

Preface

Introduction

Data-intensive computing is now starting to be considered as the basis for a new, fourth paradigm for science. Two factors are encouraging this trend. First, vast amounts of data are becoming available in more and more application areas. Second, the infrastructures allowing to persistently store these data for sharing and processing are becoming a reality. This allows unifying knowledge acquired through the three past paradigms for scientific research (theory, experiments, and simulations) with vast amounts of multidisciplinary data. The technical and scientific issues related to this context have been designated as “Big Data” challenges and have been identified as highly strategic by major research agencies.

On the other hand, the combination of the Internet and emerging technologies such as near-field communications, real-time localization, and embedded sensors, transform everyday objects into smart objects capable to understand and react to their environment. Such objects enable new computing applications, and are at the base of the vision of a global infrastructure of networked physical objects known today as the Internet of Things (IoT). The vision of an Internet of Things built from smart objects raises several important research questions in terms of system architecture, design and development, and human involvement. For example, what is the right balance for the distribution of functionality between smart objects and the supporting infrastructure? Can Cloud Computing and Big Data become true infrastructure-enablers for IoT? During the past few years, IT has seen several major paradigm shifts in technology. These shifts have been primarily in the areas of cloud computing, Big Data, mobility, and the Internet of Things. The convergence of these four areas already creates a new “platform” for enterprises to develop new business and mission capabilities. These capabilities will enable a more integrated view of IT architecture which we explore in this book.

The book presents current progress on challenges related to Big Data management by focusing on the particular challenges associated with this convergence, presenting current progresses and challenges associated with context-aware, data-intensive applications and services. A representative application category is that of Smart Environments such as Smart Cities, which cover a wide range of needs related to public safety, water and energy management, smart

buildings, government and agency administration, transportation, health, education, etc. This application class is the subject of many R&D activities and a priority for several government strategic plans for the forthcoming years. Context-aware, data-intensive applications have specific requirements including but not limited to real-time data processing, data access patterns (frequent, periodic, ad-hoc, inter-related, location, device etc.), QoS, intelligent queries, etc.

During the last years, large storage infrastructures were continuously improved in response to the increasing demands of the data intensive applications. Many works address issues related to data intensive scientific applications and Web applications in Cloud environments. In scientific applications, big data volumes collected from high throughput laboratory instruments, observation devices, LHC, and others are submitted to complex software packages for Monte Carlo simulations, extensive analyses, data mining, etc. They are also used for generating graphical outputs and visualization, and are archived in digital repositories for further use. A central topic of research is the adoption of parallel processing solutions able to respond to scalability and performance requirements. One example is the application of the MapReduce model, used initially for the processing of large volumes of textual data, for High Energy Physics data analyses and Kmeans clustering. Web applications are interactive and based on unstructured data that are processes in a write-once read-many-times manner. Research issues include the scalability, fault tolerance, and adaptability to simple data models used by Web applications. For example, the distributed storage system Bigtable is designed to scale to a very large data size. It responds to a variety of workload profiles, from throughput-oriented batch-processing jobs to latency-sensitive data delivery to end-users. Dynamo is a highly available key-value storage system that supports retrieving data by primary key. It has a clearly defined consistency window and efficiently use its resources.

Other applications with high demands for Big Data services are context-aware. Some of their requirements are similar to those of scientific and Web applications. Other needs are shared with mobile computing applications (such as battery life and limited bandwidth) and are due to the presence of mobile devices (sensors, mobile phones) in context-aware applications. Context-aware paradigm is at the base of developing many applications that include people, economy, mobility, and governance. They enrich the urban environment with situational information, which can be rapidly exploited by citizens in their professional, social, and individual activities to increase city competitiveness. In addition, user-centric approaches to services can improve the core urban systems (such as public safety, transportation, government, agency administration and social services, education, healthcare) with impact on quality of life. They allow users interact with others (people), services, and devices in a natural way yet maintaining their normal flow of activities in the real world. Context-aware applications require and use large amounts of input data, in various formats, collected from sensors or mobile users, from public open data sources, or from other applications. While several applications might share the same business and context data, the access and processing profiles can be different from one application to another. For example, public

safety applications should react rapidly to context changes to deliver real-time alerts for potential dangerous situations. Collaborative diagnosis allows physicians work together to combine their expertise for complicated diseases and lets them access more information to make a convenient assessment.

