ADVANCES IN INTELLIGENT SYSTEMS AND COMPUTING 223

Václav Snášel
Pavel Krömer
Mario Köppen
Gerald Schaefer (Eds.)

Soft Computing in Industrial Applications

Proceedings of the 17th Online World Conference on Soft Computing in Industrial Applications

Springer

Advances in Intelligent Systems and Computing

Volume 223

Series editor

Janusz Kacprzyk, Polish Academy of Sciences, Warsaw, Poland
e-mail: kacprzyk@ibspan.waw.pl

For further volumes:
http://www.springer.com/series/11156

About this Series

The series "Advances in Intelligent Systems and Computing" contains publications on theory, applications, and design methods of Intelligent Systems and Intelligent Computing. Virtually all disciplines such as engineering, natural sciences, computer and information science, ICT, economics, business, e-commerce, environment, healthcare, life science are covered. The list of topics spans all the areas of modern intelligent systems and computing.

The publications within "Advances in Intelligent Systems and Computing" are primarily textbooks and proceedings of important conferences, symposia and congresses. They cover significant recent developments in the field, both of a foundational and applicable character. An important characteristic feature of the series is the short publication time and world-wide distribution. This permits a rapid and broad dissemination of research results.

Advisory Board

Václav Snášel · Pavel Krömer
Mario Köppen · Gerald Schaefer
Editors

Soft Computing in Industrial Applications

Proceedings of the 17th Online World Conference on Soft Computing in Industrial Applications

Springer

Editors
Václav Snášel
Pavel Krömer
Faculty of Electrical Engineering
and Computer Science
Department of Computer Science
VŠB-TUO
Ostrava-Poruba
Czech Republic

Mario Köppen
Network Design and Research Center
Kyushu Institute of Technology
Fukuoka
Japan

Gerald Schaefer
Department of Computer Science
Loughborough University
Lougborough
UK

ISSN 2194-5357 ISSN 2194-5365 (electronic)
ISBN 978-3-319-00929-2 ISBN 978-3-319-00930-8 (eBook)
DOI 10.1007/978-3-319-00930-8
Springer Cham Heidelberg New York Dordrecht London

Library of Congress Control Number: 2013953229

Printed on acid-free paper

Springer is part of Springer Science+Business Media (www.springer.com)

Preface

This volume of Advances in Intelligent Systems and Computing contains accepted papers presented at WSC17, the 17th Online World Conference on Soft Computing in Industrial Applications, held from December 2012 to January 2013 on the Internet. A tradition started over a decade ago by the World Federation of Soft Computing (http://www.softcomputing.org/) again brought together researchers from over the world interested in the ever advancing state of the art in the field. Continuous technological improvements make this online forum a viable gathering format for a world class conference.

The 2012 edition of the Online World Conference on Soft Computing in Industrial Applications consisted of general track and two special sessions, namely special session on Continuous Features Discretization for Anomaly Intrusion Detectors Generation and special session on Emerging Theories and Applications in Transportation Science. The program committee received a total of 70 submissions from 25 countries, which reflects the worldwide nature of this event. Each paper was peer reviewed by typically 3 referees, culminating in the acceptance of 33 papers for publication. The organization of the WSC17 conference is entirely voluntary. The review process required an enormous effort from the members of the International Technical Program Committee, and we would therefore like to thank all its members for their contribution to the success of this conference. We would like to express our sincere thanks to the special session organizers, to the host of WSC17, VŠB-Technical University of Ostrava, and to the publisher, Springer, for their hard work and support in organizing the conference. Finally, we would like to thank all the authors for their high quality contributions. The friendly and welcoming attitude of conference supporters and contributors made this event a success!

February 2013

Václav Snášel
Pavel Krömer
Mario Köppen
Gerald Schaefer

Organization

General Chair

Václav Snášel	VŠB-Technical University of Ostrava

Program Chair

Pavel Krömer	VŠB-Technical University of Ostrava

Special Event Chair

Gerald Schaefer	Loughborough University, UK

International Advisory Board

Kalyanmoy Deb	Indian Institute of Technology Kanpur, India
Carlos M. Fonseca	University of Algarve, Portugal
Thomas Stützle	University Libre de Bruxelles, Belgium
Jörn Mehnen	Cranfield University, UK
Mario Köppen	Kyushu Institute of Technology, Japan
Xiao-Zhi Gao	Aalto University, Finland
Gerald Schaefer	Loughborough University, UK

International Technical Program Committee

Janos Abonyi	University of Pannonia, Hungary
Mohanad Alfiras	Gulf University, Bahrain
Sudhirkumar V Barai	IIT Kharagpur, India

Helio Barbosa	Laboratório Nacional de Computação Científica, Brazil
Özgür Başkan	Pamukkale University, Turkey
Miloš Besta	Google Inc., USA
Gennaro Nicola Bifulco	Università di Napoli Federico II, Italy
Leonardo Caggiani	DVT-Technical University of Bari, Italy
Piotr Cal	Wrocław University of Technology, Poland
Rodrigo T. N. Cardoso	CEFET-MG, Brazil
M. Emre Celebi	Louisiana State University in Shreveport, USA
Hilmi Berk Celikoglu	Technical University of Istanbul, Turkey
Lino Costa	University of Minho, Portugal
Keshav Dahal	University of Bradford, UK
Justin Dauwels	Nanyang Technological University, Singapore
Guy De Tre	Ghent University, Belgium
Alexandre Delbem	University of Sao Paulo, Brazil
Viviane G. Da Fonseca	INUAF-Instituto Superior D. Afonso III, Portugal
Gregorio Gecchele	University of Padova, Italy
Carlos Henggeler Antunes	University of Coimbra, Portugal
Francisco Herrera	University of Granada, Spain
Dušan Húsek	Academy of Sciences of the Czech Republic, Czech Republic
Konrad Jackowski	Wrocław University of Technology, Poland
Milica Kalić	University of Belgrade, Republic of Serbia
Petra Kersting	Technische Universität Dortmund, Germany
Frank Klawonn	University of Applied Sciences Braunschweig-Wolfenbuettel, Germany
Andrew Koh	University of Leeds, UK
Bartosz Krawczyk	Wrocław University of Technology, Poland
Renato Krohling	Federal University of Espirito Santo, Brazil
Pavel Krömer	VŠB-Technical University of Ostrava, Czech Republic
Miloš Kudělka	VŠB-Technical University of Ostrava, Czech Republic
Jouni Lampinen	University of Vaasa, Finland
Celina Leao	University of Minho, Portugal
Gabriella Mazzulla	University of Calabria, Italy
Yetis Sazi Murat	Pamukkale University, Turkey
Santosh Nanda	Eastern Academy of Science and Technology, India
Eliška Ochodková	VŠB-Technical University of Ostrava, Czech Republic
Jae Oh	Syracuse University, USA
Michele Ottomanelli	Politecnico di Bari, Italy

Suhail Owais	Applied Science University, Jordan
Ana Pereira	Polytechnic Institute of Bragança, Portugal
Jan Platoš	VŠB-Technical University of Ostrava, Czech Republic
Sebastian Polak	Jagiellonian University Medical College, Poland
Sg Ponnambalam	Monash University Sunway Campus, Malaysia
Petrica Claudiu Pop	North University of Baia Mare, Romania
Radu-Emil Precup	Politehnica University of Timisoara, Romania
Ana Maria A. C. Rocha	University of Minho, Portugal
Riccardo Rossi	University of Padova, Italy
Nicola Sacco	DIME-University of Genoa, Italy
Gerald Schaefer	Loughborough University, UK
Giovanni Semeraro	University of Bari, Italy
Sara Silva	INESC-ID, Portugal
Piotr Sobolewski	Wrocław University of Technology, Poland
Marcone Souza	Universidade Federal de Ouro Preto, Brazil
Thomas Stützle	IRIDIA, ULB, Belgium
Eiji Uchino	Yamaguchi University, Japan
Juan Velasquez	University of Chile, Chile
Milorad Vidovic	University of Belgrade, Serbia
Michał Woźniak	Wrocław University of Technology, Poland

Sponsoring Institutions

World Federation on Soft Computing
VŠB Technical University of Ostrava

Contents

Part III Emerging Theories and Applications in Transportation Science

Part I
Soft Computing in Industrial Applications

Advanced Methods for 3D Magnetic Localization in Industrial Process Distributed Data-Logging with a Sparse Distance Matrix

Abhaya Chandra Kammara and Andreas König

Abstract Wireless sensor networks/data-logging devices are increasingly applied for distributed measurement and acquiring additional contextual data. These have been applied in large scale indoor and outdoor systems with solutions based on RF, light based and ultra sound based systems. Data-loggers in liquid filled containers pose new challenges for localization because of the high reflectivity of containers and high attenuation due to the liquids obstructing communication between wireless nodes. Magnetic localization techniques have been used in many places including military research [14]. This approach was adapted for use in liquid filled containers. In this project, two prototypes, a laboratory and an industrial installation have been conceived and served for acquisition of experimental data for localization. In our paper, we exploit the sparsity met in the particular magnetic MEMS sensor swarm localization concept by introducing NLMR which is a simplified form of Sammon's mapping (NLM) and we combine it with different meta-heuristics and soft-computing techniques, e.g., gradient descent, Simulated Annealing and PSO. We compare this with Multilateration and conventional NLM localization technique. Our approach has improved the localization from a mean error of 20 cm in the first cut analysis for the industrial setup using conventional NLM down to 11 cm without and to 9 cm with apriori knowledge. Future improvements are to be expected from a thorough calibration of all system components. in [5]. The modified algorithm is capable of distributed localization producing mean localization error of 10 cm for the Warstein experiment data.

