 doi:10.20944/preprints201901.0125.v1

Article

Method of Generating Contexts based on Self-adaptive Differential Particle Swarm using Local Topology for Multimodal Optimization in the case of Multigranulation. A Case Study

Dianne Arias [1] 0000-0002-4715-5419, Yaima Filiberto [1*] 0000-0003-2279-2953 and Rafael Bello [2,] 0000-0001-5567-2638

[1] University of Camaguey 1; dianne.arias@reduc.edu.cu, yaima.filiberto@reduc.edu.cu
[2] University of Las Villas 2; rbellop@uclv.edu.cu
* Correspondence: yaima.filiberto@reduc.edu.cu; Tel.: +5353559573

Version January 11, 2019 submitted to Preprints

Abstract: Multigranulation is a new approach to the Rough Set Theory, where several separability relationships are used to obtain different granulations of the universe. The Multigranulation starts from the existence of different contexts or subsets of features to characterize the objects of the universe. In this paper, a method for the generation of contexts from the construction of similarity relations is proposed.The proposed solution was evaluated in an international database using the KNN classifier. It was also applied in the solution of a real problem in Civil Engineering specifically in Traffic Engineering, the contexts generated from the proposal used to determine the features of higher incidence in the service level of the road. The results achieved both in the international database and in the proposed application demonstrate the applicability of the proposed method.

Keywords: multigranulation; separability relationships; service level of the road

0. Introduction

Rough set theory, originated by Pawlak [1][2][3], has become a well-established mechanism for uncertainty management in a wide variety of applications related to artificial intelligence [4][5][6][7]. One of the strengths of rough set theory is that all its parameters are obtained from the given data. This can be seen in the following paragraph from [8]:
"The numerical value of imprecision is not pre-assumed, as it is in probability theory or fuzzy sets but is calculated on the basis of approximations which are the fundamental concepts used to express imprecision of knowledge". In other words, instead of using, the rough set data analysis (RSDA) utilizes solely the granularity structure of the given data, expressed as classes of suitable equivalence relations.
In the past 10 years, several extensions of the rough set model have been proposed in terms of various requirements, such as the variable precision rough set (VPRS) model [9], the rough set model based on tolerance relation [10], the Bayesian rough set model [11], the fuzzy rough set model and the rough fuzzy set model [12].
In many circumstances, we often need to describe concurrently a target concept through multi binary relations according to a users requirements or targets of problem solving, for that another extension of the RST is to use more than one separability relationship to perform the granulation of the universe, which is known as multigranulation [13][14]. In this case, from the set of predictive features A, two or more subsets $A_1, ..., A_k$, $A_i\ Am \subseteq A$, are formed of features that allow defining the separability relation. These subsets of features are called contexts [15]. Based on this multigranulation approach, different techniques for the discovery of knowledge have been formulated.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 14 January 2019 doi:10.20944/preprints201901.0125.v1

1. Multigranulation in the Rough Set Theory

Qian et al. [13] proposed multigranulation rough set (MGRS) in complete information system to more widely apply rough set theory in practical applications, in which lower/upper approximations are approximated by granular structures induced by multi binary relations. The multigranulation rough set is different from Pawlaks rough set model because the latter is constructed on the basis of a family of indiscernibility relations instead of single indiscernibility relation. In optimistic multigranulation rough set approach, the word "optimistic" is used to express the idea that in multi independent granular structures, we need only at least one granular structure to satisfy with the inclusion condition between equivalence class and the approximated target. The upper approximation of optimistic multigranulation rough set is defined by the complement of the lower approximation[37]. From the point of view of the applications of the RST, the multigranulation in the RST is very desirable in many real applications, such as analysis of data from multiple sources, discovery of knowledge to from data with large dimensions and distributive information systems. Since Qian in 2006 proposed multigranulation in the RST, the theoretical framework has been widely enriched, and many extensions of these models have been proposed and studied[35][36]. In the mutigranulation rough set theory, each of various binary relation determines a corresponding information granulation, which largely impacts the commonality between each of the granulations and the fusion among all granulations.

In their papers, Qian et al. said that the MGRS are useful in the following cases:

1. We cannot perform the intersection operations between their quotient sets and the target concept cannot be approximated by using $U/(P \bigcup Q)$ which is called a single granulation in those papers.

