RESEARCH ARTICLE OPEN ACCESS

Machine Learning in Smart Home Systems

Martín María de los Ángeles*, Martínez Natalí **
(Engineering School National University of La Pampa, Argentina.)
*** (Engineering School National University of La Pampa, Argentina.)*

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 a Machine Learning system that manages 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 proposal it is shown that is possible to improve the management of the data stream captured by the devices, providing them with interoperability through the use of metadata, so that "intelligent modules" process the acquired knowledge automatically, with little intervention from the humans. Specifically, intends to use a Processing Architecture based on Measurement Metadata (PAbMM) and AWS Machine Learning and AWS IoT technology, to intelligently establish the environmental conditions of the home or trigger alarms, in order to improve more comfort, security and reduce energy consumption, on demand or proactively.

***Keywords* -** Machine Learning, Data Stream processing, Smart Home, IoT, Metadata.

Date of Submission: 24-07-2021 Date of Acceptance: 09-08-2021

I. 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 smart home centers, to make better use of the information obtained [1].

Specifically, once the data from the sensors has been captured, it is convenient to structure them in a acknowledge 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.

In this proposal it is shown the design of software modules to support the development of smart house systems with the added value of recommendation 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 AWS Machine Learning services [2], 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.

In order to transform the data captured by the sensors into structured, normalized and appropriate information for machine learning training systems, it is proposed to use an adaptation of the Processing Architecture based on Measurement Metadata (PAbMM), already designed and prototyped by the author in [3].

Nowadays, there are processing architectures which allows the real-time data processing through configurable topologies. In this type of architectures, you can dynamically define the processing topology over the data streams and adjust it to different computation necessities, being possible delegating the data structural definition and its meaning inside of the application. In this context, I summarize to the Processing Architecture based on Measurement Metadata (PAbMM) [3], which supported by the framework for measuring and evaluating called C-INCAMI (Context-Information Need, Concept model, Attribute, Metric and Indicator) [4], incorporates metadata to the measurement process, promoting repeatability, comparability and consistency.

The data streams are structured under C-INCAMI, which allows gather data and metadata jointly inside the same stream. In this metadata, we can describe the measurement context additionally to the entity under measurement, which permits avoid analyzing the measure in isolation way. The knowledge Memory that integrates the Processing Architecture based on Measurement Metadata serve as base for to be used in recommender systems in decision making processes.

Thus, data, information, and knowledge from heterogeneous and distributed sources can be automatically and semantically processable by applications, for instance, an 'intelligent' recommendation system to support a more effective decision-making process.

This article is organized in four sections. The Section 2 outlines the Processing Architecture based on Measurement Metadata. The Section 3 outlines the Machine Learning based recommendation system proposed. The Section 4 summarizes the conclusions.

II. ARCHITECTURE BASED ON MEASUREMENT METADATA

The PAbMM is a manager of semi-structured measurements streams, enriched with metadata supported by C-INCAMI, specialized in M&E projects, which incorporates detective and predictive behavior at online with the ability to manage and provide large volumes of data on demand. Moreover, unlike other strategies for the processing of data streams, with the addition of metadata, PAbMM is reliable for guiding the processing of the measures from heterogeneous sources, analyzing each one in context of origin.

As shown on the right side of the Fig.1, the conceptual model in terms of stream processing it is as follows. The measurements are generated in the heterogeneous data sources (for example, Zigbee or Z-Wave 700 sensors), which supply a module called Measurements Adapter (MA in Fig. 1). MA incorporates together with the measured values, the metadata (e.g. metric ID, context property ID, etc.) associated to each data source into the stream in order to transmit measures to the Gathering Function (GF).

Measurements Adapters can usually be embedded in sensor devices, but also can be embedded in any computing device associated to data, just as are the home centers.

Regarding the semantic ground for measurement and evaluation (M&E), the CINCAMI conceptual framework is built on an ontology which includes the concepts and relationships needed to specify data and metadata for any M&E project; in order to promote consistence and comparability of results. Unlike other data stream processing strategies, PAbMM is able to support the appropriate processing of measures generated from heterogeneous data sources thanks to the included metadata. In this way, each measure is analyzed considering its semantic and context as per the formal definition of each M&E project

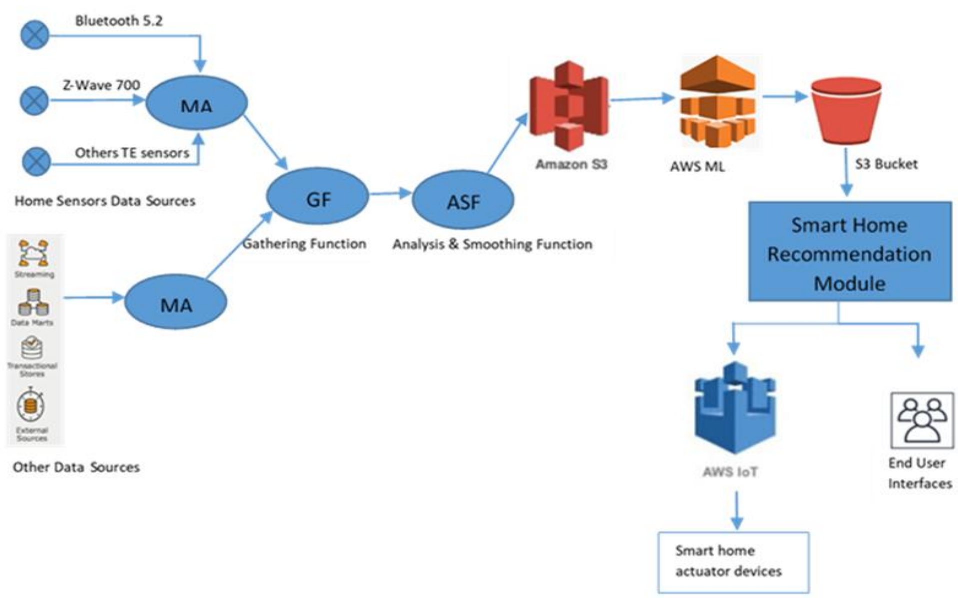

Figure 2. General architecture of the proposed ML-based recommendation system.

