Smart Home Web of Objects-based IoT Management Model and Methods for Home data mining

Jeu Young Kim, Hark-Jin Lee, Ji-Yeon Son, and Jun-Hee Park
Smart Home·Factory Ressearch Section
ETRI
Daejeon, South Korea
jeuyk, gausslee, jyson,juni@etri.re.kr

***Abstract*— Nowadays, much research in recent years has focused on IoT (Internet of Things). The home domain is the most important research area of IoT, because there expected accounts of home smart device for over 40 percent of connected device excepting mobile phone. Furthermore, enormous data is generated by home smart device. There is growing concern, but the previous works didn't address enough to manage and analyze home data. The purpose of this study is to describe and examine to manage the aggregated home IoT data based on SWO (Smart home Web of Objects), and SWO analytics platform. We shows the implementation of SWO analytics platform and a case study using real data from smart metering devices for analysis of appliance usage patterns.**

Keywords—Home IoT; Web of Objects; Analytics platform

I. INTRODUCTION

Much research in recent years has focused on Internet of Things (IoT) which connects physical objects to network and manages information of the objects. Especially, home domain is the most important field of IoT. Because the research survey [1] reported that the number of connected device will grow to nearly 8 billion devices for the year 2020 excepting mobile phone, and home devices has the biggest portion of them about 3.7 billion. It means they will generate big data such as sensory data, usage information and so on. Home service providers want to develop various and advanced service using the data, but there remains a need for an efficient method that can analyze the data. Previous works have mainly focused communication way, or the specific service of analysis. E.S.Lee, et al. [2] developed an auto-configuration system structure and protocol for Internet-capable home appliances, which supports the initial configuration and remote maintenance service of the device with only little user effort. HJ.Lee, et al. [3] proposed a three stage conversion process for interoperability among different middleware of home network that supports not only the conversion of the message format and schema, but also semantic conversion. Shih-Yeh Chen, et al. [4] developed electronic appliance recognition system by building a database mechanism, electronic appliance recognition classification, and waveform recognition. It have been carried out smart meter data analysis but they are accepted only meter data. However, IoT platform needs to accept various kinds of data and manage the data uniformed system.

Furthermore, many researches have studied IoT platform in another area. iHome Health-IoT [5] proposed platform seamlessly fuses IoT devices (e.g., wearable sensors and intelligent medicine packages) with in-home healthcare services (e.g., telemedicine) for improved user experience and service efficiency. It mainly concerned a variety of health IoT device from body to cloud, and their hardware architecture. Andrea Zanella, et al. [6] provided a comprehensive survey of the enabling technologies, protocols, and architecture for an urban IoT, and presented the technical solutions and best-practice guidelines adopted in the Padova Smart City project. They collected the temperature, humidity, light, and benzene readings, and just shown the plots. They didn't have analysis part for urban IoT yet. Hongming Cai, et al [7] proposed a platform which based on an abstract information model, information encapsulating, composing, discomposing, transferring, tracing, and interacting in Product Lifecycle Management could be carried out.

In this paper we proposed Smart home Web of Objects Analytic Platform (SWOAP) based on SWO (Smart home Web of Objects) architecture. We defined SWO concept, which is made the web (virtual) object from home device and data without any administrating. The one of big issue is to support interoperability of protocols. In order to solve it, SWO is based on standard web protocol. It is a proper technology that is open, safe, and automatic in web environment.

Fig. 1. shows total concept architecture. The base part has SWOGW (GateWay) connecting with internet and smart device. Smart devices can communicate directly to management plane by themselves. However, many kinds of legacy home device can't do it. SWOGW substitutes to communicate with management plane for legacy devices. Management plane aggregates home device information and their data using web protocol. Home device and their data are managed as object in management plane. The above of management plane is service plane. SWO service can generate new mashup service using capabilities of management plane.

We propose home analysis service, SWOAP, which combines device information, user information, stream data, and physical space information. According to spread IoT environment, data analysis area has become the most important, so this paper will introduce the model and methods centered on data analysis.

