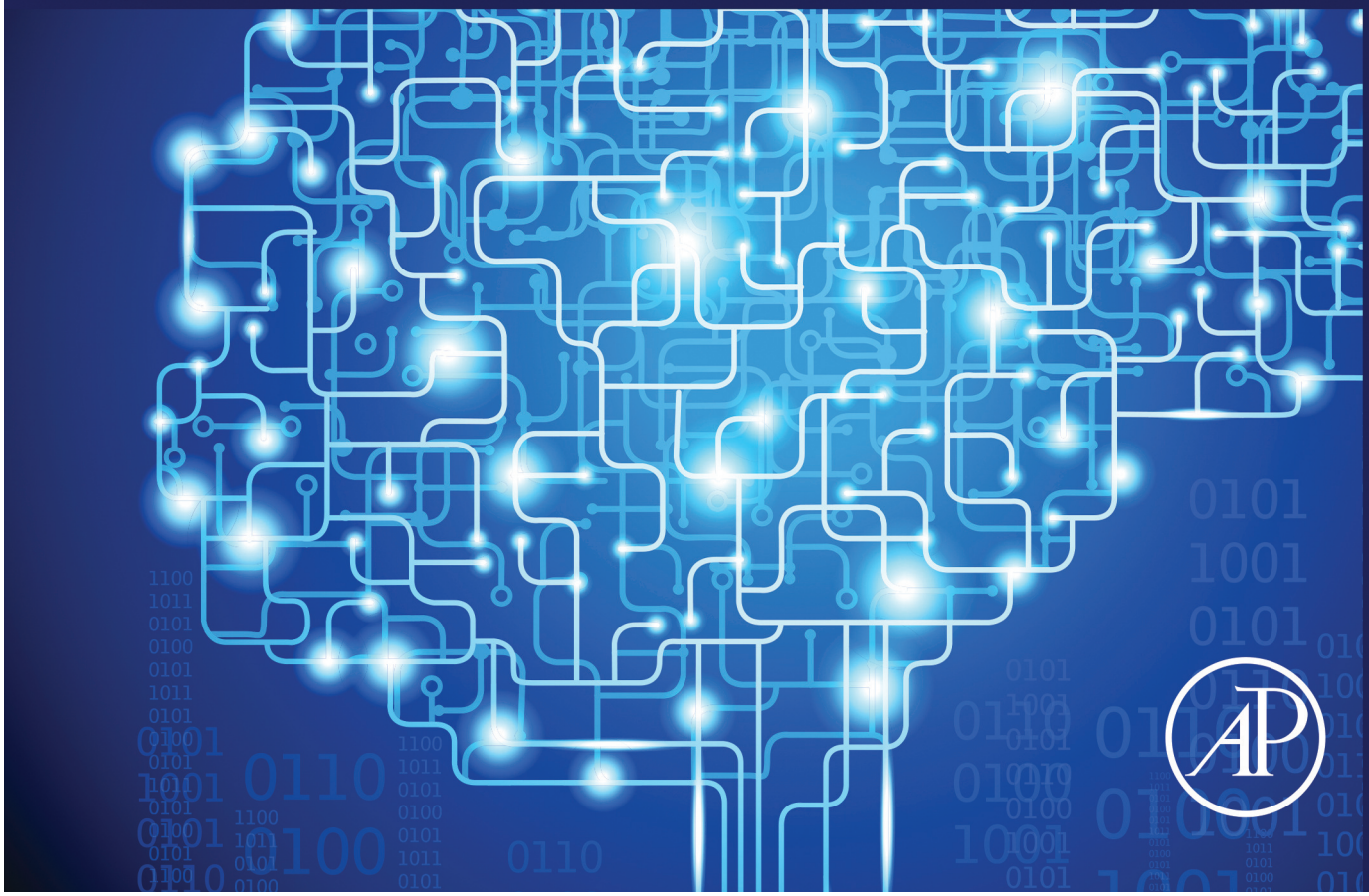
Intelligent Data Centric Systems

Series Editor Fatos Xhafa

Big Data Analytics for Sensor-Network Collected Intelligence

Edited by Hui-Huang Hsu, Chuan-Yu Chang, Ching-Hsien Hsu



Big Data Analytics for Sensor-Network Collected Intelligence

This page intentionally left blank

Big Data Analytics for Sensor-Network Collected Intelligence

Edited by

Hui-Huang Hsu
Tamkang University, Taiwan

Chuan-Yu Chang
National Yunlin University of Science and Technology, Taiwan

Ching-Hsien Hsu
Chung Hua University, Taiwan

Series Editor Fatos Xhafa
Universitat Politècnica de Catalunya, Spain

ACADEMIC PRESS
An imprint of Elsevier
elsevier.com

Academic Press is an imprint of Elsevier
125 London Wall, London EC2Y 5AS, United Kingdom
525 B Street, Suite 1800, San Diego, CA 92101-4495, United States
50 Hampshire Street, 5th Floor, Cambridge, MA 02139, United States
The Boulevard, Langford Lane, Kidlington, Oxford OX5 1GB, United Kingdom

Notices
Knowledge and best practice in this field are constantly changing. As new research and experience broaden our understanding, changes in research methods, professional practices, or medical treatment may become necessary.

Practitioners and researchers must always rely on their own experience and knowledge in evaluating and using any information, methods, compounds, or experiments described herein. In using such information or methods they should be mindful of their own safety and the safety of others, including parties for whom they have a professional responsibility.

To the fullest extent of the law, neither the Publisher nor the authors, contributors, or editors, assume any liability for any injury and/or damage to persons or property as a matter of products liability, negligence or otherwise, or from any use or operation of any methods, products, instructions, or ideas contained in the material herein.

