

Contents

About the Editors.....	xv
About the Contributors	xvii
Preface	xxiii
Acknowledgments	xxvii

CHAPTER 1	Characteristics of Intelligent Transportation Systems and Its Relationship With Data Analytics	1
	<i>Sakib M. Khan, Mizanur Rahman, Amy Apon, and Mashrur Chowdhury</i>	
1.1	Intelligent Transportation Systems as Data-Intensive Applications	1
1.1.1	ITS Data System	2
1.1.2	ITS Data Sources and Data Collection Technologies.....	3
1.2	Big Data Analytics and Infrastructure to Support ITS.....	4
1.3	ITS Architecture: The Framework of ITS Applications	9
1.3.1	User Services and User Service Requirements	10
1.3.2	Logical Architecture	11
1.3.3	Physical Architecture	11
1.3.4	Service Packages.....	12
1.3.5	Standards	13
1.3.6	Security.....	14
1.4	Overview of ITS Applications	14
1.4.1	Types of ITS Applications.....	15
1.4.2	ITS Application and Its Relationship to Data Analytics	18
1.5	Intelligent Transportation Systems Past, Present, and Future.....	21
1.5.1	1960's and 1970's	21
1.5.2	1980's and 1990's	21
1.5.3	2000's.....	22
1.5.4	2010's and Beyond	23
1.6	Overview of Book: Data Analytics for ITS Applications.....	24
	Exercise Problems	26
	References	27

CHAPTER 2	Data Analytics: Fundamentals	31
	<i>Venkat N. Gudivada</i>	
2.1	Introduction	31
2.2	Functional Facets of Data Analytics.....	32

2.2.1	Descriptive Analytics	33
2.2.2	Diagnostic Analytics	41
2.2.3	Predictive Analytics	43
2.2.4	Prescriptive Analytics	45
2.3	Evolution of Data Analytics	45
2.3.1	SQL Analytics: RDBMS, OLTP, and OLAP	46
2.3.2	Business Analytics: Business Intelligence, Data Warehousing, and Data Mining	47
2.3.3	Visual Analytics	53
2.3.4	Big Data Analytics	54
2.3.5	Cognitive Analytics	54
2.4	Data Science	55
2.4.1	Data Lifecycle	56
2.4.2	Data Quality	57
2.4.3	Building and Evaluating Models	58
2.5	Tools and Resources for Data Analytics	60
2.6	Future Directions	62
2.7	Chapter Summary and Conclusions	63
2.8	Questions and Exercise Problems	64
	References	65

CHAPTER 3 Data Science Tools and Techniques to Support Data Analytics in Transportation Applications 69

Linh B. Ngo

3.1	Introduction	69
3.2	Introduction to the R Programming Environment for Data Analytics	70
3.3	Research Data Exchange	72
3.4	Fundamental Data Types and Structures: Data Frames and List	72
3.4.1	Data Frame	73
3.4.2	List	75
3.5	Importing Data from External Files	75
3.5.1	Delimited	75
3.5.2	XML	78
3.5.3	SQL	83
3.6	Ingesting Online Social Media Data	84
3.6.1	Static Search	85
3.6.2	Dynamic Streaming	86
3.7	Big Data Processing: Hadoop MapReduce	87
3.8	Summary	90

3.9 Exercises.....	90
References	90
CHAPTER 4 The Centrality of Data: Data Lifecycle and Data Pipelines.....	91
<i>Beth Plale and Inna Kouper</i>	
4.1 Introduction	91
4.2 Use Cases and Data Variability	92
4.3 Data and Its Lifecycle	95
4.3.1 The USGS Lifecycle Model	95
4.3.2 Digital Curation Center (DCC) Curation Model.....	96
4.3.3 DataONE Model	98
4.3.4 SEAD Research Object Lifecycle Model.....	99
4.4 Data Pipelines.....	102
4.5 Future Directions.....	107
4.6 Chapter Summary and Conclusions.....	108
4.7 Exercise Problems and Questions	108
4.7.1 Exercise 1. Defining and Describing Research Data.....	108
4.7.2 Exercise 2. Mapping Research Project Onto the Lifecycle	109
4.7.3 Exercise 3. Data Organization	109
4.7.4 Exercise 4. Data Pipelines	109
References.....	110
CHAPTER 5 Data Infrastructure for Intelligent Transportation Systems.....	113
<i>Andre Luckow and Ken Kennedy</i>	
5.1 Introduction	113
5.2 Connected Transport System Applications and Workload Characteristics	114
5.3 Infrastructure Overview	115
5.4 Higher-Level Infrastructure.....	117
5.4.1 MapReduce and Beyond: Scalable Data Processing.....	117
5.4.2 Data Ingest and Stream Processing	119
5.4.3 SQL and Dataframes.....	120
5.4.4 Short-Running and Random Access Data Management.....	121
5.4.5 Search-Based Analytics	121
5.4.6 Business Intelligence and Data Science	121
5.4.7 Machine Learning	122
5.5 Low-Level Infrastructure	122
5.5.1 Hadoop: Storage and Compute Management.....	123
5.5.2 Hadoop in the Cloud.....	123
5.6 Chapter Summary and Conclusions.....	125
Exercise Problems and Questions	125
References.....	126

