A Cognitive Enabled, Edge-Computing Architecture for Future Generation IoT Environments

Franco Cicirelli, Antonio Guerrieri, Giandomenico Spezzano, Andrea Vinci

CNR - National Research Council of Italy
Institute for High Performance Computing and Networking (ICAR)
Via P. Bucci 7-11C - 87036 Rende(CS) - Italy

Email: cicirelli@icar.cnr.it, guerrieri@icar.cnr.it, spezzano@icar.cnr.it, vinci@icar.cnr.it

Abstract—Nowadays, Smart Environments (SEs) are pervasively deployed in buildings (e.g., houses, schools, and offices) and outdoor environments with the goal of improving the quality of life of their inhabitants. SEs are usually designed and developed by using well-suited architectures and platforms having the aim of simplifying and making straightforward the SE implementation. Up to now, SEs are mostly reactive and, in some ways, proactive. Current research efforts are devoted to making such environments cognitive, i.e., able to automatically adapt and adhere to the possible changes in users’ needs and behaviors. Anyway, in this field, the development of SEs is still in its infancy. In this direction, the paper proposes a novel Cognitive-enabled, Edge-based Internet of Things (CEIoT) architecture, purposely designed to develop cognitive IoT-based SEs. Such architecture wants to overcome some limitations arising during the usage of common SE platforms and architectures. CEIoT introduces some abstractions ranging from the “in-platform” implementation of decentralized cognitive algorithms to the realization of smart data aggregations.

Index Terms—Smart Environments, Cognitive Internet of Things, Architectures, Edge Computing

I. INTRODUCTION

SEs are instrumented physical environments whose goal is that of offering ICT services to their inhabitants so as to improving their quality of life and assisting people in daily-life activities [1]. An SE is a cyber-physical system able to sense and act upon the physical environment in which it operates by means of suitable sensors and actuator devices.

The design and implementation of significant SEs are a challenging and complex task [2]. This is because many issues require to be taken into account during the development of such systems. SEs are highly dynamic and naturally distributed, they gather and have to put together heterogeneous technologies and protocols with a huge number of hardware and software components. Moreover, the ever-increasing complexity of the offered services leads to the exploitation of cognitive-based technologies [3] aiming at realizing systems able to automatically adapt to changes in user’s behavior and anticipating and predicting their activities.

In order to deal with the foregoing aspects, the use of well-suited architectures, platforms, and methodological approaches becomes of utmost importance. Among others, the advantages in using them are: (i) architectures and methodological approaches promote modularity and separation of concerns and offer basic abstraction entities which can be exploited to handle complexity and reduce the efforts during the design phase; (ii) methodological approaches promote a more structured design phase and permit to exploit the competencies gained by expert designers in the domain of SEs, thus reducing the possibility of running into recurring errors; (iii) platforms permit the development of applications by exploiting ready-to-use solutions and services tailored to the exploitation of a specific architecture and methodological approach.

This paper proposes CEIoT, a Cognitive-enabled Edge-based Internet of Things architecture, which promotes the development of IoT-based SEs. CEIoT comes from the previous experiences gained by the authors in research projects [2], [4]–[7] focusing on the design and implementation of SEs like that of Smart Office [4], Smart Museum [5], Smart Street [6], and large-scale SEs in general [7]. As an original contribution, CEIoT introduces (i) some abstractions and modules aiming at exploiting cognitive computing in a distributed context made by edge-based and cloud-based computational nodes; (ii) some primitives and abstractions which can be exploited to aggregate/filter information and operate data-fusion on data coming from the sensors deployed in an SE. For the latter point, the goal is that of relieving developers from the burden of performing repeated data processing in a distributed and decentralized environment in which, e.g., the use of swarm intelligence [8] and/or gossip-based algorithms [9] can be exploited. CEIoT extends and refines an already-existing layered architecture previously proposed by the authors, upon which the iSapiens platform has been developed. iSapiens is proposed as a candidate platform to be enhanced in order to implement the CEIoT architecture.

