Using Case Based Data Stream Processing for Smart House Systems

Martínez Alejandro Maximiliano[1], Martínez Natalí [1] and Martín María de los A.[1]

[1] National University of La Pampa, 110 street 390- General Pico- La Pampa- Argentina
martinma@ing.unlpam.edu.ar

Abstract. The latest technological advances have allowed the development of smart home systems that establish a connection between humans and the devices that surround them, living in homes and working in fully automated companies. While these systems improve the quality of life of people in terms of comfort, safety and energy savings, as they automate aspects such as lighting, temperature, humidity, among others; they lack an organizational memory that manages the knowledge of the preferences, customs and behavior patterns of the inhabitants of the home or company they automate. The aforementioned lack shows an opportunity to improve these products, taking into account that the market trend in domestic devices, is to maximize automation, making them increasingly intelligent, to make decisions proactively, and to collaborate with each other and with the human being for his better quality of life. In this paper we show that it is possible to improve the management of the data stream captured by the devices, providing them with interoperability through the use of ontologies and metadata, so that "intelligent modules" process the acquired knowledge automatically, with little intervention from the humans. Specifically, we will use the Processing Architecture based on Measurement Metadata (PAbMM) and the Case based Reasoning (CBR) technology to intelligently establish the environmental conditions of the home, in order to improve comfort and reduce energy consumption.

Keywords: Data Stream Processing, Big Data, Case based Reasoning Learning, Domotic, Ontology.

1 Introduction

Home automation systems are systems capable of automating a home or building of any kind, adding aspects of safety and welfare to its inhabitants. These systems are composed of sensors that capture and send information of environment variables to a central module, which, based on the captured data, stimulates the corresponding actuators to provide comfort, safety, etc.

One way to take advantage of so much information captured by the sensors is the development of the proposed software, which can be installed in domotic home centers, to make better use of the information obtained [1].

Specifically, once the data from the sensors has been captured [2], it is convenient to structure them in acknowledge memory (also named Organizational Memory), so that later they can be exploited and used for the recommendation in subsequent decision making. For example, at the same time that the memory is recognizing the customs of the inhabitants of the house, it can trigger alarms in cases of unusual situations, as it would be if the person does not get up at the usual time, may have suffered an illness, and warn to their relatives; or recommend the closing of windows if the temperature drops below the usual, etc.

The organizational memory (OM) management represents a key asset to support decision-making processes by different organizational stakeholders [3]. The main aim of knowledge management systems is to manage, store and retrieve the organizational knowledge, so that it can be used later to learn, share knowledge, solve problems, and ultimately to support better decision-making processes. To ensure an efficient management of organizational knowledge, it is necessary to have technological platforms that support it. In the previous work, we proposed architecture based on data flow processing, for this purpose. Specifically, the Processing Architecture based on Measurement Metadata (PAbMM) [4, 5].

The PAbMM evolves the original strategy [4] incorporating support to the big data repositories in contexts of distributed computation. This implies the necessity of gather Big Data and Data Stream Processing technologies; which ensures powerful large data volumes processing, allowing efficient management of knowledge in the Organizational Memory.

The Organizational Memory that integrates the Processing Architecture based on Measurement Metadata serve as base for the organizational knowledge exchange and to be used in recommender systems in decision making processes. Therefore by having a well-developed organizational memory that supports the structuring, reusing and processing of organizational knowledge is a primary decision (and likely a success factor) to achieve such an effective management.

Nonaka and Takeuchi have said that an organization cannot create knowledge itself. Conversely, the knowledge creation basis for an organization is the individual’s tacit knowledge; and tacit knowledge is shared through interpersonal interactions [6].Therefore, in order to reach and maintain the organizational effectiveness and competitiveness, an organization needs to learn from past and present experiences and lessons learnt and to formalize organizational memories for enabling to make explicit the individual’s tacit knowledge – and why not community’s tacit knowledge as well [7].

In this paper we show the design of smart software modules to support the development of domotic systems with the added value of knowledge management software, which provides additional intelligence (to the traditional simple response to sensors).This software provide the system with a learning mechanism using an knowledge memory and reasoning based on cases, which will result in a better use of domestic resources, a higher level of comfort and security and, ultimately, better benefits for the end user.