2. In the process of some decision making, the decision or the view of each of decision makers may be independent for the same project (or a sample, object and element) in the universe. In this situation, the intersection operations between any two quotient sets will be redundant for decision making.

3. Extract decision rules from distributive information systems and groups of intelligent agents through using rough setapproaches.

Since then,many researchers have extended the classical MGRS by using various generalized binary relations.

In Fig.1 of the author Qian in [13], the bias region is the lower approximation of a set X obtained by a single granulation P $\bigcup$ Q, which is expressed by the equivalence classes in the quotient set U/(P $\bigcup$ Q), and the shaded region is the lower approximation of X induced by two granulations P+Q, which is characterized by the equivalence classes in the quotient set U/P and the quotient set U/Q together.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 14 January 2019 doi:10.20944/preprints201901.0125.v1

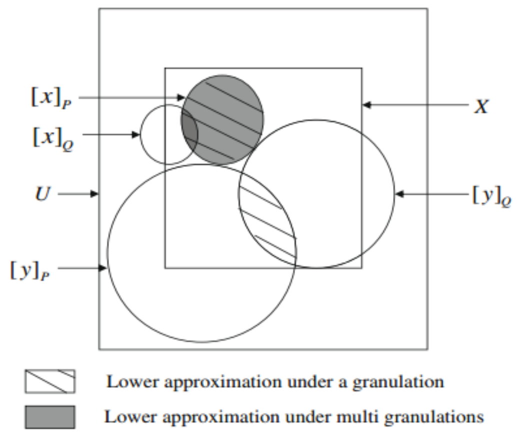

Figure 1. Difference between Pawlaks rough set model and MGRS.

From the point of view of the applications of the RST, the multigranulation in the RST is very desirable in many real applications, such as analysis of data from multiple sources, discovery of knowledge to from data with large dimensions and distributive information systems.

2. Self-Adaptive Differential Particle Swarm using a local topology for Multimodal Optimization

Particle Swarm Optimization is an effective and robust non-direct global-search method for solving challenging continuous optimization problems. The PSO meta-heuristic involves a set of particles known as swarm which explore the search space trying to locate promising regions [32]. Therefore, particles are interpreted as solutions for the optimization problem and they are represented as points in n-dimensional search space. In the case of standard PSO, each particle X_i moves through the space using its own velocity V_i, a local memory of the best position it has obtained P_i and knowledge of the best solution G found in its neighborhood. Equations 1 and 2 show how to update the particles position based on the mentioned components.

$$\begin{aligned} \vartheta_i(t+1) &= \alpha * \vartheta_i(t) + U(0, \varphi 1)(pBest(t) - x_i) \\ &\quad + U(0, \varphi 2(gBest(t) - x_i) \end{aligned} \tag{1}$$

$$x_i(t+1) = x_i(t)\vartheta_i(t+1) \tag{2}$$

In last decades Evolutionary and Swarm Intelligence algorithms have become an important improvement for both discrete and real-parameter optimization. Without niching [34] strategies they converge to a single optimum, even in multimodal search spaces where numerous global or local solutions exist. However, most real-life problems are characterized by multimodal functions. In literature several niching approaches have been proposed for computing multiple optima simultaneously, though most of them require some user-specified parameters that should be estimated in advanced (i.e. additional knowledge about problem domain is required)[34].
Then, multimodal optimization methods try to discover and maintain multiple subpopulations in a single run, where each niche corresponds to a specific peak of the fitness landscape (ideally one species per optimum). They have been developed to reduce the undesirable effects of genetic drift. In few words, niching strategies should be able to preserve the diversity in the artificial population, allowing individuals parallel convergence toward different solutions. As well, niching methods are useful to avoid stagnation or premature convergence states in global optimization problems where many sub-optimal solutions exist; offering an escaping alternative from local optima [34].
When multimodal problems are solved, the main advantage of the lbest model appears to lie in its slower convergence rate relative to the gbest model, allowing concurrently discovering several

 doi:10.20944/preprints201901.0125.v1

optima. Ironically, it is the slightly interaction among particles that is most responsible of the poor performance of the PSO based algorithms using a Ring Topology. To improve the search capability of such models a novel Differential Operator is introduced. This operator is straightforwardly inspired on the well-known differential strategy DE/current-to-rand/1 without crossover [33]. Therefore, as first step, we design a mutation operator as illustrate following equation 3.

