the IoT device can select any communication protocol

4) Optimization by feedback control
Since the Cloud sends feedback to the Cloudlet, it is possible to realize optimal data collection according to each region.

5) Advanced processing of specialized Cloudlets
Dedicated Cloudlets are created for each concept, saving platform resources and incorporating more advanced processing.

With the above advantages, the proposed method can realize more efficient, more extensible and more flexible data aggregation on Edge Computing.

4. SIMULATION

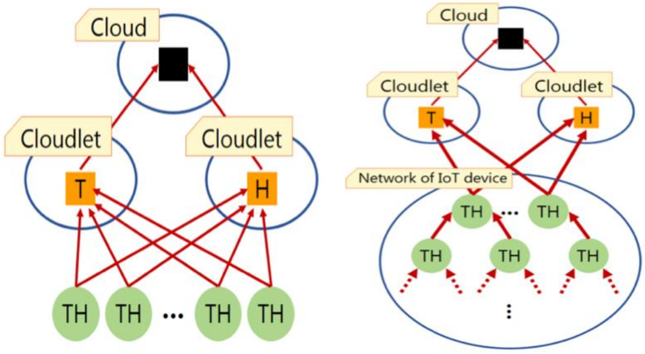

Figure 3. Model Assumed by Simulation with One-Hop to the Cloudlet (Left), and with Multi-Hop (Right)

In this section, we simulate the proposed method assuming two simple situations. The first situation on the left side of Figure 3 is when data is sent from the IoT device to the Cloudlet in one hop. In such a case, wearable devices are assumed in which the devices hardly communicate with each other. The second situation on the right side of Figure 3 is when the IoT device forms a network and sends the aggregated data to the Cloudlet. In such a case, it occurs frequently when Context aware computing [Perera, 2014] is considered by IoT, for example, environmental measurement in a sensor network is assumed.

In the simulation, it is assumed that environmental measurement is performed using two concepts. The following describes how each concept is handled by the IoT device and the Cloudlet.

Concept T: Temperature

And concatenates the received data.

Example: When {10.0} and {20.0} data are given, it becomes {10.0 20.0}.

Concept H: Humidity

And calculates the average of the received data.

Example: When {10.0} and {20.0} data are given, it becomes {15.0}.

H is aggregated in size in IoT device network, T does not change eventually.

The Cloud and each Cloudlet assume that cost n is required to process data of length n. For example, cost 2 when processing data of {10.0 20.0}, cost 1 when processing data of {15.0}. If this is expressed by order notation, the data length after the processing of the concept T is expressed as $O(n)$ where the number of data is n, and in the concept H it is $O(1)$. In the simulation, it is assumed that 10000 IoT devices transmit data 50 times. Also, it is random as to whether the IoT device actually transmits data each time.

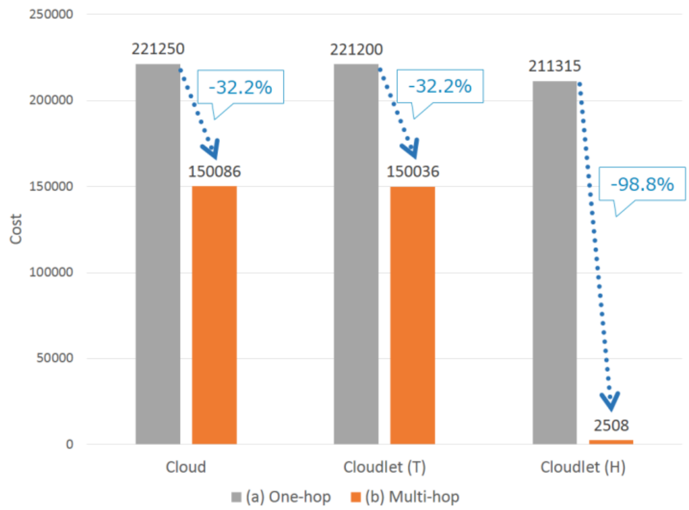

Figure 4. Result of Simulation

As a result of the simulation, the cumulative cost of the Cloud and each Cloudlet is shown in Figure 4. First of all, in all results, cost is lower in processing IoT device networks. Looking at the decrease rate, the Cloud and the Cloudlet of the concept T are -32%, whereas the Cloudlet of the concept H is -98.8%. Based on the simulation results, it turned out that there is a considerable difference in the total cost due to the concept.

Also, Basically, the traffic that divides the original cloud traffic by the number of concept flows to the network of each cloudlet. However, if there is bias in the number of packets handled by each concept, the network traffic also changes accordingly. And, network traffic to the cloud depends on how much data can be aggregated in each cloudlet.

From the simulation, we could discover the possibility of complementary use of context aware computing of IoT and proposed method. Thus, by combining the proposed method with other methods, it is possible to obtain not only the advantages shown in Section 3 but also the cost advantages.

5. CONCLUSION AND FUTURE WORK

In the proposed method, we assumed a three-layer architecture such as networks of IoT devices in the lower layer, Cloudlets in the middle layer, and a Cloud in the upper layer like the conventional Edge Computing. And, we devised connection of each layer respectively. Specifically, we designed virtualization of IoT devices based on concepts and feedback control from Cloud to Cloudlets.

By using the proposed method, various advantages were obtained besides coping with heterogeneity of data. We also found that the load of Cloud and Cloudlet can be further reduced by combining the proposed method and other methods by simulation.

Finally, the direction of how the proposed method will develop in the future is shown below,

1) Multiple concept layers

In the proposed method, virtualization and feedback control were performed based on only the concept of one layer. By preparing the concept of multiple layers, it is possible to realize finer control at the edges of the network.

2) Concept-based routing

If routing based on the concept can be realized, it is possible to perform high-speed transmission of data that does not require name resolution such as Domain Name System (DNS) and flexible network construction. Currently, Content-Centric Networking (CCN) [Ahlgren, 2012] which routes based on data name without using IP address has been proposed, so a combination of CCN and proposed method can be considered.

3) Software defined concept mapping

Introduction of Software Defined Network (SDN) [Jain, 2013] is considered as a means for realizing efficiently feedback control by the Cloud. SDN allows the Cloud to dynamically control the placement of concepts in the Cloudlet layer.

REFERENCES

Ahlgren, B. et al, 2012. A survey of information-centric networking. *IEEE Communications Magazine*, Vol. 50, Issue. 7.

Ganz, F. et al, 2011. A mediated gossiping mechanism for large-scale sensor networks. *GLOBECOM Workshops (GC Wkshps) IEEE.*

Jagadish, H. V. et al, 2014. Big data and its technical challenges. *Communications of the ACM*, Vol. 57, Issue. 7, pp. 86-94.

Jain, R. and Paul, S., 2013. Network virtualization and software defined networking for cloud computing: a survey. *IEEE Communications Magazine*, Vol. 51, Issue. 11, pp. 24-31.

Perera, C. et al, 2014. Context aware computing for the internet of things: A survey. *IEEE Communications Surveys & Tutorials*, Vol. 16, Issue. 1, pp. 414-454.

Shi, W. et al, 2016. Edge computing: Vision and challenges. *IEEE Internet of Things Journal*, Vol. 3, Issue. 5, pp. 637-646.

Whitmore, A. et al, 2015. The Internet of Things—A survey of topics and trends. *Information Systems Frontiers*, Vol. 17, Issue. 2, pp. 261-274.