This article is organized in five sections. The Section 2 outlines the case based organizational memory. The Section 3 illustrates the application of the organizational memory to a practical case: A smart house recommendation System. The Section 4 discusses related work and finally the Section 5 summarizes the conclusions.

2 Case Based Organizational Memory

One of the main goals of an organizational knowledge management strategy is to try to formalize the informal knowledge in order to allow machine-processable semantic inferences. A way of alleviating this problem from the knowledge representation standpoint is to store the knowledge in a more structured and formal way. We have followed this approach by using the case-based organizational memory strategy. It combines organizational knowledge storage technology with case-based reasoning to represent each item of informal knowledge.

In general, the organizational memories are intended to store the partial formal and informal knowledge present in an organization with automatic processing capabilities. In particular, by structuring an organizational memory in cases can also facilitate the automatic capture, recovery, transfer and reuse of knowledge for problem solving. Thus, data, information, and knowledge from heterogeneous and distributed sources can be automatically and semantically processable by web-based applications, for instance, an 'intelligent' recommendation system to support a more effective decision-making process [8].

With the aim to manage and retrieve the organizational knowledge, in the last years numerous proposals of models and tools for knowledge management and knowledge representation have arisen. However, most of them store knowledge in a non-structured or semi-structured way, hindering the semantic and automatic processing of this knowledge. In this section we specify a case-based organizational memory, so that it can be used to learn, reasoning, solve problems, and as support to better decision making as well.

A way of alleviating this problem from the knowledge representation standpoint is to store the knowledge in a more structured and formal way. We have followed this approach by using the case-based organizational memory strategy. It combines organizational knowledge storage technology with case-based reasoning to represent each item of informal knowledge. In particular, by structuring an organizational memory in cases can also facilitate the automatic capture, recovery, transfer and reuse of knowledge for problem solving.

In order to reach this aim we have constructed a common conceptualization for case-based organizational memory where concepts, attributes and their relationships should be explicitly specified; such an explicit specification of a conceptualization is one of the core steps for building ontology.

In the following sections we will describe case-based organizational memory ontology and his application to the construction of a smart house recommendation system.

2.1 Case Based Organizational Memory Ontology Overview

The organizational memory ontology aims to be at a generic level from which other representations for specific domain applications can be formulated (see Figure 1). On the one hand, the case-based organizational memory ontology is defined at a generic organizational memory level [8], and on the other hand, for characterizing the cases according to the specific knowledge domain [9] and its context [10], a domain and context ontologies should also be provided.

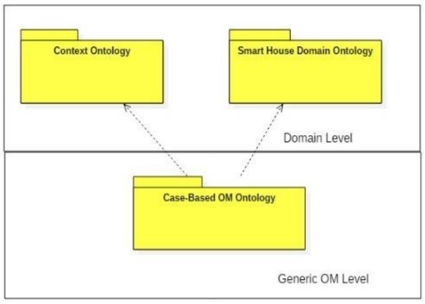

Fig. 1. The relationship between ontologies at the specific domain level and at the generic organizational memory level.

The objective of our ontology is to serve as a foundation for the organizational knowledge exchange with semantic power, which in turn facilitates the reuse, the interoperability and the automatic processing by agents [11].

The main concepts of the ontology, which are illustrated in the UML diagram of the Figure 2, are described in the following text, highlighted in italic.

An *organizational memory* is the way in which an organization stores and keeps track of what it knows, i.e., about the past, present and future knowledge. An organizational memory can have one or more *knowledge bases* which are intended to achieve different information needs of an organization –recalling that data, information and knowledge are useful assets for decision making. In addition, an organizational memory may be seen as a repository that stores and retrieves the whole specified, explicit, formal and informal knowledge present in an organization. Thus, a *knowledge base* is an organized body of related knowledge; taking into account that knowledge is a body of understanding and/or lessons learnt from skills and experiences that is constructed by people.

A type of knowledge base is a *case knowledge base* which stores the acquired knowledge in past experiences, good practices, learned lessons, heuristics, etc. to different domains; that is, it stores *cases*. A case is a contextualized knowledge item (i.e., an atomic piece of knowledge) representing an experience by means of a *problem* and its *solution*. The representation of the knowledge through cases facilitates the reuse of the knowledge acquired in past problems to be applied to a new problem in similar situations [12].