$$\widetilde{x}_i(t+1) = pBest_i(t) + F * (pBest_{r1}(t) - pBest_{r2}(t)) \tag{3}$$

Where $pBest_i(t)$ denotes the personal best position of current individual, $pBest_{r1}(t)$ and $pBest_{r2}(t)$ are the global best record achieved by two randomly selected swarm particles.
Next, a selection operation takes place, where $\widetilde{x}_i$ is accepted as current particle position if it improves the search procedures, respect to the solution generated by the PSO rules; otherwise the mutant is rejected (See in equation 4)[34].

$$x_i^{t+1} = \begin{cases} x_i^{t+1},\ sif\left(\widetilde{x}_i^{t+1}\right) \le f\left(P_i^{t+1}\right) \\ x_i^{t+1},\ sif\left(\widetilde{x}_i^{t+1}\right) > f\left(P_i^{t+1}\right) \end{cases} \tag{4}$$

Following a similar reasoning of the conventional clearing, it's used a novel diversity procedure: Heuristic Clearing. It is able to preserve the swarm diversity in lbest PSO algorithms using a Ring Topology based topology, and it does not need to be specified any niche parameter. To do that, this operator only takes into account optimal particles (See equation 5)[34].

$$\mid f(P_i) - F^* \mid < \varepsilon \tag{5}$$

In the following section we propose a method to generate contexts using multigranulation based on the rough set theory and multimodal PSO.

3. Method of Generating Contexts based on Self-adaptive Differential Particle Swarm using Local Topology for Multimodal Optimization in the case of Multigranulation.

Be a decision system $SD = A \bigcup d$ where the domain of the characteristic in $A \bigcup d$ may be discrete or continuous values, from which calculate the features weights using the PSOMulti+RST+MG method, which is a modification of PSO + RST [17]. In this case PSO Multimodal is used in order to obtain multiple maximums global (gbest) from which the set of contexts is created and the number of characteristics per context, then weights are ordered by contexts and those with a weight greater than the mean value are selected of the weight for that context. Finally the same contexts are remove. The algorithm is described below.

Algorithm 1 Pseudocode for PSOMulti+RST+MG algorithm

1. Calculate the weights (w) of features using PSOMulti+RST method.
2. Generate set of contexts using GBest(Φn)
   $C = \Phi n$
3. For each context C_i:
   Order W_i
   Select $W_{ij} \in W_{imax} \Longleftrightarrow \{W_{ij} > mean(W_i)\}$
4. Select de different context
   $\forall C_i, C_j | C_i, C_j \exists C \wedge C_i \neq C_j$

Algorithm PSOMULTI+RST+MG
Step 1: Initialize a population of particles with random positions and velocities in a D-dimensional space.

 doi:10.20944/preprints201901.0125.v1

Step 2: For each particle, evaluate the quality measure of similarity using expression 6, in D Variables.

$$\max \rightarrow \left\{ \frac{\sum\limits_{\forall x \epsilon U} \varphi(x)}{|U|} \right\} \tag{6}$$

Step 3: Compare the quality measure of the current similarity of each particle with the quality measurement of the similarity of your previous best position $pBest$. If the current value is better than that of $pBest$, then assign to $pBest$ the current value, and $pBest = x_i$, that is, the current location becomes the best one so far.
Step 4: Identify the particle in the neighborhood with the highest value for the quality of similarity measure and assign its index to the variable $gBest$ and assign the best value of the quality measure of similarity to m.
Step 5: Adjust the speed and position of the particle according to equations 1, 2 and 3 (for each dimension).
Step 6: Verify if the stop criterion is met (maximum number of iterations or if it takes five iterations without improving the quality measure of the global similarity (m)), if not, go to Step2.

