

Contents

List of Contributors	xiii
Preface	xvii

PART I BIG DATA ARCHITECTURE AND PLATFORMS

CHAPTER 1 Big Data: A Classification of Acquisition and Generation Methods	3
1 Big Data: A Classification	4
1.1 Characteristics of Big Data	4
2 Big Data Generation Methods	5
2.1 Data Sources	5
2.2 Data Types	6
3 Big Data: Data Acquisition Methods	8
3.1 Interface Methods	8
3.2 Interface Devices	11
4 Big Data: Data Management	13
4.1 Data Representation and Organization	13
4.2 Databases	16
4.3 Data Fusion and Data Integration	18
5 Summary	18
References	19
Glossary	20
CHAPTER 2 Cloud Computing Infrastructure for Data Intensive Applications	21
1 Introduction	22
2 Big Data Nature and Definition	23
2.1 Big Data in Science and Industry	23
2.2 Big Data and Social Network/Data	24
2.3 Big Data Technology Definition: From 6V to 5 Parts	24
3 Big Data and Paradigm Change	25
3.1 Big Data Ecosystem	25
3.2 New Features of the BDI	25
3.3 Moving to Data-Centric Models and Technologies	26
4 Big Data Architecture Framework and Components	27
4.1 Defining the Big Data Architecture Framework	27
4.2 Data Management and Big Data Lifecycle	28
4.3 Data Structures and Data Models for Big Data	29

4.4	NIST Big Data Reference Architecture	30
4.5	General Big Data System Requirements	33
5	Big Data Infrastructure.....	34
5.1	BDI Components	34
5.2	Big Data Stack Components and Technologies	36
5.3	Example of Cloud-Based Infrastructure for Distributed Data Processing	37
5.4	Benefits of Cloud Platforms for Big Data Applications	39
6	Case Study: Bioinformatics Applications Deployment on Cloud.....	39
6.1	Overall Description	39
6.2	UC1—Securing Human Biomedical Data	40
6.3	UC2—Cloud Virtual Pipeline for Microbial Genomes Analysis.....	41
6.4	Implementation of Use Cases and CYCLONE Infrastructure Components.....	42
7	CYCLONE Platform for Cloud Applications Deployment and Management	44
7.1	General Architecture for Intercloud and Multicloud Applications Deployment.....	44
7.2	Ensuring Consistent Security Services in Cloud-Based Applications	45
7.3	Dynamic Access Control Infrastructure.....	46
8	Cloud Powered Big Data Applications Development and Deployment Automation	47
8.1	Demand for Automated Big Data Applications Provisioning.....	47
8.2	Cloud Automation Tools for Intercloud Application and Network Infrastructure Provisioning.....	49
8.3	SlipStream: Cloud Application Management Platform.....	49
9	Big Data Service and Platform Providers.....	51
9.1	Amazon Web Services and Elastic MapReduce	52
9.2	Microsoft Azure Analytics Platform System and HDInsight.....	52
9.3	IBM Big Data Analytics and Information Management	54
9.4	Cloudera.....	54
9.5	Pentaho	54
9.6	LexisNexis HPCC Systems as an Integrated Open Source Platform for Big Data Analytics	55
10	Conclusion	57
	Acknowledgments	57
	References	58
	Glossary	60
CHAPTER 3	Open Source Private Cloud Platforms for Big Data	63
1	Cloud Computing and Big Data as a Service.....	63
1.1	Public Cloud Infrastructure	65
1.2	Advantages of the Cloud for Big Data	66

2	On-Premise Private Clouds for Big Data	66
2.1	Security of Cloud Computing Systems.....	66
2.2	Advantages of On-Premise Private Clouds	66
3	Introduction to Selected Open Source Cloud Environments	67
3.1	OpenNebula	67
3.2	Eucalyptus	69
3.3	Apache CloudStack	70
3.4	OpenStack.....	71
4	Heterogeneous Computing in the Cloud.....	75
4.1	Exclusive Mode	75
4.2	Sharing Mode	75
5	Case Study: The EMS, an On-Premise Private Cloud	76
6	Conclusion.....	78
	Disclaimer.....	78
	References	79