The structure of the paper is reported in the following. Section II introduces the background concepts and experiences which motivate CEIoT, whereas Section III introduces some related works. Section IV proposes CEIoT. Finally, conclusions and future works are reported in Section V.
II. BACKGROUND AND LESSON LEARNED

This section first introduces the iSapiens platform [6] and the architecture on top of which it was developed. The core services and offered functionalities of the platform are also described. After that, an overview of some SEs developed by using iSapiens is provided. Finally, a discussion about lessons learned related to the use of the platform and the design and implementation of real SEs is provided.

A. A Platform for Smart Environments: iSapiens

iSapiens is an IoT-based [10] platform conceived with the aim of offering specific tools for the development of distributed and pervasive SEs [11], [12] and, more in general, cyber-physical systems [13].

iSapiens, currently implemented in Java, relies on the agent metaphor and permits the exploitation of both the edge computing [14], for the execution of applications that are not very demanding in terms of computation or storage, and the cloud computing, useful when a huge amount of resources are needed. Moreover, the platform implements Virtual Objects to hide physical devices and protocols heterogeneity to the higher levels so permitting the developers to be focused only on the functionalities/services the physical devices should offer.

Agent paradigm has been chosen in iSapiens since it is well suited for the implementation of distributed and pervasive SEs [6]. Agents can execute near to the devices they need to control/manage thus implementing the edge computation and enabling real-time analysis on the data gathered on the single nodes. Furthermore, agents can execute on the cloud for implementing out-of-the-edge services (e.g., data mining or data storage) and taking advantages of the features and benefits of the cloud.

iSapiens has been enhanced with the Social Internet of Things (SIoT) paradigm [7]. The goal of SIoT is that of combining the benefits deriving from the convergence of technological solutions for the IoT and the social networks domains [15]. In particular, the exploitation of a social network in the IoT domain has proven to foster resource visibility, discovery, object reputation assessment, source crowding, and service composition.

The layers of the iSapiens architecture are shown in Figure 1 and are briefly explained in the following:

- the physical device layer, which comprehends (heterogeneous) physical objects spanning from simple sensors and actuators to smart objects. At this layer, an object/device exposes its functionalities and can expose a Social Senser, which allows interactions with SIoT [7];
- the in-network edge computation layer, which is composed of several computing nodes deployed in the physical environment that requires to be turned into an SE. Every node hosts an iSapiens server that is composed of a Virtual Object Container and an Agent Server. The Virtual Object Container is in charge of managing the local Virtual Objects, the Agent Server instead runs the agents that implement the application logic in the system, and that interact with the local Virtual Objects. All the computing nodes, where an iSapiens server run, are called iSapiens nodes. An iSapiens Node can also host a Social Object Container, i.e., the component that provides functionalities enabling the integration among physical objects and the SIoT. In particular, the Social Object Container manages Social Objects, which are the social counterpart of the Virtual Objects.
- the off-network computation layer, which provides high-level (cloud) services ranging from storage to computation and third-party services such as weather forecasting. This layer can also host the SIoT Cloud component. Through the latter component, iSapiens agents and iSapiens physical objects can inter-operate with entities belonging to other platforms, provided that these entities are SIoT compliant. The iSapiens components (i.e., Physical Objects, Virtual Objects, agents, or nodes) can be added/removed/updated at any time.

B. Developed Applications and Lessons Learned

This section gives an overview of some SEs built on top of iSapiens with the aim of highlighting commonly raised issues and the approach used to address them.

The list of the developed SEs follows:

- in [6] a Smart Street application is presented that quickly detects and reacts to anomalies by exploiting a decentralized algorithm for computing local and global aggregations on data coming from sensors deployed on a street in the city of Cosenza (Italy). In such an application, all the computation is done at the edge nodes. Local storage facilities and time series analysis are exploited for detecting anomalies.
- in [9] a distributed real-time approach for mitigating the combined sewer overflow and flooding in urban drainage networks is presented. Here, a gossip-based algorithm is exploited in order to evaluate and share an estimation of some aggregate measures at whole drainage network and sub-networks level. Such estimations permit the decentral-
ized control of some gates devoted to minimizing flooding and overflow phenomena.

