# INTELLIGENT SYSTEMS:

# TECHNIQUES AND APPLICATIONS

Edited by: Evor Hines, Mark Leeson, Manel Martínez-Ramón
Matteo Pardo, Eduard Llobet, Daciana Iliescu, Jianhua Yang

SHAKER
PUBLISHING
St. Maartenslaan 26, 6221 AX Maastricht, The Netherlands.

Printed in The Netherlands.

ISBN: 978-90-423-0345-4

Shaker Publishing BV
St. Maartenslaan 26
6221 AX Maastricht
Tel: 043-3500424
Fax: 043-3255090
http://www.shaker.nl

# INTELLIGENT SYSTEMS:

# TECHNIQUES AND APPLICATIONS

Dr Evor L. Hines, School of Engineering, University of Warwick; Coventry, CV4 7AL, UK; Email: E.L.Hines@Warwick.ac.uk; Website: http://www.eng.warwick.ac.uk/staff/elh/

Dr Mark S. Leeson, School of Engineering, University of Warwick, Coventry, CV4 7AL, UK; Email: Mark.Leeson@warwick.ac.uk

Dr Manel Martínez-Ramón, Universidad Carlos III de Madrid; Avda de la Universidad, 30; 28911 Leganés-Madrid, Spain; Email: manel@ieee.org

Dr Matteo Pardo, Sensor Laboratory, CNR – INFM; Via Valotti 9 25133 Brescia, Italy; Email:pardo@ing.unibs.it

Dr Eduard Llobet, Department of Electronics, Universitat Rovira i Virgili, Avda. Països Catalans, 26, Campus Sescelades, 43007 Tarragona, Spain; Email: eduard.llobet@urv.cat

Dr Daciana D. Iliescu, School of Engineering, University of Warwick, Coventry, CV4 7AL, UK; Email: D.D.Iliescu@warwick.ac.uk

Mr Jianhua Yang, School of Engineering, University of Warwick, Coventry, CV4 7AL, UK; Email: J.Yang.3@warwick.ac.uk

# Preface

The area of Intelligent Systems (ISs) has expanded phenomenally over the years since the 1940s; both in terms of the range of techniques and also in terms of the number of applications wherein they have often provided a competitive edge when compared with others approaches. IS includes a range of techniques that work synergistically and provides, in one form or another, flexible data/information processing capabilities for handling real life situations. IS, unlike conventional techniques, can exploit the tolerance for imprecision, uncertainty/ambiguities, approximate reasoning and partial truth in order to achieve tractability, robustness, and low cost solutions.

The techniques are in general based on biologically inspired strategies for solving problems. At the current time the major categories of IS include neural networks (NNs), fuzzy logic/systems (FL/Ss), evolutionary computation/algorithms (EC/As) (including genetic algorithms (GAs), genetic programming (GP), evolutionary strategies (ES)), support vector machines (SVM), particle swarm optimization (PSO), memetic algorithms (MAs), and ant colony optimization (ACO). In addition hybrid combinations also play a mayor role. These include neuro-fuzzy, neuro-genetic, fuzzy-genetic systems and so on.

These techniques became more and more necessary and popular to meet needs such as: 1) Handling large datasets which are very complex and comprise various forms of uncertainty, in a robust and computationally efficient manner. 2) Much of the data are inherently uncertain and noisy, thus making FLS a natural choice for processing them. 3) The learning capability of NNs, supervised, unsupervised or hybrids can be utilized effectively when extracting patterns from large datasets. This is especially true in data-rich environments where the data may be mined. 4) Many of the tasks involve search and optimization of different criteria (such as energy, alignment score and overlap strength), while requiring robust, fast and close approximate solutions. 5) Evolutionary and other search algorithms such as ACO, PSO provide effective techniques to search and explore very large and multi-modal solution spaces. 6) Additionally, many of the

requirements involve multiple, and often conflicting, objectives thus lending themselves to the application of multi-objective optimization algorithms such as multi-objective GAs.

The focus of this book is on the application of IS Techniques to solve a range of problems in the areas of Olfaction, Engineering, Telecommunications, Antennas and Medical diagnosis/imagining.