Library of Congress Cataloging-in-Publication Data
A catalog record for this book is available from the Library of Congress

British Library Cataloguing-in-Publication Data
A catalogue record for this book is available from the British Library

ISBN: 978-0-12-809393-1

Publisher: Joe Hayton
Acquisition Editor: Sonnini R. Yura
Editorial Project Manager: Ana Claudia A. Garcia
Production Project Manager: Punithavathy Govindaradjane
Cover Designer: Victoria Pearson

Typeset by SPi Global, India

Contents

PART IV BIG DATA INTELLIGENCE AND IoT SYSTEMS

List of Contributors

Ahmad Anbar
The George Washington University, Washington, DC, United States

Haytham Assem
IBM, Dublin, Ireland

Christophe Blanchet
CNRS IFB, Orsay, France

Teodora S. Buda
IBM, Dublin, Ireland

Jiannong Cao
The Hong Kong Polytechnic University, Kowloon, Hong Kong

Chuan-Yu Chang
National Yunlin University of Science and Technology, Douliu City, Yunlin County, Taiwan

Jinjun Chen
University of Technology Sydney, Broadway, NSW, Australia

Cen Chen
Hunan University, Changsha, China

Szu-Ta Chen
National Taiwan University Hospital Yun-Lin Branch, Douliu City, Yunlin County, Taiwan

Kang Chen
Southern Illinois University, Carbondale, IL, United States

Zixue Cheng
University of Aizu, Aizuwakamatsu, Japan

Cees de Laat
University of Amsterdam, Amsterdam, The Netherlands

Yuri Demchenko
University of Amsterdam, Amsterdam, The Netherlands

Mingxing Duan
Hunan University, Changsha, China

Tarek El-Ghazawi
The George Washington University, Washington, DC, United States

Weiwei W. Fang
Beijing Key Lab of Transportation Data Analysis and Mining, Beijing Jiaotong University, Beijing, China

Edmond J. Golden III
National Institute of Standards and Technology, Gaithersburg, MD, United States

