Data Management Challenges for Smart Living

Devis Bianchini(✉), Valeria De Antonellis, Michele Melchiori,
Paolo Bellagente, and Stefano Rinaldi

Department of Information Engineering, University of Brescia,
Via Branze 38, 25123 Brescia, Italy
{devis.bianchini,valeria.deantonellis,michele.melchiori,
p.bellagente,stefano.rinaldi}@unibs.it

Abstract. An information infrastructure for modern Smart Cities must be able to integrate data from multiple heterogeneous sources such as private and public energy consumption, garbage collection and environmental conditions (pollution, citizens' safety and security). In this context, citizens themselves become providers of data, in the form of comments, opinions and suggestions that should be integrated within the infrastructure. A vast amount of data must be collected, organized and analysed to extract useful insights that can be transformed into actions aimed at improving the quality of life in the city. In this paper, we discuss data management issues to be addressed for bringing benefits to different categories of stakeholders in a Smart City, ranging from citizens to the Public Administration and energy providers.

Keywords: Data management issues · Information infrastructure · Smart Cities · Brescia Smart Living

1 Introduction

Improving the quality of life of citizens and supporting Public Administrations (PA) and energy providers to deliver innnovative services through the adoption of modern technologies are primary goals of modern Smart Cities [1]. Citizens become actively part of their city life, providing suggestions, opinions and comments about administration actions (e.g., through e-Participation tools). They receive timely information about their city, the effects of PA actions, public as well as private energy consumptions, and information about the environmental conditions where they live (pollution and public security). On the other hand, PA has new tools and techniques to deeply understand dynamics of phenomena that characterise the administrated city, being able to take actions that might improve citizens daily life. Furthermore, energy providers might have the opportunity of implementing smart grids for improving their delivery services and save costs. In this framework, the national research project Brescia Smart Living (BSL) - MIUR "Smart Cities and Communities and Social Innovation" is currently being performed. The main goal of the BSL project is to move from a model based on a single monitored entity (a street, the electrical supply grid,

A. Longo et al. (Eds.): IISSC 2017/CN4IoT 2017, LNICST 189, pp. 131–137, 2018.
https://doi.org/10.1007/978-3-319-67636-4_15

the hydric system) to an integrated view of the Smart City. Every aspect is seen as part of a more complex system, and different types of data have to be collected, properly integrated and organized in order to provide new services to both citizens and PA. The effectiveness and the quality of the services is enabled by applying advanced solutions for managing large amounts of data. Data about energy consumptions on the electrical and methane supply grids, hydric system, street lighting, heating are collected from proper sensors and technological equipment of modern city, according to the Internet of Things (IoT) paradigm. These data are stored within proprietary platforms managed by single energy providers. Further information coming from external sources (e.g., weather and pollution data) are integrated and organized in integrated platforms in order to: (a) aggregate information at city level, mainly devoted to energy providers and Public Administration to give a global view of data concerning the smart city; (b) enable personalized access to data of interest for private citizens at the level of single district, building or apartment. Data from social media are integrated as well, where citizens become themselves data producers through their comments, suggestions and preferences. To meet research goals, the project is being developed over the following phases: (i) collection and identification of requirements from citizens, energy providers and Public Administration, through the submission of proper questionnaires; (ii) design and specification of functionalities, in terms of use cases focused on services provided to the actors of integrated platforms; (iii) design of data models underlying the platforms; (iv) implementation and experiments. Experiments will be performed on two districts identified in Brescia, Italy. The aim is to provide an integrated observatory over the Smart City, for different kinds of information, at different levels of aggregation, for heterogeneous although interleaved categories of users (citizens, energy providers and Public Administration). In this paper we focus on challenges that raised in the project for managing data in the context of a Smart City and we provide some hints about possible solutions to address these issues.

This paper is organized as follows: in Sect. 2 we discuss the functional architecture of Brescia Smart Living project; Sect. 3 presents data management issues; Sect. 4 lists related projects; conclusions and future directions are sketched in Sect. 5.

