MQTT-Topics Management System for Sharing of Open Data

Nasi Tantitharanukul, Kitisak Osathanunkul, Kittikorn Hantrakul, Part Pramokchon, and Paween Khoenkaw
Intelligent Innovation Laboratory (INTNIN)
Department of Computer Science, Faculty of Science, Maejo University, Chiang Mai 50290, Thailand
Email: *nasi@gmaejo.mju.ac.th, decoze@gmail.com, {kittikor, part, paween_k}@gmaejo.mju.ac.th*,

***Abstract*—At present, the sharing of open data in any city is importance. The shared data, e.g., flood monitoring data, earthquake data, or traffic data, can be used to improve the way of human life. MQTT, the lightweight messaging protocol based on the publish-subscribe concept, is the most popular protocols used to share these data. There are two elements of MQTT messaging for sharing of any data; topic and message. The first one is used to define which the subscriber are received which data and the data that we need to share are presented in the message element. To share the data with MQTT protocol, MQTT-Topic Naming Criteria of Open Data for Smart Cities (MTNC) was proposed in our previous work. In this paper, we propose MQTT-Topics Management System (MTMS) to manage the topics of the MQTT messages shared by any publisher based on MTNC. The experimental result shows that the proposed system get the high average score of usability, design and the benefits.**

***Keywords*—MQTT, Open Data, Topics Management System**

I. Introduction

Since the coming of an Internet of Things (IoTs) Technology, it allows most electronic devices or human to collect, process or monitoring data of everything in the world in real-time. Moreover, It can make a city smarter because many systems can collaborate work together via the internet communications to respond to some situations automatically. For example, the embedded systems that they can collect some data e.g., flood monitoring data, earthquake data, or traffic data, and share these data with other systems over the Internet. For sharing these data via the internet, MQTT protocol is the most popular protocols. It is because of it is the lightweight protocol, so, MQTT suit for the system that needs to share many data in real-time. There are two elements of the message send via MQTT protocol, they include topic and message. The data will be sent from a publisher to the subscribers who subscribe the topic that is published by the publisher. Meanwhile, the message element is the data that the publisher needs to share.

To share any open data via MQTT protocol, the subscribers need to know what are the topics of data that they should subscribe from the broker. The publisher can defined any topic with one or more topic levels where each topic level is separated by a forward slash, e.g., $thailand/humidity$ or $thailand/bangkok/traffic$. However, there are no standards of MQTT-Topic naming of the open data for smart cities. So, the topic can be variously defined by each publisher. Thus, our previous work, [11], proposes the standard of MQTT-Topic Naming (MTNC) to be the criteria of topic naming for sharing any open data from the devices in a smart city over the MQTT protocol.

In this paper, we propose MQTT-Topics Management System (MTMS) to manage the topics of the MQTT messages shared by any publisher based on MTNC. The proposed system can facilitate the publishers to register their topics of the data that they need to share and the subscriber can search the topics that they need to subscribe by using MTNC as the criteria.

The rest of this paper is organized as follows. Section II, and Section III is introduced the MQTT protocol, and MTNC respectively. In Section IV, the MQTT-Topics Management system is proposed. Experimental results are presented in Section V. Finally, Section VI concludes this paper.

II. MQTT Protocol

Since 2016, MQTT, the open source publish-subscribe messaging protocol, is an ISO standard (ISO/ IEC 20922: 2016), invented by Dr. Andy Stanford-Clark, and Arlen Nipper of Arcom, in 1999. It is designed for constrained devices and low-bandwidth, high latency or unreliable networks [1]. It is lightweight and simple protocol and designed to be easy to implement [9].

For the data communication over MQTT protocol, the central controller that used to distribute the message is called MQTT broker. The MQTT broker forwards, filters and prioritizes publish requests from the publishers to the subscribers as shown in Fig. 1. To communicate via MQTT protocol, the publisher (or data generator, e.g., sensor or embedded system) must define two elements include message and topic for the MQTT broker. The message element is the string data that the publisher wants to share to the subscribers via MQTT broker. Meanwhile, the topic is a string that is used by the broker to filter and decide which subscribers should receive which message.