4. Experimental results

For this study we used data sets from the UCI repository [38] ($iris$, $schizo$, $hepatitis$, $biomed$, $glass$, $analcaa_{b}ankruptcy$, $diabetes$, $liver - disorders$,$ecoli$, $vehicle$, $lung_{c}ancer$, $segment$, $new - thyroid$, $breast - w$, $bupa$). It is used to calculate the weights for KNN [23] with k = 1 the proposed method. The training and test sets were obtained, taking 75 percent of the cases for the first and 25 percent for the second, in a totally randomized manner. Following this principle of random selection the process was repeated ten times and ten training sets and ten test sets were obtained for each data set, in order to apply cross validation [24] for a better validation of the results.
The parameters used in the experimentation, for the method $PSOMulti + RST + MG$ were: TB = 40, NI = 100, $ce1$ = $ce2$ = 2 and the values of $e1$ and $e2$ for the function of similarity between attributes and for the function of similarity for the decision attribute were between 0.70-0.83 and 1.0, F=0.1. The stop condition is: when 100 iterations are reached or when in five iterations the fitness value does not improve (measure quality similarity quality).
It is used as a KNN classifier with K = 1 to make a comparison of the results obtained after the creation of the proposed method's contexts $PSOMULTI + RST + MG$ with algorithms $AdaBoostM1$[30], $RandomSubSpace$ [31] and $Bagging$ [29], implemented in the WEKA[1] tool and using the KNN as a classifier, in all cases.
For the statistical analysis of the results, the hypothesis testing techniques were used [25]. For multiple comparisons, the Friedman and Iman-Davenport tests [26] are used to detect statistically significant differences between a groups of results. The Holm test [27] is also used in order to find significantly higher algorithms.
These tests are suggested in the studies presented in [24], which states that the use of these tests is highly recommended for the validation of results in the field of automated learning. In the statistical processing of all the experimental results, the KEEL was used [28].
Table 1 shows the description of the data sets used in the experimentation, as well as the contexts obtained by the proposed method (column 4) and the number of average features for each context (column 5). Table 2 shows the results of evaluating the contexts with the KNN method for K = 1 (Knn

[1] Herramienta de código abierto escrita en Java. Disponible bajo licencia pública GNU en http://www.cs.waikato.ac.nz/ml/weka/

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 14 January 2019 doi:10.20944/preprints201901.0125.v1

algorithm $PSOMULTI + RST + MG$), as well as the results of the $AdaBoostM1$, $RandomSubSpace$ and $Bagging$ algorithms, as you can observe the proposed method obtains better results than the rest.

Table 1. Datasets

Datasets	Instances	Feature	Contexts	Features average X Contexts
iris	15	4	3	2
schizo	11	14	31	3
hepatitis	16	19	13	4
biomed	20	8	24	5
glass	22	9	34	5
analcaa-bankruptcy	5	6	10	4
diabetes	77	8	32	5
liver-disorders	35	6	19	4
ecoli	34	7	23	4
vehicle	85	18	37	5
lung-cancer	4	56	4	2
segment	231	19	13	4
new-thyroid	22	5	3	3
breast-w	70	9	30	5
bupa	35	6	13	4

Table 2. Experimental results for KNN with K=1

Datasets	AdaBoostM1	RandomSubSpace	Bagging	PSOMulti+RST+MG
iris	94.67 $\ominus$	91.33 $\ominus$	94.67 $\ominus$	95.8 $\oplus$
schizo	59.45 $\ominus$	60.45 $\ominus$	58.64 $\ominus$	92.1 $\oplus$
hepatitis	80.71 $\ominus$	82.63 $\ominus$	80.75 $\ominus$	84.8 $\oplus$
biomed	89.71 $\ominus$	90.76 $\ominus$	89.71 $\ominus$	94.6 $\oplus$
glass	72.06 $\ominus$	77.98 $\ominus$	71.58 $\ominus$	78 $\oplus$
analcaa-bankruptcy	84 $\ominus$	86 $\ominus$	88 $\ominus$	90 $\oplus$
diabetes	69.79 $\ominus$	70.3 $\ominus$	70.18 $\ominus$	73.2 $\oplus$
liver-disorders	65.2 $\ominus$	65.77 $\ominus$	63.17 $\ominus$	68.8 $\oplus$
ecoli	80.69 $\ominus$	79.46 $\ominus$	80.7 $\ominus$	81.9 $\oplus$
vehicle	70.21 $\ominus$	71.74 $\ominus$	70.57 $\ominus$	72 $\oplus$
lung-cancer	65 $\ominus$	68.33 $\ominus$	68.33 $\ominus$	70.5 $\oplus$
segment	97.23 $\ominus$	97.06 $\ominus$	97.01 $\ominus$	97.3 $\oplus$
new-thyroid	97.21 $\oplus$	97.19 $\ominus$	97.21 $\oplus$	97 $\ominus$
breast-w	95.58 $\ominus$	97.18 $\oplus$	95.72 $\ominus$	96.6 $\ominus$
bupa	63.8 $\ominus$	64.11 $\oplus$	63.55 $\ominus$	63.7 $\ominus$

Thus, $\oplus$ indicates that the accuracy is significantly better when $PSOMULTI + RST + MG$ method is used, $\ominus$ signifies that the accuracy is significantly worse and $\odot$ signifies that there is no significant differences.