- in [4] is presented the development of a Smart Office devoted to forecast some usual office activities and to appropriately adapt the office environmental conditions to such activities (e.g., in terms of thermal comfort and lightening).

The system was designed by exploiting a methodological approach suitable for the development of IoT-based ecosystems which exhibit cognitive behavior.

- in [5] an open-air Smart Museum is presented in which the statues inside the museum were equipped with “virtual senses” realized through sensors. The goal is that of making the artworks reactive to what happens around them, and, in particular, to detect dangerous situations and react to them.

C. Lessons Learned

The use of the iSapiens platform highly simplified the development of the above listed SEs. In particular, the platform supported the resolution to some frequent aspects like those related to the management of heterogeneous hardware devices and communication protocols, and the exploitation of the distributed computational resources scattered at the edge of the network and in the cloud. Moreover, the exploitation of the agent metaphor naturally permitted to design and implement dynamic systems where location-independent functionalities can be offered to the final users. Anyway, from the use of the platform, it also emerged that other important issues remain completely application dependent, and they are a matter of the software developers only. Examples of such issues are: (i) the aggregation and filtering of data coming from data sources (e.g., sensors and/or application agents), (ii) the implementation of complex agents’ behaviors (e.g., cognitive behaviours), (iii) the management and optimization of the resources available in the system.

The goal of the architecture proposed in this paper is that of fostering the development of platforms addressing the above issues so as to relieve application developers from dealing with such tasks.

III. RELATED WORK

In the last few years, the big dissemination of IoT-based Smart Environments is dramatically growing the pervasive presence of heterogeneous sensors, actuators, and smart objects among us. Such devices have to interact/cooperate among them to give particular services to users and/or to reach some specific (shared) goals. This is the reason for which research and industry have done several steps towards the development of IoT middlewares and platforms for SEs creation [16], [17].

In [18] authors implemented an SE framework that favors the interactions among Bluetooth mobile devices so to allow the dynamic discovery and exploitation of such devices.

The work in [19] presents a framework, specifically focused on the system extensibility, which supports users in the creation of SE applications that are context-aware.

In [20] authors started from the concept of active spaces to conceive a distributed middleware for the management of devices scattered in physical environments. In this work, active spaces consist of programmable environments to which such devices are connected.

In the manuscript at [21], authors presented a framework for SEs which provides several abstractions for their modeling. The purpose of this work is the realization of context-sensing and context-aware applications.

A middleware which favors the realization of smart homes based on the sharing of contextual information among the devices in such SE is presented in [22].

In [16] is introduced the Syndesi framework which has been conceived to gather user data to profile them. This profiling allows creating customized services for all users. Syndesi has been designed to support wireless sensors and actuators.

The paper in [23] shows an IoT smart-city platform for the managing and control of sensors and actuators called Sentilo. Sentilo is designed to use cloud computing and Big Data tools to manage and analyze data from distributed sensors.

In [24] is presented one of the results of the EPIC (European Platform for Intelligent Cities) project. It consists of an IoT middleware designed to tackle problems such as interoperability, heterogeneity, (re)configurability, and extensibility.

The work in [25] presents a platform for smart city environments which comprehends abstractions for the data collection and analysis, data aggregation, data mining, and storage.

Even though all these works are valuable, they all lack of specific mechanisms to support proactivity in a deployed SE. Moreover, they do not provide users with already implemented algorithms useful for developing proactive environments. In such a direction, the authors of [4] present a methodological approach for the realization of cognitive IoT SEs. Even though they work on the design of cognitive SEs, they develop cognitive behaviors on top of an already existing IoT platform.