Chu-Cheng Hsieh
Slice Technologies Inc., San Mateo, CA, United States

Ching-Hsien Hsu
Chung Hua University, Hsinchu, Taiwan

Hui-Huang Hsu
Tamkang University, Tamsui, Taiwan

Qian Huang
Southern Illinois University, Carbondale, IL, United States

Tian-Hsiang Huang
National Sun Yat-sen University, Kaohsiung, Taiwan

Chih-Chieh Hung
Tamkng University, New Taipei City, Taiwan

Pravin Kakar
Institute for Infocomm Research, Agency for Science, Technology and Research (A*STAR), Singapore

Shonali Krishnaswamy
Institute for Infocomm Research, Agency for Science, Technology and Research (A*STAR), Singapore

Chung-Nan Lee
National Sun Yat-sen University, Kaohsiung, Taiwan

Kenli Li
Hunan University, Changsha, China

Keqin Li
Hunan University, Changsha, China; State University of New York, New Paltz, NY, United States

Xiao-Li Li
Institute for Infocomm Research, Agency for Science, Technology and Research (A*STAR), Singapore

Qingyong Y. Li
Beijing Key Lab of Transportation Data Analysis and Mining, Beijing Jiaotong University, Beijing, China

Hai-Ning Liang
Xi'an Jiaotong-Liverpool University, Suzhou, China

Chen Lin
National Yunlin University of Science and Technology, Douliu City, Yunlin County, Taiwan

Xuefeng Liu
The Hong Kong Polytechnic University, Kowloon, Hong Kong

Ming Liu
Beijing Key Lab of Transportation Data Analysis and Mining, Beijing Jiaotong University, Beijing, China

Charles Loomis
SixSq Sàrl, Geneva, Switzerland

Chao Lu
Southern Illinois University, Carbondale, IL, United States

Ka L. Man
Xi'an Jiaotong-Liverpool University, Suzhou, China

Martial Michel
National Institute of Standards and Technology, Gaithersburg, MD, United States

Vijayakumar Nanjappan
Xi'an Jiaotong-Liverpool University, Suzhou, China

Minh N. Nguyen
Institute for Infocomm Research, Agency for Science, Technology and Research (A*STAR), Singapore

Declan O'Sullivan
Trinity College Dublin, Dublin, Ireland

Phyo P. San
Institute for Infocomm Research, Agency for Science, Technology and Research (A*STAR), Singapore

Olivier Serres
The George Washington University, Washington, DC, United States

Kathiravan Srinivasan
National Ilan University, Yilan City, Yilan County, Taiwan

Ming-Chun Tsai
National Sun Yat-sen University, Kaohsiung, Taiwan

Fatih Turkmen
University of Amsterdam, Amsterdam, The Netherlands

Wei Wang
Xi'an Jiaotong-Liverpool University, Suzhou, China

Junbo Wang
University of Aizu, Aizuwakamatsu, Japan

Yilang Wu
University of Aizu, Aizuwakamatsu, Japan

Chen-Ming Wu
National Sun Yat-sen University, Kaohsiung, Taiwan

Lei Xu
IBM, Dublin, Ireland

Chi Yang
University of Technology Sydney, Broadway, NSW, Australia

Jian-Bo Yang
Institute for Infocomm Research, Agency for Science, Technology and Research (A*STAR), Singapore

Zhangdui D. Zhong
Beijing Key Lab of Transportation Data Analysis and Mining, Beijing Jiaotong University, Beijing, China

This page intentionally left blank

Preface

There are three sources of information we can collect about the environment and the people in the environment: environmental sensors, wearable sensors, and social networks. Through intelligent analysis of the huge amount of sensory data, we can develop various systems to automatically detect natural and man-made events. Moreover, the systems can also try to understand people's behavior and even intention. Thus better services can be provided to people in an unobtrusive manner.

With the advances in sensor and networking technologies, we are now able to collect sensory data easily. These sensory data can be stored and processed in the cloud. Nevertheless, how to properly utilize such a huge amount of data is another essential issue. We certainly hope that advanced ICT technologies can help us perform intelligent analysis on these data and provide better services to people automatically. Exciting new systems and research results have been developed. This book aims to introduce these ambient intelligence and Internet of Things (IoT) systems, which are based on big data analytics of collected sensory data.