This book brings together research articles by eminent experts and active practitioners; reporting recent advances in integrating IS techniques, either individually or in an hybridized manner, for problem solving in order to extract more and more meaningful information and insights. Typically the data to be considered for analysis include: sequence data and structure data, and data types that are typically complex in nature requiring advanced methods to deal with them. Characteristics of the techniques presented here include the use of domain-specific knowledge for reducing the search space, dealing with uncertainty, partial truth and imprecision, efficient linear and/or sub-linear scalability, incremental approaches to knowledge discovery, and an increased level and intelligence of interactivity with human experts and decision makers.

The book has five sections: Olfaction, Engineering, Telecommunications, Antennas and Medical diagnosis/imaging. Let us now consider each one in turn.

## Section I – Olfaction

Odour classification systems used in machine olfaction, often termed electronic noses (ENs), have been gaining favour in a wide range of industrial applications. Recent advances in the field of ENs have led to new developments in sensor design, pre-processing for feature extraction and data processing techniques. As a result, the user of EN systems is provided with an increased amount of information for the discrimination of odours using multi-sensor arrays. In this section, EN applications for food analysis and gas sensing are blended with examples of novel processing and analysis methods.

Chapter 1, by Yang et al., applies a GA based sensor selection technique, known as GNMM (Genetic Neural Mathematical Method) to EN data. GNMM uses multilayer percetrons (MLPs) as the classification engine but separates the optimal sensor subset selection process from the final classification. By utilizing GAs to identify optimal sensors that are subsequently fed as inputs to MLP

classifiers, GNMM simplifies the MLP structure and makes the final classification process more efficient. The effectiveness of GNMM is demonstrated by its application to case study data which has previously been explored in the literature via alternative methods. The results generated by GNMM are generally superior to previous methods demonstrating the effectiveness of the methodology.

The application of GA and NN methods to sensor data, here to virgin olive oil, is the focus of the second chapter by García-González and Aparicio. Although sensing mechanisms and the physical-chemical phenomena that cause sensor responses have been extensively investigated, interpretation of sensor signals is arduous when dealing with the complex mixtures of compounds such as food aromas. This is because there are many interactions between volatiles and the sensitive sensor material producing a single response involving unknown synergic and masking effects. Here, an experimental procedure is proposed to determine the individual contribution of volatile compounds in the sensor response, illustrated with the examples of the aroma of virgin olive oil and metal oxide sensors. Following discussion of preliminary studies, the use of a sensor array as a non-destructive detector in parallel with the flame detector in a gas chromatography (GC-EN) is described. Gas chromatography facilitates study of the relationship between volatile compounds and sensor responses. The chapter also contains a real application of the metal oxide semiconductor sensors to quality assessment.

In the third chapter, from Vergara and Llobet, the analysis of single gases and multi-component mixtures using semiconductor gas sensor arrays and the tools of ISs is investigated. Despite the extensive literature on this subject, important problems such as lack sensor selectivity and a lack of reproducibility remain as research topics. Recent progress in methods to systematically design the use of sensor dynamics has produced significant steps forward in recent years. In this chapter, an approach adapted from system identification that selects estimates of the impulse response of the sensor-gas system from the spectral components present is presented. The selection of the former results in the determination of the optimal temperature modulating frequencies because of the direct relationship between spectral components and modulating frequencies. A practical application is included to reinforce the utility of the method, which is generally applicable.

Chapter 4, by Pardo and De Vito, takes the application of generalized linear discrimination machines or support vector machines (SVMs) as its centre of attention. SVMs have recently been employed in the EN community for pattern analysis in both classification and regression problems. This contribution provides an overview of SVM theory, a review of artificial olfaction and descriptions of interesting classification and regression case histories. In common with any other machine learning technique, SVM requires 'hyper parameter selection'. In the applications, hyper parameter tuning permits exploration of under- and over-fitting, which form the principal limitations of non-parametric learning techniques.

Chapter 5, from Vergara, Vilanova and Llobet, concludes the section with consideration of dynamic methods for improving the performance of semiconductor gas sensors. The lack of selectivity and the response drift present in semiconductor gas sensors have limited their spectrum of applications to qualitative gas analysis. IS offers tools to improve the gas sensing performance of well-known materials. Sensors in the presence of gas mixtures can be considered as dynamic systems, enabling the employment of methods that are used to study the properties of dynamic systems. This chapter reviews some of the latest developments in the field of gas sensor dynamics, and illustrates their performance for various gas analysis and electronic nose applications.

