Future Generation Computer Systems 37 (2014) 267–281

Contents lists available at ScienceDirect

Future Generation Computer Systems

journal homepage: www.elsevier.com/locate/fgcs

Intelligent services for Big Data science

C. Dobre [a,*], F. Xhafa [b]

[a] University Politehnica of Bucharest, Splaiul Independentei 313, Bucharest, Romania
[b] Universitat Politecnica de Catalunya, Girona Salgado 1-3, 08034 Barcelona, Spain

HIGHLIGHTS

- Analysis of challenges and requirements for next-generation Big Data services.
- CAPIM is a platform to collect and aggregate context information on a large scale.
- Present the implementation of an Intelligent Transportation System on top of CAPIM.
- Context-Aware Framework is designed to handle efficient storage of context data.
- Such Big Data services show different mechanisms to cope with identified challenges.

ARTICLE INFO

Article history:
Received 2 April 2013
Received in revised form
10 June 2013
Accepted 17 July 2013
Available online 9 August 2013

Keywords:
Intelligent services
Big Data
Context
Context-aware services
Mobile computing
Smart City applications
Intelligent Transport Systems

ABSTRACT

Cities are areas where Big Data is having a real impact. Town planners and administration bodies just need the right tools at their fingertips to consume all the data points that a town or city generates and then be able to turn that into actions that improve peoples' lives. In this case, Big Data is definitely a phenomenon that has a direct impact on the quality of life for those of us that choose to live in a town or city. Smart Cities of tomorrow will rely not only on sensors within the city infrastructure, but also on a large number of devices that will willingly sense and integrate their data into technological platforms used for introspection into the habits and situations of individuals and city-large communities. Predictions say that cities will generate over 4.1 terabytes per day per square kilometer of urbanized land area by 2016. Handling efficiently such amounts of data is already a challenge. In this paper we present our solutions designed to support next-generation Big Data applications. We first present CAPIM, a platform designed to automate the process of collecting and aggregating context information on a large scale. It integrates services designed to collect context data (location, user's profile and characteristics, as well as the environment). Later on, we present a concrete implementation of an Intelligent Transportation System designed on top of CAPIM. The application is designed to assist users and city officials better understand traffic problems in large cities. Finally, we present a solution to handle efficient storage of context data on a large scale. The combination of these services provides support for intelligent Smart City applications, for actively and autonomously adaptation and smart provision of services and content, using the advantages of contextual information.

1. Introduction

Every day we create 2.5 quintillion bytes of data, so much that 90% of the data in the world today has been created in the last two years alone. This data comes from sensors used to gather climate information, from posts to social media sites, digital pictures and videos, purchase transaction records, or cell phone GPS signals, to name only a few. This data is *Big Data*. Analyzing large datasets already underpins new waves of productivity growth, innovation, and consumer surplus. Big data is more than simply a matter of size; it is an opportunity to find insights in new and emerging types of data and content, to make businesses more agile and to answer questions that were previously considered beyond our reach. Until now, there was no practical way to harvest this opportunity. But today we are witnessing an exponential growth in the volume and detail of data captured by enterprises, the rise of multimedia, social media and Online Social Networks (OSN), and the Internet of Things (IoT).

Many of Big Data challenges are generated by future applications where users and machines will need to collaborate in intelligent ways together. In the near future, information will be available all around us and will be served in the most convenient way — we will be notified automatically when a congestion occurs and the car will be able to decide how to optimize our driving route, the fridge will notify us when the milk supply is out,

* Corresponding author. Tel.: +40 745174359.
E-mail addresses: ciprian.dobre@cs.pub.ro (C. Dobre), fatos@lsi.upc.edu (F. Xhafa).

0167-739X/$ – see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.future.2013.07.014

