

INTERNATIONAL JOURNAL
of
COMPUTERS, COMMUNICATIONS & CONTROL

With Emphasis on the Integration of Three Technologies

Year: 2007 Volume: II Number: 2 (April-June)

CCC Publications

www.journal.univagora.ro

EDITORIAL ORGANIZATION

Editor-in-Chief

Prof. Florin-Gheorghe Filip, *Member of the Romanian Academy*
Romanian Academy, 125, Calea Victoriei
010071 Bucharest-1, Romania, ffilip@acad.ro

Executive Editor

Dr. Ioan Dziţac
Agora University
idzitac@univagora.ro

Managing Editor

Prof. Mişu-Jan Manolescu
Agora University
rectorat@univagora.ro

Editorial secretary

Horea Oros
University of Oradea
horea.oros@gmail.com

Emma Văleanu
Agora University
evaleanu@univagora.ro

Publisher & Editorial Office

CCC Publications, Agora University
Piata Tineretului 8, Oradea, jud. Bihor, Romania, Zip Code 410526
Tel: +40 259 427 398, Fax: +40 259 434 925, E-mail: ccc@univagora.ro
Website: www.journal.univagora.ro
ISSN 1841-9836 (print version), 1841-9844 (online version)

EDITORIAL BOARD

Prof. Pierre Borne

Ecole Centrale de Lille
Cit  Scientifique-BP 48
Villeneuve d'Ascq Cedex, F 59651, France
p.borne@ec-lille.fr

Prof. Antonio Di Nola

Department of Mathematics and Information Sciences
Universit  degli Studi di Salerno
Salerno, Via Ponte Don Melillo 84084 Fisciano, Italy
dinola@cds.unina.it

Prof. Constantin Gaiandric

Institute of Mathematics of
Moldavian Academy of Sciences
Kishinev, 277028, Academiei 5, Republic of Moldova
gaiandric@math.md

Prof. Kaoru Hirota

Hirota Lab. Dept. C.I. & S.S.
Tokyo Institute of Technology,
G3-49, 4259 Nagatsuta, Midori-ku, 226-8502, Japan
hirota@hrt.dis.titech.ac.jp

Prof. Shimon Y. Nof

School of Industrial Engineering
Purdue University
Grissom Hall, West Lafayette, IN 47907, U.S.A.
nof@purdue.edu

Prof. Mario de J. P rez Jim nez

Dept. of CS and Artificial Intelligence
University of Seville
Sevilla, Avda. Reina Mercedes s/n, 41012, Spain
marper@us.es

Prof. Dr. Petre Dini

Cisco
170 West Tasman Drive
San Jose, CA 95134, USA
pdini@cisco.com

Prof.  mer Egecioglu

Department of Computer Science
University of California
Santa Barbara, CA 93106-5110, U.S.A
omer@cs.ucsb.edu

Prof. Xiao-Shan Gao

Academy of Mathematics and System Sciences
Academia Sinica
Beijing 100080, China
xgao@mmrc.iss.ac.cn

Prof. George Metakides

University of Patras
University Campus
Patras 26 504, Greece
george@metakides.net

Dr. Gheorghe P un

Institute of Mathematics
of the Romanian Academy
Bucharest, PO Box 1-764, 70700, Romania
gpaun@us.es

Prof. Dana Petcu

Computer Science Department
Western University of Timisoara
V.Parvan 4, 300223 Timisoara, Romania
petcu@info.uvt.ro

Prof. Radu Popescu-Zeletin

Fraunhofer Institute for Open Communication Systems
Technical University Berlin
Germany
rpz@cs.tu-berlin.de

Prof. Athanasios D. Styliadis

Alexander Institute of Technology
Thessaloniki
Agiou Panteleimona 24, 551 33, Thessaloniki, Greece
styl@it.teithe.gr

Prof. Horia-Nicolai Teodorescu

Faculty of Electronics and Telecommunications
Technical University "Gh. Asachi" Iasi
Iasi, Bd. Carol I 11, 700506, Romania
hteodor@etc.tuiasi.ro

Prof. Imre J. Rudas

Institute of Intelligent Engineering Systems
Budapest Tech
Budapest, Bécsi út 96/B, H-1034, Hungary
rudas@bmf.hu

Dr. Gheorghe Tecuci

Center for Artificial Intelligence
George Mason University
University Drive 4440, VA 22030-4444, U.S.A.
tecuci@gmu.edu

Dr. Dan Tufiş

Research Institute for Artificial Intelligence
of the Romanian Academy
Bucharest, "13 Septembrie" 13, 050711, Romania
tufis@racai.ro

This publication is co-financed by:

1. Agora University
2. The Romanian Ministry of Education and Research / The National Authority for Scientific Research

CCC Publications, powered by Agora University Publishing House, currently publishes the "International Journal of Computers, Communications & Control" and its scope is to publish scientific literature (journals, books, monographs and conference proceedings) in the field of Computers, Communications and Control.