A case can be seen as an ordered pair $< P, S >$, where P is the problem space, and S is the solution space. There exists a general description of problems as $P(x_1, x_2, ... , x_n)$, where each individual problem is an instance P(a1, a2,..., an); also a general description of solutions as $S(y_1, y_2, ... , y_n)$, and every individual solution $S(b_1, b_2, ... , b_n)$ is an instance of that general description. The x_i are variables that characterize the problem (*problem feature*), and the yi are variables that characterize the solution (*solution feature*), where both are features. A *feature* or attribute is a measurable physical or abstract property of an entity category. Since the stored cases refer to a specific knowledge domain, the features that characterize the problems and solutions are defined by a domain concept term; for example, the concepts coming from the Smart House domain ontology (i.e. Ambient brightness, presence, alarm status, Environmental Temperature, Environmental Pressure and so on, as we will illustrate in Section 3).

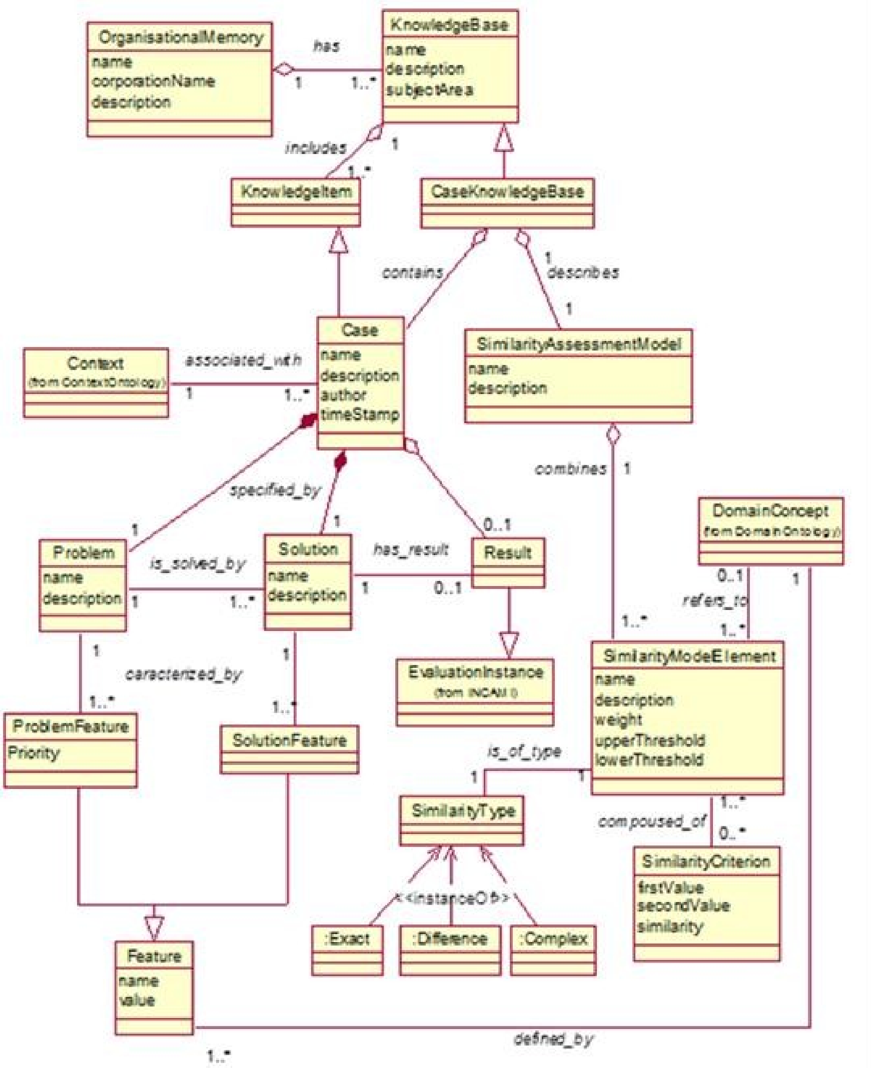

Fig. 2. The relationship between ontologies at the specific domain level and at the generic organizational memory level components.

The case-based reasoning process consists in assigning values to problem variables and finding the adequate instances for solution variables. To find the appropriate values for the instances of a solution, the similarity assessment of cases should be performed, so that for each case knowledge base a similarity assessment model should be specified. Greater detail of the functionality of the Organizational Memory can be found in [7].

