Lecture Notes in Computer Science

Edited by G. Goos and J. Hartmanis

471

Beng Chin Ooi

Fachbereichsbibliothek Informatik
TU Darmstadt
59372637

Efficient Query Processing in Geographic Information Systems

Technische Hochschule Darmstadt
FACHBEREICH INFORMATIK
BIBLIOTHEK
Inventar-Nr.: 8500
Sachgebiete: H.2.4
Standort: 1

Springer-Verlag
Berlin Heidelberg New York London
Paris Tokyo Hong Kong Barcelona

69,-

Table of Contents

1 Introduction 1

1.1 Geographic Information Systems 1

1.2 Motivation 2

1.3 Indexing Structures 4

1.4 Query Optimization 5

1.5 Thesis Synopsis 6

2 Related Work and GEOQL 9

2.1 Geographic Information Systems 9

2.1.1 The Nature of Data 9

2.1.2 On the Design of a GIS 10

2.1.3 Existing Systems 11

2.2 Query Languages 14

2.2.1 QUEL Extensions 14

2.2.2 QBE Extensions 16

2.2.3 SQL Extensions 17

2.2.4 GEOQL — An Augmented SQL 18

2.3 The Need for Efficient Spatial Indexing Mechanisms 22

2.4 Point Structures 23

2.4.1 Introduction 23

2.4.2 Quad-Trees 23

2.4.2.1 Point Quad-Trees 23

2.4.2.2 Region Quad-Trees 25

2.4.3 The kd-Trees 28

2.5 Access Structures for Extended Spatial Objects 32

2.5.1 Bounding Rectangles 32

2.5.2 Search Types 33

2.5.3 Spatial Indexing Methods 33

2.5.4 Corner Stitching 34

2.5.5 Cell Methods based on Dynamic Hashing 35

2.5.5.1 The Grid File 35
2.5.5.2 The EXCELL Method 41
2.5.5.3 PLOP-Hashing 42
2.5.6 Quad-Tree Based Structures 42
2.5.7 Locational Keys 44
2.5.8 Matsuyama's kd-Tree 47
2.5.9 The 4-D-Tree 49
2.5.10 The R-Tree 49
2.5.11 The R^+-Tree 52
2.5.12 The Cell Tree 57
2.5.13 Summary 57
2.6 Optimizations 60
2.6.1 Introduction to Query Optimization 60
2.6.2 Definitions 62
2.6.3 Optimization Strategies 63
2.6.3.1 Query Transformation 64
2.6.3.2 Query Evaluations 71
2.6.3.3 Optimization in System R 73
2.6.3.4 Optimization in INGRES — Query Decomposition 74
2.6.4 Global Optimization Strategies 74
2.6.5 Extensible DBMS and Resident Memory Systems — Current Trends 75
2.6.6 Optimization in GIS 77
2.6.7 Summary 78
2.7 Conclusion 78

3 The Spatial kd-Tree 80
3.1 A New-Structure — The Spatial kd-Tree 80
3.2 Searching 86
3.3 Insertion 88
3.4 Deletion 90
3.5 Directory Paging 96
3.5.1 Page Structure 97
3.5.2 Paging Strategy 98

3.6 Static Tree Construction 100
3.7 Data Storage Structure 105
3.8 Spatial Primitive Operations 107
3.9 Set Operations 111
3.10 Summary 115

4 Performance Analysis and Case Studies 116
4.1 Empirical Analysis 116
4.1.1 Quantities of Interest 116
4.1.2 Parameters 117
4.1.3 Simulation Methodology 117
4.2 A Variant of the skd-Tree 118
4.3 Evaluation with Non-uniform Data 121
4.3.1 The Skew Factor 122
4.3.2 Data Sets 122
4.3.3 Empirical Results 123
4.4 Further Comparisons 129
4.4.1 Object Mapping Structures 130
4.4.2 Object Duplication Structures 132
4.5 Summary 135

5 Query Optimization 136
5.1 Motivation 136
5.2 Extending Techniques 137
5.2.1 Overview of An Extended System Architecture 137
5.2.2 Extended Optimization 138
5.3 Classification of GEOQL Queries 139
5.4 An Introduction to the Extended Optimization 141
5.4.1 Grouping of Predicates 142
5.4.2 Answer Construction 142
5.4.2.1 The Not Operator 143
5.4.2.2 The And Operator 143
5.4.2.3 The Or Operator 143
5.5 Optimization Strategy 144
5.5.1 Parsing 144

5.5.2 Logical Transformation 145
5.5.3 Decomposition 148
5.5.4 Subquery Sequencing 153
5.5.4.1 Results Formation 153
5.5.4.2 SQP Formulation 159
5.5.4.3 Cost Estimation 163
5.6 Conclusions 164
6 Implementation and Experiments 166
6.1 Justification 166
6.2 RTB and ISQL 166
6.3 The GEOQL Database 168
6.4 The Experimental System 169
6.5 Discussion 176
6.5.1 Extensibility 176
6.5.2 Limitations 177
6.6 Conclusion 177

7 Conclusions 178

References 181
Appendix A 196
Appendix B 202
Appendix C 207