## Section II – Engineering

The considerable breath of application areas in which IS methods are applicable is particularly demonstrated by this section. The chapters develop techniques that range from river flow to yield prediction, incorporating image and signal processing in addition to classification techniques.

Chapter 6, from Yang et al., presents the application of a new technique to the prediction of the longitudinal dispersion coefficient in rivers. The longitudinal dispersion coefficient is a key variable in describing longitudinal transport in a river. In recent years, NNs have become established as tools for environmental modellers, overcoming some of the difficulties associated with traditional statistical approaches. Little attention has been paid to date, however, to the task of selecting the most appropriate NN inputs. In this chapter, the use of a GA based variable selection technique in applied to longitudinal dispersion

coefficient prediction. The GNMM of chapter one is adapted to this purpose with extremely good results, which are favourably compared with principal component analysis (PCA) and the use of a self-organizing map (SOM).

Remote sensing offers the seventh chapter, from Camps-Valls et al., in particular the classification of the images obtained in this area. Remote sensing of the Earth is an active research field of science concerned with acquisition of data by airborne or satellite sensors, and its subsequent transmission and processing. Remote sensing imaging instruments acquire information in several spectral bands that ultimately characterize the pixel in the acquired image. Here the latest developments based on kernel methods for remote sensing image classification are summarised. In particular, the work focuses on various types of SVM, kernel-based fusion of heterogeneous information, multi-temporal classification kernel schemes and the specific problem of change detection with kernels. Methods are analyzed theoretically and through numerical and visual inspection of the classification maps in a number of real scenarios.

There follows a chapter from Sloper, Miotto and Hines, which considers the use of computational intelligence in the context of a large data acquisition project at the European CERN facility. The ATLAS particle physics experiment encompasses a trigger and data acquisition system (TDAQ), which is large, heterogeneous and distributed. This chapter presents the design and development an automatic system for detecting and appropriately dealing with the different types of errors, which may occur in TDAQ. The size and complexity of the system introduces considerable uncertainty and “noise” in the form of irrelevant data or incomplete information. The outputs from fuzzy C-means (FCM), PCA and SOM analysis feed MLP based NNs that are trained to recognize error scenarios with extremely reliable results.

In chapter nine, Li et al. consider the application of unsupervised learning to failure modes in fibre reinforced plastic. The different types of damage occurring in composite materials may be discriminated by analysing acoustic emission (AE). The problem is then the decision as to which failure mechanism produces which AE signature. Here this question is elegantly addressed by obtaining the two most significant frequency components from the AE signal via a fast Fourier transform (FFT). The data are then analyzed using a SOM and FCM, giving successful classification and the identification of plausible physical categories.

Section III – Telecommunications

This section illustrates the range of application areas where IS techniques have been implemented in telecommunications applications. The need to allocate wavelengths or frequencies with constraints leads naturally to difficult computational problems, ideal for stochastic optimisation. In this part, applications in the radio frequency and optical domains are presented since the appropriate techniques map easily from one physical layer to another with adjustment of the fitness function and constraints.

Chapter 10, by An, Hines and Leeson, demonstrates the use of GA and fuzzy logic (FL) methods for cellular radio network optimisation. A cellular radio typically comprises mobile stations and fixed based stations. To achieve high quality mobile communications the coverage of the latter is limited to cells in small geographic areas. This concept produces two technically challenging engineering problems, namely that of channel allocation and call blocking need to be addressed for optimum system performance. The former is formulated as a signal-to-interference ratio (SIR) maximization problem that is implemented using GAs, whereas the latter is tackled using FL with a simple channel borrowing technique. It is shown that the GA optimized channel allocation scheme consistently optimises the network and that the FL system produces a substantial reduction in blocked calls.

Chapter 11, by Kavian et al., addresses the design of resilient optical core networks using GAs. A key issue in such networks is the assignment of wavelengths to meet the routing needs of the traffic to be conveyed. This is a difficult mathematical problem and thus ideal for GA application, and the inclusion of a back up transmission path to cover fibre failure increases the complexity. In the chapter, lightpaths are encoded into chromosomes made up of a fixed number of genes equal to the number of entries in the traffic demand matrix. Each gene represents one valid path, producing a coding via a variable length binary string. In each new generation, each member of the population represents a set of valid but possibly incompatible paths and some must be discarded. The best paths according to fitness function are assigned the minimum number of wavelengths within the problem constraints. The method is shown to be efficient and able to design survivable optical mesh transport networks.