IJCCC is listed in the Directory of Open Access Journals <http://www.doaj.org>

IJCCC is indexed in The Collection of Computer Science Bibliographies

<http://liinwww.ira.uka.de/bibliography/Misc/ijccc.html>,

Open J-Gate <http://www.openj-gate.com> and in

FIZ KARLSRUHE's informatics portal www.io-port.net/index_eng.html

Copyright © 2006-2007 by CCC Publications

Contents

Variable Selection and Grouping in a Paper Machine Application	
By Timo Ahola, Esko Juuso, Kauko Leiviskä	111
Modelling of Wastewater Treatment Plant for Monitoring and Control Purposes by State – Space Wavelet Networks	
By Adam Borowa, Mietek A. Brdys, Krzysztof Mazur	121
Predictive Control of a Wastewater Treatment Process	
By Sergiu Caraman, Mihaela Sbarciog, Marian Barbu	132
Self-Organizing Maps for Analysis of Expandable Polystyrene Batch Process	
By Mikko Heikkinen, Ville Nurminen, Yrjö Hiltunen	143
Blind Steganalysis: Estimation of Hidden Message Length	
By Sanjay Kumar Jena, G.V.V. Krishna	149
An Enhancement of the Random Sequence 3-Level Obfuscated Algorithm for Protecting Agents Against Malicious Hosts	
By Kamalrulnizam Abu Bakar, Bernard S. Doherty	159
Ant systems & Local Search Optimization for flexible Job Shop Scheduling Production	
By Nouredine Liouane, Ihsen Saad, Slim Hammadi, Pierre Borne	174
Fault Detection for Large Scale Systems Using Dynamic Principal Components Analysis with Adaptation	
By Jesús Mina, Cristina Verde	185
Robust PID Decentralized Controller Design Using LMI	
By Danica Rosinová, Vojtech Veselý	195
Author index	205

Variable Selection and Grouping in a Paper Machine Application

Timo Ahola, Esko Juuso, Kauko Leiviskä

Abstract: This paper describes the possibilities of variable selection in large-scale industrial systems. It introduces knowledge-based, data-based and model-based methods for this purpose. As an example, Case-Based Reasoning application for the evaluation of the web break sensitivity in a paper machine is introduced. The application uses Linguistic Equations approach and basic Fuzzy Logic. The indicator combines the information of on-line measurements with expert knowledge and provides a continuous indication of the break sensitivity. The web break sensitivity defines the current operating situation at the paper mill and gives new information to the operators. Together with information of the most important variables this prediction gives operators enough time to react to the changing operating situation.

Keywords: variable selection, grouping, paper machine, web breaks

1 Introduction

Data-driven modelling requires always variable selection or grouping. In small systems, expert knowledge gives a clear basis for the variable selection since possible interactions and causal effects are known fairly well. In these cases, few modelling alternatives can be compared interactively. Variable selection becomes important when the number of variables increases, especially when normal process data is used. As a model should include a reasonable number of variables, a modular approach based on variable grouping provides a better process insight, which makes the model assessment easier.

In practical cases, variable selection is necessary either because it is computationally infeasible to use all available variables, or because of estimation problems when limited data samples with a large number of variables are present.

Variable grouping means finding feasible groups and combinations of variables for modelling. It is closely connected to data clustering since the interactions can depend on the operating area. In large-scale systems, the number of possible variable combinations becomes easily very large, e.g. the case models of the web break indicator included originally 24 variables, which mean 2024 alternative three variable combinations. The newest version has 72 variables leading to 59,640 three variable groups, 1,028,790 four variable groups and 13,991,544 five variable groups. Most of these alternatives are useless, and therefore, methods for selecting reasonable variables for modelling are crucial.

There is a lot of recent literature on variable selection and both model and data-based techniques are in use. Spectroscopic data, multi-sensor systems, multivariate analysis and modelling of large-scale systems seem to require efficient methods for variable selection. Four different methods for variable selection – genetic algorithm, iterative PLS, uninformative variable elimination by PLS and interactive variable selection for PLS - in partial least square (PLS) regression are studied and compared to a calibration made with manually selected wavelengths in [1]. The application is NIR analysis of pharmaceutical tablets. It has been found that multiresolution analysis (Haar wavelets) pre-processing before variable selection leads to simpler models with lower errors than single-wavelength selection in NIR data [2].