PART II BIG DATA PROCESSING AND MANAGEMENT

CHAPTER 4	Efficient Nonlinear Regression-Based Compression of Big Sensing Data on Cloud.....	83
1	Introduction	83
1.1	Motivation	84
1.2	Organization of the Chapter.....	85
2	Related Work and Problem Analysis.....	86
2.1	Related Work.....	86
2.2	Problem Analysis: Real-World Requirements for Nonlinear Regression.....	87
3	Temporal Compression Model Based on Nonlinear Regression	88
3.1	Nonlinear Regression Prediction Model.....	89
4	Algorithms	91
4.1	Algorithm for Nonlinear Regression.....	91
4.2	Nonlinear Regression Compression Algorithm Based on MapReduce	92
5	Experiments	94
5.1	Experiment Environment and Process.....	94
5.2	Experiment for the Compression With Nonlinear Regression.....	95
5.3	Experiment for Data Loss and Accuracy.....	96
6	Conclusions and Future Work.....	97
	References	97

CHAPTER 5	Big Data Management on Wireless Sensor Networks.....	99
1	Introduction	99
2	Data Management on WSNs.....	101
2.1	Storage	102
2.2	Query Processing.....	102
2.3	Data Collection.....	105
3	Big Data Tools	106
3.1	File System	106
3.2	Batch Processing	107
3.3	Streaming Data Processing	108
4	Put It Together: Big Data Management Architecture	109
4.1	Batch Layer	109
4.2	Serving Layer	110
4.3	Speed Layer.....	110
5	Big Data Management on WSNs.....	111
5.1	In-Network Aggregation Techniques and Data Integration Components	111
5.2	Exploiting Big Data Systems as Data Centers	112
6	Conclusion	114
	References	114
	Glossary	116
CHAPTER 6	Extreme Learning Machine and Its Applications in Big Data Processing.....	117
1	Introduction	118
1.1	Background.....	118
1.2	Artificial Neural Networks.....	118
1.3	Era of Big Data	120
1.4	Organization	121
2	Extreme Learning Machine.....	121
2.1	Traditional Approaches to Train ANNs.....	121
2.2	Theories of the Extreme Learning Machine	123
2.3	Classical ELM	125
2.4	ELM for Classification and Regression.....	126
2.5	ELM for Unsupervised Learning	129
3	Improved Extreme Learning Machine With Big Data	131
3.1	Shortcomings of the Extreme Learning Machine for Processing Big Data ...	131
3.2	Optimization Strategies for the Traditional Extreme Learning Machine	131
3.3	Efficiency Improvement for Big Data	132
3.4	Parallel Extreme Learning Machine Based on MapReduce	134
3.5	Parallel Extreme Learning Machine Based on Apache Spark	140

4	Applications.....	144
4.1	ELM in Predicting Protein Structure	144
4.2	ELM in Image Processing.....	145
4.3	ELM in Cancer Diagnosis.....	145
4.4	ELM in Big Data Security and Privacy.....	146
5	Conclusion	146
	References	146
	Glossary	150

PART III BIG DATA ANALYTICS AND SERVICES

CHAPTER 7 Spatial Big Data Analytics for Cellular Communication Systems..... 153

1	Introduction	153
2	Cellular Communications and Generated Data	154
3	Spatial Big Data Analytics.....	155
3.1	Statistical Foundation for Spatial Big Data Analytics	155
3.2	Spatial Pattern Mining From Spatial Big Data Analytics	156
4	Typical Applications	161
4.1	BS Behavior Understanding Through Spatial Big Data Analytics	161
4.2	User Behavior Understanding Through Spatial Big Data Analytics	163
5	Conclusion and Future Challenging Issues.....	164
	Acknowledgments	164
	References	164
	Glossary	166

CHAPTER 8 Cognitive Applications and Their Supporting Architecture for Smart Cities..... 167

1	Introduction	167
2	CSE for Smart City Applications.....	168
2.1	Architecture Specification.....	169
2.2	Big Data Analysis and Management	171
3	Anomaly Detection in Smart City Management	171
3.1	Related Work to Anomaly Detection	172
3.2	Challenges and Benefits of Anomaly Detection in Smart Cities.....	173
4	Functional Region and Socio-Demographic Regional Patterns	
	Detection in Cities.....	174
4.1	Discovering Functional Regions.....	174
4.2	Deep Learning and Regional Pattern Detections.....	176
5	Summary.....	181
	References	181
	Glossary	184