2.2 Similarity Assessment Model Representation

Most of the case-based reasoning applications have been focused on problems of specific domains. However, in order to be useful to an organization, a case-based reasoning system should be fitted in with the main knowledge sources that can stem from diverse domains, and so the similarity functions appropriate to each case knowledge base.

As we can see in Figure 2, each case knowledge base has a similarity assessment model that specifies it. On the one hand, the similarity assessment model class represents the general description of problems as $P(x_1, x_2 \ldots x_n)$, i.e. the problem structure, by combining several similarity model elements -one per each problem feature x_i.

On the other hand, it is a function with associated similarity elements that models the similarity assessment of cases. In turn the similarity model element is a function with associated similarity criterion that models the similarity assessment of a feature. We propose a model to define the case structure indicating the features that characterize it and the possible similarity models.

Usually, the similarity between a recovered case R and a new case C is defined as the sum of the similarities among its constituent feature values multiplied by their weights, i.e. the so-called Nearest Neighbor formula:

$$Similarity(R, C) = \sum_{f \in F} w_f * sim_f(f_R, f_C) \qquad (1)$$

Where w_f is the weight of the feature f, and sim_f is the similarity measurements function to the feature. Therefore, in order to provide an appropriate representation of the similarity model, it is necessary to represent the weights that model the relative relevance, and the similarity function type for each specific feature. The weights are represented as an attribute inside each similarity element, and the similarity type is restricted to be one out of three general types, namely: Exact, Difference and Complex.

The inclusion of these three types of functions is based on the analysis of numerous investigation works in the case-based reasoning area, as well as taking into account they cover the similarity representation needs of most cases in the Software and Web Engineering area. Particularly:

The Exact similarity function. Returns 1 if two feature values are the same and 0 otherwise.

The Difference similarity function. Is inversely proportional to the difference between the feature values. It can only be applied when it is possible to define the value difference; for instance, between numerical values the difference similarity function returns 1 if both features are equals, and return $1/|f_c-f_n|$ in other case.

The Complex similarity function. Solves the similarity for all those situations where the two previous functions are not applicable; for example, the semantic difference between two synonymous terms that is neither completely the same nor completely different. If the number of a feature values is finite, it is feasible to have beforehand the similarity measure values for all possible values' pairs. In our model, these parameters are represented in the Similarity Criterion class, which is defined as the assessment pattern used to determine the semantic similarities between two feature values.

Ultimately, an exhaustive glossary of terms, attributes and relationships are shown in [8], where the terminology for the case-based organizational memory ontology is explicitly described.

3 A Smart House Recommendation System

We will elaborate and implement a case-based knowledge base and its similarity assessment model for a specific domain: a smart house recommendation system.

This case base stores a body of related knowledge about the lighting status and energy consumption of a home, based on a variety of data, processed by the PAbMM architecture and taking the domotic sensors as a data source.

So, and in terms of PAbMM, each sensor is a heterogeneous data source that produces and stores the data streams; in this way and from the processed data, you can build knowledge management systems to support the decision making in the smart house domain, based in the experience stored through the Organizational Memory.

3.1 Design of the Model for the light intensity knowledge base

To illustrate the knowledge base and for easier understanding, a simplified model of the Case structure is shown. This case knowledge base characterizes the problem situation through various characteristics including Natural brightness of the environment, (measured in lumens), usual task which is being carried out at that time, the day of the week and if the device was operated manually or automatically (See Table 1). In addition, the context data of location and time is taken into account.

Analogously, the LightPercentage feature; which indicates the Percentage of light level to be dimmed, will characterize the solution. For each feature that characterizes a case, we should establish its weight and its similarity function type (see Table 1). These design decisions could be made by an expert taking into account which features are considered more relevant from the similarity point of view to evaluate in the end the global similarity of two cases.

Table 1. Example of similarity assessment model for the smart house case base.

Feature	Description	Type	Wht
NaturalBrightness	Natural brightness of the environment, (measured in lumens)	Difference	0.40
DayMoment	Indicates what usual task is being carried out at that time. (Possible values are *sleeping, getting up, breakfast, working, lunch, dining, and going to bed*).	Exact	0.25
DayOfWeek	Indicates the day of the week (Possible values are *Monday, Tuesday, etc.*)	Exact	0.15
Manual	Indicates if the device was operated manually or automatically.	Exact	0.20

Once defined the case structure and its similarity assessment model, each case is stored with all the feature values that characterize it and its solution accordingly. Two case examples are shown in Figure 3.