A. C. Kammara (✉) · A. König
Institute of Integrated Sensor Systems, TU Kaiserslautern, 67663 Kaiserslautern, Germany
e-mail: abhay@eit.uni-kl.de

A. König
e-mail: koenig@eit.uni-kl.de

V. Snášel et al. (eds.), *Soft Computing in Industrial Applications*,
Advances in Intelligent Systems and Computing 223, DOI: 10.1007/978-3-319-00930-8_1,

1 Introduction

Sensors and Sensing systems are getting ubiquitous in industry and homes alike. Providing contextual information is essential and advantageous in industry and in ambient intelligence systems. Data-loggers have become prominent in such setups. Localization of such mobile sensors are traditionally done using RF, ultrasonic, IR and other methods. This spectrum of approaches is not suitable in our case, because of the high attenuation and reflections caused in liquid containers of industrial processes. Magnetic localization on the contrary can be a feasible method for localization in this situation.

Magnetic localization techniques have been in use for a long period of time from their introduction in 1962 [1]. They have been used in head tracking applications [3] in military [14], for silent localization of underwater sensors [9], in location and orientation tracking [8], in medical systems [12, 13]. The approaches vary from using Internal (mostly in medical approaches) and External (Using artificially generated magnetic fields or earth magnetic field). In this project quasi-DC fields with artificially generated magnetic fields in coils is used. The idea behind such a localization approach for data-logging had its patent filed [21] on 18.05.2010. Similar idea has been used recently for data-logging in [20]. Interestingly an approach similar to this project is also being used in Indoor positioning systems [10, 11].

In the following section the particular project approach, providing the data for the localization algorithm experiments reported in this work, will be described in detail. Based on the data from measurement, localization algorithms are employed to estimate the coordinates of WSN nodes. Commonly, algorithms from, e.g., multi-dimensional-scaling (MDS), are employed. These basically are fine, but in their majority base on the assumption of a densely populated distance matrix and require substantial post-processing for the final coordinate determination. In the regarded research project, inter node communication is practically unavailable, so

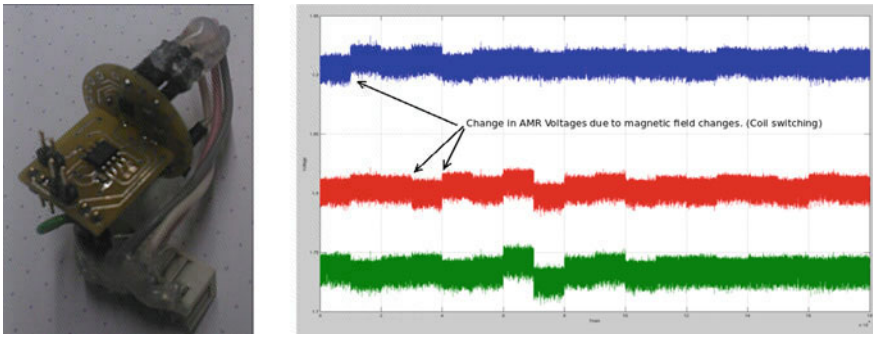

Fig. 1 3D AMR sensor node (*left*), noisy raw data obtained from 3D AMR node in a scaled down ISE lab setup (*right*)

that the resulting distance matrix is sparse, i.e., having only anchor to sensor node non-zero entries. Thus, in this paper we investigate the recall extension of Sammon's non-linear mapping (NLMR) for localization purposes and compare it with multilateration based on data acquired in the industrial target environment. Also, we enhance NLMR with different soft computing techniques and show that we can get more than competitive results, than original Sammon's mapping (NLM) with much less computational effort.

2 The Localization System Setup

In our localization concept and implementation triaxial Anisotropic Magneto-resistive (AMR) sensors complementing the data-loggers which will be deployed in the liquid containers. A magnetic coils system has been mounted on the container hull producing quasi-DC magnetic fields controlled by a central control unit synchronized to the sensors clock. This is used to produce voltage values in the sensors, which can be converted to distances. These distances can be used to compute the location of the sensor node.

2.1 Sensor Node

In the project, artificially generated magnetic fields are sensed by a triaxial Anisotropic Magneto-resistive (AMR) sensor. AMR sensor makes use of a magneto-resistive effect to detect magnetic fields. In the project, AMR sensor type AFF755B from Sensitec Gmbh was used to design a proprietary 3D sensor. The Fig. 1 shows a PCB AMR sensor. There is currently a MEMS prototype which will be used in future experiments.

2.2 Magnetic Field Generation

There are different types of coils that could be used for magnetic field generation. In this project, circular coils of container specific diameters are used to generate the fields. The coils should be positioned in such a way that there are at least 4 magnetic fields observed by the sensor at any point in the cylinder. Generally localization is done in such a way the distance between the coil and the sensor is much greater that the radius of the coil.

Magnetic fields are not simultaneously produced from all the coils. The magnetic field are generated one coil at a time. Each coil is switched-on in both directions. Ternary switching allows a differential and energy-aware measurement, eliminating static magnetic offsets and reducing the flipping of AMR sensors to a

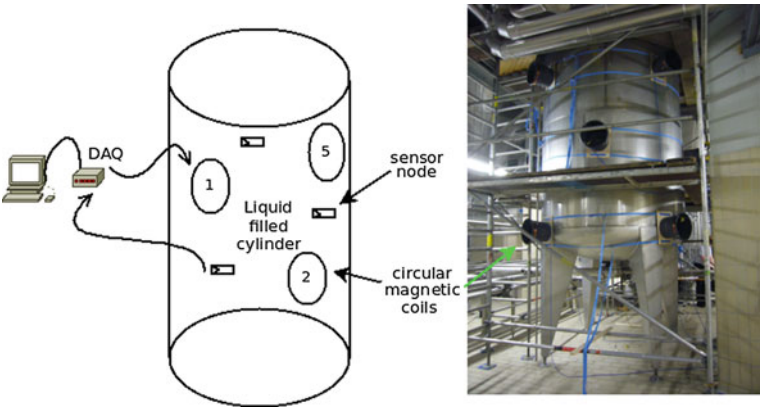

Fig. 2 Sketch of the measurement system with sensor and coils (*left*), photo of industrial container with coils (*right*)

minimum. In the experiments a DAQ board (DT9816) was used to control the coils and in data acquisition as shown in the Fig. 2. In the final planned setup the control circuitry is different for the coil switching and reading from AMR. The clock synchronization errors can be resolved by using synchronization methods described in [19].

2.3 Distance Calculation

Three distances are obtained for each activation of the coil. These voltages are converted into distances using the formulae given below. The angle between the node and the sensor is not considered since it is an unknown. However there are techniques which could be used as mentioned in [18].

$$d = \left(\left(\frac{\frac{1}{2} \times \mu_0 \times n \times R^2 \times I}{B_M} \right) - R^2 \right)^{\frac{1}{2}} \tag{1}$$

where,

$$Bm = \frac{V_M}{S \times V_S \times G}, V_M = \frac{V_i^p - V_i^n}{2}, i = \{x, y, z\}$$

where S is the sensitivity, V_S is sensor voltage, G is the gain of the amplifier, n is the number of windings, R denotes the radius of the coil and $\mu_0 = 4 \times \pi \times 1e^{-7}$.

