

Paving the Way for a Real-Time Context-Aware Predictive Architecture

David Corral-Plaza

School of Engineering, University of Cádiz,
Avda. de la Universidad de Cádiz 10, 11519, Puerto Real, Cádiz (Spain)
david.corral@uca.es

Supervisors: Guadalupe Ortiz, Juan Boubeta-Puig
Department of Computer Science and Engineering, University of Cádiz,
Avda. de la Universidad de Cádiz 10, 11519, Puerto Real, Cádiz (Spain)
{guadalupe.ortiz, juan.boubeta}@uca.es

Abstract. Internet of Things society generates and needs to consume huge amounts of data in a demanding context-aware scenario. Such exponentially growing data sources require the use of novel processing methodologies, technologies and tools to facilitate data processing in order to detect and prevent situations of interest for the users in their particular context. To solve this issue, we propose an architecture which making use of emerging technologies and cloud platforms can process huge amounts of heterogeneous data and promptly alert users of relevant situations for a particular domain according to their context. Last, but not least, we will provide a graphical tool for domain experts to easily model, automatically generate code and deploy the situations to be detected and the actions to be taken in consequence. The proposal will be evaluated through a real case study related to air quality monitoring and lung diseases in collaboration with a doctor specialist on lung diseases of a public hospital.

Keywords: Prediction, Context-awareness, Internet of Things, Event-driven Service-oriented Architectures.

1 Motivation and Problem Statement

Nowadays, society has numerous methods for consumption, production and exchange of huge amounts of data. This data is mainly originated in the multitude of existent devices and channels: social networks, smart devices such as smartphones, tablets or wearables, Internet of Things (IoT) devices, et cetera. Such exponentially growing data sources require the use of processing methodologies, and tools different from conventional ones in order to fast profit from the relevant information obtained. Even more, in the smart world [1] context awareness [2] plays a highly relevant role which must be taken into account when providing the information to the final recipient.

Therefore, it is vital to be able to process such large amounts of data together with users' context [3], in order to detect real-time situations of interests that are relevant

for a particular user, context and domain. Moreover, we have to follow real-time analysis not only to detect situations of interest, but also to predict them so that we can take necessary actions in advance to prevent user from undesirable situations.

The problem is that, currently, an efficient integration of all the stages implied in this process is not provided. We can find systems to process big amounts of data in real time [4, 5], systems which are aware of the user's context [6], and systems which apply prediction algorithms to vast amounts of data [7]. However, when trying to integrate all of them, we have several problems: (1) the *variety* of data formats required to be managed by an unique system, given the heterogeneity of the IoT sensing data; (2) the handicap of *predicting* situations of interest from such heterogeneous data, which is related to different domains and contexts; (3) the *volume* and *velocity* of this data, becoming a problem to achieve system scalability and performance gains; and (4) the *domain experts* that do not have enough technical knowledge to define what they need to detect or predict (the *data value*) with current tools. These problems lead us to the research challenges explained in the following section.

2 Research challenges

We will deal with the following research challenges:

1. First of all, data will have to be gathered and processed in real time. As discussed before, the source where these data come from will be very varied — for example, social networks, smartphone sensors, IoT sensors, et cetera —; therefore, we will need a means to homogenize all these data with an appropriate structure to be used in the data analysis systems and an easy procedure to add new sources.
2. Besides, we need a system which able to constantly detect the user context and to provide a consistent response according to real-time circumstances of a given user.
3. Not only that, a collaborative architecture is required so that several parties can provide context data and relevant events to enrich the relevant available data.
4. Suitable methodologies and tools to process and correlate huge amounts of data in order to detect and predict situations of interest in real time for a particular user in a particular domain. These should be integrated in our system architecture. It is essential that the architecture is efficient and scalable [8].
5. Finally, domain experts should be able to easily design the situations to be detected and predicted in real time, as well as the actions to be taken consequently, depending on a particular set of incoming data and context information.

3 Proposed Solution

We envision a holistic Event-Driven Service-Oriented Architecture (ED-SOA). Such envisioned architecture, shown in **Fig. 1**, should contain the following modules:

- Firstly, *Data producers* should gather data from several sources (databases, IoT sensors, social networks, et cetera) and send them to the data collector.

- Secondly, *Data Collector* follows the necessary transformations so that the information received can be used in the following phases of our solution. It is an intermediate layer that performs a process of homogenization since information will most probably be received in different formats and structures in most scenarios.
- Thirdly, *Data Processing* should provide Complex Event Processing (CEP), context-awareness and prediction module. Initially we bet for FIWARE [9], a Platform as a Service (PaaS) development tool in the cloud.
- Fourthly, we have *Data Consumers*, which can be databases, end users or additional endpoints which pave the way for the collaborative architecture. Such data consumers communicate with the previous module through a REST interface, however additional protocols might be required.
- Finally, in the bottom of **Fig. 1**, we can see the graphical modeling tool for pattern and actions definition for FIWARE. Such tool is expected to be an extension of MEdit4CEP [10], with the goal of easily modeling the different events and patterns expected in a specific application domain and being able to deploy them, automatically, in the processing system that is currently running in FIWARE.