The Holm test was applied, with respect to the general accuracy of the KNN, and it is corroborated that the results are significantly higher when the contexts obtained by the $PSOMULTI + RST + MG$ method are used. Tables 3 and 4 show the results of the statistical tests related to this result.

P-values obtained in by applying post hoc methods over the results of Friedman procedure. Average ranks obtained by each method in the Friedman test.

Table 3. Average Rankings of the algorithms (Friedman)

Algorithm	Ranking
AdaBoostM1	3.1667
RandomSubSpace	2.3667
Bagging	3.0667
PSOMulti+RST+MG	1.4

Friedman statistic (distributed according to chi-square with 3 degrees of freedom): 17.94.
P-value computed by Friedman Test: 0.000453.

Iman and Davenport statistic (distributed according to F-distribution with 3 and 42 degrees of freedom): 9.281596.
P-value computed by Iman and Daveport Test: 0.000078909633.

Table 4. Post Hoc comparison Table for $\alpha = 0.05$ (FRIEDMAN)

i	algorithm	$z = (R_0 - R_i)/SE$	p	Holm
3	AdaBoostM1	3.747666	0.000178	0.016667
2	Bagging	3.535534	0.000407	0.025
1	RandomSubSpace	2.05061	0.040305	0.05

5. Applications of the Method in the Solution of a Real Problem

In this section a real problem related with the branch of the Civil Engineering is solved, using the following procedure:
Step 1: Build the decision system for the application domain
Step 2: Calculate the weight using the quality of similarity measure (using PSOMulti+RST+MG)
Step 3: Generate set of contexts using the weight calculated in **Step 2**
Step 4: Apply the weights per contexts in the classification with KNN.

The concept of "Level of Service" it was presented as a means to quantify or to classify the operational quality of the service offered by a road to drivers and users. It defined "Level of Service" like a qualitative measure that describes the operational conditions inside the current of the traffic and their perception for the driver and the passenger. A definition of level of service generally describes these conditions in such terms as speed and time of journey, maneuver freedom, interruptions of the traffic, comfort, comfort and security [39].

In the level of service it influences the intensity of the vehicular interaction, the conditions of the road and their environment, and the quality of the regulation and signaling of the road. They have been defined six levels of service for each type of road; assigning them of the letter "A" to the "F", being the level of service "A" the one that represents the best operation conditions and the level of service "F", the worst conditions [39]. The problem is to predict the Level of Service. The description of the dataset is shown in Table 5.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 14 January 2019 doi:10.20944/preprints201901.0125.v1

Table 5. Description of the data-set used in the experiment.

Attributes	Description
Cant-carril	Number of lanes.
Senti-circula	Sense of the way.
Anch-carril	Rail width in meters.
Tipo-separa-central	Type of central separator.
Anch-separa	Width of the central separator in meters.
Estad-pavi	Good, regular or bad.
Parad-omni	Yes or no.
Parq-lateral	Yes or no.
Anch-acera	Width of the sidewalk in meterss.
VHMD	Hourly volume of maximum demand.
Cant-bici	Number of bicycles.
Porc-bici	Percentage of bicycles.
Cant-coche	Amount of animal traction vehicles.
Porc-coche	Percent of animal traction vehicles.
Cant-moto	Number of motorcycles.
Porc-moto	Percent of motorcycles.
Cant-vehi-ligero	Number of light vehicles.
Porc-vehi-ligero	Percent of light vehicles.
Cant-vehi-pesado	Number of heavy vehicles.
Porc-vehi-pesado	Percent of heavy vehicles
Vmedia	Average speed of travel of light vehicles.
Cat-via	Main artery, minor artery or collectors.
level-serv	Level of service of the road (A-F).