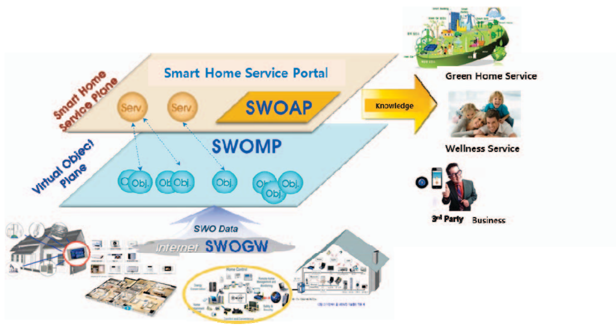

Fig. 1. The concept architecture of Smart home Web of Objects

Section 2 shows SWOAP system architecture. Section 3 explains implementation and use case. Section 4 concludes this work and future works.

II. SYSTEM ARCHITECTURE

SWOAP is analytic platform based on SWO. SWOAP consists of two parts, resource core layer, and knowledge building framework layer (Fig.2.). Resource core layer aggregates smart appliance and legacy appliance information through management platform. It is created each domain resource by resource core layer and built relationships between resources. This is proposed our view to take various data. Knowledge Framework Layer conducts home data analysis according to resource concept. The more detailed explanation is on under sections.

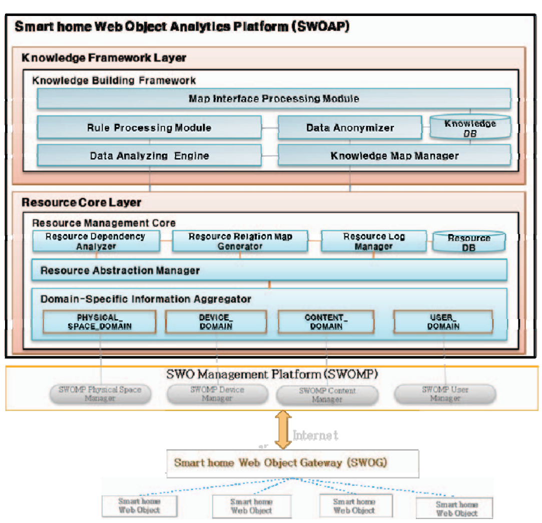

Fig. 2. The system architecture

A. SWOA Protocol

The analytics platform aggregates home data from management platform. We defined a procedure the aggregating sequence of device information shows Fig.3. First, SWOAP requests all of home site id to management platform. After then, in order to get device information in each home, SWOAP requests all of device id to management platform about each home. Finally, SWOAP requests device information using device id that get previous step. The APIs were defined by standard specification based on JSON in this study.

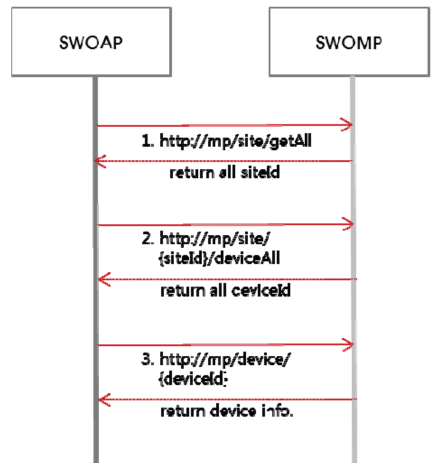

Fig. 3. The procedure of aggregating SWO information by management protocol

The one of the most difficult problems is real-time event receiving in web environment. Using polling API, web socket, WS-event, and so on are used the way, but they aren't stable enough. We defined light weight event protocol. This event protocol just notify event occurrence from management platform and SWO to SWOAP. The defined protocols are on Table 1. Events are defined SWO add, SWO delete, and dedicated events in each SWO, for example, changing on/off status of light from the light device. When receiving event notices, SWOAP requests detailed information using management platform APIs.