IBM Engineers have so far developed the Watson [26] platform for realizing cognitive computing. Watson has been used to develop many applications requiring high power computing. As an example, in [27], it has been used to make cognitive computing on big data related to life science research. Even though it is a very powerful solution, Watson is not usable in every context since it is mostly cloud-based. Moreover, it is a licensed, closed, and expensive product and not always it can be considered as the best solution for prototypes or research products. At the best of our knowledge, Watson is the only platform, to date, allowing the fast prototyping of cognitive applications. This paper presents an open architecture that wants to overcome the limitations of Watson by introducing abstractions and modules aiming at exploiting cognitive computing in a distributed context made by both edge- and cloud-based computational nodes. Moreover, the proposed architecture introduces some abstractions which can be exploited to aggregate/filter information and operate data-fusion on data coming from the sensors deployed in an SE.

IV. THE COGNITIVE ARCHITECTURE

As highlighted in sections II and III, there is a lack of platforms and architectures which provide components and
libraries purposely tailored to solve some of the previously-identified common tasks. For instance, there is the need for specific high-level components for supporting the developers in the realization of cognitive systems, and for the implementation of decentralized estimators on data gathered at global network or sub-networks level [28].

In this paper, we propose a Cognitive-enabled, Edge-based Internet of Things (CEIoT) architecture, whose main components are depicted in Figure 2. Figure 3 highlights, instead, a view of CIoT which is focused on computation location and system modules.

A. A General IoT Edge-Based Architecture

In order to highlight the edge nature of the architecture and the related issues, the bottom of Figure 3 lists the computing locations where the operations of the upper components are executed. Each location, namely Single Node, Networked Edge Nodes, and Cloud, offers different capabilities and rises different challenges.

Single Node. When a component needs to compute or provide facilities on a single edge node, it exploits data stored only at that node, it interacts with other components hosted at the same node and has a limited amount of memory and computational resources available.

Networked Edge Nodes. The operation of a component may need to involve a set of networked edge nodes, as an instance for executing operations in a decentralized fashion in order to achieve scalability. When a component has this requirement, it has to take into account the latency introduced by the communication between different nodes, the inherently distributed nature of the node network, and the topology of the network itself.

Cloud. Finally, a component can execute or offer services at the cloud level. Here, a large amount of computational and memory resources is available for performing high-demanding tasks or storing historical data. The communication latency introduced by the distance from the physical environment is greater with respect to the edge nodes, and the provided services need to be designed with scalability and elasticity properties in mind.

B. The Architecture System Modules

Given the computation locations, a set of modules is defined, each furnishing facilities for taking into account a specific issue with the aim of ease the developing of cognitive IoT environments. For each module, Figure 3 highlights the computation locations involved.

The Data Management module offers services for managing data storage, persistence, sharing, and permissions on all the possible locations. Services provided on a single node are related to information that needs to be stored on a single node and to be available by entities residing on the same node. Here, issues about limits in storage capacity of the edge nodes arise and need to be taken into account. When considering the networked edge nodes level, solutions need to be envisaged about distributed and decentralized storage, replica management, data querying, data retrieving, and bandwidth consumption. Finally, data management in the cloud offers virtually unlimited storage capabilities, but communication latency needs to be taken into account.

The Region Aggregation module offers high-level services related to aggregation and clustering of data/entities located on different edge-nodes or in the cloud. Here, the provided services and results are furnished in a transparent fashion with respect to the data producers identity, their location, and the edge network topology. As an example, an application entity can be interested in a region which computes and makes available, every five minutes, an estimation of the average of the temperature produced by the entities in its neighborhood. The application entity may be interested only on the aggregation results during time, or may be interested in communicating information with the involved entities without taking into account their identity, execution location and so on. This module needs to exploit algorithms and techniques which permit decentralized aggregation and data stream clustering, and can require computing on the networked edge nodes or in the cloud depending from data storage location, precision requirements or algorithm complexity.