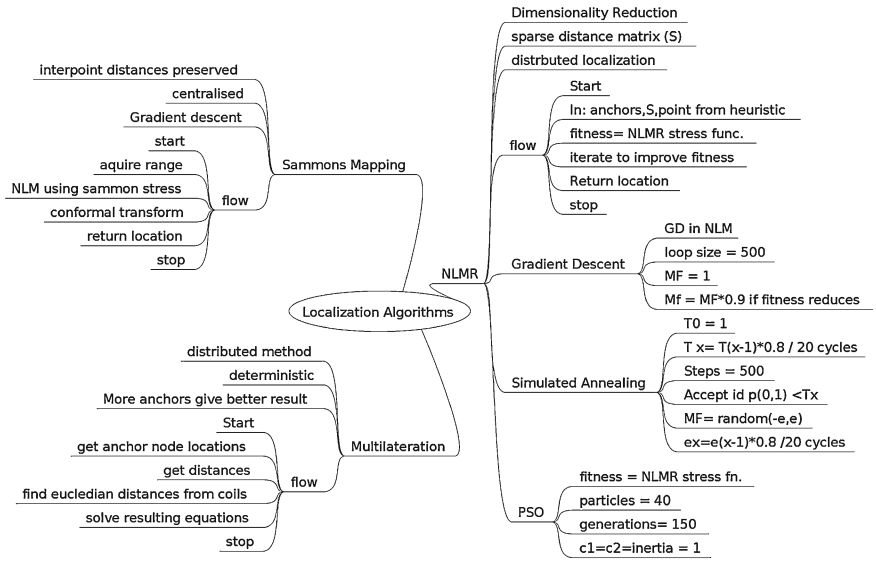

Fig. 3 Survey of localization algorithms and parameter settings

3 Localization

3.1 MDS Mapping

MDS mapping is a popular technique used in localization. There are many types of MDS mapping, we are interested in the non linear Sammon's mapping (NLM). In NLM localization is done based on reducing the cost function E(m)

$$E = \frac{1}{\sum_{i=1}^{N-1}\sum_{j=i+1} N d_{ij}^{*}} \sum_{i=1}^{N-1}\sum_{j=i+1}^{N} \frac{\left(d_{ij}^{*} - d_{ij}\right)^2}{d_{ij}^{*}} \tag{2}$$

where N is number of nodes, d_{ij}^{*} is Euclidean distance between X_i and X_j in Higher dimension and d_{ij} is Euclidean distance between Y_i and Y_j in lower dimension. In the localization problem there is no dimensionality reduction. The main disadvantage of this method in our scenario is that we do not have inter sensor distances. A modification to Sammon's mapping that limits the involved number of distances seems promising [5] Fig. 3.

3.2 NLMR

Sammons recall or NLMR was described in [5] for dimensionality reduction. This technique is a simplification of Sammons mapping (NLM) where inter point

distances are ignored and a set of code-book vectors (previously mapped points) are used to find the positions of all points. This is much less resource consuming as compared to Sammons mapping (NLM) [5]. Here we have a cost function E(m) described by,

$$E_i(m) = \frac{1}{c}\sum_{j=1}^{K}\frac{(d_{Xij} - d_{Yij}(m))^2}{d_{Xij}} \tag{3}$$

where,

$$d_{Xij} = \sqrt{\sum_{q=1}^{m}(v_{iq}^r - v_{jq}^t)^2}, c = \sum_{j=1}^{K} d_{Xij}$$

d_{Xij} is the distance between recall and training data in high dimensional space and K is the number of code book vectors. The distances in the new space are found using gradient descent technique

Gradient Descent: In the Gradient Descent approach we make use of the NLMR cost function and the following equations.

$$y_{iq}(m+1) = y_{iq}(m) - MF \times \Delta y_{iq}(m) \tag{4}$$

with,

$$\Delta y_{iq}(m) = \frac{\left(\frac{\partial E_i(m)}{\partial y_{iq}(m)}\right)}{\left(\frac{\partial E_i^2(m)}{\partial y_{iq}(m)^2}\right)}, 0 < MF \leq 1 \tag{5}$$

Where $y_{iq}(m+1)$ is the new position, MF is the magic factor which reduces with time, $y_{iq}(m)$ is the current position and E(m) is the cost function at the current position.

Modification: The specialty of NLMR approach is that it matches perfectly with our requirements. We do not obtain an inter-point distance between the nodes. We have a set of mapped positions (coils). We have to make some changes to the original algorithm to make it work for localization.

In our case we do not require a dimensionality reduction. So the known locations (magnetic coils) will act as the trained data set. We do not have any location for the unknown value, however we have the distance information from the known locations which can be directly given to the algorithm.

Unlike Sammon's mapping we do not have to do a reverse mapping (conformal transform) after we get our output (Table 1).

After making these changes, we can make use of any soft-computing technique to generate the values to be given to the Sammon's recall function. In our approach we tried Gradient descent, simulated annealing and PSO.

NLMR - Gradient descent: The gradient descent approach makes use of Quasi- Newton method similar to the method described in Eq. 4, [5]. We make use

of Magic factor initialized with 1 and reduced when the new stress is not as good as the old stress. Only better solutions are accepted in this approach.

NLMR - Simulated Annealing: We use the basic simulated annealing where we start with a relatively high temperature ($T_0 = 1$) which is reduced ($Tx = T_{(x-1)} * 0.8$) over the number of cycles and reduce the chances of choosing a bad solution as the temperature decreases (Accept any solution if p(0,1)$leq\ t_x$). The new solutions are found by a markov chain shown in Eq. 4 with a random MF between $-e_0$ and e_0 where e (energy factor) reduces over time. The algorithm runs for 500 iterations to get the best solutions. The number of iterations required was found heuristically.

NLMR - Particle swarm optimization: Standard particle swarm optimization described in [4] is used with $C_1 = C_2 = 1$ and the algorithm has no inertia. Having no inertia helped in faster convergence of the algorithm. 40 particles were used with 150 generations found the best results in the experiments.The standard PSO was unable to converge to a solution using Sammon's mapping and required special approaches to reach convergence [15]. However, in the modified NLMR method the standard PSO is able to converge easily.

Multilateration: Multilateration is used in wireless sensor networks and a standard technique for efficient and effective localization. It serves as a reference in our work. Traditionally the time difference of arrival of the signal is used in multilateration technique, we use the distances obtained with our magnetic localization system. In our approach we make use of the linear least squares method (Moore-Penrose pseudo-inverse) to get the best fit solution.

Coil Selection: In the previous works of this project coil selection was used by taking a heuristic of the closest distances (effectively closest coils (anchors)) available. However in our data analysis we found that erroneous distances detected by the sensor was also present in closer coils. In our analysis we compared the distances from the sensor and the distances calculated from the ground truth value. In Fig. 4, three coils with their distance errors are plotted for all our data. The peaks correspond to distance errors due to angle (which is not honored in our current formula for distance calculation). In Fig. 4, we can see the original positions where the sensors were placed. Even though coil 12 was furtherest from

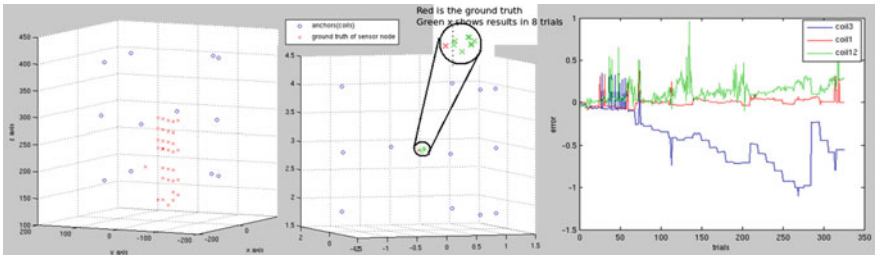

Fig. 4 Plot showing the ground truth and anchor (coils) locations (*left*), results for 8 trials on one ground truth location (*middle*), variations between expected distances and acquired distances for 325 experiments (*right*)

Table 1 Parameters used in the industrial setup

trials	pos	trials/pos	coils	coil dia.	windings	V-Sensor	I-coils
325	40	2 to 20	12	0.25 m	230	3.7V	3A

Table 2 Results making use of all available anchors (coils) (top), results using anchors (coils) selected by apriori knowledge (1,2,4,5,6,7,8,9) (bottom), all values are in meters