2 General Architecture

Figure 1 shows an overview of the general architecture adopted in the BSL project. Data are collected from both energy consumption domain (electricity, heating, hydric and natural gas supply grids) and urban services domain (garbage collection and security). This information is stored within domain-specific platforms, owned by the energy and service providers that are participating to the project. Both historical data and (near) real-time data collected from sensors installed on-field (home automation hardware equipment, new generation meters for supply grids, wearable devices for security and safety monitoring) have been considered. The architecture also includes data coming from external sources, for weather forecasting and air pollution estimation.

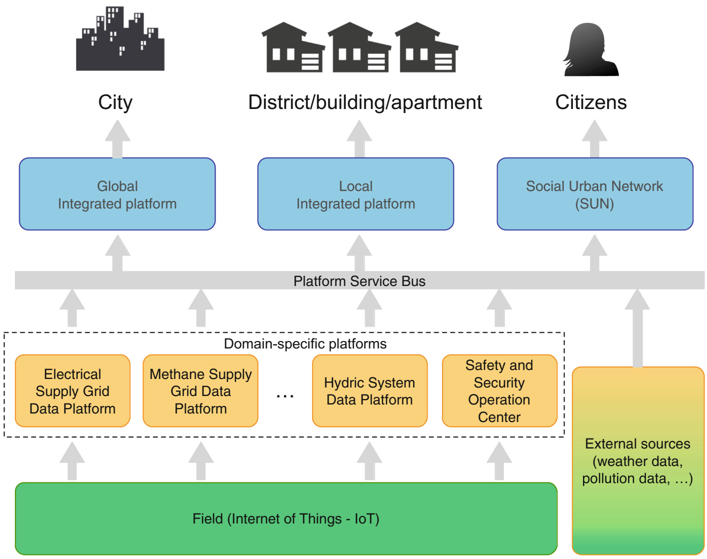

Fig. 1. Overview of the Brescia Smart Living general architecture.

Data coming from domain specific platforms and external datasources are organized within integrated platforms through a Platform Service Bus, that is in charge of managing message exchange between all the architectural components. Communication between domain-specific datasources and integrated platforms may be either synchronous (for instance, concerning historical data of supply grids) or asynchronous (e.g., events raised from home automation hardware or wearable devices, to be promptly showed on the platforms).

The visualisation of average energy consumptions (properly aggregated to preserve citizens' privacy), the visualisation of information about the pollution, the statistical data on crime rate, information about the status of garbage collection points are in charge of the *Global Integrated platform*. Its role is to provide a global view at city level about consumption and services for citizens and Public Administration. The platform also supports PA to take decisions given the current environmental conditions (e.g., weather and pollution conditions, to take effective and timely actions to preserve the health of fragile citizens).

Citizens can register themselves and access services of the *Local Integrated platform*. This platform enables the visualisation of the personal energy consumptions, as well as comparison of these data against benchmarks and average values locally at district/building/apartment level. It also enables citizens of smart homes to monitor, control and analyse data collected by smart plugs and the new generation meters on supply grids.

Finally, mobile applications and a dashboard of proper Key Performance Indicators will be designed and implemented on top of the integrated platforms.

3 Data Management Challenges in Brescia Smart Living

The aim of the integrated platforms introduced above is to provide citizens, energy providers and PA with a tool for collecting, integrating and visualizing heterogeneous information available in the context of a Smart City, at different levels of aggregation, in order to help them to take decisions in their daily life and stimulate their virtuous behaviour in using private and public resources. Several interdisciplinary aspects should be considered. Among the issues concerning data management, we mention the following ones.

(Big) data collection, organization and aggregation. The quantity and velocity of data to be collected pose not trivial problems for their efficient collection and organization, in order to speed up data access and exploration [2]. In the BSL project, data collected on-field through the domain-specific platforms present a poor structure (they are basically schemaless) and must be stored and indexed efficiently. At their finest granularity, they are stored within a document-oriented NoSQL database (using JSON documents). The Local Integrated platform is based on this database. Aggregated information, both at city level and at district level, are computed and stored within a relational database, that provides a more structured organization of information to enable exploration and visualisation facilities [3].