The data used for the study were been of counts carried out in different schedules in urban roads in Cuba, in the main arteries of the city of Camaguey.

To predict service levels of a road allows the engineers to base the decisions that propose in this respect of necessities of new roads, their physical and geometric characteristics assisting at the wanted levels of service. They also allow to fix corrections in existent roads, impacting on the organization of the traffic or envelope the characteristics of the road with the objective of elevating the quality of the operational level. The parameters used in the experimentation, for the method $PSOMulti + RST + MG$ were: TB = 40, NI = 100, $ce1$ = $ce2$ = 2 and the values of $e1$ and $e2$ for the function of similarity between attributes and for the function of similarity for the decision attribute were 0.75 and 1.0, F=0.1. The stop condition is: when 100 iterations are reached or when in five iterations the fitness value does not improve (measure quality similarity quality). An experimental study for the data-set traffic is performed (Table 7).

Table 6. Results of the general classification accuracy for level of service with 1NN.

Dataset	AdaBoostM1	RandomSubSpace	Bagging	PSOMulti+RST+MG
Transito3	60	57.5	57.5	63.2

6. Conclusions

In this paper a new method of generating contexts based on similarity relationships for multigranulation using Self-adaptive Differential Particle Swarm using Local Topology for Multimodal Optimization is proposed. The main contribution is the construction of similarity relations based on the quality of similarity measure of Rough Sets Theory as a function of membership to build contexts for multigranulation. This measure calculates the degree of similarity in a decision system in which the feature may have discrete or continuous values. The contexts obtained were evaluated in international databases with the k-NN. The results achieved were significantly superior to the compared methods, which shows the effectiveness of the proposed method. The effectiveness of the method was demonstrated not only in international databases but also in the solution of real problems related to the branches of Civil Engineering (problem of prediction of the level of service in urban roads).

References

1. Pawlak,Z. Rough Sets. *International Journal of Computer and Information Sciences* **1982**, *11*, 341-356.
2. Komorowski, J. and Pawlak, Z. and Polkowski, L. and Skowron, A. Rough sets: A tutorial. *Rough fuzzy hybridization: A new trend in decision-making* **1999**, *11*, 3-98.
3. Greco, S. and Matarazzo, B. and Slowinski, R. Rough sets theory for multicriteria decision analysis. *European journal of operational research* **2001**, *129*, 1-47.
4. Duntsch,I., Gediga,G. Uncertainty measures of rough set prediction. *Artificial Intelligence* **1998**, *106*, 109-137.
5. Jensen,R., Shen,Q. Fuzzy-rough sets assisted attribute selection. *IEEE Transactions on Fuzzy Systems* **2007**, *15*, 73-89.
6. Jeon,G., Kim,D. Jeong,J. Rough sets attributes reduction based expert system in interlaced video sequences. *IEEE Transactions on Consumer Electronics* **2006**, *52*, 1348-1355.
7. Liang,J.Y., Dang,C.Y., Chin,K.S., Yam Richard,C.M. A new method for measuring uncertainty and fuzziness in rough set theory. *International Journal of General Systems* **2002**, *31*, 331-342.
8. Pawlak,Z. Rough Sets: Theoretical Aspects of Reasoning about Data. *System Theory, Knowledge Engineering and Problem Solving* **1991**, *9*.
9. Ziarko,W. Variable precision rough sets model. *Journal of Computer System Science* **1993**, *46*, 39-59.
10. Skowron,A., Stepaniuk,J. Tolerance approximation spaces. *Fundamenta Informaticae* **1996**, *27*, 245-253.
11. Slezak,D., Ziarko,W. The investigation of the Bayesian rough set model. *International Journal of Approximate Reasoning* **2005**, *40*, 81-91.
12. Dubois,D., Prade,H. Rough fuzzy sets and fuzzy rough sets. *International Journal of General Systems* **1990**, *17*, 191-209.
13. Qian,Y.H., Liang,J.Y., Yao,Y.Y., Dang,C.Y. MGRS: a multi-granulation rough set. *Information Sciences* **2010**, *180*, 949-970.
14. Lin,G., Liang,J., Qian,Y. Multigranulation rough sets: From partition to covering. *Information Sciences* **2013**, *241*, 101-118.
15. Intan,R., Mukaidono,M. Multi-rough Sets Based on Multi-contexts of Attributes. *Lectures Notes on Artificial Intelligence* **2003**, *2639*, 279-282.
16. Filiberto,Y., Bello,R., Caballero,Y., Larrua,R. A method to build similarity relations into extended Rough Set Theory. *10th International Conference on Intelligent Systems and Applications* **2010**, *10384*, 978-1-4244-8135-4.
17. Filiberto,Y., Bello,R., Caballero,Y., Larrua,R. Using PSO and RST to predict the resistant capacity of connections in composite structures. *Studies in Computational Intelligence* **2010**, *284*, 359-370.
18. Qian, Y.H. and Liang, J.Y. Combination Entropy and Combination Granulation in Incomplete Information System. **2006**.
19. Qian,Y., Liang,J., Dang,C. Knowledge structure, knowledge granulation and knowledge distance in a knowledge base. *International Journal of Approximate Reasoning* **2009**.
20. Epitropakis,M.G., Plagianakos,V.P. and Vrahatis,M.N. Finding multiple global optima exploiting differential evolutions niching capability. *IEEE Symposium on Differential Evolution* **2011**, 1-8.
21. Bird,S. and Li,X. Adaptively choosing niching parameters in a PSO. *8th Annual Conference on Genetic and Evolutionary Computation* **2006**, 3-10.
22. Liu,Y., Ling,X., Shi,Z., Mingwei,L.V., Fang,J. and Zhang,L. A survey on Particle Swarm Optimization algorithms for multimodal function optimization. *Journal of Software* **2011**, *12*, 2449-2455.
23. Zaldivar,J.M. Estudio e incorporación de nuevas funcionalidades al k-NN Workshop v1.0. *Computación* **2008**.
24. Demsar,J. Statistical comparisons of clas-sifiers over multiple data sets. *Journal of Machine Learning Research* **2006**, *7*, 1-30.
25. Sheskin,D. Handbook of parametric and nonparametric statistical procedures. *Chapman and hall* **2003**.
26. Iman,R. and Davenport,J. Approximations of the critical region of the friedman statistic. *Communications in Statistics* **1980**, *9*, 571-595.
27. Holm,S. A simple sequentially rejective multiple test procedure. *Journal of Statistics* **1979**, *6*, 65-70.