All Coils	**NLMR-GD**	**NLMR-SA**	**NLMR-PSO**	**Multilateration**
LEμ	0.17951	0.17914	**0.11429**	0.16702
LEσ	0.13097	0.12557	**0.06580**	0.15194
Max LE	0.76031	**0.62029**	0.69258	1.31961
Min LE	0.02230	0.01720	**0.00911**	0.01302
μ X	−0.09350	−0.09780	**−0.00121**	−0.05915
μ Y	0.08262	0.08639	**0.01846**	−0.02365
μ Z	0.02610	0.02201	0.02336	**−0.00190**
σ X	0.14338	0.13490	**0.06336**	0.12351
σ Y	0.08881	0.08591	**0.07171**	0.11815
σ Z	0.06942	**0.06904**	0.08571	0.13306
Max err X	0.67247	0.55810	0.50856	**0.48092**
Max err Y	0.58066	0.55335	**0.50618**	0.60638
Max err Z	**0.20882**	0.21477	0.30020	1.15277
Min err X	0.00017	**0.00001**	0.00018	0.00043
Min err Y	0.00140	0.00060	0.00065	**0.00020**
Min err Z	0.00026	0.00026	0.00021	**0.00011**
Sel. Coils	**NLMR-GD**	**NLMR-SA**	**NLMR-PSO**	**Multilateration**
LE μ	0.10106	**0.09345**	0.09534	0.13912
LE σ	**0.05822**	0.05861	0.06129	0.11206
Max LE	0.69430	**0.68370**	0.70647	0.80797
Min LE	0.01774	0.01126	0.01239	**0.00573**
μ X	**−0.00238**	0.01458	0.00813	−0.03728
μ Y	**0.00126**	0.00150	0.00686	−0.03760
μ Z	**0.00323**	0.02655	−0.00776	0.03375
σ X	0.04970	**0.04145**	0.04564	0.07312
σ Y	0.06513	0.06770	**0.06290**	0.09209
σ Z	0.08290	**0.07033**	0.08144	0.11892
Max err X	0.19471	**0.13390**	0.17319	0.19803
Max err Y	0.52350	0.53103	0.49736	**0.49735**
Max err Z	**0.50800**	0.57788	0.58183	0.78424
Min err X	0.00043	**0.00010**	0.00043	0.00012
Min err Y	**0.00003**	0.00007	0.00065	**0.00003**
Min err Z	0.00033	0.00021	**0.00009**	0.00018

Table 3 Parameters used in the ISE lab experimental setup

trials	pos	trials/pos	coils	coil dia.	windings	V-Sensor	I-coils
56	56	1	6	0.13 m	100	5V	5A

Table 4 First-cut results from ISE lab orthogonal demonstrator

	NLMR-GD	NLMR-SA	NLMR-PSO	Multilateration
LE μ	0.07169	**0.07137**	0.08715	0.12526
LE σ	**0.03312**	0.03429	0.03428	0.07147
μ X	−0.05512	0.05434	**0.05112**	0.10509
μ Y	0.03315	**0.03307**	0.03383	−0.05274
μ Z	0.00532	−0.00639	−0.00744	**−0.00212**
σ X	**0.02857**	0.02953	0.03727	0.07489
σ Y	0.02062	**0.02057**	0.02612	0.03253
σ Z	0.02881	0.02978	0.05372	**0.01735**

any of the data points while coil 3 was close to most of them coil 3 has a higher error as compared to coil 12. The reason behind this maybe due to an unfortunate angle with respect to the data points. Here, we remove 4 coils depending on such a study (coils 3,10,11,12) and observe the improvement in results. More research on coil selection will be done in our future work (Tables 2, 3).

4 Results

The data acquisition based on the container of Fig. 2 (right), was confined to a cubic volume in the container of 350 cm × 350 cm × 250 cm. The acquired data was first-cut analyzed in prior project work, employing the conventional NLM localization method. These previous results serve as the baseline to comparatatively evaluate the suggested new methods. In the tables below we use Localization Error (LE) which is the distance between the ground truth and obtained values. The parameter settings of all conducted experiments are summarized in Fig. 3. In addition to the *stale* database from the industrial environment, we extracted fresh data from the ISE lab setup, which has just six smaller scaled coils in an orthogonal arrangement of 150 cm × 150 cm × 150 cm. Raw data in Fig. 2 was obtained here. In a first step, 56 different locations with just one trial each were sampled in a single Z-plane and the data was subject to multilateration and the suggested methods. The results given in Table 4 in comparison to Multilateration confirm the viability of our approach. The ISE lab demonstrator is reshaped to cylindrical shape and campaigns with different sensor heads including MEMS and wireless sensor will follow-up.

5 Conclusion

In our work, we have presented an adaptation of NLMR, a simplified Sammon's mapping for dimensionality reduction and used it for localization with a sparse distance matrix, e.g., for liquid filled containers or general indoor localization.

This was applied to industrial data available as benchmark and obtained comparable results to Multilateration. The results were also better than those achieved by conventional NLM in first-cut analysis with the same data. We also presented simulated annealing and particle swarm optimization to reduce the NLMR cost function and these methods provide better results than the commonly used gradient descent method [2,5]. The standard NLMR PSO even provides results for all coils selected comparable to results of the other methods when coils giving unreliable information are removed.

Acknowledgments This work was partly supported by the Federal Ministry of Education and Research (BMBF) in the program mst-AVS, in the project ROSIG grant no. 16SV3604 of the PAC4PT consortium (Partners were UST GmbH, IMST GmbH, Krohne, Warsteiner, and microTEC Ges. für Mikrotechn. GmbH (Coord.). The industrial environment data, employed here for the algorithmic studies, has been acquired and first-cut analysed by conventional NLM in the project work by S. Carrella and D. Groben. All algorithms were implemented using python (numpy, scipy, and matplotlib packages) and Matlab.

References

1. Kalmus, H.P.: A new guiding and tracking system. IRE Trans. Aerosp. Navig. Electron. **9**, 7–10 (1962)
2. Sammon, J.W.: A nonlinear mapping for data structure analysis, IEEE Trans. Comput. C-18, 401–409 (1969)
3. Raab, F.E., et al.: Magnetic position and orientation tracking system, IEEE Trans. Aerosp. Electron. Syst. **5**, 709–717 (1979). ISSN 18237843
4. Kennedy, J., Eberhart, R.C.: Particle swarm optimization. Proc, IEEE International Conference on Neural Networks, IJCNN (1995)
5. König, A.: Interactive visualisation and analysis of hierarchical neural projections for data mining. In IEEE TNN, Special Issue on Neural Networks for, Data Mining and Knowledge Discovery, pp. 615–624, 11(3), May, 2000
6. König, A.K.: "Dimensionality reduction techniques for interactive visualization, exploratory data analysis, and classification", Pattern Recognition in Soft Computing Paradigm, World Scientific, N.R. Pal (eds.), 2: 1–37, 2001
7. Paperno, E., Sasada, I., Leonovich, E. A new method for magnetic position and orientation tracking. IEEE Trans. Magn. **37**(4), JULY 2001
8. Prigge, E.A.: A positioning system with no line-of-sight restrictions for cluttered environments. Stanford University, Dissertation (2004)
9. Callmer, J., Skoglund, M.: F. Silent localization of underwater sensors using magnetometers. EURASIP J. Adv. Sig. Process, Gustafsson (2010)
10. Blankenbach, J., Norrdine, A.: Position estimation using artificial generated magnetic fields, : International Conference on Indoor Positioning and Indoor Navigation (IPIN), 15–17 September 2010. Zürich, Switzerland (2010)
11. Blankenbach, J., Norrdine, A., Hellmers, H.: Adaptive signal processing for a magnetic indoor positioning system geodetic institute, Technische Universität Darmstadt - Short paper IPIN
12. Placidi, G., Franchi, D., Maurizi, A., Sotgiu, A.: Review on patents about magnetic localisation systems for in-vivo catheterizations INFM c/o Dept. of Health Sci., Uni. of LAquila, Via Vetoio Coppito 2, 67100 LAquila, Italy

13. Ascension Technology Corporation Products Application: [Online]. http://www.ascension-tech.com/medical/pdf/TrakStarWRTSpecSheet.pdf, checked Nov. 5 2012
14. Polhemus: [Online]. Available: http://www.polhemus.com/?page=Military_Why_Magnetic_Tracking, checked Nov. 5 2012
15. Iswandy, K., önig, A.K.: Soft-computing techniques to advance non-linear mappings for multi-variate data visualization and wireless sensor localization. e-Newsletter IEEE SMC Soc., Issue #29, Dec. 2009
16. Carrella, S., Iswandy, K., Lutz, K., önig, A.K.: 3D-Localization of low-power wireless sensor nodes based on AMRSensors in industrial and AmI applications, VDE Verlag GmbH Berlin Offenbach, pp. 522–529, 2010
17. Iswandy, K., Carrella, S., önig, A.K.: Localization system for low power sensor nodes deployed in liquid-filled industrial containers based on magnetic sensing. Tagungsband XXIV Messtechnisches Symposium des AHMT, 23.-25. September, pp. 108–121, 2010
18. Carrella, S., Iswandy, K., önig, A.K.: A system for localization of wireless sensor nodes in industrial applications based on sequentially emitted magnetic fields sensed by tri-axial AMR sensors
19. Carrella, S., Iswandy, K., önig, A.K.: System for 3D localization and synchronization of embedded wireless sensor nodes based on AMR sensors in industrial environments, Proceedings Sensor+test 2011
20. Reinecke, S., öpping, U.P., Hampel, U.: Autonome sensorpartikel zur räumlichen Parametererfassung in groβskaligen Behältern, Sensor & Test 2012
21. Pending Patent Application: Method and apparatus for determining the spatial coordinates of at least one sensor node in a container, filing date: 18.05.2010, Int. Pub. 24.11.2011, Int.No.: WO 2011/144325 A2.