User-driven information service design. Supplied information should be accessed through proper services built on top of the underlying resources. Services design must be driven by the needs of the final users, ranging over PA, citizens and categories of users with specific requirements, such as fragile subjects. In the BSL project, service design started from a classification of users into stakeholders (Public Administration, energy providers) and citizens. The latter ones have been in turn distinguished according to the typology of collected data (traditional meters, new generation electric and methane meters, smart meters for home automation). Considering different categories of users, services designed for the BSL project range from data visualisation services to decision support services: (i) to help stakeholders to take actions in the context of the smart city; (ii) to enable citizens to compare their personal data with average values at building, district and city level.

Data privacy and ownership aspects. The above mentioned issues necessarily require to pay attention to the ownership of data, their retention and the protection of sensitive information. Proper data access mechanisms have to be defined for the different categories of users, according to recent national and European laws that address data privacy and preservation [4]. Within the BSL project, these aspects will be considered according to the European GDPR 2016/679 (General Data Protection Regulation). In particular, this regulation focuses on principles like privacy by design, that requires that data protection is included in the development of business processes for products and services, the right to be forgotten applied to personal data and the principle of data portability.

Data Security and Integrity. The integrity of the information should be always guaranteed or, at least, problems must be detected. For example, sensors deployed all around the city might get uncalibrated for some reasons. The system must be able to detect and manage this wrong data to avoid a decision-making process on wrong basis. A full digital description of the plants (including information about measurement accuracy of each device, certificate of calibration when available) and an history of measurements are needed to achieve this target. Using this set of data, it is possible to design algorithms able to analyze the system behavior and to provide suitable metrics that can highlight malfunctions as well as any data integrity issues, in a similar way as Intrusion Detection Systems (IDS) in cybersecurity identify network attacks.

System scalability and interoperability. The system must have the ability to integrate new data-collecting sub-systems or to create new sub-systems exploiting the existing hardware. A city grows in long periods and the data-collecting technologies advance faster, therefore managing the heterogeneity of the technologies over time is fundamental to build a robust system. Interoperability between different systems into the smart city must be achieved [6]. In the BSL project, to address this issue, a new generation of Discovery Services should be developed. As the Domain Name System (DNS) records matching information about IPs and domains, a Service Discovery System (SDS) for IoT will record the topology and the characteristics of the network nodes, including their communication, measurement and actuation capabilities. In this way, a new node is able to recover the resources it needs, ensuring flexibility, interoperability and reuse of existing hardware [5].

4 Related Work

Compared to recent and on-going Smart Cities projects, BSL provides a wider spectrum of services (as summarised in the previous sections), considering not only the energy consumption data, but also crime rate data, as well as data about environmental conditions. Moreover, BSL brings together the integration platforms at multiple levels, thus giving relevance both to the Smart Cities services meant for PA and the ones designed for citizens (also considering support for decisions devoted to fragile users).

Smart Cities projects. The Optimus project [7] aims to support Public Administration to optimise energy consumption. The focus here is on a semantic-based data acquisition module, to integrate heterogeneous data sources into a relational database, and Decision Support System to enforce an energy manager while taking his/her decisions. The project has been tested on three pilot cases in Italy (city of Savona), Spain (Sant Cugat del Vallès) and The Netherlands (Zaanstad). Similarly, the BESOS project [8] is focused on the implementation of distributed Energy Management Systems (EMS) for energy saving. BESOS is mainly devoted to PA and energy service companies, while the citizens' involvement is more marginal compared to the BSL project. The SFpark project [9] is focused on public transportation for the city of San Francisco. The project aims to provide advanced data mining and planning facilities for the PA to improve urban public mobility services. Primary goal of the Res Novae project [10] was to provide an integrated platform for visualising and monitoring the energy consumption at a city scale. The Res Novae project has been tested in the city of Bari, Italy. The ROMA project [11] aims to integrate data from heterogeneous sources (security, mobility and weather) to increase city resilience and support the PA in management of emergency situations.