Wavelength selection for process monitoring has also been done using genetic algorithms (GA) coupled with a curve resolution method (OPA) [3]. Variable selection is also an important topic in using multiway methods in modelling NIR spectra from a pharmaceutical batch process [4]. NIR analysis of sugar cane juice has utilized partial least squares (PLS) pruning for variable selection [5]. UV-VIS and NIR spectrometry of oils takes advantage of the successive projections algorithm (SPA) in large-scale variable selection [6].

Quantitative structure–activity relationship (QSAR) studies require also sophisticated methods for variable selection. There is a report on applying multi-objective genetic programming (GP) to the HEPT data and constructing the nonlinear QSAR model using counter-propagation (CP) neural network with the selected variables [7]. Particle swarm optimizer (PSO) applies for the same purpose and the comparison with GP is in [8]. Norinder [9] reports also on the use of support vector machine (SVM) in QSAR. Statistical parametric mapping (SPM), relying on the general linear model and classical hypothesis testing, is a benchmark tool for assessing human brain activity using data from fMRI experiments [10]. Prediction-based variable selection has reported to give 82 % success rate in Quantitative Structure–Property Relationship models (QSPR) based on in vivo blood–brain permeation data [11].

In multi-sensor systems, variable selection problem originates from two reasons: the high dimensionality in the data used is due to a high number of sensors or many features extracted, or both. Fuzzy ARTMAP classifier analyses the results from a 12-element gas sensor array [12, 13]. Fast wavelet transform is useful in feature selection before calibration in stripping voltammetry [14].

Principal component analysis (PCA) is a well-known method for variable selection. Testing of loadings and their estimated standard uncertainties are used to calculate significance on each variable for each component [15]. Variable selection can also mean identifying a k -subset of a set of original variables that is optimal for a given criterion that adequately approximates the whole data set [16]. The application of Principal Component Regression to the trajectories of the process variables (block-wise PCR) has given straightforward results without requiring a deep knowledge of the process [17]. In this case, variable selection methods and technical information of the process has allowed the process variables most correlated with the final quality be revealed.

Genetic algorithms (GAs) have been proposed recently for many applications including variable selection for multivariate calibration, molecular modelling, regression analysis, model identification, curve fitting, and classification. GAs are also incorporated with Fisher discriminant analysis (FDA) for key variable identification for trouble-shooting problems of the Tennessee Eastman process [18]. GA and simulated annealing have also been combined for reduction in the number of variables in neural network models [19]. Two other approaches for the selection of variables in neural networks are in [20] and [21].

This paper is organised as follows: Section 2 concerns with knowledge-based variable selection and grouping, Section 3 with variable grouping with data analysis and Section 4 with model-based variable selection. The case-based reasoning system for evaluating paper machine web breaks is shortly revisited in Section 5.

2 Knowledge-based variable selection

Knowledge can be used in decreasing the number of variables. For example, if we have a case with 10 process variables and group them in all possible groups with three, four and five variables, we end up to 582 groups. If we can, based on the process knowledge, include variable 1 in all groups with three variables, variable 10 in all groups with four variables, and variables 5 and 6 in all groups with five variables, we have 176 groups to analyse. This means that using process knowledge has decreased the number of alternatives by 70 percent.

Some variable combinations should be avoided, e.g. calculated variables should not be used together with the variables used in calculating them. Also a group containing a controlled variable and its set-point is not usually appropriate. These problems are avoided by defining the inappropriate groups as non-groups, i.e. as variables groups, which should not be a part of any acceptable variable group.

3 Variable grouping with data analysis

Correlation is a statistical technique which can show whether and how strongly pairs of variables are related. Binary correlations and their combinations are used in pruning the set of acceptable groups defined by the domain expertise. For forecasting models, input variables should have a high correlation with the output variables, but a low one with each other. For case detection, causality is not always as clear: there is nor necessarily any definite output variable i.e. also groups where several variables have a high correlation between each other are acceptable. This sets new requirements for the model assessment.

In practical cases, the results from correlation analysis are improved with appropriate filtering and using correct time delays between the variables. Calculation of moving averages, medians and value ranges includes already a time delay, which depends on the calculation window and the applied methodology.

Nonlinear scaling is the essential feature in using Linguistic Equations method [22]. It improves the correlation analysis of curvilinear relationships, since the correlation analysis is a linear method. Finding patterns in data with high dimension is difficult. However, in data sets with many variables, groups of variables often move together as they are measuring the same phenomena. A host of clustering approaches helps in digging out these interactions.

As shown in Introduction, Principal Component Analysis (PCA) is a conventional method to decrease the dimensionality in data without losing the information stored in the correlated variables. It searches for new fewer linear combinations of the original variables that explain the most of the variance of the original data. These linear combinations can be viewed as a linear transformation to the hyperplane defined by the principal components or a rotation and a stretch that transform original data to a new bias.