CHAPTER 6	Security and Data Privacy of Modern Automobiles.....	131
	<i>Juan Deng, Lu Yu, Yu Fu, Oluwakemi Hambolu, and Richard R. Brooks</i>	
6.1	Introduction	131
6.2	Connected Vehicle Networks and Vehicular Applications.....	132
6.2.1	In-Vehicle Networks	132
6.2.2	External Networks.....	133
6.2.3	Innovative Vehicular Applications	133
6.3	Stakeholders and Assets.....	135
6.4	Attack Taxonomy	137
6.5	Security Analysis.....	137
6.5.1	Network and Protocol Vulnerability Analysis	138
6.5.2	Attacks.....	140
6.6	Security and Privacy Solutions	146
6.6.1	Cryptography Basics	147
6.6.2	Security Solutions for Bus Communications	148
6.6.3	WPAN Security and Privacy	152
6.6.4	Secure VANETs.....	153
6.6.5	Secure OTA ECU Firmware Update	155
6.6.6	Privacy Measurement of Sensor Data	157
6.6.7	Secure Handover	158
6.7	Future Research Directions	158
6.8	Summary and Conclusions.....	159
6.9	Exercises.....	159
	References	159
CHAPTER 7	Interactive Data Visualization	165
	<i>Chad A. Steed</i>	
7.1	Introduction	165
7.2	Data Visualization for Intelligent Transportation Systems	167
7.3	The Power of Data Visualization.....	167
7.4	The Data Visualization Pipeline	169
7.5	Classifying Data Visualization Systems	171
7.6	Overview Strategies	172
7.6.1	Data Quantity Reduction	173
7.6.2	Miniaturizing Visual Glyphs	174
7.7	Navigation Strategies	175
7.7.1	Zoom and Pan	176
7.7.2	Overview + Detail.....	176
7.7.3	Focus + Context	177

7.8	Visual Interaction Strategies	177
7.8.1	Selecting	177
7.8.2	Linking	178
7.8.3	Filtering	178
7.8.4	Rearranging and Remapping	179
7.9	Principles for Designing Effective Data Visualizations	179
7.10	A Case Study: Designing a Multivariate Visual Analytics Tool	181
7.10.1	Multivariate Visualization Using Interactive Parallel Coordinates	182
7.10.2	Dynamic Queries Through Direct Manipulation	182
7.10.3	Dynamic Variable Summarization via Embedded Visualizations	183
7.10.4	Multiple Coordinated Views	183
7.11	Chapter Summary and Conclusions	185
7.12	Exercises	186
7.13	Sources for More Information	187
7.13.1	Journals	187
7.13.2	Conferences	187
	References	187

CHAPTER 8 Data Analytics in Systems Engineering for Intelligent Transportation Systems 191

Ethan T. McGee and John D. McGregor

8.1	Introduction	191
8.2	Background	192
8.2.1	Systems Development V Model	192
8.2.2	Continuous Engineering	194
8.2.3	AADL	195
8.3	Development Scenario	202
8.3.1	Data Analytics in Architecture	202
8.3.2	The Scenario	203
8.4	Summary and Conclusion	209
8.5	Exercises	209
8.6	Appendix A	211
8.6.1	EMV2 Error Ontology	211
	References	213