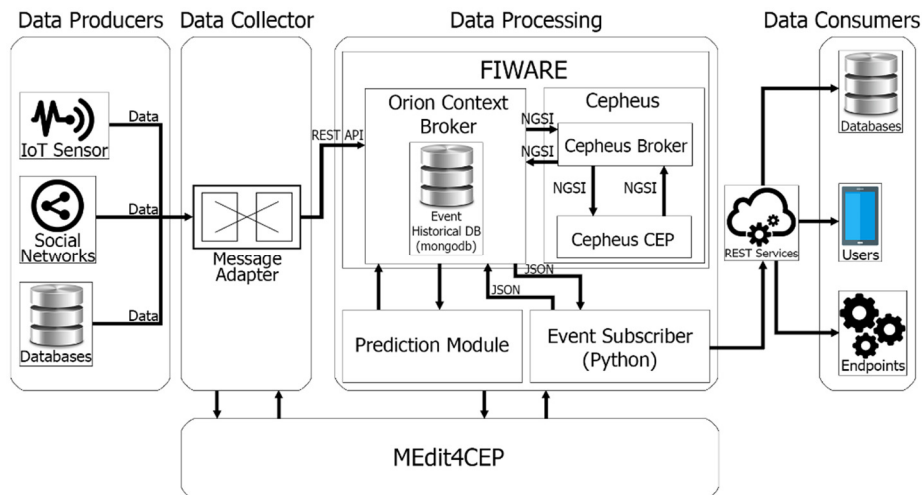


Fig. 1. Proposed architecture

4 Preliminary Results

By the time being, we have implemented and tested part of the proposed architecture:

1. The data producers integrated are databases, IoT sensors, IoT platforms and message queues. Integrating social networks is our following step to be done here.
2. Currently, data collector is being improved; we initially implemented such data collector and homogenization process through an Enterprise Service Bus (ESB);

however our experience, supported also with other results in the research group [3, 11], shows that the ESB decreases the system performance.

3. Among the different enablers existing in the FIWARE catalog, we already started using Orion Context Broker and Cepheus: Orion has become the brain of our application and control all the information that the different modules receive and emit inside FIWARE. Cepheus is a CEP engine which will let us define the event patterns to be monitored. We did not start yet working on the prediction module.
4. In the fourth stage we included databases and mobile applications and we are working on the inclusion of message queues to foster the collaboration [11].
5. The work on the domain experts graphical tool is not yet started.

In order to test the proposed architecture, we particularized it for air quality domain. We have analyzed and reported air observations made around the Andalusian territory for several months. For this purpose we have connected the data provided from Andalusian air quality stations as the main data source for our system, we have created an Android app which monitors the user context and we have provided personalized alerts according to the sensed air quality and user context. We also measured the performance of the system and even though the system was clearly efficient for the case study in question. As a result, we detected that our initial implementation was not scalable enough, reason why we are working now on improving it.

5 Related Work

Concerning prediction architectures where SOA and CEP are used, R. Mousheimish et al. [7] propose one which is able to generate CEP rules for any person, without the knowledge of a domain expert. The authors extend their work in [12], for an artworks transportation case study and use an additional framework to monitor relevant values and predict if they are going to take undesired values according to a previously specified Service Level Agreement. The main limitation of their proposals is that they do not take into account several inputs sources.

Concerning proposals based on FIWARE, Fazio et al. [6] make use of FIWARE to design a real e-health remote patient monitoring architecture, which allow caregivers to improve remote assistance to patients at home. Wolfert et al. [4] describe how FIWARE has been applied to the smart farming domain, given support to tackle the data chain of big data applications: data capture, data storage, data transfer, data transformation, data analytics and data marketing. Fernández et al. [5] have developed SmartPort, a FIWARE-based platform for sensor data monitoring in a seaport located in Gran Canaria, Spain. Comparing these proposals with our proposed work, most of them do not benefit from using the CEP technology and Cepheus component to automatically detect meaningful patterns in real time; Fazio et al.'s one is an exception. In addition, they do not provide prediction mechanisms.

6 Evaluation Plan

As could not be otherwise, during the development of the PhD, we will continuously review the literature related to the problem stated and the proposed solutions.

On the other hand, we will carry out different experiments in different application domains to evaluate in an empirical way the performance of the proposed architecture and solution. Among others, we will go on with the air quality scenario.

In this same line, we expect to test a system to detect air quality alerts with people suffering from all types of lung diseases in collaboration with Dr. Carmen Maza Ortega, a specialist in lung diseases at the *Hospital Universitario de Puerto Real* (Spain), where additional sources for information can be integrated in the system.

7 Planned Timeline

In this section we schedule the tasks to be performed during the thirty-six months expected for the development of the PhD, which we roughly identify between November 2017 and November 2020 (see **Fig. 2**).

