
Contents

1	Introduction	1
1.1	Overview	2
1.1.1	Conventional Data Warehouses	2
1.1.2	Spatial Databases and Spatial Data Warehouses	4
1.1.3	Temporal Databases and Temporal Data Warehouses	5
1.1.4	Conceptual Modeling for Databases and Data Warehouses	6
1.1.5	A Method for Data Warehouse Design	7
1.2	Motivation for the Book	8
1.3	Objective of the Book and its Contributions to Research	11
1.3.1	Conventional Data Warehouses	12
1.3.2	Spatial Data Warehouses	13
1.3.3	Temporal Data Warehouses	13
1.4	Organization of the Book	14
	Review Questions	16
2	Introduction to Databases and Data Warehouses	17
2.1	Database Concepts	18
2.2	The Entity-Relationship Model	19
2.3	Logical Database Design	23
2.3.1	The Relational Model	23
2.3.2	The Object-Relational Model	32
2.4	Physical Database Design	37
2.5	Data Warehouses	40
2.6	The Multidimensional Model	43
2.6.1	Hierarchies	44
2.6.2	Measure Aggregation	46
2.6.3	OLAP Operations	47
2.7	Logical Data Warehouse Design	49
2.8	Physical Data Warehouse Design	51
2.9	Data Warehouse Architecture	55

- 2.9.1 Back-End Tier 56
- 2.9.2 Data Warehouse Tier 57
- 2.9.3 OLAP Tier 58
- 2.9.4 Front-End Tier 58
- 2.9.5 Variations of the Architecture 59
- 2.10 Analysis Services 2005 59
 - 2.10.1 Defining an Analysis Services Database 60
 - 2.10.2 Data Sources 61
 - 2.10.3 Data Source Views 61
 - 2.10.4 Dimensions 62
 - 2.10.5 Cubes 64
- 2.11 Oracle 10g with the OLAP Option 66
 - 2.11.1 Multidimensional Model 67
 - 2.11.2 Multidimensional Database Design 68
 - 2.11.3 Data Source Management 69
 - 2.11.4 Dimensions 70
 - 2.11.5 Cubes 71
- 2.12 Conclusion 73
- Review Questions 73

- 3 Conventional Data Warehouses 77**
 - 3.1 MultiDim: A Conceptual Multidimensional Model 78
 - 3.2 Data Warehouse Hierarchies 81
 - 3.2.1 Simple Hierarchies 83
 - 3.2.2 Nonstrict Hierarchies 89
 - 3.2.3 Alternative Hierarchies 95
 - 3.2.4 Parallel Hierarchies 95
 - 3.3 Advanced Modeling Aspects 99
 - 3.3.1 Modeling of Complex Hierarchies 99
 - 3.3.2 Role-Playing Dimensions 101
 - 3.3.3 Fact Dimensions 103
 - 3.3.4 Multivalued Dimensions 103
 - 3.4 Metamodel of the MultiDim Model 108
 - 3.5 Mapping to the Relational and Object-Relational Models 109
 - 3.5.1 Rationale 109
 - 3.5.2 Mapping Rules 110
 - 3.6 Logical Representation of Hierarchies 114
 - 3.6.1 Simple Hierarchies 114
 - 3.6.2 Nonstrict Hierarchies 122
 - 3.6.3 Alternative Hierarchies 125
 - 3.6.4 Parallel Hierarchies 125
 - 3.7 Implementing Hierarchies 126
 - 3.7.1 Hierarchies in Analysis Services 2005 126
 - 3.7.2 Hierarchies in Oracle OLAP 10g 128
 - 3.8 Related Work 130

3.9	Summary	132
	Review Questions	134
4	Spatial Data Warehouses	137
4.1	Spatial Databases: General Concepts	138
4.1.1	Spatial Objects	138
4.1.2	Spatial Data Types	138
4.1.3	Reference Systems	140
4.1.4	Topological Relationships	140
4.1.5	Conceptual Models for Spatial Data	142
4.1.6	Implementation Models for Spatial Data	142
4.1.7	Models for Storing Collections of Spatial Objects	143
4.1.8	Architecture of Spatial Systems	144
4.2	Spatial Extension of the MultiDim Model	145
4.3	Spatial Levels	147
4.4	Spatial Hierarchies	147
4.4.1	Hierarchy Classification	147
4.4.2	Topological Relationships Between Spatial Levels	153
4.5	Spatial Fact Relationships	156
4.6	Spatiality and Measures	157
4.6.1	Spatial Measures	157
4.6.2	Conventional Measures Resulting from Spatial Operations	160
4.7	Metamodel of the Spatially Extended MultiDim Model	161
4.8	Rationale of the Logical-Level Representation	163
4.8.1	Using the Object-Relational Model	163
4.8.2	Using Spatial Extensions of DBMSs	164
4.8.3	Preserving Semantics	165
4.9	Object-Relational Representation of Spatial Data Warehouses	166
4.9.1	Spatial Levels	166
4.9.2	Spatial Attributes	168
4.9.3	Spatial Hierarchies	169
4.9.4	Spatial Fact Relationships	174
4.9.5	Measures	176
4.10	Summary of the Mapping Rules	178
4.11	Related Work	179
4.12	Summary	182
	Review Questions	183
5	Temporal Data Warehouses	185
5.1	Slowly Changing Dimensions	186
5.2	Temporal Databases: General Concepts	189
5.2.1	Temporality Types	189
5.2.2	Temporal Data Types	190
5.2.3	Synchronization Relationships	191