A topic can consists more than one topic levels where each level is separated by a forward slash. For example, the carbon monoxide (CO) sensors that are located in a shopping mall's parking lot, it can publish its carbon monoxide level using this hierarchical topic:

$$ShoppingMall/parkingLot/CO\ .$$

From the decision of the MQTT broker, the subscribers can receive only the message with the same topic that they subscribe. Fig. 1 is the example, when the publisher publishes the message with topic $data2$, only one subscriber can get this

978-1-5090-5210-3/17/$31.00 ©2017 IEEE

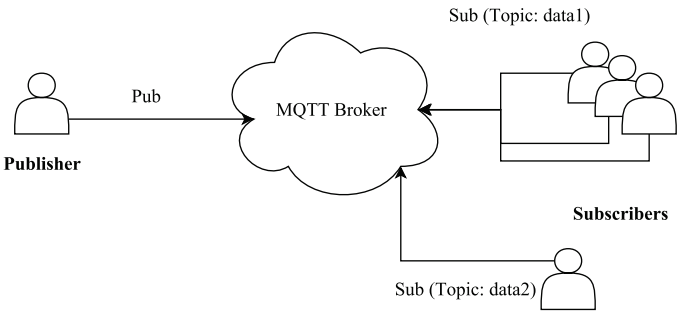

Fig. 1: The data communication over MQTT protocol.

message but if the publisher publishes the message with topic $data1$, there are three subscribers get the message from the broker.

At present, there are many systems that use MQTT protocol to be the core of data transportation instead of the HTTP protocol. It is because MQTT protocol consumes less power than HTTP protocol per hour of operation while the messages sent by MQTT in one hour is ten times more than HTTP protocol. That is MQTT protocol can reduce the power consumption and also faster than HTTP [2].

III. MQTT-Topic Naming Criteria of Open Data for Smart Cities (MTNC)

In [11], MQTT-Topic Naming Criteria of Open Data for Smart Cities (MTNC) are proposed. MTNC is the Criteria for the publishers to define the topic of their data before sharing the data via MQTT protocol. Moreover, MTNC facilitates any subscriber to find their need data easily.

Since, MQTT does not support labeling messages with types or metadata [13], so, lots of data shared over MQTT protocol have the variety of topics and cannot presume by subscribers. In [11] they found that the MQTT Topic can present some metadata of its message. For example, the topics, $Japan/averageTemperature$ and $carbondioxide/China$ can present the meaning of its message i.e., the first topic may be the data about the average of temperatures in Japan and the last topic may be the amount of carbon dioxide in China. So, in [11] they proposed the MTNC to be the Criteria for the publishers to define their topic name before sharing their data via MQTT protocol. It makes all topics named by using MTNC can show their metadata by themselves.

A. Topic Elements Based on MTNC

To define the topic's name base on MTNC, the topic's name should be included three attributes; Objective, Location, and Owner.

1) Objective: There are three components of the objective attributed; main objective, sub-objective, and user-defined objective. The main objective and sub-objective are defined by [11]. There are 4 groups of the main objective; Environment, Energy, Human, and Living, and each group has its own sub-objective, as shown in Fig. 2.

Main Objective	Sub Objective	Main Objective	Sub Objective
Environment	Air	Human	Birth
	Water		Death
	Soil		Population
	Waste		Dint
	Forest	Living	Transport
Energy	Electric		Emergency
	Gas		Healthcare
	Oil		Education
	Light		Safety

Fig. 2: The objectives of MQTT-Topic for smart cities.

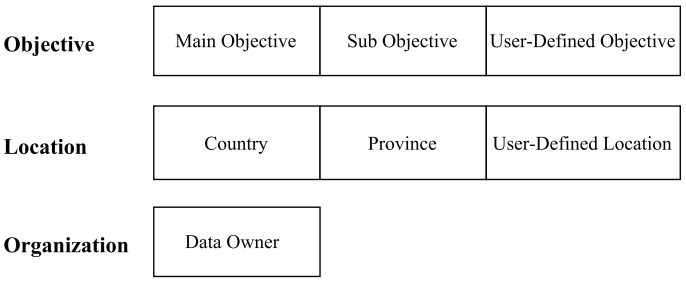

Fig. 3: Three attributes of MQTT-Topic.

