

Contents

Preface	xi
1 Introduction	1
1.1 Peer-to-Peer Systems	1
1.1.1 Concepts	2
1.1.2 Architectures	3
1.2 Peer-to-Peer Data Management	11
1.2.1 Architectures	13
1.2.2 Design Dimensions	15
1.2.3 Overview of the Lecture	17
2 Structured Peer-to-Peer Databases	19
2.1 Range Partitioning	19
2.1.1 Static Range Partitioning	21
2.1.2 Dynamic Range Partitioning	22
2.1.3 Load Balancing	23
2.1.4 Overlay Network Construction	28
2.1.5 Efficient Range Query Processing	33
2.2 Complex Query Processing	35
2.2.1 Managing Multiple Attributes	36
2.2.2 Join Queries	38
2.2.3 Aggregation Queries	40
2.3 Web Data Management	50
2.3.1 Triple Indexing	51
2.3.2 Schema-based Indexing	53
3 Peer-to-peer Data Integration	55
3.1 Peer-to-peer Data Integration	55
3.2 Query Rewriting	57
3.2.1 Attribute Mapping	57
3.2.2 Mapping Tables	58
3.3 View-based Data Integration	61

3.3.1	Global-as-View Mappings	61
3.3.2	Global-and-Local as View Mappings	63
3.4	Incomplete and Incorrect Mappings	67
3.4.1	Handling Incomplete Mappings	68
3.4.2	Discovering Mappings	68
3.4.3	Correcting Mappings	70
4	Peer-to-peer Retrieval	75
4.1	Full Text Search	75
4.1.1	Bloom Filters	80
4.1.2	Shortening of Posting Lists	81
4.1.3	Caching	83
4.1.4	Fagin's Algorithm	84
4.1.5	Peer Summaries	86
4.2	Structured Document Search	87
4.2.1	Path Indexing	88
4.2.2	Filtering	92
5	Semantic Overlay Networks	95
5.1	Concept Spaces	95
5.2	Semantic Overlay Network Construction	99
5.3	Protocol-driven SON	100
5.4	Mapping-based SON	103
5.4.1	Mapping Categories	103
5.4.2	Mapping Concept Vectors	104
6	Conclusion	109
A	Notational Conventions	113
	Bibliography	117
	Author's Biography	133
	Index	135