The theme of this book is closely related to two hot topics: the Internet of Things and big data analytics. Systems and technologies introduced in the book can be used as supplementary materials for courses involving these two topics. Researchers, professionals, and practitioners in related fields can also find useful information and technologies for their work. There are four parts of this book: big data architecture and platforms; big data processing and management; big data analytics and services; and big data intelligence and IoT systems. Each part includes three or four chapters. Here we briefly introduce each of the 14 chapters.

Part I: Big Data Architecture and Platforms

1. **Big Data: A Classification of Acquisition and Generation Methods**
   Vijayakumar Nanjappan, Hai-Ning Liang, Wei Wang, Ka L. Man
   This chapter points out that it is very difficult to store, process, and analyze huge amounts of data using conventional computing methodologies and resources. The authors classify the data into digital and analog, environmental and personal. Data types and formats as well as input mechanisms are also highlighted. These will help us understand the active and passive methods of data collection and production.
2. **Cloud Computing Infrastructure for Data Intensive Applications**
   Yuri Demchenko, Fatih Turkmen, Cees de Laat, Ching-Hsien Hsu, Christophe Blanchet, Charles Loomis
   This chapter proposes a cloud-based big data infrastructure (BDI). The general architecture and functional components of BDI are described in detail. BDI is supported by the definition of the big data architecture framework (BDAF). Two case studies in bioinformatics are illustrated in the chapter to provide examples of requirements analysis and implementation.
3. **Open Source Private Cloud Platforms for Big Data**
   Martial Michel, Olivier Serres, Ahmad Anbar, Edmond J. Golden III, Tarek El-Ghazawi
   This chapter tells us that it is beneficial to use private clouds, especially open source clouds, for big data. Security, privacy, and customization are the major concerns. The chapter introduces the most prominent open source clouds in view of big data processing. A case study using an On-Premise Private Cloud is also presented to demonstrate the implementation of such an environment.

Part II: Big Data Processing and Management

4. **Efficient Nonlinear Regression-Based Compression of Big Sensing Data on Cloud**
   Chi Yang, Jinjun Chen
   This chapter proposes a compression method for big sensing data based on a nonlinear regression model. It improves the effectiveness and efficiency for processing real-world big sensing data. Regression design, least squares, and triangular transform are discussed in this chapter. It is demonstrated that the model achieves significant storage and time performance gains over other compression models.
5. **Big Data Management on Wireless Sensor Networks**
   Chih-Chieh Hung, Chu-Cheng Hsieh
   This chapter gives an overview of data management issues and solutions in wireless sensor networks. There are two possible models: centralized and decentralized. Data management can be centralized for the benefit of computation, or decentralized for energy saving. Three major issues for data management in both models are introduced: storage, query processing, and data collection. Some case studies are also discussed.
6. **Extreme Learning Machine and Its Applications in Big Data Processing**
   Cen Chen, Kenli Li, Mingxing Duan, Keqin Li
   This chapter first reviews the extreme learning machine (ELM) theory and its variants. Due to its memory-residency and high space/time complexity, the traditional ELM cannot train big data efficiently. Optimization strategies are necessary to solve this problem. Thus, parallel ELM algorithms based on MapReduce and Spark are described. Finally, practical applications of the ELM for big data are also presented in this chapter.