2) Location: There are three components of the location attribute including; country, province, and user-defined location. The alpha-3 code of ISO 3166 standard is used to represent country attribute. The province attribute is the full name of provinces of the data are generated and the user-defined location is specific location that the publisher can define by themselves. In [11], they use the tree roman characters for Thailand's province abbreviation. For example, the topic $THA/BKK/CentralWorldHotel$ means the data are generated at Central World Hotel, Bangkok, Thailand.
Note: $CentralWorldHotel$ can be other names such as the name of a building, place, etc., that the data are generated.

3) Owner: For the owner attribute, it is individual naming from each publisher. On the other hand, the publishers can define by themselves.

Fig. 3 shows the components of each attribute in MQTT-Topic in MTNC. The topic including,

Objective:$MainObjective/SubObjective/X$
Location: $Country/Province/X$
Owner: X

Where X is the user-defined data and the order of the attributes is $Objective/Location/Owner$.

From MTNC, the reverse of natural reading of topic $Environment/air/carbonDioxide/THA/CMI/MaejoUniversity/INTNINLab$ means the data generated by INTNINLab at Maejo University, Chiang Mai, Thailand, about the carbon dioxide of the air in the field of environment (the values of this topic is presented in its message). When we read from the topic, we can guess that this data is the data about the volume of carbon dioxide in the air at Maejo University.

IV. MQTT-Topics Management System (MTMS)

This section we propose MQTT-Topics Management System (MTMS) to manage the topics of the open data that any publisher would like to publish. There are 4 important components of the proposed system include; broker, web application, daemon, and database. Fig 4 is the overview of the proposed system.

A. Web Application Component

The main components of the web application are; users registration, topic registration, and topic searching. From Fig 4, any publisher can register to the proposed system and determine the topic of data that they need to share into the proposed system. When a publisher registers a topic into the web application then the topic is collected into the database. Meanwhile, any subscriber can use the proposed system to find the topic of data that they interested. In the process of topic registration, the web application creates the labels of data by split a topic using the symbol "/" into many words and use them as the label of each data to be used in the searching method. Since the topic named with MTNC has 7 tiers, so each topic has 7 words; 3 words from objective attribute, 3 words from location attribute, and 1 word from owner attribute. We implement our web application based on the Bootstrap framework to make the web application responsive and look well in the mobile phone of any user. The topic registration page and the topic searching page are shown in Fig 5 and Fig 6 respectively.

B. Broker Component

In our proposed system, Mosquitto, an Open Source MQTT Broker is selected to build our broker component. Mosquitto provides a lightweight method of carrying out messaging using a publish/subscribe model. It suitable for "Internet of Things" messaging such as with low power sensors or mobile devices such as phones, embedded computers or microcontrollers [1]. From Fig 4, After the subscribers search the topic's name of the data that they need to subscribe by using the web application component, the subscribers can connect to the broker and subscribe for the needed data with these topics. When some publishers publish some message on this topics, then, the broker publish these messages to the subscribers who subscribe for the messages with the topic that is the same as the topic that the publishers publish to the broker. Moreover, the broker will publish the message to the daemon too (if the daemon subscribes for this topic).

C. Daemon Component (* Optional)

In the proposed system, the daemon acts like a subscriber, implemented with Python language. It is the option when we need to collect the data from the publishers into our database. If we need all of data, the daemon is subscribed all the topics registered in the proposed system. The daemon is necessary if the data is importance and needed to be collected e.g., the data about the rainfall and all of them can be used to predict flood or landslide.

D. Database Component

For the database, it can be divided into two parts; the data of users and the data that are published by the users. The data that we need to collect are written to the database by the daemon.