In the twelfth chapter, from Van Vaerenbergh and Santamaría, the focus on radio frequency transmission, in particular multiple-input multiple-output or

MIMO systems. MIMO systems offer significant increases in spectral efficiency over traditional single-input single-output (SISO) transmission. A number of computationally efficient algorithms for time-invariant MIMO symbol detection exist but their application in time-varying environments is difficult because they need perfect receiver channel state information. The clustering approach to this problem exploits the fact that when channels are excited by signals from a finite alphabet, the noisy observations tend to be clustered around a finite number of so-called channel states. Once these have been estimated (typically by applying any supervised clustering technique based on a known training sequence), channel equalization reduces to a classification problem, which can efficiently be solved using a radial basis function network (RBFN). The time-varying MIMO case produces non-convex clusters shapes that overlap, meaning that conventional clustering algorithms fail. Here, a new version of a spectral clustering technique is utilised to deliver better results than state-of-the-art MIMO decoding techniques for time-correlated channels.

Chapter 13 from Ren et al. considers the routing of traffic in ad hoc networking. A mobile ad hoc network (MANET) consists of a collection of wireless nodes, which dynamically form a temporary network, without using any existing infrastructure of wired networks or centralized administration. Here, the use of swarm intelligence (SI) to tackle traffic routing is the central thrust. Following a comparison between existing methods, SILS, a hybrid routing protocol for MANETs based on SI is presented. The lack of link stability, arising from the constant node movement, is identified as key to the optimisation of a routing protocol. Preliminary results show that SILS performs well in comparison with the popular routing techniques employed today. Further work to optimise the parameters of SILS and to introduce GA assisted methods is discussed.

Section IV – Antennas

The design of the antennas needed for modern communication systems offers substantial opportunity to apply IS methods. Traditionally, an approximate design has been produced and then refinements made by empirical methods and laboratory tests. IS offers the prospect of producing designs that are close to finished, requiring only minimal final adjustments in the field. A flavour of the range of work in this field is presented within this section.

The synthesis of antenna arrays provides the fourteenth chapter, by González-Ayestarán and Las-Heras. They address the design problem of calculating the feeding values that must be applied to the ports of an antenna array in order to obtain a specified radiation pattern, known as synthesis. This problem is typically solved using antenna models that do not take into account the real properties of the elements of the array, assuming a certain degree of error in the solution because of the complexity of developing accurate models able to include the real antenna behaviour in the different synthesis schemes. Support vector regression (SVR) can be used to develop such accurate models from experimental or simulated data accounting for all the real properties of the elements of the array. Furthermore, these SVR-based models can be easily included in different synthesis schemes or in novel SVR-based schemes able to solve synthesis problems with specifications given in different manners. Support vector multi-regression (SVMR) can also be used to develop efficiently 3D models of the arrays. SVR is shown to be a flexible and powerful tool for the modelling and synthesis of array antennas.

In Chapter 15, Randazzo and Pastorino review the use of SVMs for the estimation of the directions of arrival of electromagnetic waves. Determining the incident angles of waves impinging on receiving antennas is of fundamental importance not only for the development of smart antennas for wireless communications but also in many other applications related to imaging and security. The main advantage of the SVM based approach is related to the ability of the system to determine incident angles not included in the original training set. This property is a consequence of the excellent generalization capability of SVMs, which have been proven in several applications related to antennas and in other areas of applied electromagnetics. The chapter presents a brief review of the mathematical formulation of the estimation algorithm and then discusses recent relevant results from the literature.

Kernel Methods for smart antennas by Martínez-Ramón et al. are the focus of this the sixteenth chapter. Adaptive array processing improves performance of receivers in time varying scenarios. Linear methods provide easy to implement, adaptive and low computational burden solutions. For that reason, methods such as least mean squares (LMS) and recursive least squares (RLS) are widely used in practice. Nonlinear algorithms are usually impractical due to their high complexity and because they often need large training datasets for their parameters to be adjusted. Also, it may be hard to obtain adaptive versions of

them. Kernel methods, together with SVMs provide nonlinear, low complex machines that have good generalization performance. Here we present a nonlinear and complex-valued SVM that is applied to antenna array processing, and its adaptive version, and we compare it with the standard methods in several multiuser varying scenarios.