Part III: Big Data Analytics and Services

7. **Spatial Big Data Analytics for Cellular Communication Systems**
   Junbo Wang, Yilang Wu, Hui-Huang Hsu, Zixue Cheng
   This chapter surveys methodologies of spatial big data analytics and possible applications to support the cellular communication (CC) system. The CC system provides the most popular way to connect people. However, it still faces challenges, such as unbalanced crowd communication behavior and video transmission congestion. Spatial big data analytics can help the CC system to provide services with better quality of service (QoS). Challenging issues are highlighted in this chapter.
8. **Cognitive Applications and Their Supporting Architecture for Smart Cities**
   Haytham Assem, Lei Xu, Teodora S. Buda, Declan O'Sullivan
   This chapter proposes a cognitive architecture to enable big data applications with sensory data for smart cities. It deals with organization, configuration, security, and optimization. This chapter also reviews related work on location-based social networks and presents a novel approach to detect urban patterns, especially anomalies. This is essential for better understanding of human activities and behaviors.
9. **Deep Learning for Human Activity Recognition**
   Phyo P. San, Pravin Kakar, Xiao-Li Li, Shonali Krishnaswamy, Jian-Bo Yang, Minh N. Nguyen
   This chapter presents a systematic feature learning method for the problem of human activity recognition (HAR). It adopts a deep convolutional neural network (CNN) to automate feature learning from raw inputs. It is not necessary to handcraft features in advance. Such a

unification of feature learning and classification results in mutual enhancements. This is verified by comparing experimental results with several state-of-the-art techniques.

10. **Neonatal Cry Analysis and Categorization System Via Directed Acyclic Graph Support Vector Machine**
    Szu-Ta Chen, Kathiravan Srinivasan, Chen Lin, Chuan-Yu Chang
    This chapter introduces a neonatal cry analysis and categorization system. From the cry of the newborn, the system can identify different types of feelings such as pain, sleepiness, and hunger. The sequential forward floating selection (SFFS) algorithm is used to choose the discriminative features. The selected features are then used to classify the neonatal cries by the directed acyclic graph support vector machine (DAG-SVM). The system is useful for parents and nursing staff.

Part IV: Big Data Intelligence and IoT Systems

11. **Smart Building Applications and Information System Hardware Co-Design**
    Qian Huang, Chao Lu, Kang Chen
    This chapter emphasizes that a comprehensive understanding of information system hardware is necessary when designing efficient smart building applications. The necessity and importance of application and hardware co-design are discussed in this chapter. A case study is also given to show that application and hardware co-design optimize the smart building design from a system perspective.
12. **Smart Sensor Networks for Building Safety**
    Xuefeng Liu, Jiannong Cao
    This chapter presents the design and implementation of effective and energy-efficient structural health monitoring (SHM) algorithms in resource-limited wireless sensor networks (WSNs). Compared to traditional wired transmission, WSNs are low cost and easy to deploy for building monitoring. Distributed versions of SHM algorithms can help overcome the bandwidth limitation. A WSN-Cloud system architecture is also proposed for future SHM.
13. **The Internet of Things and Its Applications**
    Chung-Nan Lee, Tian-Hsiang Huang, Chen-Ming Wu, Ming-Chun Tsai
    This chapter first compares two lightweight protocols for the Internet of Things (IoT): MQ telemetry transport (MQTT) and the constrained application protocol (CoAP). Both protocols reduce the size of the packet and the over-loading of the bandwidth, thus saving battery power and storage space. The major techniques for big data analytics are then introduced. Finally, intelligent transportation systems and intelligent manufacturing systems are presented as examples.
14. **Smart Railway Based on the Internet of Things**
    Qingyong Y. Li, Zhangdui D. Zhong, Ming Liu, Weiwei W. Fang
    This chapter discusses the framework and technologies for a smart railway based on Internet of Things (IoT) and big data. The architecture of a smart railway, including the perception and action layer, the transfer layer, the data engine layer, and the application layer, is presented first. A case study on intelligent rail inspection is then introduced. This chapter shows that a smart railway is promising in improving traditional railway systems.

ACKNOWLEDGMENTS

This book is a part of the book series "Intelligent Data-Centric Systems." First of all, we would like to thank the series editor, Prof. Fatos Xhafa, for his encouragement and guidance in developing this book. We gratefully acknowledge all the contributing authors of the chapters. This book would not have been possible without their great efforts. We are also indebted to Ms. Ana Claudia Garcia, the editorial project manager, and the whole production team at Elsevier for their continuous help in producing this book. Finally, we thank our families for their love and support.

Hui-Huang Hsu, Chuan-Yu Chang, Ching-Hsien Hsu
September 2016