CASE 1

PROBLEM:

NaturalBrightness	120
DayMoment	Breakfast
DayOfWeek	Monday
Manual	No

SOLUTION: *LightPercentage: 60*

CASE 2

PROBLEM:

NaturalBrightness	250
DayMoment	Working
DayOfWeek	Tuesday
Manual	No

SOLUTION: *LightPercentage: 90*

Figure 3. Example of representation of two stored past cases.

A new decision in Illumination managing can benefit from the case-based organizational memory by recovering the smart house information of the most similar brightness conditions. Let us suppose that the case base stores the two cases shown in Figure 3, among others, and a change is detected in the environmental context variables. In order to reduce costs and improve comfort, we can take advantage of the recorded knowledge by retrieving and reusing the most similar past experience. Therefore, the Table 2 shows the similarity calculation of each feature of the new case compared to the previous past ones, i.e., “Case 1” and “Case 2”. Hence the global similarity calculations give as outcomes:

$$Similarity(Case\ 1,\ New) = 0.4 * (\frac{1}{30}) + 0.25 * 1 + 0.15 * 0 + 0.2 * 1 = 0.582$$

$$Similarity(Case\ 2,\ New) = 0.4 * (\frac{1}{100}) + 0.25 * 0 + 0.15 * 0 + 0.2 * 1 = 0.204$$

Table 2. Example of similarity assessment between the previous two past cases and the new.

Feature	Case 1	Case 2	New Case	Similarity Case1/New	Similarity Case2/New
NaturalBrightness	120	250	150	0.132	0.004
DayMoment	Breakfast	Working	Breakfast	0.25	0
DayOfWeek	Monday	Tuesday	Sunday	0	0
Manual	No	No	No	0.2	0.2

Resulting “Case 1” as the most similar to the new case, and therefore the light will be dimmed to 60 percent.

4 Related Work

In our specific work, we have illustrated the use of the case-based reasoning approach to develop a case-based recommendation system for smart house control, using the advantage of having the PAbMM architecture, which manage large volumes of structured data together with their metadata. The main goal is to exploit the knowledge that con be extracted from organizational memory under a key-value structure (i.e. case-solution structure) stored in the Big Data repository, allowing to incorporate more experience for recommending the courses of actions to the decision-making process.

Firstly, from the organizational memory point of view, there are numerous proposals in the knowledge management area, for example the ones documented in [13, 14]. Most of them capture and store the knowledge in repositories of documents like manuals, memos, and text files systems, etc. where structured or semi-structured storage strategies are seldom used. These approaches usually do not employ powerful mechanisms of semantic and automatic knowledge processing based on ontologies therefore causing very often loss of time and high investment in human resources.

Secondly, from the stream processing architecture point of view, there are recent works which make focus on the data stream processing from a syntactic point of view. In this context, the data model of the stream is based in a key-value structure and incorporates techniques for the adaptive management of high-rate streams [15]. Our architecture incorporates metadata based on an Ontology, facilitating consistent and comparable analysis from the statistical point of view, being able to fire alarms based in the interpretation of the decision criteria of the indicators whose value was got from the data.

5 Conclusions

The case based knowledge management represents a key asset to support a more effective decision-making process by different stakeholders [8]. Also, cases are a type of knowledge that facilitates capture, transfer, reasoning and reuse through automatic processing. In this direction, we development an Case-based knowledge memory, which can be installed in domotic home centers, to make better use of the so much information captured by the sensors.

In the previous work [16, 17], we had specified a case-based organizational memory for the stream processing architecture based on measurement metadata. In this paper we deepen the former research and considerably expand the development of the proposed technologies, applying them to home systems. As a result, we obtained a robust and mature smart software modules to be installed in the home centrals.

This software provides the system with a learning mechanism using a knowledge memory and reasoning based on cases, which will result in a better use of domestic resources, a greater level of comfort and safety and, in short, better performance towards the end user.

The aforementioned modules were installed in a FIBARO [18] smart home system. Currently, the system is in the learning stage, memorizing cases through the processing of data streams captured from the sensors. After six months of knowledge acquisition, based on information obtained by different sensors, a recommendation system based on this knowledge memory will be activated.

