

Contents

- 1 Introduction 1**
 - 1.1 Introduction 1
 - 1.2 Data Types with Examples 3
 - 1.3 Pattern Recognition: Preliminaries 4
 - 1.3.1 Data Acquisition 6
 - 1.3.2 Feature Selection 7
 - 1.3.3 Classification 7
 - 1.3.4 Clustering 8
 - 1.3.5 Distance Measures in Clustering 9
 - 1.3.6 Model Selection 10
 - 1.3.7 Model Order Selection 10
 - 1.4 Robustness to Outliers and Missing Values 11
 - 1.5 Fuzzy Set-Theoretic Approach: Relevance and Features 12
 - 1.6 Applications of Pattern Recognition and Learning 13
 - 1.7 Summary and Scope of the Book 14
- 2 Some Single- and Multiobjective Optimization Techniques 17**
 - 2.1 Introduction 17
 - 2.2 Single-Objective Optimization Techniques 18
 - 2.2.1 Overview of Genetic Algorithms 19
 - 2.2.2 Simulated Annealing 23
 - 2.3 Multiobjective Optimization 25
 - 2.3.1 Multiobjective Optimization Problems 26
 - 2.3.2 Various Methods to Solve MOOPs 28
 - 2.3.3 Recent Multiobjective Evolutionary Algorithms 29
 - 2.3.4 MOOPs and SA 33
 - 2.4 An Archive-Based Multiobjective Simulated Annealing Technique: AMOSA 36
 - 2.4.1 Introduction 36
 - 2.4.2 Archived Multiobjective Simulated Annealing (AMOSA) 37
 - 2.4.3 Archive Initialization 38

2.4.4	Clustering the Archive Solutions	38
2.4.5	Amount of Domination	40
2.4.6	The Main AMOSA Process	41
2.4.7	Complexity Analysis	45
2.5	Simulation Results	46
2.5.1	Comparison Measures	47
2.5.2	Comparison of Binary-Encoded AMOSA with NSGA-II and PAES	48
2.5.3	Comparison of Real-Coded AMOSA with the Algorithm of Smith et al. and Real-Coded NSGA-II	51
2.5.4	Discussion on Annealing Schedule	56
2.6	Discussion and Conclusions	57
3	Similarity Measures	59
3.1	Introduction	59
3.2	Definitions	60
3.2.1	Need for Measuring Similarity	60
3.3	Similarity/Dissimilarity for Binary Variables	61
3.4	Distance for Nominal/Categorical Variable	63
3.4.1	Method 1: Each Category Is Represented by a Single Binary Variable [278]	64
3.4.2	Method 2: Each Category Is Represented by Several Binary Variables [278]	64
3.5	Distance for Ordinal Variables	65
3.5.1	Normalized Rank Transformation	66
3.5.2	Spearman Distance	67
3.5.3	Footrule Distance	67
3.6	Distance for Quantitative Variables	68
3.6.1	Euclidean Distance	68
3.6.2	Minkowski Distance of Order λ	68
3.6.3	City Block Distance	69
3.6.4	Chebyshev Distance	69
3.6.5	Canberra Distance	70
3.6.6	Bray-Curtis Distance	70
3.6.7	Angular Separation	70
3.6.8	Correlation Coefficient	71
3.6.9	Mahalanobis Distance	72
3.7	Normalization Methods	72
3.8	Summary	73
4	Clustering Algorithms	75
4.1	Introduction	75
4.2	Preliminaries	76
4.2.1	Definition of Clustering	76
4.2.2	Some Clustering Techniques	76

4.3	Partitional Clustering Techniques	77
4.3.1	K -Means Clustering Technique	77
4.3.2	K -Medoid Clustering Technique	79
4.3.3	Fuzzy C -Means Clustering Algorithm	79
4.4	Distribution-Based Clustering Approach	80
4.5	Hierarchical Clustering Techniques	82
4.6	Density-Based Clustering Techniques	83
4.7	Grid-Based Clustering Techniques	85
4.8	Some Recent Clustering Techniques	85
4.9	Some Evolutionary Approaches to Clustering	87
4.9.1	Algorithms for a Fixed Value of the Number of Clusters	88
4.9.2	Algorithms with Variable Number of Clusters	89
4.10	MOO and Clustering	90
4.11	Summary	91
5	Point Symmetry-Based Distance Measures and Their Applications to Clustering	93
5.1	Introduction	93
5.2	Some Existing Symmetry-Based Distance Measures	94
5.3	A New Definition of the Point Symmetry Distance [27]	100
5.4	Some Properties of $d_{ps}(\bar{x}, \bar{c})$	102
5.5	Kd-Tree-Based Nearest Neighbor Computation	103
5.6	GAPS: The Genetic Clustering Scheme with New PS-Distance [27]	104
5.6.1	Chromosome Representation and Population Initialization	104
5.6.2	Fitness Computation	106
5.6.3	Selection	107
5.6.4	Crossover	107
5.6.5	Mutation	108
5.6.6	Termination	109
5.6.7	Complexity Analysis	109
5.7	On the Convergence Property of GAPS	110
5.7.1	Preliminaries	110
5.7.2	Convergence Proof	112
5.8	Experimental Results of GAPS	113
5.8.1	Data Sets Used	113
5.8.2	Implementation Results	114
5.8.3	Summary	121
6	A Validity Index Based on Symmetry: Application to Satellite Image Segmentation	125
6.1	Introduction	125
6.2	Some Existing Cluster Validity Indices	126
6.2.1	BIC Index	126
6.2.2	Calinski-Harabasz Index	127