Neural Network Ensemble Based on Feature Selection for Non-Invasive Recognition of Liver Fibrosis Stage

Bartosz Krawczyk, Michał Woźniak, Tomasz Orczyk, Piotr Porwik, Joanna Musialik and Barbara Błońska-Fajfrowska

Abstract Contemporary medicine concentrates on providing high quality diagnostic services, yet it should not be forgotten that the comfort of the patient during the examination is also of high importance. Therefore non-invasive methods that allows to precisely predict the state of the disease are currently one of the key issues in the medical business. The paper presents a novel ensemble of neural networks applied to recognition of liver fibrosis stage from indirect examination method. Several neural network models are build on the basis of outputs of different feature selection algorithms. Then an ensemble pruning procedure with the usage of diversity measures is conducted in order to eliminate redundant predictors from the pool. Finally the weights of classifiers in the fusion process are assessed to establish their influence on the output of the whole ensemble. Proposed method is compared with several state-of-the-art ensemble methods. Extensive experimental investigations, carried out on a dataset collected by authors, show that the proposed method achieve a satisfactory level of the

B. Krawczyk (✉) · M. Woźniak
Department of Systems and Computer Networks, Wroclaw University of Technology, Wyb Wyspianskiego 27, 50-370 Wroclaw, Poland
e-mail: bartosz.krawczyk@pwr.wroc.pl

M. Woźniak
e-mail: michal.wozniak@pwr.wroc.pl

T. Orczyk · P. Porwik
Institute of Computer Science, University of Silesia, Bedzinska 39, 41-200 Sosnowiec, Poland
e-mail: tomasz.orczyk@us.edu.p

P. Porwik
e-mail: piotr.porwik@us.edu.pl

J. Musialik
Department of Gastroenterology and Hepatology, Medical University of Silesia, Katowice, Poland

B. Błońska-Fajfrowska
Department of Basic Biomedical Science, Medical University of Silesia, Sosnowiec, Poland

V. Snášel et al. (eds.), *Soft Computing in Industrial Applications*, Advances in Intelligent Systems and Computing 223, DOI: 10.1007/978-3-319-00930-8_2,

fibrosis level recognition, outperforming other machine learning algorithms and thus may be used as a real-time medical decision support system for this task.

1 Introduction

Liver fibrosis is a condition where fibrous scare tissue accumulates in the liver. It is a common complication of many diseases but in this research we use medical data of patients infected with liver hepatitis type B and C virus (HBV/HCV). Early detection of liver fibrosis is very important as the condition may stay completely hidden for months or even years, but untreated may lead to liver cirrhosis and in consequence to patient's decease.

In most cases the condition stays in so called compensated state, so no visible changes nor dysfunctions might be observed, but although most medical examination results are within their normal results, some slight discrepancies may be observed and used to evaluate the liver fibrosis stage [13]. In this research we will be using blood test results.

Mentioned results coming from the non-invasive biomedical examination are then treated as an input for the machine learning algorithms. Our aim is to create an accurate medical decision support tool that will allow for an automatic classification of patients under the observation.

Among many machine learning techniques it is impossible to select the best one for the task at hand without any a priori knowledge about the data [18]. Therefore standard procedure concentrates on building several predictors, testing their performance and selecting the single best model according to some criterion. Yet it should be beard in mind that in many cases different models may contribute uniquely to the analysed task. That is why the Multiple Classifier Systems (MCS) are one of the major research directions in machine learning [14]. They propose to utilize more than one model in hope to exploit their strengths while reducing their drawbacks. It has been shown many times that a ensemble may achieve a better accuracy than any of the individual members taking part in it.

In this work we propose an ensemble based on neural networks, as those classifiers were one of the first to highlight the effectiveness of the MCS approach [6]. We address three main issues in our compound model—assuring base diversity among individual models, discarding the redundant predictors and creating a fusion methodology that exploits the classifier local competencies. Individual models are build on the basis of outputs of different feature selection algorithms. This way we achieve both—the initial diversity (as different methods return different reduced feature spaces) and the decreased complexity of the model (due to the reduction of the feature space). Then ensemble pruning, based on the diversitymeasure, is conducted to discard predictors that are similar to each other and therefore add no value to the committee. Finally a trained fusion procedure,

based on individual discriminant functions, is performed in order to boost the quality of the proposed MCS.

The proposed method is compared to other state-of-the-art ensembles in order to asses its quality for the task of liver fibrosis recognition.

2 Liver Fibrosis Recognition

The only method of liver fibrosis stage recognition giving a 100 % accuracy is an autopsy and this is due to the fact that the condensation of scare tissue within the liver may vary in different regions of the organ. For the same reason the most common examination method—liver biopsy doses not guarantee the correct diagnosis. This method is unfortunately not only inaccurate, but also may lead to serious health complications including risk of patient's death.

There are two (or three) common description methods for liver biopsy samples. One used in the article is METAVIR [2] (4 stages of fibrosis) and the other are: Histological Activity Index (HAI Score) also known as Knodell Score [9] (3 stages of fibrosis) and it's modified version called Ishak Score [7] (6 stages of fibrosis). The METAVIR has been specifically designed and validated for patients with hepatitis C. All these systems rely on a histological image of the liver and in consequence their quality depends on a sample size and doctor's experience.

Due to facts presented in the first paragraph also some non invasive examination methods have been developed. Most common one is APRI-test [17], but also ELF-Test [12] and FIBRO-Test [4] have been developed by medical companies. All these methods are blood test based, but the first one is very general and can be used only to detect advanced fibrosis or cirrhosis. Two other are more specific, but also more expensive for patients. All blood test based methods try to detect some dependencies between liver functionality and blood test results, so they are indirect and non-invasive methods. This is very important, because in opposite to liver biopsy, they may be repeated in regular periods of time without a harm for a patient.

For the purpose of presented research we acquired medical data records from 103 real patients of the Branch of the Gastroenterology and Hepatology of the Independent Public Central Hospital of the Silesian Medical University. Table 1 presents number of examined patients for each fibrosis stage (F0..F4) and Table 2 presents characteristics of acquired medical records.

Table 1 Number of patients with given fibrosis stage [n (%)]

*F*0	*F*1	*F*2	*F*3	*F*4
2 (2 %)	34 (33 %)	5 (5 %)	16 (15 %)	46 (45 %)

Table 2 Blood test results characteristics [mean (std. deviation)]

HB (g/l)	14 (1.80)
RBC ($10^6/\mu l$)	4 (0.69)
WBC ($10^3/\mu l$)	6 (2.35)
PLT ($10^3/\mu l$)	161 (73.08)
PT (sec.)	13 (9.99)
PTP (%)	91 (17.23)
APTT (sec.)	37 (7.29)
INR	1 (0.16)
ASPT (IU/l)	69 (55.14)
ALAT (IU/l)	77 (65.43)
ALP (IU/l)	105 (57.70)
BIL (mg/dl)	2 (2.42)
GGTP (IU/l)	94 (101.31)
KREA (mg/dl)	1 (0.24)
GLU (mg/dl)	93 (17.44)
Na (mmol/l)	138 (3.28)
K (mmol/l)	5 (5.74)
Fe (mmol/l)	92 (63.51)
CRP (IU/l)	5 (28.77)
TG (mg/dl)	107 (53.76)
CHO (mg/dl)	191 (53.23)
Ur. acid (mg/dl)	6 (1.35)
TP (g/dl)	7 (0.84)
TIBC	316 (95.02)
Neutr ($10^3/\mu l$)	3.42 (1.35)
Lymph ($10^3/\mu l$)	2.05 (0.55)
Mono ($10^3/\mu l$)	0.58 (0.19)
Eos ($10^3/\mu l$)	0.17 (0.13)
Baso ($10^3/\mu l$)	0.03 (0.02)
Albu (%)	59.3 (7.04)
Glb. 1(%)	2.9 (1.31)
Glb. 2 (%)	8.8 (2.57)
Glb. (%)	10.8 (1.63)
Glb. (%)	18.4 (6.89)

3 Neural Network Ensemble

The introduced method of classifier ensemble design consists of three main steps:

- Building the pool of individual classifiers.
- Pruning the acquired pool by discarding redundant predictors.
- Using a sophisticated trained fuser to deliver the final output of the ensemble.