CHAPTER 9 Data Analytics for Safety Applications 215

Yuanchang Xie

9.1	Introduction	215
9.2	Overview of Safety Research	215
9.2.1	Human Factors	215
9.2.2	Crash Count/Frequency Modeling	216

9.2.3	Before and After Study	217
9.2.4	Crash Injury Severity Modeling	217
9.2.5	Commercial Vehicle Safety	218
9.2.6	Data Driven Highway Patrol Plan	218
9.2.7	Deep Learning from Big and Heterogeneous Data for Safety	219
9.2.8	Real-Time Traffic Operation and Safety Monitoring	219
9.2.9	Connected Vehicles and Traffic Safety	220
9.3	Safety Analysis Methods	221
9.3.1	Statistical Methods	221
9.3.2	Artificial Intelligence and Machine Learning	225
9.4	Safety Data	227
9.4.1	Crash Data	228
9.4.2	Traffic Data	228
9.4.3	Roadway Data	229
9.4.4	Weather Data	230
9.4.5	Vehicle and Driver Data	230
9.4.6	Naturalistic Driving Study	230
9.4.7	Big Data and Open Data Initiatives	231
9.4.8	Other Data	233
9.5	Issues and Future Directions	233
9.5.1	Issues With Existing Safety Research	233
9.5.2	Future Directions	234
9.6	Chapter Summary and Conclusions	235
9.7	Exercise Problems and Questions	236
	References	237

CHAPTER 10 Data Analytics for Intermodal Freight Transportation

	Applications	241
	<i>Nathan Huynh, Majbah Uddin, and Chu Cong Minh</i>	
10.1	Introduction	241
10.1.1	ITS-Enabled Intermodal Freight Transportation	241
10.1.2	Data Analytics for ITS-Enabled Intermodal Freight Transportation	242
10.2	Descriptive Data Analytics	242
10.2.1	Univariate Analysis	242
10.2.2	Bivariate Analysis	247
10.3	Predictive Data Analytics	249
10.3.1	Bivariate Analysis	249
10.3.2	Multivariate Analysis	253
10.3.3	Fuzzy Regression	256
10.4	Summary and Conclusions	259

10.5	Exercise Problems	260
10.6	Solution to Exercise Problems	261
	References	261

CHAPTER 11 Social Media Data in Transportation 263

Sakib M. Khan, Linh B. Ngo, Eric A. Morris, Kakan Dey, and Yan Zhou

11.1	Introduction to Social Media	263
11.2	Social Media Data Characteristics	264
11.2.1	Volume and Velocity	265
11.2.2	Veracity	266
11.2.3	Variety	266
11.2.4	Value	266
11.3	Social Media Data Analysis	267
11.4	Application of Social Media Data in Transportation	270
11.4.1	Transportation Planning	270
11.4.2	Traffic Prediction	270
11.4.3	Traffic Management During Planned Events	271
11.4.4	Traffic Management During Unplanned Events	271
11.4.5	Traffic Information Dissemination	272
11.5	Future Research Issues/Challenges for Data Analytics-Enabled Social Media Data	272
11.5.1	Social Media: A Supplemental Transportation Data Source	272
11.5.2	Potential Data Infrastructure	273
11.6	Summary	277
11.7	Conclusions	277
11.8	Exercise Problems	278
	References	278

CHAPTER 12 Machine Learning in Transportation Data Analytics 283

*Parth Bhavsar, Ilya Safro, Nidhal Bouaynaya, Robi Polikar,
and Dimah Dera*

12.1	Introduction	283
12.2	Machine Learning Methods	284
12.2.1	Supervised Learning	284
12.2.2	Unsupervised Learning	285
12.3	Understanding Data	286
12.3.1	Problem Definition	286
12.3.2	Data Collection	287
12.3.3	Data Fusion	288
12.3.4	Data Preprocessing	289

12.4	Machine Learning Algorithms for Data Analytics.....	290
12.4.1	Regression Methods.....	290
12.4.2	Decision Trees.....	293
12.4.3	Neural Networks	295
12.4.4	Support Vector Machine.....	297
12.4.5	Clustering	298
12.4.6	Evaluation	299
12.5	An Example.....	300
12.6	Summary.....	303
12.7	Questions and Solutions.....	303
	References	304
	Appendix	305

Index	309
-------------	-----