 doi:10.20944/preprints201901.0125.v1

28. Alcalá,J., Fernández,A. KEEL Data-Mining Software Tool: Data Set Repository, Integration of Algorithms and Experimental Analysis Framework. *Journal of Multiple-Valued Logic and Soft Computing* **2010**.
29. Breiman,L. Bagging predictors. *Machine learning* **1996**, *24*, 123-140.
30. Freund,Y. and Schapire,R.E. Experiments with a new boosting algorithm. **1996**, 148-156.
31. Kam,T.H. The random subspace method for constructing decision forests. *IEEE transactions on pattern analysis and machine intelligence* **1998**, *20*, 832-844.
32. Kennedy,J. and Eberhart,R. Particle Swarm Optimization. *IEEE Int. Conference on Neural Networks* **1995**, 1942-1948.
33. Storn,R. and Price,K. Differential Evolution - A simple and efficient heuristic for global optimization over continuous spaces. *J. of Global Optimization* **1997**, 341-359.
34. Napoles,G., Grau,I., Bello,R., Falcon,R. and Abraham,A. Self-adaptive differential particle swarm using a ring topology for multimodal optimization. *Intelligent Systems Design and Applications* **2013**, 35-40.
35. Qian,Y.H., Dang,C.Y., Liang,J.Y. MGRS in incomplete information systems. *IEEE Conference on Granular Computing* **2007**, 163-168.
36. Qian,Y.H., Liang,J.Y., Dang,C.Y. Incomplete multigranulation rough set. *IEEE Transactions on Systems* **2010**, *20*, 420-431.
37. Qian,Y., Zhang,H., Sang,Y. and Liang,J. Multigranulation decision-theoretic rough sets. *International Journal of Approximate Reasoning* **2014**, *55*, 225-237.
38. Asuncion,A. and Newman,D. UCI machine learning repository." A study of the behaviour of several methods for balancing machine learning training data. *SIGKDD Explorations* **2007**, *6*, 20-29.
39. Reyes M. Cardenas J. Cal, R. Ingenieria de transito. *fundamentos y aplicaciones* **2013**.