V. Experimental Results

In this section, we present the experimental results to evaluate our proposed work. The experiments are conducted on 100 users, divided into two groups, 50: 50. The first group is determined as the publishers and the last group is the subscribers. Each publisher is assigned to register their 10 topics of data about their city based on MTNC into the proposed system. Then, the subscribers are assigned to find some topics of data about the city 10 times, where the subscribers do not know about the MTNC. The experimental results show that more than 70% the subscribers can discover their needed data when they select the data by reading only the topics of data.

Finally, the usability, design and the benefits of the proposed software are evaluate from the 100 users. The result shows that the proposed software has the high average score of usability, design and the benefits, shown in Fig 7. From the result the average score of usability, design and the benefits of the proposed software are more than 4.00 from 5.00 with at most 0.78 of standard deviation.

VI. Conclusion and Future Work

Since MQTT, the lightweight messaging publish/subscribe protocol can be used to share any dynamic data, the sharing of data, e.g., flood monitoring data, earthquake data, or traffic data via MQTT protocol, can be used to improve the way of human life. However, there are variety of topic naming when the publishers shared their data over the MQTT protocol. To create the standard of topic naming MTNC was proposed in our previous work. In this work, we implement the MQTT-Topics Management System (MTMS) to manage the topics of the open data that any publisher would like to publish based on MTNC. The experimental results show that more than 70% the subscribers can discover their needed data when they select the data by reading only the topics of data although they do not know about MTNC. Moreover, the average score of usability, design and the benefits of the proposed software are more than 4.00 from 5.00 with at most 0.78 of standard deviation.

Acknowledgment

This research is partially supported by Maejo University. We would like to thank the colleagues at INTNIN Laboratory, Faculty of Science, Maejo University for their support.

References

[1] mqtt.org. http://mqtt.org. [Online; accessed 20-July-2016].

[2] M. H. Amaran, N. A. M. Noh, M. S. Rohmad, and H. Hashim. A comparison of lightweight communication protocols in robotic applications. *Procedia Computer Science*, 76:400 – 405, 2015.

[3] A. Antoni, M. Marjanovi, P. Skoir, and I. P. arko. Comparison of the cupus middleware and mqtt protocol for smart city services. In *2015 13th International Conference on Telecommunications (ConTEL)*, pages 1–8, July 2015.

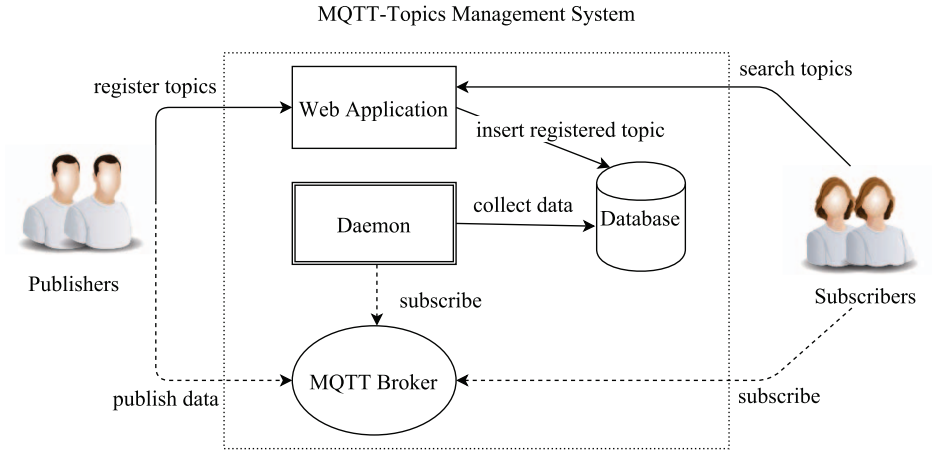

Fig. 4: The OCM Platform.