Thus, an aim of this book is to become a state-of-the-art reference discussing progress made, as well as prompting future directions on the theories, practices, standards, and strategies that are related to the emerging computational technologies and their association with supporting the Internet of Things advanced functioning for organizational settings including both business and e-science. Apart from interoperable and intercooperative aspects, the book deals with a notable opportunity namely, the current trend in which a collectively shared and generated content is emerged from Internet end-users. Specifically, the book presents advances on managing and exploiting the vast size of data generated from within the smart environment (i.e. smart cities) toward an integrated, collective intelligence approach. We believe interfunctionality between e-infrastructures will enable the store and generation of a vast amount of data, which if combined and analyzed through an agile, synergetic, collaborative, and collective intelligence manner will make a difference in the organizational settings and their user communities.

The book also presents methods and practices to improve large storage infrastructures in response to increasing demands of the data-intensive applications. We address here the category of context-aware applications, which have specific requirements and a special profile. Context data can be used not only to accurately understand the semantics of business data in the benefit of applications, but it can be exploited for improving the performance and facilitate the management of Big Data store service. Of particular interest are the features of a Big Data store Cloud service able to support context-aware applications and their applicability in bringing the Internet of Things to the functionality of an integrated collective intelligence approach on the future Internet.

Who Should Read the Book?

The content of the book offers state-of-the-art information and references for work undertaken in the challenging areas of IoT and Big Data. The book is intended for those interested in joining interdisciplinary and transdisciplinary works in the areas of Smart Environments, Internet of Things and various computational technologies for the purpose of an integrated collective computational intelligence approach into the Big Data era. These tools include but not limited to the use of distributed data capturing, crowd sourcing, context awareness and event the data-driven paradigm, integration and their analytics and visualizations.

Thus, the book should be of particular interest for:

Researchers and doctoral students working in the area of IoT sensors, distributed technologies, collective and computational intelligence, primarily as a reference publication. The book should be also a very useful reference for all researchers and doctoral students working in the broader fields of computational and collective intelligence, emerging and distributed technologies.

Academics and students engaging in research informed teaching and/or learning in the above fields. The view here is that the book can serve as a good reference offering a solid understanding of the subject areas.

Professionals including computing specialists, practitioners, managers, and consultants who want to understand and realize the aspects (opportunities and challenges) of using computational technologies for the Big Data coupled with Internet of Things advanced functioning and its integrated collective intelligence approach in various organizational settings such as in smart environments.

Book Organization and Overview

The book contains 19 self-contained chapters that were very carefully selected based on peer review by at least two expert and independent reviewers. The book is organized into three parts according to the thematic topic of each chapter. The following three parts reflect the general themes of interest to the IoT and Big Data communities.

Part I: Foundations and Principles

The part focuses on presenting the foundations and principles of on-going investigations, and presents an analysis of current challenges and advances related to Big Data management. The part consists of seven chapters. In particular:

“[Big Data Platforms for the Internet of Things](#)” discusses the challenges, state of the art, and future trends in context-aware environments (infrastructure and services) for the Internet of Things. It makes a critical analysis of current opportunistic approaches using the elements of a newly defined taxonomy. The chapter reviews state-of-the-art Big Data platforms, from an Internet of Things perspective.

“[Improving Data and Service Interoperability with Structure, Compliance, Conformance and Context Awareness](#)” revisits the interoperability problem in the IoT context, focusing on structure, compliance, conformance, and context-awareness particular challenges. The chapter proposes an architectural style, structural services, which combines the service modelling capabilities of SOA with the flexibility of structured resources in REST.