Enabling technological infrastructure. The research topics addressed within Brescia Smart Living and related projects also rely on the availability of integration platforms and infrastructures that are able to address the data management issues we underlined in this paper. Available platforms must be based on standards, be scalable and modular, flexible enough to allow ease extension of functional and non functional requirements imposed by the dynamic environment of Smart Cities. Oracle and IBM proposed their own solutions for implementing Smart Cities projects. In particular, the Oracle Smart City Platform provides a set of front-office functions over multiple communication channels (e.g., telephone, web, chat), big data management solutions and analytical tools. It has been configured for several projects (such as the SFpark project mentioned above). IBM proposed a platform that integrates many tools for managing data in the context of smart cities, like the Intelligent Operations Center IOC [12], providing functions to visualise data on a tabular, graphical and map-based interface, SPSS [14] for stochastic and predictive analysis of data and CPLEX [13] for solving optimisation problems. IOC has been applied in the Res Novae project and has been chosen also for implementing the global integration platform within the BSL project. Compared to these Smart Cities platforms, middleware solutions provide communication drivers, data management facilities and APIs, but require additional development efforts to create new applications and services on top of them. Indra proposed a cross-platform and multi-device middleware called SOFIA2 [15]. This middleware is devoted to the development of smart applications that use real-time information according to a big data approach. It has been applied in two pilot projects for the cities of La Coruña and Turin. Tridium Niagara Framework [16] offers a development platform that connects and translates data from nearly any device or system, managing and optimizing performance when dealing with heterogeneous data formats and protocols. It also enables the development of software objects to manage data at cloud and edge computing.

5 Conclusions and Future Directions

Design and implementation of the Brescia Smart Living project poses complex data management issues. The overall target of the project is to develop and innovate services through the integration of heterogeneous data at various levels of aggregation and provided to citizens and PA to take decisions in their daily life and stimulate virtuous behaviour. Future efforts will be devoted to the extension of the infrastructure with new services (mobility, healthcare and education services), as well its application to new districts and smart cities.

Acknowledgments. The BSL consortium is leaded by A2A and includes as partners: Beretta Group, Cauto, Cavagna Group, the Municipality of Brescia, University of Brescia, Enea, STMicroelectronics and an association of private companies (see www.smartcityitalia.net/projects/brescia-smart-living/).

References

1. Khatoun, R., Zeadally, S.: Smart cities: concepts, architectures, research opportunities. Commun. ACM **59**(8), 46–57 (2016)
2. Chauhan, S., Agarwal, N., Kar, A.: Addressing big data challenges in smart cities: a systematic literature review. Info **18**(4), 73–90 (2016)
3. Bagozi, A., Bianchini, D., De Antonellis, V., Marini, A., Ragazzi, D.: Summarisation and relevance evaluation techniques for big data exploration: the smart factory case study. In: Dubois, E., Pohl, K. (eds.) CAiSE 2017. LNCS, vol. 10253, pp. 264–279. Springer, Cham (2017). doi:10.1007/978-3-319-59536-8_17
4. Li, Y., Dai, W., Ming, Z., Qui, M.: Privacy protection for preventing data over-collection in smart city. IEEE Trans. Comput. **65**(5), 1339–1350 (2016)
5. Bellagente, P., Ferrari, P., Flammini, A., Rinaldi, S.: Adopting IoT framework for Energy Management of Smart Building: a real test-case. In: 2015 IEEE 1st International Forum on Research and Technologies for Society and Industry (RTSI), Turin, Italy, pp. 138–143 (2015). doi:10.1109/RTSI.2015.7325084, ISBN 978-1-4673-8166-6
6. Ahlgren, B., Hidell, M., Ngai, E.: Internet of Things for smart cities: interoperability and open data. IEEE Internet Comput. **20**(6), 52–56 (2016)
7. OPTIMising the energy USe in cities with smart decision support systems. http://optimus-smartcity.eu
8. Building Energy decision Support systems fOr Smart cities. http://besos-project.eu
9. San Francisco Park project. http://sfpark.org
10. Res Novae Project. http://resnovae-unical.eu
11. Resilience enhancement Of a Metropolitan Area. http://www.progetto-roma.org
12. IOC - Intelligent Operations Center. http://www-03.ibm.com/software/products/it/intelligent-operations-center
13. C-PLEX Optimizer. http://www-01.ibm.com/software/commerce/optimization/cplex-optimizer/
14. SPSS. http://www-01.ibm.com/software/it/analytics/spss/
15. Indra, SOFIA2 Web site. http://sofia2.com/
16. Tridium Niagara Framework. http://www.tridium.com/en/products-services/niagaraax