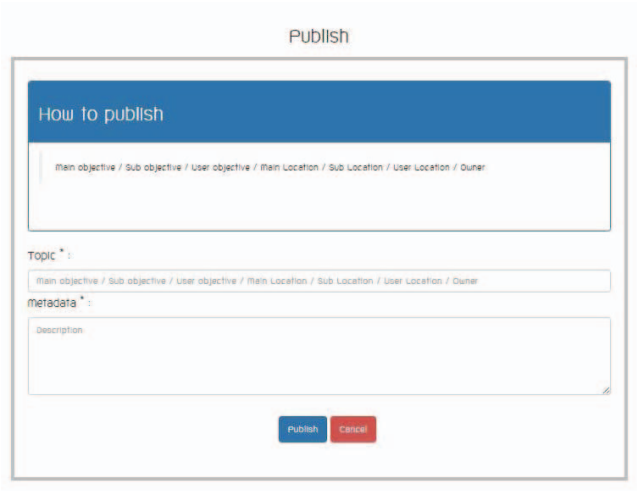

Fig. 5: The topic registration page.

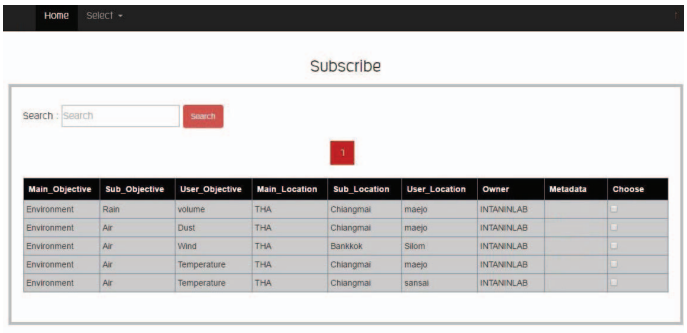

Fig. 6: The topic searching page.

	Usability	Design	Benefits
Average Score	4.20	4.30	4.38
SD	0.78	0.68	0.73

Fig. 7: The average score of usability, design and the benefits of the proposed software.

[4] C. Doblander, K. Zhang, and H. A. Jacobsen. Publish/subscribe for mobile applications using shared dictionary compression. In *2016 IEEE 36th International Conference on Distributed Computing Systems (ICDCS)*, pages 775–776, June 2016.

[5] A. Domingo, B. Bellalta, M. Palacin, M. Oliver, and E. Almirall. Public open sensor data: Revolutionizing smart cities. *IEEE Technology and Society Magazine*, 32(4):50–56, 2013.

[6] D. Evans. The internet of things how the next evolution of the internet is changing everything. *White Paper, Cisco Internet Business Solutions Group (IBSG)*, 2011.

[7] S. Gupta, A. Padhy, A. Adhikari, and A. Dutta. A semantic web and linked data based framework for smart city data management. In *2016 13th International Conference on Electrical Engineering/Electronics, Computer, Telecommunications and Information Technology (ECTI-CON)*, pages 1–6, June 2016.

[8] C. Harrison, B. Eckman, R. Hamilton, P. Hartswick, J. Kalagnanam, J. Paraszczak, and P. Williams. Foundations for smarter cities. *IBM J. Res. Dev.*, 54(4):350–365, July 2010.

[9] ISO/IEC20922:2016. *Information technology – Message Queuing Telemetry Transport (MQTT) v3.1.1*. 2016.

[10] N. Tantitharanukul, J. Natwichai, and P. Boonma. A trail-based approach for job scheduling in distributed systems with workflows. In *Ninth International Conference on Digital Information Management (ICDIM 2014)*, pages 57–62, Sept 2014.

[11] N. Tantitharanukul, K. Osathanunkul, K. Hantrakul, P. Pramokchon, and P. Khoenkaw. Mqtt-topic naming criteria of open data for smart cities. In *The 20th International Computer Science and Engineering Conference 2016 (ICSEC)*, 2016.

[12] URBACT. Opencities final report. April, 2011.

[13] O. Vermesan and P. Friess, editors. *Internet of Things: From Research and Innovation to Market Deployment*. River Publishers Series in Communication. River, Aalborg, 2014.

[14] M. N. E. v. S. . Wendy Carrara, Wander Engbers. Analytical report 4 open data in cities, european data portal. 2016.