Section V – Medical diagnosis/imaging

Medical data are well known to be noisy in the sense that there are often incomplete records, wrongly entered numbers or some other outcome of the process of human interaction in a hospital setting. Historically, health professionals have applied medical expertise to make diagnostic decisions in the presence of imperfect information. Modern IS methods are now available to aid this process and some of these are presented in this section.

Chapter 17, by Empson et al., outlines an approach to find optimal clusters in patients who have non-specific low back pain (NSLBP). The results of SOM training using clinical data from a study are presented. Clustering is then performed using both the k-means (KM) and FCM clustering algorithms to enable extraction of the patient subgroup memberships from the SOM. The clustering results are shown to have better agreement than when the data is clustered directly, and the clusters found fall into clinically recognizable groups.

The eighteenth chapter, from Alonso-Atienza et al., tackles the detection of ventricular fibrillation in its early stages to substantially reduce the risk of sudden cardiac death. This, in turn, allows ‘recovering shock therapy’ to be administered in a timely manner. A number of classification methods have previously been proposed, to characterise ventricular fibrillation by temporal, spectral or time-frequency measurements. There is thus a large amount of information about arrhythmia conveyed by the variables present. Here, special attention is given to the feature selection problem to discover the most informative features to allow effective pathology discrimination. There is a summary of filter and wrapper methods from the literature, and a comparative analysis of several representative databases. The results obtained are good, and show that significant information about ventricular fibrillation may be retrieved by a moderate number of features. Further work on increasing the detection accuracy of these schemes, and to establish a clinical correspondence of these relevant variables is also discussed.

The provision of a decision support tool for the anaesthetist, and modelling the knowledge of the anaesthetist in evaluating patients during pre-operative assessment are the subjects of the final chapter, by Folland, Hines and Morgan. An anaesthetist must decide prior to a surgical operation whether the general physical state of the patient is robust enough for the stresses that are a natural part of surgery. The work is based on an evolutionary fuzzy mixture model (EFMM) based system that encapsulates the decision strategy of an anaesthetist in determining patient suitability for surgical anaesthesia. Methods for automatically identifying cardio-respiratory pathology and for fusing this information with other patient information are demonstrated. It is shown that a genetic program can evolve the fuzzy system such that classification accuracies in excess of ninety per cent can be attained.

In summary, the chapters on IS techniques and their application to a range of problems provides a representative selection of the available methods and their effectiveness in real domains. While the field is rapidly evolving with the availability of new data and new tools, these chapters clearly indicate the importance and potential benefit of synergistically combining the potentials of the various techniques in order to be better able to face new challenges. The book will be useful to graduate students and researchers in computer science, bioinformatics, computational and molecular biology, electrical engineering, system science and information technology; both as text and reference book for some parts of the curriculum. Moreover, researchers and practitioners in diverse industries, and in research and development laboratories will also benefit from this book.

We take this opportunity to thank all the authors for contributing chapters related to their current research work that provide the state of the art in advanced methods for solving such problems.

*Evor Hines*
*Mark Leeson*
*Manel Martínez-Ramón*
*Matteo Pardo*
*Eduard Llobet*
*Daciana Iliescu*
*Jianhua Yang*
*June, 2008*

# ACKNOWLEDGEMENTS

We would like to thank our respective universities for their encouragement and support to produce this book. We would also like to thank Jianhua Yang for his tireless efforts in the design work on the book cover and in the typesetting of this manuscript; and to Vera van Ratingen of Shaker Publishing BV for her support in the production of this book.

In addition last but not least we would like to thank our respective families and friends for their patience, understanding and support during the preparation of this manuscript.

*Evor Hines*
*Mark Leeson*
*Manel Martínez-Ramón*
*Matteo Pardo*
*Eduard Llobet*
*Daciana Iliescu*
*Jianhua Yang*

# Table of Contents

## Section II: Intelligent Systems and Engineering



## Section III: Intelligent Systems and Telecommunications

## Section IV: Intelligent Systems and Antennas



## Section V: Intelligent Systems and Medical diagnosis/imaging