These homogenized data are reported to the central gathering function (Gathering Function -GF in Fig.1). In turn, GF introduces streams into a buffer organized by trace groups, this organization allows further consistent statistical analysis at trace group level, without representing additional processing loads. Within each trace group, the organization of measures is tracked by metric. Also, homogeneous comparisons of attributes can be made for different trace groups.

After the measures are organized in the buffer, the Analysis & Smoothing Function (ASF in Fig. 1) component performs a set of statistical analysis on the stream in order to detect deviations or problems with data, considering its formal definition. Then a descriptive, correlation and principal components analysis is applied. These techniques allow detecting inconsistent situations, trends, correlations, and/or identifying system components that incorporate more variability.

Once the data streams has been analyzed to facilitate the processing of so much information from heterogeneous sources, the structured data streams, is incorporated into a data repository S3 with measurements and metadata, and remains available to meet requests for services associated with data on historical measurement.

From a technological point of view, is possible the management of the S3 repositories [5] in the distributed computation context, the data provision to thirds by subscription, and maintaining the real-time measurement processing in parallel. Thus, we prioritize the use of the AWS Cloud computing technology with the aim of promoting the extensibility, dynamism and broadcast of the architecture.

Thus, PAbMM can be viewed as a simple topology[6], which allows us to manage the changes, make the adjustments/fixes that being necessaries, and make easier its interoperability. In service terms, our focus is in: a) Collect the measurements in structured and interoperable way by C-INCAMI, b) Store all the measures under a single repository, c) Use the stored measures to answer queries or to support a cluster/classification analysis, d) Provide data in real-time and historical on demand, e) Get and maintain a smart house knowledge memory.

III. MACHINE LEARNING BASED RECOMMENDATION SYSTEM

To implement the proposed idea an AWS Machine Learning pipeline is used, as shown in the fig. 1. One of the reasons for choosing this solution is that smart home users commonly have no knowledge in information technology or machine learning processes. They are people with no computer science background. Therefore, automation was a key design decision to be able to deliver not just a packaged solution, but also an automated pipeline without requiring manual intervention every week. AWS services enable the generation of powerful training models transparently for the end user.

The first node in the pipeline is the S3 container with the data streams obtained from the home sensors and stored in a structured way by the PAbMM architecture. This large data set will serve as input to the AWS services from which our models will be trained.

The next step in the pipeline is a SageMaker [7] instance of AWS ML. This platform allows to build, train and implement machine learning models at any scale. This is where data engineering and machine learning take place.

An additional S3 bucket was enough to store our temporary and final outputs after running our models. In this way, the generation of predictions based on the AML model are stored in an S3 bucket for later use by the smart home recommendation system.

A recommendation module, which can be installed in smart home centers or implemented as a cloud service, will be responsible for generating the recommendations and/or acting when necessary, in order to improve the comfort and safety of the home and reduce electricity consumption costs.

This module will consult as input information a configuration profile, which can be modified according to the recommendation preferences of the different users (for example: recommendation level, enabling actions, saving preferences, etc.) in addition to the predictions generated by the ML model. In each run cycle of the model, this module will be able to present the recommendations to the user to act manually on the devices through their controllers, or better yet, act directly on the controllers using the AWS IoT service [8].

IV. CONCLUSIONS

This report presents an innovative smart home idea that provides original solutions to solve the challenges faced by smart home users. The main challenge that this proposal solves is to make the devices “learn” from the preferences, customs and use of devices of the inhabitants of the home. Although ML technology has made important advances in recent years, to obtain good recommendations it is necessary to provide the ML service with standardized data, consistently labeled and corrected so that they can be compared and obtain good predictions[9,10].

IMHO, there are currently no smart home devices that are provided with software to support the stated challenge, be it for consistent data collection, such as smart home centers that have machine learning based recommendation systems.

In this work a comprehensive solution to such challenges was shown, illustrated in the architecture of figure 1. The first part of the solution referred to the acquisition of data using the PAbMM Architecture, is a work elaborated by the author that has been prototyped and successfully applied in different knowledge management solutions [1].

The following was a discussion of how AWS cloud services can be leveraged to empower smart home devices with greater automation through machine learning. The proposed AWS ML pipeline is a first version of the design in order to illustrate the idea, but if this solution is implemented, a more detailed review of the architecture shown should be done. The proposed recommendation module can be installed in smart home centers, to make better use of the so much information captured by the sensors.

REFERENCES

[1] M. Martín and M Diván, "Applications of Case Based Organizational Memory Supported by the PabMM Architecture" in Advances in Science, Technology and Engineering Systems Journal. ISSN 2415-6698., 2017.

[2] Amazon Machine Learning site. https://docs.aws.amazon.com/machine-learning/index.html.

[3] 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.

[4] 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.

[5] Amazon Simple Storage Service site. https://docs.aws.amazon.com/s3/index.html.

[6] 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.

[7] Amazon SageMaker site. https://docs.aws.amazon.com/sagemaker/index.html.

[8] Amazon IoT site. https://docs.aws.amazon.com/iot/.

[9] 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.

[10] M Lee, M Lee, S Hur, and I Kim, Load Adaptive and Fault Tolerant Distributed Stream Processing System for Explosive Stream Data, Transactions on Advanced Communications Technology, vol. 5, *no. 1*, pp. 745-751, 2016.