“[Big Data Management Systems for the Exploitation of Pervasive Environments](#)” focuses on Big Data management systems for the exploitation of pervasive environments. The chapter presents an original solution that uses Big Data technologies for redesigning an IoT context-aware application for the exploitation of pervasive environment addressing problems and discussing the important aspects of this solution.

“[On RFID False Authentications](#)” presents security challenges in the Internet of Things. Within this context, the chapter offers a state-of-the-art review of existing authentication protocols. The authors give a necessary and sufficient condition for false authentications prevention, and propose a naive semaphore-based solution which revises the pattern by adding semaphore operations so as to avoid false authentications.

“[Adaptive Pipelined Neural Network Structure in Self-Aware Internet of Things](#)” is concerned with self-healing systems that can auto-detect, analyze, and fix or reconfigure issues associated with their self-behavior and performance. Self-healing processes should occur in real-time to restore the desired functionality as soon as possible. Adaptive neural networks are proposed as a solution to some of these challenges; monitoring the system and environment, mapping a suitable solution and adapting the system accordingly.

“[Spatial Dimensions of Big Data: Application of Geographical Concepts and Spatial Technology to the Internet of Things](#)” highlights the Spatial Dimensions of Big Data, and presents a case study on the application of geographical concepts and spatial technology to the Internet of Things and Smart Cities. By applying spatial relationships, functions, and models to the spatial characteristics of smart objects and the sensor data, the flows and behavior of objects and people in Smart Cities can be efficiently monitored and orchestrated.

“[Fog Computing: A Platform for Internet of Things and Analytics](#)” discusses another potential foundation of IoT platforms: fog computing. The chapter examines disruptions due to the explosive proliferation of endpoints in IoT, and proposes a hierarchical distributed architecture that extends from the edge of the network to the core called Fog Computing. In particular, the chapter discusses a new dimension that IoT adds to Big Data and Analytics: a massively distributed number of sources at the edge.

Part II: Advanced Models and Architectures

This part focuses on presenting theoretical and state-of-the-art models, architectures, e-infrastructures, and algorithms that enable the inter-cooperative and inter-operable nature of the Internet of Things for the purpose of collective and computational intelligence, in the Big Data context, and consists of six chapters. In particular:

“**Big Data Metadata Management in Smart Grids**” focuses on the Big Data management in Smart Grids. In such environments, data are collected from various sources, and then processed by different intelligent systems with the purpose of providing efficient system planning, power delivery, and customer operations. Three important issues in managing and solutions to overcome them are discussed. Concrete examples from the offshore wind energy are used to demonstrate the solutions.

“**Context-Aware Dynamic Discovery and Configuration of ‘Things’ in Smart Environments**” proposes the Context-Aware Dynamic Discovery of Things (CADDOT) model. A tool is presented, SmartLink, used to establish direct communication between sensor hardware and cloud-based IoT middleware platforms. The prototype tool is developed on an Android platform. Global Sensor Network (GSN) is used as the IoT middleware for a proof-of-concept validation.

“**Simultaneous Analysis of Multiple Big Data Networks: Mapping Graphs into a Data Model**” proposes a new model to map web multinet network graphs in a data model. The result is a multidimensional database that offers numerous analytical measures of several networks concurrently. The proposed model also supports real-time analysis and online analytical processing (OLAP) operations, including data mining and business intelligence analysis.

“**Toward Web Enhanced Building Automation Systems**” reviews strategies to provide a loosely coupled Web protocol stack for interoperation between devices in a building. Making the assumption of seamless access to sensor data through IoT paradigms, the chapter provides an overview of some of the most exciting enabling applications that rely on intelligent data analysis and machine learning for energy saving in buildings.