3.1 Creating the Pool of Classifiers

One of the most important steps in the ensemble design is preparation of the individual classifiers that are used as base models for the committee [16]. Such models should be complementary to each other, exhibiting at the same time high accuracy and high diversity. By fulfilling this requirements the ensemble classifier may outperform any single model from the pool. Additionally ensemble methods allow, in a natural way, to exploit information coming from different sources—and that is why we have decided to use this approach for our application.

It is a common knowledge that there is no single optimal approach for feature selection task and results obtained on the basis of different methods may differ significantly. Therefore instead of selecting a single best feature selection method we use several of them to reduce the dimensionality of the feature space. Then, on each of their individual output, a neural networkclassifier is build. Therefore for L used feature selection methods we construct a pool of L individual classifiers:

$$\Pi^{\Psi} = \{\Psi_1, \Psi_2, ..., \Psi_L\}. \tag{1}$$

This way we use all the reduced feature spaces in hope that they will be complementary to each other and provide a valuable contribution to the ensemble.

3.2 Ensemble Pruning

In the MCS design it is assumed that not all of the L models in Π^{Ψ} should be used as ensemble members. There are several different ways in the literature on how to select valuable members to the committee. Among them diversity measures are considered to be one of the most popular [3, 11]. They are based on the idea that ideal ensemble consists of classifiers of high individual accuracy and high diversity. Classifiers with low accuracy but high diversity will produce output of low quality, while adding similar members to the committee will only increase the computational complexity of the model.

Among diversity measures there are two major types: pairwise and non-pairwise. The former ones shows how two classifiers differ from each other, while the latter ones measure the diversity of the whole ensemble. Those two groups have different advantages and weaknesses and it is up to the researcher to select them according to his experience.

For measuring the diversity of whole ensemble we used the *entropy measure*. The highest diversity among classifiers for a particular object $x_j \in X$ is equal to the L/2 of the votes in x_j with the same value (0 or 1) and the other L [L/2] with the alternative value. Denote by $l(x_j)$ the number of classifiers that correctly recognize given sample. With this we can describe entropy-based diversity as:

$$E = \frac{1}{N}\sum_{j=1}^{N}\frac{1}{L-\lceil L/2\rceil}min\{l(z_j), L-l(z_j)\}. \quad (2)$$

E varies between 0 and 1, where 0 indicates no difference and 1 indicates the highest possible diversity.

An exhaustive search is performed to find the pruned pool of K classifiers exhibiting the highest diversity. Pruning ensemble for the proposed method is necessary as it is very likely that some of the used feature selection methods return similar feature subsets, thus leading to creation of similar classifiers. Using diversity-based pruning this situation can be dealt with, as redundant models that have no contribution to the ensemble are eliminated.

3.3 Fusion of Individual Classifiers

Classifier fusion algorithms can make decisions on the basis of class labels given by individual classifiers or they can construct new discriminant functions on the basis of individual classifier support functions. The first group includes voting algorithms [8], while the second group is based on discriminant analysis. The main form of discriminants is the posterior probability typically associated with probabilistic pattern recognition models, although outputs of neural networks or other functions whose values are used to establish the decision of the classifier (so called support functions) are also employed. The design of improved fusion classification models, especially trained fusers, is the focus of current research [19].

Assume that we have K classifiers $\Psi^{(1)}$, $\Psi^{(2)}$, ..., $\Psi^{(K)}$ in a pool after the pruning procedure. For a given object $x \in \mathcal{X}$, each individual classifier decides for class $i \in \mathcal{M} = \{1, ..., M\}$ based on the values of discriminants. Let $F^{(l)}(i, x)$ denote a function that is assigned to class i for a given value of x, and that is used by the l-th classifier $\Psi^{(l)}$. The combined classifier Ψ uses the decision rule:

$$\Psi(x) = i \quad if \quad \hat{F}(i, x) = \max_{k\in M}\hat{F}(k, x), \quad (3)$$

where

$$\hat{F}(i, x) = \sum_{l=1}^{K} w^{(l)}F^{(l)}(i, x) \quad and \quad \sum_{i=1}^{K} w^{(l)} = 1. \quad (4)$$

The weights can be set depending on the classifier and class number: weight $w^{(l)}(i)$ is assigned to the l-th classifier and the i-th class, and given classifier weights assigned to different classes may differ [20].

The used type of trained fuser we employ is a neural fuser can be implemented as a one-layer perceptron [23]. The values of support functions given by each of the base classifiers serve as input, while the output is the weighted support for each

of the classes. One perceptron fuser is constructed for each of the classes under consideration. The perceptron may be trained with any standard procedure used in neural network learning; the input weights established during the learning process are then the weights assigned to each of the base classifiers. This method is a quite fast and it exploits the advantage of well-developed training algorithm for searching the solution space.

4 Experimental Results

4.1 Set-up

Eight different feature selection algorithms were used, namely: ReliefF [22], Fast Correlation Based Filter [21], Genetic Wrapper [5], Simulated Annealing Wrapper [5], Forward Selection [5], Backward Selection [5], Quick Branch and Bound [5] and Las Vegas Incremental [5]. Therefore the pool consisted of eight different neural networks.

Neural network architecture was as follows: the number of neurons in the input layer was equal to the number of selected features, the number of output neurons was equal to the number of classes and the number of hidden neurons was equal to the half of the sum of number of neurons from the former layers, as suggested in. Quickprop algorithm was used as a training procedure.

Genetic Wrapper used population equal to 50 with 200 iterations, probability of cross-over equal to 0.7 and probability of mutation 0.3.

As reference methods we have selected most popular ensembles—Bagging, Boosting, Random Forest and Random Subspace—as they were used in our previous work [10]—in there one may also find the details of used parameters for these ensemble classifiers. Additionally we have compared our method with the single best classifier from the pool (i.e. built on the basis of the most effective feature selection algorithm), all classifiers from the pool (i.e. without the pruning procedure) and with simple majority voting (i.e. without the trained fuser). By this we can establish the influence of three steps in our proposed ensemble on the final accuracy.

The combined 5x2 CV F test [1] was carried out to asses the statistical significance of obtained results.

All experiments were carried out in the R environment [15] and computer implementations of the classification methods used were taken from dedicated packages built into the above mentioned software. This ensured that results achieved the best possible efficiency and that performance was not diminished by a bad implementation.

Table 3 Accuracy of the investigated methods [%]

NNE[1]	*Bagg*[2]	*Boost*[3]	*RandS*[4]	*RandF*[5]	*SB*[6]	*All*[7]	*MV*[8]
90.12	80.50	84.92	88.54	87.02	83.12	88.54	86.45
2,3,4,5,6,7,8	–	2,6	2,3,5,6,8	2,3,6	2	2,3,6,7	2,3,6

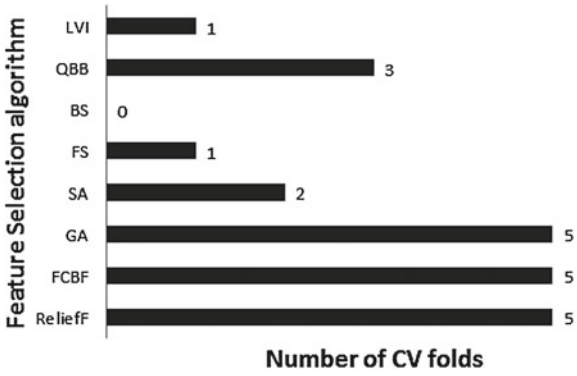

Fig. 1 Number of CV folds in which a neural network based on a given feature selection method was selected to the ensemble

4.2 Results

Results are presented in the Table 3, with small numbers under accuracy indicating from which classifiers this method is significantly better. *NNE* stands for the proposed method, *Bagg* for Bagging, *Boost* for Boosting, *RandS* for Random Subspace, *RandF* for Random Forest, *SB* for Single Best model, *All* for not pruned ensemble and *MV* for ensemble with simple majority voting scheme.