TABLE I. SWOA EVENT PROTOCOL

Num	Message ID	Description	Direction
1	1001	SWO Add	SWOMP → SWOAP
2	1002	SWO Add Response	SWOAP →SWOMP
3	1003	SWO Delete	SWOMP → SWOAP
4	1004	SWO Delete Response	SWOAP → SWOMP
5	1005	SWO Event	SWOMP → SWOAP
6	1006	SWO Event Response	SWOAP → SWOMP

B. Resource Core Layer

Resource core layer is charge of management home data which is used all around the analytics platform. Basically home data is classified in domains and abstracted a resource. The resource has specific data structure for each domain. The one of features on our platform is building the relation information between resources. The relation information is made of implicit knowledge through internal analyzing. For example, a device domain resource and a physical domain resource can have LOCATED relationships between them, if the device located on the physical space. The more detailed resource architecture described elsewhere [8] [9]. For the proposed system, we extended content domain to include stream data. The stream data is collected configured time unit like one minute, one hour, or so on, because it is continuously delivered in very short time interval. Resource core layer manages the metadata of content resource, and real data is stored in database. In content domain, it can have PRIVEOUS and NEXT relationships according to timestamp between content resources. Using the diverse information model in resource core layer, we could deal with home data what we need. Basically, home platform needs to manage based on home address. Resource core layer can support to build the relationship between the physical space domain and the other resource domain. Furthermore, we can arrange the resource group centered on another domain, such as device group in a same network area, or in a same user, or so on.

C. Knowledge Framework Layer

This framework carries out model driven analysis. We proposed the model description schema based on PMML (Predictive Modeling Markup Language). PMML [10] is the leading standard for statistical and data mining models and supported by over 20 vendors and organizations. With PMML, it is easy to develop a model on one system using one application and deploy the model on another system using another application. PMML uses XML to represent mining models. The structure of the models is described by an XML Schema.

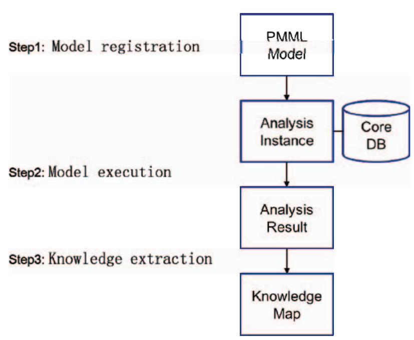

Fig. 4. Model-driven anlalysis process

At the model registration step, model developer described analysis model using PMML and PMML extension schema which was defined additional field in our study. The extension part applies our resource concept to analysis model. It accepts the resource type, data of the resource, home information and so on. Consequently, SWOAP cans analysis data of the resource according to model definition. At the execution step, it makes an execution instance using analysis model and received analysis rule from service developer. Analysis rules mean specific analysis condition depends on service such as number of cluster, target device, target home, or so on. It gets raw data from core database and after then start analysis. At the knowledge extraction step, it builds home knowledge combining analysis result and execution instance information. The extracted home knowledge is formatted by home knowledge RDF schema and stored in RDF store. The application can query some home knowledge from this home knowledge map.

III. Implementation

A. Testbed and test scenario

Our scenario is discovering eating habit by analyzing cook device usage. Fig.5. shows our testbed. Above of the figure is software structure and under of the figure is physical structure. In order to aggregate real data, we installed smart metering device to 10 households. After then, a ricepot plug in the smart metering device. The smart metering device became SWO by metering agent, and they send their basic description and their power usage data periodically. In this environment, our platform has 10 resources of device domain and 10 resources of physical space domain based on home address at least, and the power usage data become numerous content resources. Moreover, the platform builds the relation objects among them. We provide web service APIs that can request aggregated resource information, sensory data of resource, and so on.

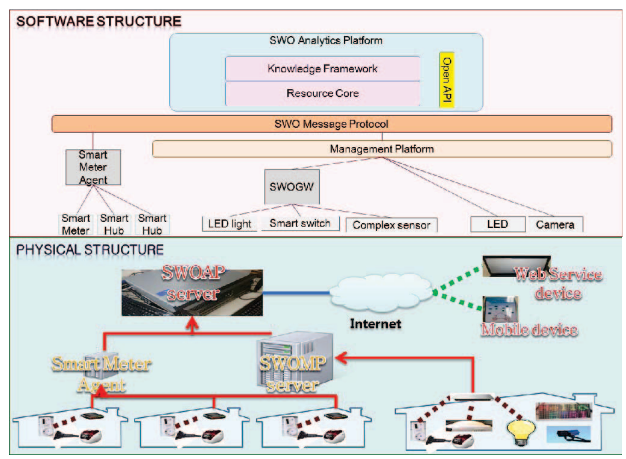

Fig. 5. Ricepot usage pattern analysis Testbed

Fig. 6. visualizes the aggregated power usage data of a smart meter connecting the ricepot data in our platform. In that, we could infer when it shows the high power usage bar, ricepot worked cook mode, when it shows the lower power usage bar than above it, it worked warm mode, and when it shows the lowest, it stopped. We analyzed whole power usage data based on these usage characters of ricepot.