- 5.2.4 Conceptual and Logical Models for Temporal Databases 193
- 5.3 Temporal Extension of the MultiDim Model 194
 - 5.3.1 Temporality Types 194
 - 5.3.2 Overview of the Model 196
- 5.4 Temporal Support for Levels 199
- 5.5 Temporal Hierarchies 200
 - 5.5.1 Nontemporal Relationships Between Temporal Levels .. 200
 - 5.5.2 Temporal Relationships Between Nontemporal Levels .. 202
 - 5.5.3 Temporal Relationships Between Temporal Levels 202
 - 5.5.4 Instant and Lifespan Cardinalities 203
- 5.6 Temporal Fact Relationships 205
- 5.7 Temporal Measures 206
 - 5.7.1 Temporal Support for Measures 206
 - 5.7.2 Measure Aggregation for Temporal Relationships 211
- 5.8 Managing Different Temporal Granularities 211
 - 5.8.1 Conversion Between Granularities 212
 - 5.8.2 Different Granularities in Measures and Dimensions... 212
 - 5.8.3 Different Granularities in the Source Systems and in
the Data Warehouse 214
- 5.9 Metamodel of the Temporally Extended MultiDim Model.... 215
- 5.10 Rationale of the Logical-Level Representation 217
- 5.11 Logical Representation of Temporal Data Warehouses 218
 - 5.11.1 Temporality Types 218
 - 5.11.2 Levels with Temporal Support 220
 - 5.11.3 Parent-Child Relationships 224
 - 5.11.4 Fact Relationships and Temporal Measures 230
- 5.12 Summary of the Mapping Rules 232
- 5.13 Implementation Considerations 233
 - 5.13.1 Integrity Constraints 234
 - 5.13.2 Measure Aggregation 238
- 5.14 Related Work 241
 - 5.14.1 Types of Temporal Support 242
 - 5.14.2 Conceptual Models for Temporal Data Warehouses 242
 - 5.14.3 Logical Representation 244
 - 5.14.4 Temporal Granularity 246
- 5.15 Summary 246
- Review Questions 248

- 6 Designing Conventional Data Warehouses 251**
 - 6.1 Current Approaches to Data Warehouse Design 252
 - 6.1.1 Data Mart and Data Warehouse Design 252
 - 6.1.2 Design Phases 254
 - 6.1.3 Requirements Specification for Data Warehouse Design. 254
 - 6.2 A Method for Data Warehouse Design 256
 - 6.3 A University Case Study 257

6.4	Requirements Specification	259
6.4.1	Analysis-Driven Approach	259
6.4.2	Source-Driven Approach	267
6.4.3	Analysis/Source-Driven Approach	271
6.5	Conceptual Design	271
6.5.1	Analysis-Driven Approach	272
6.5.2	Source-Driven Approach	281
6.5.3	Analysis/Source-Driven Approach	284
6.6	Characterization of the Various Approaches	286
6.6.1	Analysis-Driven Approach	286
6.6.2	Source-Driven Approach	288
6.6.3	Analysis/Source-Driven Approach	289
6.7	Logical Design	289
6.7.1	Logical Representation of Data Warehouse Schemas	289
6.7.2	Defining ETL Processes	293
6.8	Physical Design	294
6.8.1	Data Warehouse Schema Implementation	294
6.8.2	Implementation of ETL Processes	300
6.9	Method Summary	301
6.9.1	Analysis-Driven Approach	302
6.9.2	Source-Driven Approach	302
6.9.3	Analysis/Source-Driven Approach	303
6.10	Related Work	304
6.10.1	Overall Methods	306
6.10.2	Requirements Specification	307
6.11	Summary	311
	Review Questions	312
7	Designing Spatial and Temporal Data Warehouses	315
7.1	Current Approaches to the Design of Spatial and Temporal Databases	316
7.2	A Risk Management Case Study	316
7.3	A Method for Spatial-Data-Warehouse Design	318
7.3.1	Requirements Specification and Conceptual Design	318
7.3.2	Logical and Physical Design	329
7.4	Revisiting the University Case Study	332
7.5	A Method for Temporal-Data-Warehouse Design	333
7.5.1	Requirements Specification and Conceptual Design	334
7.5.2	Logical and Physical Design	341
7.6	Method Summary	345
7.6.1	Analysis-Driven Approach	345
7.6.2	Source-Driven Approach	346
7.6.3	Analysis/Source-Driven Approach	347
7.7	Related Work	349
7.8	Summary	350

- Review Questions 351
- 8 Conclusions and Future Work** 353
 - 8.1 Conclusions 353
 - 8.2 Future Work 356
 - 8.2.1 Conventional Data Warehouses 356
 - 8.2.2 Spatial Data Warehouses 357
 - 8.2.3 Temporal Data Warehouses 359
 - 8.2.4 Spatiotemporal Data Warehouses 360
 - 8.2.5 Design Methods 361
 - Review Questions 362
- A Formalization of the MultiDim Model** 363
 - A.1 Notation 363
 - A.2 Predefined Data Types 363
 - A.3 Metavariables 364
 - A.4 Abstract Syntax 365
 - A.5 Examples Using the Abstract Syntax 367
 - A.5.1 Conventional Data Warehouse 367
 - A.5.2 Spatial Data Warehouse 369
 - A.5.3 Temporal Data Warehouse 372
 - A.6 Semantics 374
 - A.6.1 Semantics of the Predefined Data Types 375
 - A.6.2 The Space Model 375
 - A.6.3 The Time Model 379
 - A.6.4 Semantic Domains 380
 - A.6.5 Auxiliary Functions 380
 - A.6.6 Semantic Functions 383
- B Graphical Notation** 391
 - B.1 Entity-Relationship Model 391
 - B.2 Relational and Object-Relational Models 393
 - B.3 Conventional Data Warehouses 394
 - B.4 Spatial Data Warehouses 396
 - B.5 Temporal Data Warehouses 397
- References** 399
- Glossary** 419
- Index** 433