The Cognitive module offers high-level services related to the approximation of the model underlying the behavior of an environment, a data source or the impact of specific actions on the environment itself. In other words, this module furnishes services aiming at permitting a system to learn the behavior of environments and users, and the way they react to events and actions. This can be exploited for forecasting, anomaly detection, user’s preferences understanding, or for planning actions which can change the environmental status to the desired one. This module offers services which can be based on unsupervised, semi-supervised learning techniques on data streams, time-series analysis, deep learning, and reinforcement learning techniques [29], [30]. It can involve computation on a single node or in the cloud depending from data location, time and precision requirements. While the approximation of a model is, in general, a task requiring a lot of computational and storage resources, the evaluation of a model can fit the computational resources available in a single node.

The Virtual Objects module offers a set of abstractions for coping with the devices heterogeneity of the IoT environments. The goal here is to furnish to the other components a common interface for exploiting devices and smart objects, which is independent of the plethora of different communication protocols, capabilities and behavioral constraints of different vendors devices. It works at a single node level, as it is responsible for mediates the connection between physical devices and computational nodes.

The Social Internet of Things module is devoted to permitting other components the exploitation of the Social Internet of Things paradigm, which leverages on technologies of both the Internet of Things and Social Network domains. Such module offers services for the dynamic creation, management and querying of social networks of things, discovered through the exploitation of opportunistic relationships based on mu-
tual communication, ownership, cooperation, and location. It operates on both single node and cloud levels, where the former is exploited for logging and listening communication between entities, which permits the creation of opportunistic relationships, and the latter for storing and querying the created social graphs [7].

The **Mobility Management** module offers solutions for the management of issues related to mobile devices and entities. Mobile entities rise problem involving naming, addressing policies, and management of dynamic systems where entities can join/leave the system with high frequency.

The **User Interface** module provides services for managing interaction with a user or a group of users. Such module explicitly maps the communication between a smart environment and the users, managing common behaviors and offering high-level functionalities to users.

The **Runtime Services** module offers functionalities related to the execution and the management of the application entities, spanning from the management of entity life-cycles, to the management of the communication between entities, the entity discovery services (e.g. yellow or white pages), and the entities execution environment.

The **Application Entities** can exploit all the previous modules, and are devoted to executing the application behavior of the smart environment, implementing the functionalities which directly impact the user and the smart environment.

Besides the listed module, a **Cross Platform Interoperability** component is required for providing abstractions and services managing, for each module, interoperability issues between a given platform and others. Such component should be rooted in the principles of voluntary interoperability, and methodologies like Inter-Meth [31].

**V. Conclusions and Future Work**

This paper proposed CEIoT, a Cognitive-enabled Edge-based Internet of Things architecture, purposely developed for the design of Smart Environments (SEs) which are able to exhibit cognitive behaviors. As an original contribution, CEIoT offers well-suited abstractions supporting cognitive computing in a distributed context, both for edge-based and cloud-based computational nodes, as well as some abstractions devoted to aggregate/filter information and operate data-fusion on data coming from sensors deployed in an SE. CEIoT comes from the experience gained in the context of research projects finalized to the realization of SEs and tries to overcome the limitations arisen in the use of architectures and platforms already known in the literature. As a reference platform, iSapiens was chosen as a candidate for implementing and supporting the core functionalities of CEIoT. On-going and future works are devoted to implement and support CEIoT concepts on top of the iSapiens platform and to exploit CEIoT functionalities in a significant case study in order to better highlight CEIoT benefits and effectiveness.

**Acknowledgment**

This work has been partially supported by the “COGITO - Sistema dinamico e cognitivo per consentire agli edifici di apprendere ed adattarsi” project, funded by the Italian Government (ARS01_00836) and by the “GLAMOUR - Green Learning and Adaptive Multi-interface IoT enabled devices through social inteReActions”, funded by POR Calabria FESR-FSE, Italy (CUP: J28C17000250006).

**References**