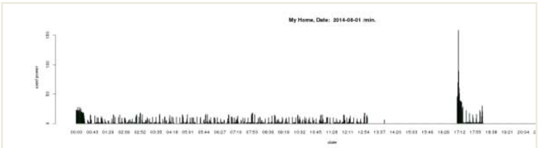

Fig. 6. The aggregated power usage data

The above of Fig. 7. shows the result of clustering analysis about power usage data. We set features that are time and power usage value. As a result, we discovered three clusters, and defined the status reference data of ricepot using the clusters. Fig.x shows the result of ricepot usage patterns. We applied association rules algorithm to the usage pattern analysis. We divided a bulk of steam data into units per one week which can be dealt a transaction on association rules. The platform executed the analysis using model, transactions, and previous result. The under of Fig. 7. visualizes the result. The rows mean the transaction, and the columns mean the item. The usage patterns are made by frequencies of items on transactions.

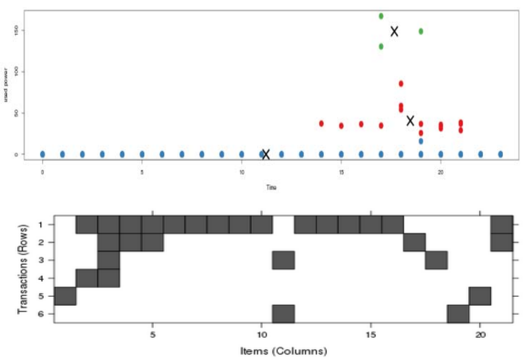

Fig. 7. The result of ricepot usage pattern analysis

The previous result can be queried on the knowledge map. Because all results of the analysis are stored on knowledge map in RDF form by the platform. Through the analysis we can get how many and when people cook for a week. Although the result looks simple, it will become useful information about home and residents if various analysis results are accumulated on knowledge map.

B. Evaluation

A simple comparison of home data technologies from different perspectives is given in Table II. to compare the data model characters. We selected five features. First feature is target service domain. Commonly, data model tends to design dependent on what for service. Second feature is dealing with how many kinds of data. Third feature is data expendability. It means the proposed system can accept new kind of home data. Forth feature is whether provide relation information between data or not. The final feature is whether support multi user or not.

TABLE II. THE COMPAISION OF HOME DATA MODEL

	Target service domain	**How many kinds of data**	**Data Expend ability**	**Data relation**	**Multi user**
MARJORIE SKUBIC, et al.[11]	Health-care	1	X	O	O
Shuai Zhang, et al.[12]	Home activity	1	X	X	O
Xiao Qin, et al. [13]	Energy saving	2	O	O	X
This paper	Home activity	4	O	O	O

The compared paper mainly studied home data analysis methods. They basically collected and mined the sensor data for each service domain. However, excepting third paper, the others concerned one kind of data – sensor type, and can't support new kind of data. Our study already has 4 kinds of data structure, and it can be expended freely about new resource domain. Many of studies developed to discover and manage about data relationship. This feature is the important point that makes the system to be intelligent. Our study also designed the relation object, and can add new relation definition. Supporting multi user is also important feature, especially home domain. Our study defined user domain, and managed user resource.

IV. CONCLUSION

A few researchers are addressed IoT management and data mining in home domain. This paper proposed not only the home IoT management model and methods but also model driven home data mining procedure. We defined content resources for dealing the stream data. We showed the use case that analysis ricepot usage patterns in proposed platform. In order to does it, smart metering devices and their data had to become the resources, and the result of analysis be stored in some database has the query methods.

