Contents

1 Swarm Intelligence in Data Mining
Crina Grosan, Ajith Abraham and Monica Chis 1
1.1 Biological Collective Behavior 1
1.2 Swarms and Artificial Life 4
1.2.1 Particle Swarm Optimization (PSO) 4
1.2.2 Ant Colonies Optimization 9
1.3 Data mining 10
1.3.1 Steps of Knowledge Discovery 10
1.4 Swarm Intelligence and Knowledge Discovery 11
1.5 Ant Colony Optimization and Data mining 15
1.6 Conclusions 16
References 16

2 Ants Constructing Rule-Based Classifiers
David Martens, Manu De Backer, Raf Haesen, Bart Baesens, Tom Holvoet 21
2.1 Introduction 21
2.2 Ant Systems and Data Mining 22
2.2.1 Ant Systems 22
2.2.2 Data Mining 24
2.2.3 Data Mining with Ant Systems 25
2.3 AntMiner+ 27
2.3.1 The Construction Graph 28
2.3.2 Edge Probabilities 29
2.3.3 Heuristic Value 29
2.3.4 Pheromone Updating 30
2.3.5 Early Stopping 32
2.4 Distributed Data Mining With AntMiner+:
a Credit Scoring Case 33
2.5 Experiments and Results 34
2.5.1 Experimental Set-Up 34
2.5.2 Datasets 35

Credit Scoring 35
Toy Problems 36
2.5.3 Software Implementation 38
2.5.4 Discussion 39
2.6 Conclusion and Future Research 40
References 41

3 Performing Feature Selection with ACO
Richard Jensen 45
3.1 Introduction 45
3.2 Rough Feature Selection 47
3.2.1 Theoretical Background 47
3.2.2 Reduction Method 48
3.3 Fuzzy-Rough Feature Selection 50
3.3.1 Fuzzy Equivalence Classes 50
3.3.2 Fuzzy Lower and Upper Approximations 51
3.3.3 Fuzzy-Rough Reduction Method 52
3.3.4 A Worked Example 53
3.4 Ant-based Feature Selection 56
3.4.1 ACO Framework 57
3.4.2 Feature Selection 58
Selection Process 59
Complexity Analysis 59
Pheromone Update 60
3.5 Crisp Ant-based Feature Selection Evaluation 60
3.5.1 Experimental Setup 61
3.5.2 Experimental Results 62
3.6 Fuzzy Ant-based Feature Selection Evaluation 63
3.6.1 Web Classification 63
System Overview 63
Experimentation and Results 65
3.6.2 Systems Monitoring 66
Comparison of Fuzzy-Rough Methods 68
Comparison with Entropy-based Feature Selection 69
Comparison with the use of PCA 70
Comparison with the use of a Support Vector Classifier 70
3.7 Conclusion 71
References 72

4 Simultaneous Ant Colony Optimization Algorithms for Learning Linguistic Fuzzy Rules
Michelle Galea, Qiang Shen 75
4.1 Introduction 75
4.2 Background 76
4.2.1 Fuzzy Rules and Rule-Based Systems 76

Fuzzy Sets and Operators 77
Linguistic Variables and Fuzzy Rules 78
Classification using Fuzzy Rules 79
A Rule-Matching Example 80
4.2.2 Ant Colony Optimization and Rule Induction 81
4.3 Simultaneous Fuzzy Rule Learning 84
4.3.1 Why Simultaneous Rule Learning 84
4.3.2 *FRANTIC-SRL* 86
Rule Construction 86
Heuristic 87
Pheromone Updating 88
Transition Rule 88
Rule Evaluation 89
4.4 Experiments and Analyses 90
4.4.1 Experiment Setup 90
The Datasets 90
Other Induction Algorithms 91
FRANTIC-SRL Parameters 92
4.4.2 Saturday Morning Problem Results 93
4.4.3 Water Treatment Plant Results 93
4.5 Conclusions and Future Work 95
References 97

5 Ant Colony Clustering and Feature Extraction for Anomaly Intrusion Detection
Chi-Ho Tsang, Sam Kwong 101
5.1 Introduction 101
5.2 Related Works 103
5.3 Ant Colony Clustering Model 104
5.3.1 Basics and Problems of Ant-based Clustering Approach 104
5.3.2 Measure of Local Regional Entropy 106
5.3.3 Pheromone Infrastructure 107
5.3.4 Modified Short-term Memory and α-adaptation 109
5.3.5 Selection Scheme, Parameter Settings and Cluster Retrieval 110
5.4 Experiments and Results 111
5.4.1 Dataset Description and Preprocessing 111
5.4.2 Metrics of Cluster Validity and Classification Performance 112
5.4.3 Cluster Analysis on Benchmark Datasets 114
5.4.4 ACCM with Feature Extraction for Intrusion Detection 116
5.5 Conclusions 120
5.6 Future Works 121
References 121

