

Contents

Contributors	xi
List of Figures	xiii
List of Tables	xix
Preface	xxi
Acronyms	xxv

1. Overview of Data-Driven Solutions	1
<i>Yinhai Wang and Ziqiang Zeng</i>	
1.1 General Background	1
1.1.1 Government Investment	2
1.1.2 Academic Community Research Trend	3
1.1.3 Transportation Industry Involvement	3
1.2 Data-Driven Innovation in Transportation Science	4
1.3 Methodologies for Data-Driven Transportation Science	5
1.4 Applications in Data-Driven Transportation Science	6
1.5 Overview and Roadmap	7
References	9
2. Data-Driven Energy Efficient Driving Control in Connected Vehicle Environment	11
<i>Xuwei Qi, Guoyuan Wu, Kanok Boriboonsomsin and Matthew J. Barth</i>	
2.1 Introduction	13
2.2 Background and State of the Art	14
2.2.1 PHEV Modeling	14
2.2.2 Operation Mode and SOC Profile	14
2.2.3 EMS for PHEVs	15
2.2.4 PHEVs' SOC Control	16
2.3 Problem Formulation	17
2.3.1 Data-Driven On-Line EMS Framework for PHEVs	17
2.3.2 Optimal Power-Split Control Formulation	19
2.4 Data-Driven Evolutionary Algorithm (EA) Based Self-Adaptive On-Line Optimization	20
2.4.1 Optimality and Complexity	23

2.4.2	SOC Control Strategies	23
2.4.3	EDA-Based On-Line EMS Algorithm With SOC Control	25
2.4.4	Synthesized Trip Information	27
2.4.5	Off-Line Optimization for Validation	28
2.4.6	Real-Time Performance Analysis and Parameter Tuning	28
2.4.7	On-Line Optimization Performance Comparison	29
2.4.8	Analysis of Trip Duration	31
2.4.9	Performance With Charging Opportunity	33
2.5	Data-Driven Reinforcement Learning-Based Real-Time EMS	34
2.5.1	Introduction	34
2.5.2	Dynamic Programming	36
2.5.3	Approximate Dynamic Programming and Reinforcement Learning	37
2.5.4	Reinforcement Learning-Based EMS	38
2.5.5	Action and Environmental States	39
2.5.6	Reward Initialization (With Optimal Results From Simulation)	40
2.5.7	Q-Value Update and Action Selection	41
2.5.8	Validation and Testing	42
2.5.9	Model Without Charging Opportunity (Trip Level)	42
2.5.10	Model With Charging Opportunity (Tour Level)	44
2.6	Conclusions	47
	References	47
3.	Machine Learning and Computer Vision-Enabled Traffic Sensing Data Analysis and Quality Enhancement	51
	<i>Guohui Zhang and Yinhai Wang</i>	
3.1	Introduction	51
3.1.1	Significance of Vehicle Classification Volumes	51
3.1.2	Research Motivation	52
3.1.3	Research Objectives	53
3.2	State of the Art and Practice	53
3.2.1	Single-Loop Vehicle Length Estimation and Machine Learning Application	53
3.2.2	Computer Vision-Based Traffic Detection	54
3.3	Methodology	55
3.3.1	Machine Learning Approach for Vehicle Classification Volume Estimation	55
3.3.2	Computer Vision Algorithms to Measure Vehicle Classification Volumes	59
3.4	Experimental Tests and Discussions	66
3.4.1	ANN Approach Performance Evaluation	66
3.4.2	VVDC System Performance Evaluation	71
3.5	Conclusions	76
	References	77