The recommendation system will shut down the electrical devices when they are not needed, taking into account previous experiences acted upon by the users. Currently, energy consumption is being measured, in order to calculate future energy savings.

References

1. M. Martín and M. Diván, Case Based Organizational Memory for Processing Architecture based on Measurement Metadata. Proc. 5th International Conference on Reliability, Infocom Technologies and Optimization (ICRITO), Noida, India, 2016.
2. M Diván and L Olsina, Process View for a Data Stream Processing Strategy based on Measurement Metadata. Electronic Journal of Informatics and Operations Research, vol. 13, no. 1, pp. 16-34, June 2014.
3. M. Martín and L. Olsina, Added Value of Ontologies for Modeling an Organizational Memory, in IGI Global (Ed). Building Organizational Memories: Will You Know What You Knew, pp. 127-147.
4. M. Diván and L. Olsina, Process View for a Data Stream Processing Strategy based on Measurement Metadata. Electronic Journal of Informatics and Operations Research, vol. 13, nº 1, pp. 16-34, June 2014.
5. M. Diván, M. Martín et al., Towards a Data Processing Architecture for the Weather Radar of the INTA Anguil. Proceedings of International Workshop on Data Mining with Industrial

Applications. IEEE, Edited by S. Gómez, E. Hochsztain & D. Romero. Pp. 72-78. Asunción, Paraguay, 2015.
6. Nonaka I.; Takeuchi H. Knowledge-Creating Company: How Japanese Companies Create the Dynamics of Innovation, Oxford University Press (1995).
7. M Martín, Organizational Memory Based on Ontology and Cases for Recommendation System, PhD Thesis, Computer Science School, National University of de La Plata (Argentina)., 2010.
8. Martín, M., Olsina, L. Ontology for a Case-based Organizational Memory, In Proceedings del XI Workshop Iberoamericano de Ingeniería de Requisitos y Ambientes Software (IDEAS), Recife, Brasil. pp. 351-356, ISBN 97-8857084-13-46 (2008).
9. L. Olsina and M. Martín, Ontology for Software Metrics and Indicators. Journal of Web Engineering (JWE), vol. 3, nº 4, pp. 262-281, 2004.
10. Molina H.; Olsina L. Towards the Support of Contextual Information to a Measurement and Evaluation Framework. In proceed. of the IEEE Computer Society, 6th Int'l Conference on the Quality of Information and Communication Technology (QUATIC'07) Lisbon, Portugal, pp. 154-163. (2007).
11. H. Chen and Z. Wu, On Case-Based Knowledge Sharing in Semantic Web, of XV International Conference on Tools with Artificial Intelligence, California, 2003.
12. Schwabe, G.; Hausmann, D. The Use of Software-Tools for Case-Based-Reasoning Methods. Department of Computer Science, Dresden University of Technology. Dresden, Germany. Disponible en www.informatik.uni-bremen.de/~hausmann/papers/use_of_tools.ps. (2004).
13. S. Lindstaedt, M. Strohmaier, H. Rollett, J. Hrastnik, K. Bruhnsen, G. Droschl and M. Gerold, KMap: Providing Orientation for Practitioners when Introducing Knowledge Management of Practical Aspects of Knowledge Management: 4th International Conference, PAKM 2002 Vienna, Austria, December 2--3, 2002 Proceedings, vol. 2569, D. Karagiannis y U. Reimer, Edits., Berlin, Heidelberg, Springer Berlin Heidelberg, 2002, pp. 2-13.
14. W. Karl-Heinz, A Case Based Reasoning Approach for Answer Reranking in Question Answering, CoRR, vol. abs/1503.02917, 2015.
15. Sankar, K.; Simon C.K. Foundations of Soft Case-Based Reasoning. Wiley Series on Intelligent Systems. ISBN 0471086350. (2004).
16. M. Martín and M. Diván, Case Based Organizational Memory for Processing Architecture based on Measurement Metadata. Proc. 5th International Conference on Reliability, Infocom Technologies and Optimization (ICRITO), Noida, India, 2016.
17. M Diván and M Martín, "Applications of Case Based Organizational Memory Supported by the PabMM Architecture" in Advances in Science, Technology and Engineering Systems Journal. ISSN 2415-6698., 2017.
18. FIBARO website. www.fibaro.com