6 Particle Swarm Optimization for Pattern Recognition and Image Processing
Mahamed G.H. Omran, Andries P. Engelbrecht, Ayed Salman 125
6.1 Introduction 125
6.2 Background 126
6.2.1 The clustering problem 126
The *K*-means Algorithm 128
The Fuzzy C-means Algorithm 129
Swarm Intelligence Approaches 130
6.2.2 Color Image Quantization 130
6.2.3 Spectral Unmixing 132
Linear Pixel Unmixing (or Linear Mixture Modeling) 132
Selection of the End-Members 133
6.3 Particle Swarm Optimization 134
6.4 A PSO-based Clustering Algorithm with Application to Unsupervised Image Classification 135
6.4.1 Experimental Results 137
6.5 A PSO-based Color Image Quantization (PSO-CIQ) Algorithm 138
6.5.1 Experimental Results 140
6.6 The PSO-based End-Member Selection (PSO-EMS) Algorithm 141
6.6.1 The Generation of Abundance Images 143
6.6.2 Experimental results 143
6.7 Conclusion 148
References 148

7 Data and text mining with hierarchical clustering ants
Hanene Azzag, Christiane Guinot, Gilles Venturini 153
7.1 Introduction 153
7.2 Biological and computer models 154
7.2.1 Ants based algorithms for clustering 154
7.2.2 Self-assembly in real ants 155
7.2.3 A computer model of ants self-assembly for hierarchical clustering . 155
7.2.4 Self-assembly and robotics 157
7.3 Two stochastic and deterministic algorithms 158
7.3.1 Common principles 158
7.3.2 Stochastic algorithm: AntTree$_{STOCH}$ 158
7.3.3 Deterministic algorithm with no thresholds and no parameters : AntTree$_{NO-THRESHOLDS}$ 161
7.3.4 Properties 162
7.4 Experimental results with numeric, symbolic and textual databases 164
7.4.1 Testing methodology 164
7.4.2 Parameters study 166
7.4.3 Tested algorithms 166
7.4.4 Results with numeric databases 168
7.4.5 Results with symbolic databases 168

7.4.6 Processing times . . . 169
7.4.7 Comparison with biomimetic methods . . . 170
7.4.8 Comparative study on textual databases . . . 172
7.5 Real world applications . . . 175
7.5.1 Human skin analysis . . . 175
7.5.2 Web usage mining . . . 177
7.5.3 Generation and interactive exploration of a portal site . . . 179
7.6 Incremental clustering of a large data set . . . 182
7.6.1 Principles of AntTree$_{INC}$ . . . 182
7.6.2 Results with incremental and large data sets . . . 184
7.7 Conclusions . . . 186
References . . . 186

8 Swarm Clustering Based on Flowers Pollination by Artificial Bees
Majid Kazemian, Yoosef Ramezani, Caro Lucas, Behzad Moshiri . . . 191
8.1 Introduction . . . 191
8.2 Clustering . . . 192
8.2.1 What is clustering? . . . 192
8.2.2 Why swarm intelligence? . . . 193
8.2.3 Swarm clustering . . . 193
8.2.4 Some artificial models . . . 194
8.3 FPAB . . . 195
8.3.1 FPAB underlying algorithms . . . 196
Picking up pollen . . . 197
Pollinating . . . 198
Natural selection . . . 198
Merge algorithm . . . 199
8.4 Experimental results . . . 199
8.5 Conclusion and future works . . . 200
References . . . 201

9 Computer study of the evolution of 'news foragers' on the Internet
Zsolt Palotai, Sándor Mandusitz, András Lőrincz . . . 203
9.1 Introduction . . . 203
9.2 Related work . . . 204
9.3 Forager architecture . . . 205
9.3.1 Algorithms . . . 206
9.3.2 Reinforcing agent . . . 208
9.3.3 Foragers . . . 209
9.4 Experimental results . . . 210
9.4.1 Environment . . . 210
9.4.2 Time lag and multiplication . . . 210
9.4.3 Compartmentalization . . . 211
9.5 Discussion . . . 214
9.6 Conclusions . . . 217

References 217

10 Data Swarm Clustering
Christian Veenhuis, Mario Köppen 221
10.1 Introduction 221
10.2 Data Clustering 223
10.3 Flock Algorithms 223
10.4 Particle Swarm Optimization 225
10.5 Data Swarm Clustering 226
10.5.1 Initialization 227
10.5.2 Iteration 228
10.5.3 Cluster Retrieval 234
10.6 Experimental Setup 234
10.6.1 Synthetical Datasets 234
10.6.2 Real Life Datasets 236
10.6.3 Parameters 236
10.7 Results 237
10.8 Conclusion 240
References 241

11 Clustering Ensemble Using ANT and ART
Yan Yang, Mohamed Kamel, Fan Jin 243
11.1 Introduction 243
11.2 Ant Colony Clustering Algorithm with Validity Index (ACC-VI) 245
11.2.1 Ant Colony Clustering Algorithm 245
11.2.2 Clustering Validity Index 247
11.2.3 ACC-VI Algorithm 248
11.3 ART Algorithm 249
11.4 Clustering Ensemble Model 253
11.4.1 Consensus Functions 253
11.4.2 ART Ensemble Aggregation Model 253
11.5 Experimental Analysis 255
11.5.1 Artificial Data Set (2D3C) 256
11.5.2 Real Data Set (Iris) 256
11.5.3 Reuter-21578 Document Collection 258
11.6 Conclusions 262
Acknowledgements 262
References 262

Index 265