4.	Data-Driven Approaches for Estimating Travel Time Reliability	81
	<i>Shu Yang and Yao-Jan Wu</i>	
4.1	Introduction	81
4.1.1	Significance of Travel Time Reliability	81
4.1.2	Definition of TTR	82
4.1.3	Motivation and Research Questions	83
4.1.4	Chapter Organization	83
4.2	State of the Art and the Practice	84
4.2.1	Probability Distribution Family Selection for Travel Time Distribution	84
4.2.2	Data Size Selection in Estimating TTR	85
4.2.3	Freeway TTR Measures	86
4.3	Estimating Freeway TTR and Its Accuracy	87
4.3.1	TTR Measures	87
4.3.2	Insensitivity of Probability Distribution Family Selection	90
4.3.3	Introduction to the Bootstrap	94
4.3.4	Accuracy of TTR Measures	98
4.3.5	Optimal Quantity of Travel Time	100
4.4	From Segment-Based TTR to OD-Based TTR	101
4.4.1	Significance of OD-Based TTR	102
4.4.2	OD-Based TTR Measurement	102
4.4.3	OD-Based TTR Information Delivery	106
4.5	Conclusion and Recommendations	106
	References	108
5.	Urban Travel Behavior Study Based on Data Fusion Model	111
	<i>Meng Li, Mingqiao Zou and Huiping Li</i>	
5.1	Introduction	111
5.2	Research Background	113
5.3	Agent-Based Traveler Behavior Model	115
5.3.1	Travel Behavior Data Collection	115
5.3.2	Model Development	117
5.3.3	Policy and Scenario Analysis	124
5.4	Behavior Model in Cooperation of VMS and Traffic Signal	126
5.4.1	Drivers' Diversion Model	126
5.4.2	Cooperative Mechanism of VMS and TSC	129
5.4.3	Applications	131
5.5	Conclusions	134
	References	134

6. Urban Travel Mobility Exploring With Large-Scale Trajectory Data	137
<i>Jinjun Tang</i>	
6.1 Introduction	137
6.2 Transportation Demand Analysis and Attractiveness Modeling	139
6.2.1 Data Source	139
6.2.2 Distribution Pattern of Demand	139
6.2.3 Clustering Based on DBSCAN	142
6.2.4 Attractiveness Model for Choosing Pick-Up Clusters	145
6.3 Trips Distribution Analysis	147
6.3.1 Distance Distribution	148
6.3.2 Travel Time Distribution	149
6.3.3 Average Speed Distribution	153
6.4 Traffic Distribution Based on Entropy-Maximizing Model	155
6.5 Network Construction and Dynamic Characteristics	158
6.5.1 Degree and Strength Distribution	158
6.5.2 Degree vs Strength Distribution	159
6.5.3 $k_j^{\text{out}} k_j^{\text{in}}$ vs w_{ij} Correlation	161
6.5.4 Betweenness vs Strength and Clustering Coefficient	162
6.5.5 Network Construction and Structure Entropy	164
6.6 Spatial-Temporal Properties of Urban Travel	166
6.6.1 Traffic Zone Identification	166
6.6.2 Travel Pattern Analysis	168
6.6.3 Hotspot Analysis	169
6.7 Conclusions	171
References	173
7. Public Transportation Big Data Mining and Analysis	175
<i>Xiaolei Ma and Xi Chen</i>	
7.1 Introduction	175
7.2 Public Transportation Big Data Preprocessing Method	178
7.2.1 Public Transportation Smart Card Data Cleaning	178
7.2.2 GPS Data Cleaning	180
7.3 Application of Public Transportation Data in Planning	181
7.3.1 Extraction of the Commuting Characteristics of Public Transportation Passengers	183
7.3.2 Identification of Commuters and Estimation of Their Places of Work and Residence	184
7.4 Application of Public Transportation Data in Operation and Management	189
7.4.1 Prediction Model for Public Transportation Bus Arrival Times	190
7.4.2 Case Study	191

