Editorial: Special issue HAIS12-IGPL

The 14 papers included in this special issue represent a selection of extended contributions presented at the 7th International Conference on Hybrid Artificial Intelligent Systems, HAIS 2012, held in Salamanca, Spain, 28–30 March 2012, and organized by the BISITE and the GICAP research groups.

The International Conference on HAIS 2012 has become a unique, established and broad inter-disciplinary forum for researchers and practitioners who are involved in developing and applying symbolic and sub-symbolic techniques aimed at the construction of highly robust and reliable problem-solving techniques to present the most relevant achievements in this field.

The papers are organized as follows.

In the first contribution by Oliver Kramer *et al.*, they answer the question if the recognition rate of ensemble classifiers is significantly better than the recognition rate of the native classifiers. They analyse two types of bagging classifiers, that is: (i) support vector machine and nearest neighbour ensembles and (ii) random forests. Authors compare their performance in terms of accuracy and robustness on a Nonintrusive Appliance Load Monitoring Systems (NIALM) data set recorded in a field study. The experimental analysis concentrates on recognition rates w.r.t. various training set sizes, on the influence of neighbourhood sizes and the numbers of decision trees in random forest ensembles. It turns out that the decision tree ensembles belong to the best classifiers in the employed scenarios.

In the paper by Chira and Pop, they consider the capacitated fixed-charge transportation problem in a two-stage supply chain network, in which the manufacturer can ship to any of the existing m distribution centres and each of the distribution centres can ship to any of the n customers aiming at minimizing the cost of the physical distribution flow. Due to the complexity of the problem, they propose an improved hybrid algorithm combining the Nearest Neighbour search heuristic with a powerful local search procedure.

The extensive computational experiments on benchmark instances from the literature show that their hybrid algorithm is competitive in comparison with the known heuristics published to date.

In the contribution by Ferreiro *et al.*, they have developed and applied a strategy to detect and isolate process and/or sensor faults by combining a neural network-based functional approximation procedure associated with an online identification algorithm, both processed by recursive rule-based techniques using parity space approaches. A case study dealing with the supervision of a solar volumetric receiver was performed using the proposed intelligent techniques. The conducted study produced reliable and acceptable Intelligent Fault Detection and Isolation (IFDI) results on the basis of heuristic knowledge-based rules.

In the paper by Andreica and Chira, they investigate the evolution and dynamics of small-world networks for the density classification task in Cellular Automatas (CAs). Both un-weighted and node-weighted networks are evolved to serve as the CA topology in density classification using the majority rule over the neighbouring nodes. Allowing a weight for each node of the network, offers the means to induce flexibility in the application of the majority rule.

Computational experiments indicate a better performance of node-weighted networks compared to their weight-free versions for the density classification task. The best-evolved networks were analysed in terms of their tolerance to dynamic network changes understood as node additions and removals. Results indicate a good performance and robustness of the obtained small-world networks for CA density problem.

In the contribution by Golinska-Pilarek and Muñoz, they present a hybrid approach based on the combination of qualitative reasoning, quantitative data and logical methods. In their paper, they also introduce a new propositional dynamic logic Quantum Mechanics (QM) for representation and reasoning with relative movements of objects. In this way, it is possible to infer additional information about movements by using axioms and the logic apparatus. A sound and complete deduction system in dual tableaux style for the logic QM is presented. The system can be used for verification of validity of formulas of the logic in question.

In the paper by Napierala and Stefanowski, the role of classification strategies in rule set-based component classifiers inside bagging is studied. They argue that introducing abstaining in bagging, that is, allowing component classifiers to refrain from predicting class labels in ambiguous situations, improves classification of imbalanced data. Comparative experiments with five different strategies and two rule induction algorithms confirm this hypothesis.

The paper by Ougiaroglou *et al*. presents an adaptive hybrid and cluster-based method for speeding up the k-NN classifier. The proposed method reduces the computational cost as much as possible while maintaining classification accuracy at high levels. The method is based on the well-known k-means clustering algorithm and consists of two main parts: (i) a pre-processing algorithm that builds a two level, cluster-based data structure and (ii) a hybrid classifier that classifies new items by accessing either the first or the second level of the data structure. The proposed approach was tested on seven real-life data sets and the experiential measurements were statistically validated by the Wilcoxon signed ranks test. The results show that the proposed classification method can be used either to achieve high accuracy with slightly higher cost or to reduce the cost at a minimum level with slightly lower accuracy.