6.2.3	Silhouette Index	127
6.2.4	DB Index	127
6.2.5	Dunn's Index	128
6.2.6	Xie-Beni Index	128
6.2.7	PS Index	129
6.2.8	I -Index	129
6.2.9	CS-Index	129
6.3	<i>Sym</i> -Index: The Symmetry-Based Cluster Validity Index	130
6.3.1	The Cluster Validity Measure	130
6.3.2	Mathematical Justification	134
6.3.3	Interaction Between the Different Components of <i>Sym</i> -Index	136
6.4	Experimental Results	140
6.4.1	Data Sets	140
6.4.2	Comparative Results	140
6.4.3	Analysis of Results	143
6.5	Incorporating d_{ps} in Some Existing Cluster Validity Indices	147
6.6	Point Symmetry-Based Cluster Validity Indices	148
6.6.1	Symmetry-Based Davies-Bouldin Index (<i>Sym-DB</i> Index)	148
6.6.2	Symmetry-Based Dunn's Index (<i>Sym-Dunn</i> Index)	149
6.6.3	Symmetry-Based Generalized Dunn's Index (<i>Sym-GDunn</i> Index)	149
6.6.4	New Symmetry Distance-Based PS-Index (<i>Sym-PS</i> Index)	150
6.6.5	Symmetry-Based Xie-Beni Index (<i>Sym-XB</i> Index)	151
6.6.6	Symmetry-Based FS-Index (<i>Sym-FS</i> Index)	151
6.6.7	Symmetry-Based K -Index (<i>Sym-K</i> Index)	151
6.6.8	Symmetry-Based SV-Index (<i>Sym-SV</i> Index)	152
6.7	Experimental Results	152
6.7.1	Discussion of Results	152
6.8	Application to Remote Sensing Imagery	155
6.8.1	Simulated Circle Image (SCI)	156
6.8.2	SPOT Image of Kolkata	156
6.9	Discussion and Conclusions	160
7	Symmetry-Based Automatic Clustering	165
7.1	Introduction	165
7.2	Some Existing Genetic Algorithm-Based Automatic Clustering Techniques	166
7.3	Description of VGAPS	168
7.3.1	Chromosome Representation and Population Initialization	168
7.3.2	Fitness Computation	168
7.3.3	Genetic Operations and Terminating Criterion	168

7.4	On the Convergence Property of VGAPS	171
7.4.1	To Check Whether the Mutation Matrix Is Positive	172
7.4.2	Conditions on Selection	173
7.5	Data Sets Used and Implementation Results	174
7.5.1	Data Sets Used	174
7.5.2	Results and Discussions	174
7.6	Extending VGAPS to Fuzzy Clustering	178
7.6.1	Fuzzy Symmetry-Based Cluster Validity Index	178
7.6.2	Fuzzy-VGAPS Clustering	179
7.7	Implementation Results and Comparative Study	182
7.7.1	Discussion of Results	184
7.7.2	Application to MR Brain Image Segmentation	189
7.7.3	Experimental Results	190
7.8	Summary	194
8	Some Line Symmetry Distance-Based Clustering Techniques	197
8.1	Introduction	197
8.2	The Line Symmetry-Based Distance	197
8.2.1	Definition	198
8.3	GALSD: The Genetic Clustering Scheme with Line Symmetry-Based Distance	199
8.3.1	String Representation and Population Initialization	199
8.3.2	Fitness Computation	199
8.3.3	Genetic Operators	200
8.4	Experimental Results	200
8.4.1	Data Sets Used	200
8.4.2	Discussion of Results	202
8.4.3	Computation Time	203
8.5	Application to Face Recognition	204
8.5.1	Human Face Detection Algorithm	205
8.5.2	Experimental Results	207
8.6	A Generalized Line Symmetry-Based Distance and Its Application to Data Clustering	209
8.7	Implementation Results	210
8.7.1	Data Sets Used	211
8.7.2	Discussion of Results	211
8.8	Discussion and Conclusions	215
9	Use of Multiobjective Optimization for Data Clustering	217
9.1	Introduction	217
9.2	MOPS: Multiobjective Clustering Using Point Symmetry Distance	219
9.2.1	Selection of a Solution from the Archive	220
9.2.2	Experimental Results	220
9.3	VAMOSA: Symmetry-Based Multiobjective Clustering Technique for Automatic Evolution of Clusters	221

9.3.1	Data Sets Used for Experiment	222
9.3.2	Experimental Results	223
9.4	A Generalized Automatic Clustering Algorithm in a Multiobjective Framework	227
9.4.1	<i>GenClustMOO</i> : Multiobjective Clustering Technique . . .	228
9.4.2	Subcluster Merging for Objective Function Calculation . .	232
9.5	Experimental Results	233
9.5.1	Data Sets Used	233
9.5.2	Discussion of Results	235
9.6	Discussion and Conclusions	242
References		245
Index		259

Unsupervised Classification

Similarity Measures, Classical and Metaheuristic
Approaches, and Applications

Bandyopadhyay, S.; Saha, S.

2013, XVIII, 262 p., Hardcover

ISBN: 978-3-642-32450-5