Contents

Foreword **xix**

1 Introduction **5**
1.1 Motivations 5
1.2 Contributions 6
1.3 Structure of the Document 7

2 Preliminaries **9**
2.1 Aggregation Functions 9
2.2 Time Series 13
2.2.1 Time Series Distances 13
2.2.2 Time Series Forecasting 15
2.3 Re-identification Methods 17
2.3.1 Distance Based Record Linkage 18
2.3.2 Probabilistic Record Linkage 19
2.4 Microdata Protection Methods 20
2.4.1 Rank Swapping 23
2.4.2 Microaggregation 24
2.5 Information Loss and Disclosure Risk 29
2.5.1 Information Loss Measures 31
2.5.2 Disclosure Risk Measures. 33
2.5.3 Score Computation 33
2.6 Data Sets Description 33
2.6.1 CASC Data Sets 34
2.6.2 UCI Data Sets 35

3 Microaggregation Analysis **41**
3.1 Attribute Selection in Multivariate Microaggregation 41
3.1.1 Specific Measures for Microaggregation 42
3.1.2 Strategies to Group Attributes in Microaggregation . . . 43
3.1.3 Motivating Examples 44
3.1.4 Experiments with Real Data Sets 46
3.1.5 Attribute Selection Consequences 51
3.2 Modeling Projections in Microaggregation 52

Copia gratuita. Personal free copy http://libros.csic.es

3.2.1 Algorithm Description . . . 52
3.2.2 Experiments . . . 54
3.3 Improving Microaggregation for Complex Record Anonymization 56
3.3.1 One Dimension Microaggregation . . . 56
3.3.2 Experimental Results . . . 60

4 Specific Disclosure Risk Measures 69
4.1 Rank Swapping Record Linkage . . . 69
4.1.1 Algorithm Description . . . 70
4.1.2 New Rank Swapping Methods . . . 72
4.1.3 Experimental RS-RL Results . . . 74
4.1.4 Specific Record Linkage Methods for the New Variations of Rank Swapping . . . 80
4.2 Alignment Record Linkage . . . 82
4.2.1 Algorithm Description . . . 83
4.2.2 Experimental A-RL Results . . . 84
4.3 Projected Record Linkage . . . 87
4.3.1 Traditional Disclosure Risk Evaluation for Multivariate Microaggregation . . . 87
4.3.2 The Projected Record Linkage Technique . . . 89
4.3.3 Consequences of Pro-RL in Multivariate Microaggregation 90

5 Record Linkage using Fuzzy Integrals 97
5.1 An Alternative Disclosure Risk Scenario . . . 97
5.1.1 Aggregation Functions for Building Representatives . . . 100
5.1.2 Example . . . 101
5.2 Experiments . . . 103
5.2.1 Preprocessing . . . 103
5.2.2 Tests . . . 104
5.2.3 Results . . . 104

6 Time Series Protection 111
6.1 Time Series Protection . . . 111
6.1.1 Time Series Microaggregation . . . 112
6.2 Time Series Information Loss Measures . . . 113
6.2.1 Information Loss Computation . . . 113
6.3 Time Series Disclosure Risk Measures . . . 115
6.3.1 Time Series Normalization . . . 115
6.3.2 Time Series Re-identification . . . 118
6.3.3 Time Series Interval Disclosure . . . 118
6.3.4 The Computation of the DR Measures . . . 119
6.4 Final Trade-off Evaluation . . . 119
6.5 Experiments . . . 120
6.5.1 Data Protection . . . 120
6.5.2 Results . . . 121

© CSIC © del autor o autores / Todos los derechos reservados

7 Conclusions and Future Directions **125**
7.1 Summary of Contributions 125
7.2 Conclusions 126
7.3 Future Directions 127

Copia gratuita. Personal free copy http://libros.csic.es