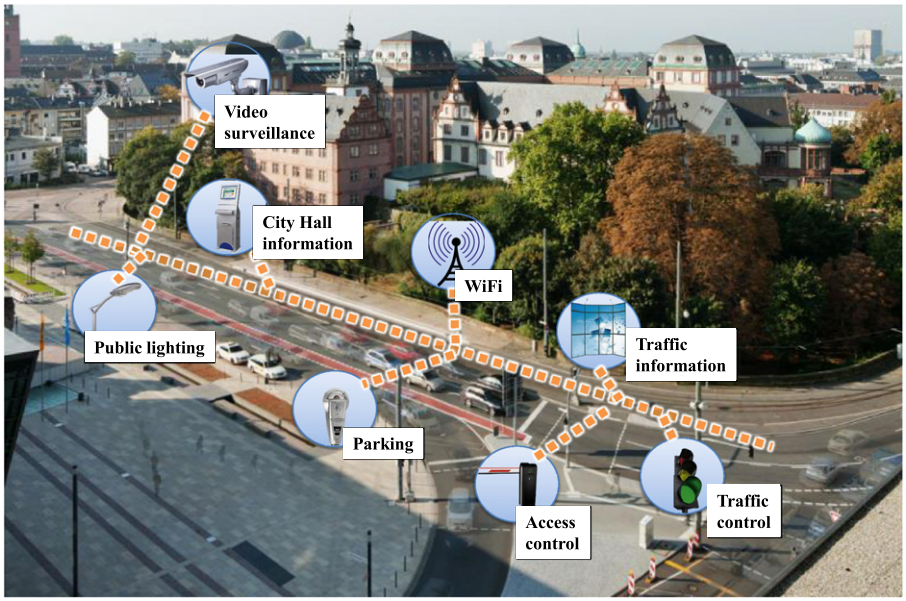

Fig. 1. The vision of integrated community-enabled intelligent services.

etc. Technology becomes more and more part of our daily life. New technologies have finally reached a stage of development in which they can significantly improve the lives of any city's inhabitants. Our cities are fast transforming into artificial ecosystems of interconnected, interdependent intelligent digital "organisms". They are transforming into *smart cities*, as they benefit more and more from intelligent applications designed to drive a sustainable economic development and an incubator of innovation and transformation that merges the virtual world of Mobile Services, IoT and OSN with the physical infrastructures of Smart Building, Smart Utilities (i.e., electricity, heating, water, waste, transportation, and unified communication & collaboration infrastructure). The transformation of the metropolitan landscape is driven by the opportunity to embed intelligence into any component of our towns and connect them in real-time, merging together physical world of objects, humans and virtual conversation and transactions.

There are already examples of Smart Cities like Malaga, Amsterdam, or Boston. In Europe projects such as "European Initiative on Smart Cities" (SETIS) ignited many local Public administrations to launch new initiatives to take advantage of the opportunity of funding for Smart Cities [1]. Several ICT companies (e.g., IBM: Smart Planet, Accenture: Intelligent City Network, CISCO: Connected Urban Development, Ericsson Smart City, etc.) and research institutions (e.g., MIT: Smart City SENSEable lab, Terreform One, etc.) already offer services and solutions' components that can help to build more livable, sustainable cities by innovative ICT usage. In many towns such as Boulder or Amsterdam many ICT companies are working together with utilities, universities and other organization to provide integrated solutions. In 2016, it is estimated that $39.5 billion will be spent on smart city technology, up from $8.1 billion in 2010 [2].

We came a long way since the original "wired city" vision advocated in the 1980s [3]. The notion of wiring the city then was needed to support networking very diverse activities and routine services such as those provided by municipalities – libraries, welfare services – over WANs. Today, wireless devices incur relatively low costs. And, they come equipped with ever more sensors and storage and processing capabilities. As a consequence, they can replace the traditional wired-sensor infrastructure and act as virtual sensors that can capture information over very small time scales and very fine levels of spatial resolution. Such devices that range from purpose-built sensors to individual hand-held devices that are as mobile as those using them provide massive capability to store and transmit data that pertains to movement and activity levels across space and time. Some of the most elaborate applications involve transport, but other services can easily be integrated all-together as well (see Fig. 1).

However, today there is no unique model for a Smart City, and each city approaches the concept in its own particular way, with different projects and objectives. One of the obvious but much misunderstood features of these new urban technologies is the fact that *they produce massive streams of data in real time and space*. We are just beginning to grasp the nature of this 'Big Data'. So far, most of the datasets from which scientists and researchers were able to extract real meaning are quite small in comparison with the sort of data that can be captured by smart city applications. Imagine the movements of people in a large city like London, for example, where there are something like 3 million travelers a day using some form of public transport. If the municipality could capture data about locations of these travelers to extract meaningful information, with technology available today probably much of this data can be reduced or aggregated. Thus, sequences with evident meaning could be extracted. But things become more complicated if we think that such kind of data is available continuously. Over sufficiently long periods of time, one can begin to extract changes to the structure and form of the city and the way people behave. But this yields for solutions to store and manage tremendous amounts of data. And all this data will probably need to be supplemented with all kinds of other information (e.g., transport data, data relating to social and economic interactions).

We have barely begun to get a sense of the dimensions of this kind of data, of the privacy implications, of ways in which we can code it with respect to meaningful attributes in space and time. As we move into an era of unprecedented volumes of data and computing power, the benefits are not for business alone. Data can help citizens access government, hold it accountable and build new services to help themselves. In one sense, all this is part of a world that is fast becoming digital in all its dimensions. People will develop more easily their understanding and design ideas using digital representations and data. This vision will support the development of the new ideas for the future of urban and social life.

In this paper we present our solutions designed to support next-generation Big Data applications. The contribution of this paper is as follows: we present first CAPIM, a platform designed