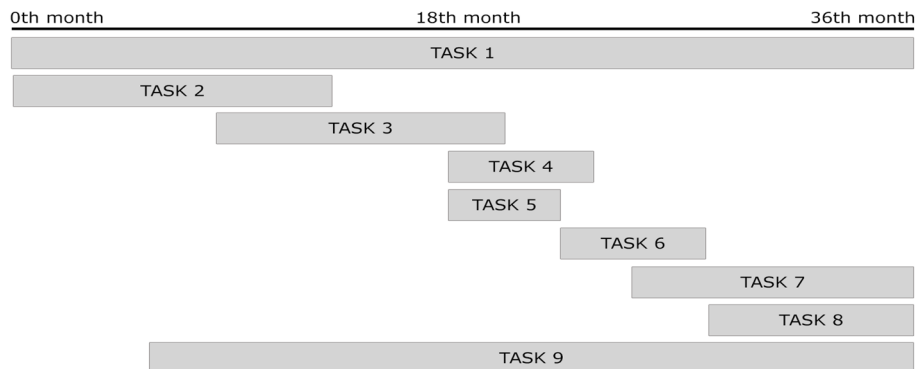


Fig. 2. Estimated timeline for the PhD

Task 1. Reviewing literature related to problem statement and proposed solution.

Task 2. Studying emerging and well-established technologies and tools for data processing with the aim of detecting and predicting particular situations of interest.

Task 3. Developing the proposed architecture according to previous tasks.

Task 4. Evaluating the architecture through a case study in a relevant domain.

Task 5. Facilitating the addition of new sources and formats.

Task 6. Incorporating MEdit4CEP or an alternative tool in the architecture to permit pattern and action graphical design and code generation and deployment.

Task 7. Evaluating the proposed architecture thorough a real case study.

Task 8. Writing and defending the PhD

Task 9. Disseminating the research results in conferences and journals.

8 Conclusions

We have outlined a starting PhD focused on providing real-time context-aware detection and prediction under the smart world demanding scenario. We envision a solution based on an ED-SOA in the cloud combining several cutting-edge technologies which will provide us with early context-aware notifications and predictions in a particular domain of application. The system envisioned is designed to be scalable and highly configurable, so that it can integrate an undefined and heterogeneous number of data sources. Moreover, the system will be easily applicable to several domains, and domain experts will be provided with a graphical modeling tool which will prevent them from coding and configuration issues, facilitating the architecture widespread use.

References

1. Liu, H., Ning, H., Mu, Q., Zheng, Y., Zeng, J., Yang, L.T., Huang, R., Ma, J.: A review of the smart world. *Future Gener. Comput. Syst.* (2017).
2. Abowd, G.D., Dey, A.K., Brown, P.J., Davies, N., Smith, M., Steggles, P.: Towards a Better Understanding of Context and Context-Awareness. Presented at the 1st International Symposium on Handheld and Ubiquitous Computing, Karlsruhe, Germany (1999).
3. Garcia de Prado, A., Ortiz, G., Boubeta-Puig, J.: CARED-SOA: a Context-AwaRe Event-Driven Service Oriented Architecture. *IEEE Access.* 5, 1–18 (2017).
4. Wolfert, S., Ge, L., Verdouw, C., Bogaardt, M.-J.: Big Data in Smart Farming – A review. *Agric. Syst.* 153, 69–80 (2017).
5. Fernández, P., Santana, J.M., Ortega, S., Trujillo, A., Suárez, J.P., Domínguez, C., Santana, J., Sánchez, A.: SmartPort: A Platform for Sensor Data Monitoring in a Seaport Based on FIWARE. *Sensors.* 16, 417 (2016).
6. Fazio, M., Celesti, A., Márquez, F.G., Glikson, A., Villari, M.: Exploiting the FIWARE cloud platform to develop a remote patient monitoring system. In: 2015 IEEE Symposium on Computers and Communication (ISCC). pp. 264–270 (2015).
7. Mousheimish, R., Taher, Y., Zeitouni, K.: autoCEP: Automatic Learning of Predictive Rules for Complex Event Processing. In: Sheng, Q.Z., Stroulia, E., Tata, S., and Bhiri, S. (eds.) *Service-Oriented Computing*. pp. 586–593. Springer International Publishing, Cham (2016).
8. Papazoglou, M.: Service-oriented computing: concepts, characteristics and directions. Presented at the July 6 (2012).
9. Fiware, <http://www.fiware.org/>.
10. Boubeta-Puig, J., Ortiz, G., Medina-Bulo, I.: MEdit4CEP: A model-driven solution for real-time decision making in SOA 2.0. *Knowl.-Based Syst.* 89, 97–112 (2015).
11. Garcia-de-Prado, A., Ortiz, G., Boubeta-Puig, J.: COLLECT: COLLaborativE ConText-aware service oriented architecture for intelligent decision-making in the Internet of Things. *Expert Syst. Appl.* 85, 231–248 (2017).
12. Mousheimish, R., Taher, Y., Zeitouni, K., Dubus, M.: PACT-ART: Enrichment, Data Mining, and Complex Event Processing in the Internet of Cultural Things. Presented at the (2016).