CHAPTER 9	Deep Learning for Human Activity Recognition	186
1	Introduction	186
2	Motivations and Related Work	189
3	Convolutional Neural Networks in HAR.....	190
3.1	Temporal Convolution and Pooling.....	190
3.2	Architecture	192
3.3	Analysis	194
4	Experiments, Results, and Discussion	195
4.1	Experiment on OAR Dataset.....	195
4.2	Experiment on Hand Gesture Dataset.....	198
4.3	Experiment on REALDISP Dataset.....	199
4.4	Computational Requirements	201
4.5	Future Directions	202
5	Conclusion	202
	References	202
	Glossary	204
CHAPTER 10	Neonatal Cry Analysis and Categorization System Via Directed Acyclic Graph Support Vector Machine	205
1	Introduction	205
2	Neonatal Cry Analysis and Categorization System.....	206
2.1	Cry Signal Preprocessing	206
2.2	Feature Extraction—Essential Features	210
2.3	Selection of Features	214
2.4	Categorization and Validation	214
3	Experimental Results and Discussion	216
3.1	Environment of the Experiments	216
3.2	Experiment 1: Neonatal Cry Analysis and Categorization—Employing 15 Extracted Features	217
3.3	Experiment 2: Neonatal Cry Analysis and Categorization—Deploying the Selected Four Features	218
3.4	Experiment 3: Comparison of Neonatal Cry Analysis and Categorization Between Male and Female Babies.....	219
3.5	Experiment 4: Comparison of Proposed System With Y. Abdulaziz's Approach.....	219
4	Conclusion	220
	Acknowledgment.....	221
	References	221

PART IV BIG DATA INTELLIGENCE AND IoT SYSTEMS

CHAPTER 11 Smart Building Applications and Information System Hardware Co-Design.....	225
1 Smart Building Applications.....	225
1.1 The Ever-Increasing Need for Smart Buildings	225
1.2 Smart Building Applications.....	227
2 Emerging Information System Hardware	230
2.1 Overview.....	230
2.2 Examples	230
3 Big Data Application and Information Hardware Co-Design.....	235
3.1 Motivation and Challenge	235
3.2 Case Study and Discussion	237
4 Conclusions.....	239
References	239
Glossary	240
CHAPTER 12 Smart Sensor Networks for Building Safety	241
1 Introduction	241
2 Related Works	244
3 Background: Modal Analysis.....	245
3.1 Modal Parameters.....	246
3.2 The ERA.....	247
4 Distributed Modal Analysis	249
4.1 Stage 1: Try to Distribute the Initial Stage of Modal Analysis Algorithms...	249
4.2 Stage 2: Divide and Conquer.....	250
5 A Multiscale SHM Using Cloud.....	252
6 Conclusion	254
Acknowledgments	254
References	254
Glossary	255
CHAPTER 13 The Internet of Things and Its Applications	256
1 Introduction	256
2 Collection of Big Data From IoT	257
2.1 MQ Telemetry Transport	259
2.2 Constrained Application Protocol.....	261
2.3 MQTT vs. CoAP	262

3 IoT Analytics.....	263
3.1 Related Works	263
3.2 Outlier Detection for Big Data	265
3.3 Island-Based Cloud GA	268
4 Examples of IoT Applications	268
4.1 Applications on Intelligent Transportation Systems.....	269
4.2 Applications on Intelligent Manufacturing Systems	272
5 Conclusions.....	277
References	277
Glossary	279
CHAPTER 14 Smart Railway Based on the Internet of Things.....	280
1 Introduction	280
2 Architecture of the Smart Railway	281
2.1 Overview.....	281
2.2 Perception and Action Layer	282
2.3 Transfer Layer	284
2.4 Data Engine Layer.....	284
2.5 Application Layer.....	286
3 IRIS for Smart Railways.....	287
3.1 Rail Defects	287
3.2 The State-of-the-Art for Rail Inspection.....	289
3.3 Rail Inspection Based on the IoT and Big Data.....	294
4 Conclusion	295
Acknowledgment.....	296
References	296
Glossary	297
Index	299

Lis

Ahmad
The G

Haytha
IBM, D

Christo
CNRS

Teodor
IBM, D

Jiannor
The H

Chuan-
Nation

Jinjun
Univer

Gen C
Hunar

Szu-Ta
Nation

Kang
South

Zixue
Univer

Cees
Univer

Yuri D
Univer

Ming
Hunar

Tarek
The G

Wei
Beijin
Univer

Edmo
Nation

Chu-
Slice