The study by Fernandez-Ceniceros *et al.* proposes the use of a HAIS to accurately predict the response of a particular type of steel connection: the bolted lap joint. The proposed hybrid system is composed of hard and soft computing components. First, a set of 800 Finite Element (FE) simulations of different joint configurations was conducted to generate the training and testing data sets for the development of the system. Secondly, a multilayer perceptron network ensemble model was trained and tested. Additionally, a procedure based on genetic algorithms was included to optimize simultaneously both the settings of the model and the number of inputs variables involved in the process. The optimized ensemble model is compared to other soft computing alternatives and shows higher generalization capacity when it deals with testing data. Finally, the results support the use of this hybrid system to create prediction models with similar performance to the FE method in terms of accuracy but with a greatly reduced computational effort.

The paper by Álvarez *et al.* introduces a methodology for intelligent character recognition based on a new feature extraction technique that relies on the number of vertical strokes and the relative positions of the horizontal strokes connected to them. They also propose a new grammar to feed these features into a Deterministic Finite Automaton (DFA) that will generate string representations of every character. Their approach allows the construction of Knowledge Bases and their use with other writers without any previous training. The experiments with several writers and, hence, several Knowledge Bases, show promising performance rates, especially for the recognition of characters written by the owner of the training set.

The research by Villar *et al*. describes two very simple and intuitive techniques to deal with the two former problems making use of heuristics. The underlying idea is to avoid or to limit the surplus computation needed to reduce the problems' effects with simple human-like rules. In the case of bloat, a heuristic to deal with incoherent node sequences is proposed; in the case of over-fitting, the models are allowed a small error and those performing with higher bias are then penalized. These simple techniques are evaluated in a carefully designed test bed, which enables the analysis of their

behaviour. Results show the proposed bloat preventing heuristic, which enhances the results both in the genotype and phenotype landscapes, whereas the over-fitting technique slightly improves the evolutionary process.

In the paper by Martín del Rey *et al.*, a novel encryption scheme for RGB digital images is introduced. The proposed protocol consists of two iterative phases: the confusion phase ruled by the 2D chaotic Cat map, and the diffusion phase governed by a reversible memory cellular automata over F82. This algorithm is shown to be secure against the more important cryptanalytic attacks on image encryption methods (statistical attacks, sensitivity dependence, differential attack, etc.).

The contribution by Zapater *et al.*, proposes the usage of clustering-based outlier detection techniques coupled with a Trust and Reputation System (TRS) engine to detect anomalies in Data Centres. They show how Self Organizing Maps (SOM) and Growing Neural Gas (GNG) can be applied to detect cooling and workload anomalies, respectively, in a real Data Centre scenario with very good detection and isolation rates, in a way that is robust to the malfunction of the sensors that gather server and environmental information.

In the paper by Aboul Ella Hassanien *et al.*, they propose an approach based on rough sets and a local transfer function classifier for heart valve disease detection. To achieve this objective, and to increase the efficiency of the predication model, a Boolean reasoning discretization algorithm is introduced to discrete the heart signal data set, then the rough set reduction technique is applied to find all reducts of the data which contain the minimal subset of attributes that are associated with a class label for classification. Then, the rough sets dependency rules are generated directly from all generated reducts. A rough confusion matrix is used to evaluate the performance of the predicted reducts and classes. Finally, a local transfer function classifier was employed to evaluate the ability of the selected descriptors to discriminate whether they represent healthy or unhealthy.

In the final contribution by Simic *et al.*, they present biological intelligence for modelling and optimization on vehicle routing problem (VRP) of logistics distribution. The aim of this research is to create a novel hybrid model including genetic and firefly algorithms in routing heterogeneous fleet of vehicles in logistics distribution system. The proposed hybrid model in routing heterogeneous fleet VRP is tested on realistic data set in Serbian company 'Carnex'. The experimental results of this method are compared to the empirical results and show that the experimental results obtained by new hybrid genetic—firefly model are better than empirical results in the company.

The guest editors wish to thank Professor Dov Gabbay (Editor-in-Chief of *Logic Journal of the IGPL*) for providing the opportunity to edit this special issue. We would also like to thank the referees who have critically evaluated the papers within the short time. Finally, we hope the reader will share our joy and find this special issue very useful.

Acknowledgements

HÉCTOR QUINTIÁN
Universidad de Salamanca, Spain
e-mail: hector.quintian@usal.es

EMILIO CORCHADO
Universidad de Salamanca, Spain
e-mail: escorchado@usal.es

ANDRE DE CARLVALHO
University of Sao Paulo, Brazil
e-mail: andre@icmc.usp.br

AJITH ABRAHAM
Machine Intelligence Research Labs,Washington, DC, USA
e-mail: ajith.abraham@ieee.org

MICHAŁ WÓZNIAK
Wroclaw University of Technology, Poland
e-mail: michal.wozniak@pwr.wroc.pl

MANUEL GRAÑA
University of the Basque Country, San Sebastian, Spain
e-mail: ccpgrrom@si.ehu.es

SUNG-BAE CHO
Yonsei University, Seoul, Korea
e-mail: sbcho@cs.yonsei.ac.kr