How many times a neural network classifier was selected as the ensemble member for 5x2 CV is presented in Fig. 1.

4.3 Results Discussion

The proposed neural network ensemble, based on feature selection methods, outperforms all the previously used MCS for this problem. Interesting conclusions arise from the analysis of the differences between our model and three simplified versions of it. The weakest results are returned by single best model approach, which highlights the usefulness of utilizing more than one classifier to fully exploit the outputs of feature selection methods. Second biggest accuracy boost lies in the used fuser—trained fusion of individual classifiers allows to derive an optimal linear combination of them. Finally the pruning step has the smallest but still statistically significant impact on the ensemble design. This may be due to the fact

that the trained fuser search procedure worked better with the reduced pool of classifiers.

As the ensemble based on all three steps was statistically better than models with one of the steps removed one may conclude that all of them have an impact on the quality of the proposed MCS and should not be discarded.

5 Conclusions

The presented paper shows that, despite some problems (like the fact that it is not easy to get blood test results of patients with diagnosed chronic hepatitis, infected with HCV that have no other medical conditions and are not under any medical therapy, or that blood test results which were available for research were inconsistent, i.e. some patients have one set of blood tests, while other patients have a set of other blood tests) it is possible to reach similar or even lower error level than commercial tests. Applying the proposed MCS based on a neural networks coupled with different feature selection strategies, ensemble pruning and trained fusion lead to high recognition accuracy, that outperformed esteemed off-the-shelf ensemble classification methods. Additionally we proved that each of the three steps embedded in the proposed committee design has an important impact on the quality of the final prediction and thus should not be omitted. The proposed method will be used in practice by several leading Polish hospitals.

Our future works will concentrate on the problem of imbalanced class distribution among biopsy patients and possible presence of class label noise in the data.

References

1. Alpaydin, E.: Combined 5 x 2 cv f test for comparing supervised classification learning algorithms. Neural Comput. **11**(8), 1885–1892 (1999)
2. Bedossa, P., Poynard, T.: An algorithm for the grading of activity in chronic hepatitis c. the metavir cooperative study group. Hepatology **24**, 289–293 (1996)
3. Bi, Y.: The impact of diversity on the accuracy of evidential classifier ensembles. Int. J. Approximate Reasoning **53**(4), 584–607 (2012)
4. BioPredictive. http://www.biopredictive.com/intl/physician/fibrotest-for-hcv/view?set_language=en
5. Guyon, I., Gunn, S., Nikravesh, M., Zadeh L. (eds.): Feature Extraction, Foundations and Applications. Springer, Heidelberg (2006)
6. Hansen, L.K., Salamon, P.: Neural network ensembles. IEEE Trans. Pattern Anal. Mach. Intell. **12**(10), 993–1001 (1990)
7. Ishak, K., Baptista, A., Bianchi, L., Callea, F., De Groote, J., Gudat, F., Denk, H., Desmet, V., Korb, G., MacSween, R.N., et al.: Histological grading and staging of chronic hepatitis. Hepatology **22**, 696–699 (1995)
8. Kittler, J., Alkoot, F.M.: Sum versus vote fusion in multiple classifier systems. IEEE Trans. Pattern Anal. Mach. Intell. **25**(1), 110–115 (2003)

9. Knodell, R.G., Ishak, K.G., Black, W.C., Chen, T.S., Craig, R., Kaplowitz, N., Kiernan, T.W., Wollman, J.: Formulation and application of a numerical scoring system for assessing histological activity in asymptomatic chronic active hepatitis. Hepatology **1**, 431–435 (1981)
10. Krawczyk, B., Woźniak, M., Orczyk, T., Porwik, P., Musialik, J., Błońska-Fajfrowska, B.: Classification techniques for non-invasive recognition of liver fibrosis stage. J. Med. Inform. Technol. **20**, 121–127 (2012)
11. Krawczyk, B., Woźniak, M.: Combining diverse one-class classifiers. In: Corchado, E., Snasel, V., Abraham, A., Wozniak, M., Grana, M., Cho, S-B. (eds.) Hybrid Artificial Intelligent Systems, volume 7209 of Lecture Notes in Computer Science, pp. 590–601. Springer, Berlin (2012)
12. Siemens Medical. http://www.medical.siemens.com/webapp/wcs/stores/servlet/PSGenericDisplay~q_catalogId~e_-111~a_langId~e_-111~a_pageId~e_103713~a_storeId~e_10001.htm
13. Orczyk, T., Pałys, M., Porwik, P., Musialik, J., Błońska-Fajfrowska, B.: Simple and non-invasive liver fibrosis stage prediction method. J. Med. Inform. Technol. **17**, 227–232 (2011)
14. Rokach, L.: Pattern Classification Using Ensemble Methods. Series in Machine Perception and Artificial Intelligence. World Scientific Publishing, Singapore (2010)
15. R Development Core Team: R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria (2008)
16. Ting, Kai, Wells, Jonathan, Tan, Swee, Teng, Shyh, Webb, Geoffrey: Feature-subspace aggregating: ensembles for stable and unstable learners. Mach. Learn. **82**, 375–397 (2011)
17. Wai, C.T., Greenson, J.K., Fontana, R.J., Kalbfleisch, J.D., Marrero, J.A., Conjeevaram, H.S., Lok, A.S.: A simple noninvasive index can predict both significant fibrosis and cirrhosis in patients with chronic hepatitis c. Hepatology **38**, 518–526 (2003)
18. Wolpert, DH.: The supervised learning no-free-lunch theorems. In: Proceedings of the 6th Online World Conference on Soft Computing in Industrial Applications, pp. 25–42 (2001)
19. Michal Wozniak. Experiments with trained and untrained fusers. In Emilio Corchado, Juan Corchado, and Ajith Abraham, editors, Innovations in Hybrid Intelligent Systems, volume 44 of Advances in Soft Computing, pages 144–150. Springer Berlin / Heidelberg, 2007.
20. Wozniak, Michal, Zmyslony, Marcin: Combining classifiers using trained fuser—analytical and experimental results. Neural Netw. World **13**(7), 925–934 (2010)
21. Yu, L., Liu, H.: Feature selection for high-dimensional data: a fast correlation-based filter solution. In: Proceedings of the Twentieth International Conference on Machine Learning, vol. 2, pp. 856–863 (2003)
22. Yu, L., Liu, H.: Efficient feature selection via analysis of relevance and redundancy. J. Mach. Learning Res. **5**, 1205–1224 (2004)
23. Zmyslony, M., Krawczyk, B., Wozniak, M.: Combined classifiers with neural fuser for spam detection. In: Herrero, A., Snasel, V., Abraham, A., Zelinka, I., Baruque, B., Quintin, H, Calvo, JL., Sedano, J., Corchado, E. (eds.) International Joint Conference CISIS12-ICEUTE12-SOCO12 Special Sessions, volume 189 of Advances in Intelligent Systems and Computing, pp. 245–252. Springer, Heidelberg (2012)

Cooperative and Non-cooperative Equilibrium Problems with Equilibrium Constraints: Applications in Economics and Transportation

Andrew Koh

Abstract In recent years, a plethora of multi-objective evolutionary algorithms (MOEAs) have been proposed which are able to effectively handle complex multi-objective problems. In this paper, we focus on Equilibrium Problems with Equilibrium Constraints. We show that one interpretation of the game can also be handled by MOEAs and then discuss a simple methodology to map the non-cooperative outcome to the cooperative outcome. We demonstrate our proposed methodology with examples sourced from the economics and transportation systems management literature. In doing so we suggest resulting policy implications which will be of importance to regulatory authorities.

1 Introduction

This focus of this paper is on hierarchical optimization problems. Figure 1 shows the structure of a single-leader follower game/bilevel optimization problem [8] which has attracted attention in the Evolutionary Computation community in recent years.[1] In this paper we study a generalized variant of such problems known as Equilibrium Problems with Equilibrium Constraints (EPECs) as illustrated in

The author is grateful for financial support by the Engineering and Physical Sciences Research Council of the UK under Grant EP/H021345/1.

[1] E.g. a Special Session on Bilevel Optimization was convened at the 2012 IEEE Congress on Evolutionary Computation (CEC) (June 10–15) in Brisbane, Australia.