7.5 Introduction of a Public Transportation Big Data Platform Based on E-Science	192
7.5.1 Main Functions of the Public Transportation Big Data Platform	194
7.5.2 Functions of the Public Transportation Big Data Platform	195
7.6 Conclusions	199
Acknowledgment	199
References	199
8. Simulation-Based Optimization for Network Modeling With Heterogeneous Data	201
<i>Xiqun (Michael) Chen</i>	
8.1 Introduction	201
8.2 Literature Review	203
8.3 Simulation-Based Optimization	204
8.3.1 Framework	204
8.3.2 Design of Experiments (DoE)	205
8.3.3 Surrogate Models	205
8.3.4 Link-Based and Path-Based MFD	207
8.3.5 Calibration and Exploitation	208
8.4 Application	209
8.4.1 Heterogeneous Data	209
8.4.2 Simulation Network	209
8.4.3 Validation	211
8.4.4 SBO Results	216
8.5 Conclusions	221
Acknowledgments	221
References	223
9. Network Modelling and Resilience Analysis of Air Transportation: A Data-Driven, Open-Source Approach	227
<i>Xiaoqian Sun and Sebastian Wandelt</i>	
9.1 Introduction	227
9.2 Data Preparation	228
9.3 Air Transportation Network Modeling	229
9.4 Air Transportation Network Analysis	234
9.4.1 Centralities	234
9.4.2 Robustness Curves	237
9.4.3 Air-Side Accessibility of Nodes	239
9.4.4 Communities	239
9.4.5 Airline Networks	241
9.4.6 Multiple Airport Regions	242
9.5 Conclusions	244
References	244

10. Health Assessment of Electric Multiple Units	247
<i>Tianyun Shi, Haiyan Shen, Li Li, Peng Sun and Ge Guo</i>	
10.1 Introduction	247
10.2 Data Source and Structure	249
10.3 Health Assessment of EMU	253
10.3.1 Feature Layer Data Fusion	253
10.3.2 Decision-Making Level Data Fusion	254
10.4 Data Application and Analysis	256
10.4.1 Feature Layer Health Data Analysis	256
10.4.2 Decision-Making Level Health Data Analysis	258
10.5 Conclusion and Outlook	262
References	262
Index	265

Contributors

Numbers in Parentheses indicate the pages on which the author's contributions begin.

Matthew J. Barth (11), Department of Electrical and Computer Engineering; College of Engineering-Centre for Environmental Research and Technology (CE-CERT), University of California, Riverside, CA, United States

Kanok Boriboonsomsin (11), College of Engineering-Centre for Environmental Research and Technology (CE-CERT), University of California, Riverside, CA, United States

Xi Chen (175), School of Transportation Science and Engineering, Beihang University, Beijing, People's Republic of China

Xiqun (Michael) Chen (201), College of Civil Engineering and Architecture, Zhejiang University, Hangzhou, People's Republic of China

Ge Guo (247), Institute of Computing Technology, China Academy of Railway Sciences, Beijing, People's Republic of China; Department of Civil and Environmental Engineering, University of Washington, Seattle, WA, United States

Meng Li (111), Department of Civil Engineering, Tsinghua University, Beijing, People's Republic of China

Huiping Li (111), Department of Civil Engineering, Tsinghua University, Beijing, People's Republic of China

Li Li (247), Institute of Computing Technology, China Academy of Railway Sciences, Beijing, People's Republic of China

Xiaolei Ma (175), School of Transportation Science and Engineering, Beihang University, Beijing, People's Republic of China

Xuewei Qi (11), Department of Electrical and Computer Engineering; College of Engineering-Centre for Environmental Research and Technology (CE-CERT), University of California, Riverside, CA, United States

Haiyan Shen (247), Institute of Computing Technology, China Academy of Railway Sciences, Beijing, People's Republic of China

Tianyun Shi (247), Institute of Computing Technology, China Academy of Railway Sciences, Beijing, People's Republic of China

Xiaoqian Sun (227), National Key Laboratory of CNS/ATM, School of Electronic and Information Engineering, Beihang University, Beijing, People's Republic of China

Peng Sun (247), Institute of Computing Technology, China Academy of Railway Sciences, Beijing, People's Republic of China