

Spatial Network Big Databases

KwangSoo Yang · Shashi Shekhar

Spatial Network Big Databases

Queries and Storage Methods



Springer

KwangSoo Yang
Computer Science Department
Florida Atlantic University
Boca Raton, FL
USA

Shashi Shekhar
Department of Computer Science
and Engineering
University of Minnesota
Minneapolis, MN
USA

ISBN 978-3-319-56656-6 ISBN 978-3-319-56657-3 (eBook)
DOI 10.1007/978-3-319-56657-3

Library of Congress Control Number: 2017937267

© Springer International Publishing AG 2017

This work is subject to copyright. All rights are reserved by the Publisher, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed.

The use of general descriptive names, registered names, trademarks, service marks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.

The publisher, the authors and the editors are safe to assume that the advice and information in this book are believed to be true and accurate at the date of publication. Neither the publisher nor the authors or the editors give a warranty, express or implied, with respect to the material contained herein or for any errors or omissions that may have been made. The publisher remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Printed on acid-free paper

This Springer imprint is published by Springer Nature
The registered company is Springer International Publishing AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

Preface

Spatial Network Big Data (SNBD) refers to spatial network datasets whose size, variety, or update rate exceeds the capacity of commonly used spatial network computing and spatial network database technologies to learn, manage, and process with reasonable effort. SNBD has the potential to transform society via next-generation routing services, emergency and disaster response, and discovery of potentially useful patterns embedded in these datasets. The use of SNBD is rapidly expanding into the transportation arena to improve the management and security of transportation infrastructure and enable data-driven decision-making. However, there are significant challenges to the use of SNBD because current methods, models, and algorithms do not always scale and/or perform well when storing, managing, and analyzing large volumes of data. Interestingly, most of the SNBD collected today are not used at all, and data that are used are not fully exploited. In addition, SNBD datasets tend to be used mostly for real-time control or anomaly detection, rather than optimization or prediction based on historical data, which is where their greatest value lies. Investigating novel SNBD storage and processing platforms, such as database systems and data analytics tools, is critical to realize the full value of SNBD.

Developing Spatial Network Big Database Systems (SNBDS) requires overcoming three key challenges. First, it requires new data models to represent the complex and interrelated structure of SNBD. Second, fully exploiting SNBD requires scalable query processing and optimization methods, which are currently lacking. Finally, SNBDS require I/O efficient storage and access methods that leverage scalability and efficiency of big data query processing. These challenges lead us to rethink both existing theories and models. This book presents a collection of concepts, algorithms, and techniques that effectively harness the power of SNBD. Reading this book is a first step toward understanding the immense challenges and novel applications of SNBD database systems. This book is organized in seven chapters. After reviewing some preliminaries in Chap. 1, we

introduce basic graph algorithms in Chap. 2. Chapters 3–5 formally model spatial network query problems and explore algorithms that minimize the computational cost for query processing. Chapter 6 introduces strategies to develop I/O efficient storage and access methods. Chapter 7 summarizes the book’s major themes.

Boca Raton, FL, USA
January 2017

KwangSoo Yang

Acknowledgements

I would like to express my sincere appreciation and thanks to my advisor Prof. Shashi Shekhar, Department of Computer Science, University of Minnesota. His enthusiasm, inspiration, and great efforts helped me to grow as a research scientist. I would like to thank Springer for their advice and support to publish this book. I would like to express my gratitude to the University of Minnesota Spatial Computing Research Group for their brilliant comments and suggestions. I would like to thank Kim Koffolt for improving the readability of this book. Lastly, I thank my mother, sister, brother, and the Lundberg family for everything they have done for me.

Contents

1 Spatial Network Big Databases: An Introduction	1
1.1 Spatial Network Big Data	1
1.2 Application Domain	2
1.3 Spatial Network Big Database Management Systems	2
1.4 Computational Challenges	6
References	7
2 Basic Graph Algorithms	9
2.1 A Brief Introduction to Graph Theory	9
2.2 Network Representations	9
2.2.1 Node-Node Adjacency Matrix	10
2.2.2 Node-Edge Incidence Matrix	10
2.2.3 Adjacency List	11
2.2.4 Forward Star	12
2.3 Shortest Paths	12
2.3.1 Single-Source Shortest Path (SSSP)	13
2.3.2 All-Pairs Shortest Paths (APSP)	14
2.4 Block Decomposition	15
2.5 Maximum Network Flow	16
2.5.1 Augmenting-Path Algorithm	16
2.5.2 Push-Relabel Algorithm	18
2.6 Bipartite Weighted Matching	19
2.7 Graph Partitioning	22
References	24
3 Capacity Constrained Network Voronoi Diagrams	27
3.1 Introduction	27
3.1.1 Application Domains	27
3.1.2 Problem Definition	28
3.1.3 Problem Hardness	29

3.1.4	Literature Review	30
3.1.5	Outline of the Chapter	31
3.2	Algorithms for Capacity Constrained Network Voronoi Diagram	32
3.2.1	Pressure Equalizer (PE) Algorithm	32
3.2.2	PE-BTCC Algorithm	35
3.2.3	PE-Minor Algorithm	42
3.3	Case Study with Brooklyn, NY Road Network	43
3.4	Summary	44
	References	44
4	Distance-Constrained k Spatial Sub-networks	47
4.1	Introduction	47
4.1.1	Application Domain	47
4.1.2	Problem Definition	48
4.1.3	Problem Hardness	49
4.1.4	Literature Review	49
4.1.5	Outline of the Chapter	50
4.2	Algorithm for Distance-Constrained k Spatial Sub-networks	50
4.3	Case Study with Chicago Road Network	54
4.4	Summary	54
	References	55
5	Evacuation Route Planning	57
5.1	Introduction	57
5.1.1	Application Domain	57
5.1.2	Problem Definition	58
5.1.3	Literature Review	58
5.1.4	Outline of the Chapter	59
5.2	Algorithms for Evacuation Route Planning	60
5.2.1	Capacity Constrained Route Planner Algorithm	60
5.2.2	Dartboard Network Cuts for Evacuation Route Planning Algorithm	62
5.3	Experimental Analysis	67
5.3.1	Experiment Design	68
5.3.2	Experimental Results and Analysis	69
5.4	Summary	71
	References	71
6	Storage Schemes for Spatio-Temporal Network Datasets	73
6.1	Introduction	73
6.1.1	Application Domains	73
6.1.2	Basic Concepts	74
6.1.3	Problem Statement	77

6.1.4	Literature Review	77
6.1.5	Outline of the Chapter	78
6.2	Lagrangian-Connectivity Partitioning Approaches for SSTN	79
6.2.1	LCP-G1S for LGetOneSuccessor()	79
6.2.2	LCP-GVS for LGetAllSuccessors()	81
6.2.3	Algorithm for LCP-GVS	87
6.3	Cost Models	90
6.4	Experimental Analysis	91
6.4.1	Experimental Design	91
6.4.2	Experimental Results and Analysis	93
6.5	Summary	96
	References	97
7	Summary	99
7.1	Capacity Constrained Network Voronoi Diagram	100
7.2	Distance-Constrained k Spatial Sub-networks	100
7.3	Evacuation Route Planning	100
7.4	Storage Schemes for Spatio-Temporal Network Datasets	101