A. Koh (✉)
Institute for Transport Studies, University of Leeds, Leeds, LS2 9JT, Leeds, UK
e-mail: a.koh@its.leeds.ac.uk
URL: http://www.its.leeds.ac.uk

V. Snášel et al. (eds.), *Soft Computing in Industrial Applications*,
Advances in Intelligent Systems and Computing 223, DOI: 10.1007/978-3-319-00930-8_3,

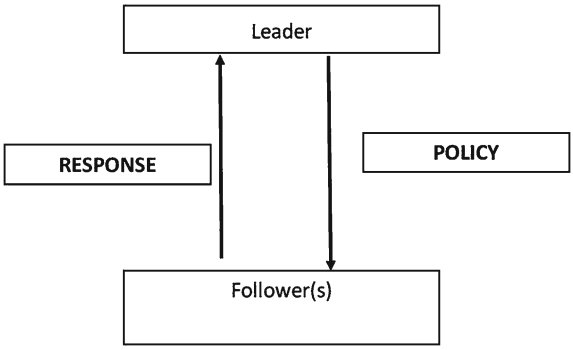

Fig. 1 Bilevel program or mathematical problem with equilibrium constraint (MPEC)

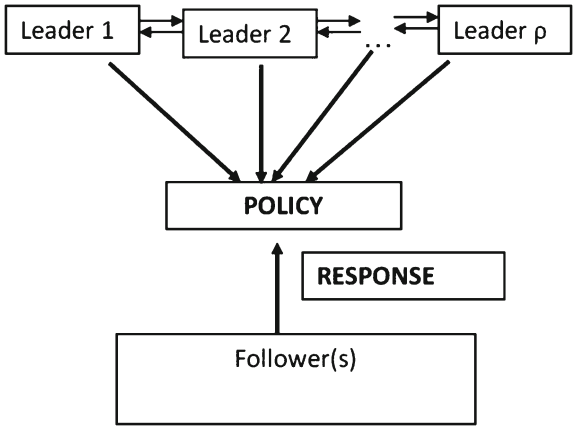

Fig. 2 Equilibrium problem with equilibrium constraint (EPEC)

Fig. 2. Both are hierarchical games where the followers take the leader's variables as given and their responses are subsequently imposed as a nonlinear binding constraint on the actions of the leader(s). The difference is that EPECs are characterized by the presence of multiple leaders.

In this multi-leader generalization of the classic Stackelberg [18] game, researchers have conjectured that there could be two possible behaviors of the leaders [11, 15]. At one extreme, they could cooperate and such a postulate leads naturally to a Multi-Objective EPEC (hereinafter termed MOPEC) [20]. At the other extreme, these leaders could act engage in a non-cooperative Nash [14] game amongst themselves thereby resulting in a Non Cooperative EPEC (NCEPEC).

Under either postulate of leader behavior, we argue that meta-heuristics offer a powerful solution methodology for EPECs that are usually tackled using tools of generalized calculus [11]. For MOPECs, population based MultiObjective Evolutionary Algorithms (MOEAs) are particularly suitable due to their inherent ability to identify multiple Pareto Optimal solutions in a single run [2]. For the latter class of NCEPECs, a Differential Evolution [16] based algorithm exploiting

a concept from [10] has been proposed in [9]. In this paper we suggest a methodology that maps the Non Cooperative outcome to the Cooperative outcome by modification of the algorithm proposed in [9].

The rest of this paper is structured thus. We give an overview of notation used in this paper in the next section. In Sect. 3 fundamental notions of Multi-Objective optimization are reviewed before an evolutionary algorithm for solving MO problems is given. Section 4 outlines a solution algorithm for NCEPECs. Section 5 discusses a simple method to map the Non-Cooperative outcome to the Cooperative outcome. Section 6 illustrates the concepts with numerical examples. Section 7 summarizes and provides directions for further research.

2 Notation

In this paper we consider ρ-person games. Focusing on the leaders, each game is defined by a tuple $\{N, X_i, U_i\}$ where N is the set of leaders $\{1, 2, \ldots, \rho\}$, X_i is the strategy/action space for leader $i, i \in N$ and U_i is the payoff function (or reward), $U_i : \mathbb{R}^\rho \to \mathbb{R}^1$, that a leader gets by playing an action/strategy, dependent on the actions which all others take. The collective action of all leaders, often referred to as a strategy profile, is denoted by $\boldsymbol{x} = [x_1, \ldots, x_i, \ldots, x_\rho]^\mathsf{T}$. It is convenient to write x_{-i} when referring to the strategies of every leader *excluding* the leader i i.e. $x_{-i} = [x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_\rho]^\mathsf{T}$. With a slight abuse of notation, we also write $\boldsymbol{x} = [x_i, x_{-i}]^\mathsf{T}$. Note that $[x_i, x_{-i}]$ does not mean that the components of $\boldsymbol{x}$ are reordered, so that strategies of leader i becomes the first block. Unless otherwise specified, all vectors are assumed to be column vectors.

The response of the followers that affects the actions of the leaders is assumed to take the form of a Variational Inequality (VI) constraint that defines equilibrium in some parametric system. Following [12] we assume that the solution of this VI exists and is unique for a given a vector of the leaders' strategies.

3 Cooperative EPECs (MOPECS)

A generic MOPEC is shown in Eq. 1. Except for the variational inequality constraint, this problem takes the form of a generic multi objective optimization problem conventionally handled by MOEAs [1, 2].

$$\text{Program} \quad \text{MOPEC} \quad \left\{ \max_{\boldsymbol{x} \in X} (U_1(\boldsymbol{x}, \boldsymbol{y}), \ldots, U_\rho(\boldsymbol{x}, \boldsymbol{y}))^\mathsf{T} \right. \tag{1a}$$

where for given $\boldsymbol{x}$, $\boldsymbol{y}$ is the unique solution of the Variational Inequality in 1b:

$$L(\boldsymbol{x}, \boldsymbol{y})^\mathsf{T}(\boldsymbol{y} - \boldsymbol{y}^*) \geq 0, \forall \boldsymbol{y} \in \Upsilon(\boldsymbol{x}) \tag{1b}$$

MOEAs apply stochastic operators to a parent population to evolve a fitter child population to solve multi-objective problems. During the selection phase, a comparison is made between a chromosome $\boldsymbol{a}$ from the parent population and a chromosome $\boldsymbol{b}$ from the child population on the basis of fitness and the weaker of the two is discarded. Since one of the tasks of an MOEA is to identify the entire Pareto front [2], fitness is assigned based on Pareto Domination: $\boldsymbol{a}$ Pareto Dominates $\boldsymbol{b}$ if $\boldsymbol{a}$ is no worse off than $\boldsymbol{b}$ in all objectives *and* $\boldsymbol{a}$ is strictly better than $\boldsymbol{b}$ in at least one objective ([2], Definition 2.5, pp. 28).

Algorithm 1 Multi-Objective Self Adaptive Differential Evolution (MOSADE) [5]

1. Evaluate initial population P of $|P|$ random individuals.
2. Set archive $\mathcal{A}$ to $\emptyset$
3. While stopping criterion not met, do:
   For each individual $P_i, i \in \{1, \ldots, |P|\}$ repeat:
   (a) Use DE to create candidate C_i from parent P_i.
   (b) Evaluate C_i by solving lower level VI 1b
   (c) If P_i dominates C_i, discard C_i else go to Step 4
4. Compare C_i with each member of $\mathcal{A}$,
   (a) if maximum size of $\mathcal{A}$ reached, choose between C_i or each member of $\mathcal{A}$ depending on which occupies the less crowded region of function space
   (b) if C_i dominates any $\mathcal{A}$, remove the member of $\mathcal{A}$ so dominated, accept C_i into $\mathcal{A}$
   (c) if C_i is dominated by any $\mathcal{A}$, reject C_i
5. Update DE control parameters as described in [5]

Algorithm 1 outlines the Multi-Objective Self Adaptive Differential Evolution (MOSADE) Algorithm [5] that was used to generate the Pareto Fronts for the MOPECs to be described in Sect. 6. MOSADE uses an archive to store solutions as they are discovered during the search process. To evaluate the candidate, it is necessary to solve the lower level VI problem in Eq. 1b to maintain the leader-follower paradigm implicit in such hierarchical optimization problems [8].

4 Non Cooperative EPECs (NCEPECS)

In the NCEPEC each leader i treats his competitor's strategic variables as exogenous when maximizing his payoff as in Eq. 2.

$$\forall i \in N, \quad \text{Player } i \text{ solves: } \left\{ \max_{\boldsymbol{x} \in X} U_i(\boldsymbol{x}, \boldsymbol{y}) \right. \tag{2}$$

where for given $\boldsymbol{x}$, $\boldsymbol{y}$ is the unique solution of the Variational Inequality in 1b.

It can be shown that the solution of Eq. 2, if one exists, is a Nash Equilibrium (NE) which is obtained when the condition in Eq. 3 is satisfied [14].