Many of previous research did home data analysis, but they almost had dependency on specific service, or device. We designed the resource model based on domain that can accept various source, and model-driven analysis platform based on resource. Furthermore, the result of analysis is stored on knowledge map. Knowledge map consists of resource information, relation information, and result of analysis related to what user/service want to discover knowledge. As time goes on, knowledge map has cumulative effects about home and residents. Our future works is to research to infer new knowledge using knowledge map.

ACKNOWLEDGMENT

This work was supported by the IT R&D program of MOTIE/KEIT [10047233, Development of Smart Home Web of Objects Architecture].

REFERENCES

[1] Harbor Research, “What Exactly Is The "Internet of Things?”, March, 2014.

[2] Eun-Seo Lee, Hark-Jin Lee, Kwangil Lee, and Jun-Hee Park, "Automating Configuration System and Protocol for Next-Generation Home Appliances," ETRI Journal, vol. 35, no. 6, Dec. 2013, pp. 1094-1104.

[3] LEE, Hark-Jin; LEE, Eun-Seo; LEE, Kwangil; PARK, Jun-Hee; RYOU, Jae-Cheol, “MAHI: A Multiple Stage Approach for Home Network Interoperability”, IEICE Trans. Commun. vol. 97 (12) p. 2689-2697

[4] Shih-Yeh Chen; Chin-Feng Lai; Yueh-Min Huang; Yu-Lin Jeng, "Intelligent home-appliance recognition over IoT cloud network," Wireless Communications and Mobile Computing Conference (IWCMC), 2013 9th International , vol., no., pp.639,643, 1-5 July 2013.

[5] Geng Yang; Li Xie; Mantysalo, M.; Xiaolin Zhou; Zhibo Pang; Li Da Xu; Kao-Walter, S.; Qiang Chen; Li-rong Zheng, "A Health-IoT Platform Based on the Integration of Intelligent Packaging, Unobtrusive Bio-Sensor, and Intelligent Medicine Box," Industrial Informatics, IEEE Transactions on , vol.10, no.4, pp.2180,2191, Nov. 2014.

[6] Zanella, A.; Bui, N.; Castellani, A.; Vangelista, L.; Zorzi, M., "Internet of Things for Smart Cities," Internet of Things Journal, IEEE , vol.1, no.1, pp.22,32, Feb. 2014.

[7] Hongming Cai; Li Da Xu; Boyi Xu; Cheng Xie; Shaojun Qin; Lihong Jiang, "IoT-Based Configurable Information Service Platform for Product Lifecycle Management," Industrial Informatics, IEEE Transactions on , vol.10, no.2, pp.1558,1567, May 2014.

[8] Ji-Yeon Son; Jun-Hee Park; Kyeong-Deok Moon; Young-Hee Lee, "Resource-aware smart home management system by constructing resource relation graph," Consumer Electronics, IEEE Transactions on , vol.57, no.3, pp.1112,1119, August 2011.

[9] Jeu Young Kim; Ji Hyun Lee; Ji-Yeon Son; Jun-Hee Park, "Resource relation map based fault diagnosis and fault tolerance methods for home network environments," Network Operations and Management Symposium (APNOMS), 2013 15th Asia-Pacific , vol., no., pp.1,5, 25-27 Sept. 2013.

[10] PMML, www.dmg.org/

[11] Skubic, M.; Guevara, R.D.; Rantz, M., "Automated Health Alerts Using In-Home Sensor Data for Embedded Health Assessment," Translational Engineering in Health and Medicine, IEEE Journal of , vol.3, no., pp.1,11, 2015.

[12] Shuai Zhang, Sally I. McClean,Member, IEEE, and BryanW. Scotney, “Probabilistic Learning From Incomplete Data for Recognition of Activities of Daily Living in Smart Homes”, IEEE TRANSACTIONS ON INFORMATION TECHNOLOGY IN BIOMEDICINE, VOL. 16, NO. 3, MAY 2012.

[13] Xiao Qin, Susan Lysecky, and Jonathan Sprinkle, “A Data-Driven Linear Approximation of HVAC Utilization for Predictive Control and Optimization”, IEEE TRANSACTIONS ON CONTROL SYSTEMS TECHNOLOGY, VOL. 23, NO. 2, MARCH 2015.