“**Intelligent Transportation Systems and Wireless Access in Vehicular Environment Technology for Developing Smart Cities**” focuses on Intelligent Transport Systems (ITS) and wireless communications as enabling technologies for Smart Cities. After reviewing main advances and achievements, the chapter highlights major research projects developed in Europe and the USA. The chapter presents the main contributions that ITS can provide in the development of Smart Cities as well as the future challenges.

“**Emerging Technologies in Health Information Systems: Genomics Driven Wellness Tracking and Management System (GO-WELL)**” presents GO-WELL, an example of future personal health record (PHR) concept. GO-WELL is based on clinical envirogenomic knowledge base (CENG-KB) to engage patients for predictive care. The chapter describes concepts related to the inclusion of personalized medicine, omics revolution, incorporation of genomic data into medical decision processes, and the utilization of enviro-behavioral parameters for disease risk assessment.

Part III: Advanced Applications and Future Trends

The part focuses on presenting cutting-edge Internet of Things related applications to enable collective and computational intelligence, as well as prompting future developments in the area. The part consists of six chapters. In particular:

“[Sustainability Data and Analytics in Cloud-based M2M Systems](#)” analyzes data and analytics applications in M2M systems for sustainability governance. The chapter presents techniques for M2M data and process integration, including linking and managing monitored objects, sustainability monitoring data and analytics applications, for different stakeholders who are interested in dealing with large-scale monitoring data in M2M environments. It presents a cloud-based data analytics system for sustainability governance that includes a Platform-as-a-Service and an analytics framework. The authors also illustrate their prototype on a real-world cloud system for facility monitoring.

“[Social Networking Analysis](#)” makes an analysis of the role of social network analysis on applications dealing with studies of kinship structure, social mobility, science citations, use of contacts among members of nonstandard groups, corporate power, international trade exploitation, class structure, and many other areas. Building a useful understanding of a social network is to sketch a pattern of social relationships, kinships, community structure, interlocking dictatorships, and so forth for analysis.

“[Leveraging Social Media and IOT to Bootstrap Smart Environments](#)” deals with the realization of Smart Environments, and presents lightweight Cyber Physical Social Systems to include building occupants within the control loop to allow them some control over their environment. The authors define the concept of citizen actuation, and present an experiment where they obtain a reduction in average energy usage of 26 %. The chapter proposes the Cyber-Physical Social System architecture, and discusses future research in this domain.

“[Four-Layer Architecture for Product Traceability in Logistic Applications](#)” describes the design of an auto-managed system for the tracking and location of products in transportation routes, called Transportation Monitoring System (TMSsystem). A four-layer system is proposed to provide an efficient solution for the Real-Time Monitoring (RTM) of goods. Two Web Services are proposed, Location Web Service and Check Order Web Service, so that customers can easily access information about the shipment of their orders. Finally, a Web Application is developed to access those Web Services.

“[Disaster Evacuation Guidance Using Opportunistic Communication: Potential for Opportunity-Based Service](#)” proposes a disaster evacuation guidance using opportunistic communication where evacuees gather location information of impassable and congested roads by disaster into their smartphones by themselves, and also share the information with each other by short-range wireless communication between nearby smartphones. The proposed guidance is designed not only to navigate evacuating crowds to refuges, but to rapidly aggregate the disaster information.

“[iPromotion: A Cloud-Based Platform for Virtual Reality Internet Advertising](#)” presents a large-scale platform for distributing Virtual Reality advertisements over the World Wide Web. The platform aims at receiving and transmitting large amounts of data over mobile and desktop devices in Smart City contexts, is based on a modular and distributed architecture to allow for scalability, and incorporates content-based search capabilities for VR scenes to allow for content management.

Nik Bessis
Ciprian Dobre

Big Data and Internet of Things: A Roadmap for Smart
Environments

Bessis, N.; Dobre, C. (Eds.)

2014, XIX, 470 p. 169 illus., Hardcover

ISBN: 978-3-319-05028-7