Principal components are calculated by defining the eigenvectors of the covariance matrix or utilizing the singular value decomposition. Usually, only few first principal components (2 or 3) are used \hat{U} they are enough to explain most of the variance in the data set. There are also extensions in the basic methods that apply for analyzing time trajectories.

4 Model-based variable selection

Isokangas and Ruusunen [23] describe the automated procedure for finding interactions between variables from large datasets. This occurs systematically by constructing simple dynamic model candidates with complete input combinations for data segments of the varying and sliding window size. The final analysis goes on according to the structure properties of the best candidate models.

Model candidate construction, validation and testing proceed in the following way: the half of all available data is used in training and validation so that model candidates are constructed systematically from the beginning of data with selected data window size. After a data window has been used for training, the window of the same size is taken for validation. The procedure uses a partly overlapping data window. For example, if the data window is 400 minutes, first models are constructed using training data from 1 - 400 minutes and data from 401 - 800 for validation of a model candidate under evaluation. Next, all model candidates are constructed using training data from 201 - 600 and validation data from range 601 - 1000 minutes. To define the right size of training data, different window sizes are systematically tested at this stage. Models are evaluated with the correlation coefficient and RMS-error measure using validation data. Best models are further tested with independent testing data, which is another half of available data.

5 Paper mill example

Paper web breaks commonly account for 2-7 percent of the total production loss, depending on the paper machine type and its operation. This could mean 1.5 million euros lost annually at a single paper machine. According to statistics only 10-15 percent of web breaks have a distinct reason. The most of the indistinct breaks are due to dynamical changes in the chemical process conditions.

The main area of interest in the indicator development is the paper making process before the actual paper machine. This includes also the short circulation and the wet end of the paper machine. In this area, the paper making process is typically non-linear with many, long delays that change in time and with process conditions, there are process recycles at several levels, there are closed control loops, there exist factors that can not be measured and there are interactions between physical and chemical factors. Also several different paper grades are produced with different production conditions and operating parameters.

This Section shows how to combine on-line measurements and expert knowledge in paper machine modelling in developing the sensitivity indicator for paper web breaks [24]. The indicator would give the process operators a continuous indication of the web break sensitivity in an easily understandable way. Being able to indicate the break risk would give a possibility to react on changes of the break sensitivity in time and therefore avoid breaks.

5.1 Experimental data

The actual measurements from a paper machine were used. The main interest was in paper machine variables and the variables just before the paper machine. The final selection of variables used expert knowledge and altogether 73 variables (72 variables + information on the break occurrence) were studied. These variables were supposed to influence on paper web breaks.

The measurements were collected from the mill automation system during normal operation and no special test runs were made. The measurements were used as such to retain their information content, and, on the other hand, to keep the application as simple as possible. Only a simple filtering was added to the indicator software to make rapid changes slower and to cut the outliers from the data. The measurement data was divided into periods of 24 hours. Further the data sets were classified into five categories, depending on how many breaks there were in one day: no breaks (0), a few breaks (1 – 2), normal (3 – 4), many breaks (5 – 6) and a lot of breaks (> 6).

5.2 Reasons for web breaks

Different statistical methods were used, but reliable correlations between single variables and web breaks did not exist. Therefore, the only way to proceed was the classification and modelling of break situations to find out differences between operating situations leading to breaks. Case-Based Reasoning was used for the identification of different operating situations instead of trying to predict a single break occurrence.

Identified operating situations contain information about how many breaks there will be in the near future and this information is given to the process operators as the web break sensitivity. The identification is performed using Linguistic Equation approach and Fuzzy Logic [24].

5.3 Correlation analysis

Before modelling, correlation analysis was used in order to find out binary interactions between different process variables. The basic tool used for these analyses was Microsoft Excel spreadsheet. The correlation exceeding 0.6 was considered worth mentioning. According to this analysis, correlation

varies quite a lot with time. The variation in correlation rates is due to the usage of normal on-line measurements, which include the effects of different control operations.

The most important result of this analysis was that interactions vary in different operating situations, and the number of breaks also varies with time, and this was the basis for different case models. Due to different interactions, also different variables became important in different operating situations.

5.4 Model-based variable grouping

The web break sensitivity indicator was developed as a Case-Based Reasoning type application with Linguistic Equations approach and Fuzzy Logic [24]. The case base contains case models with different number of breaks. A new case is presented to the system as a collection of on-line measurements. The indicator compares the new case to the examples in the case base and uses the information of the best fitting case to calculate the predicted break sensitivity. As output the system gives numerical value for the predicted amount of breaks [24, 25]. Figure 1 shows the principal structure of the case base, and Figures 2 and 3 the different stages of Case-Based Reasoning.

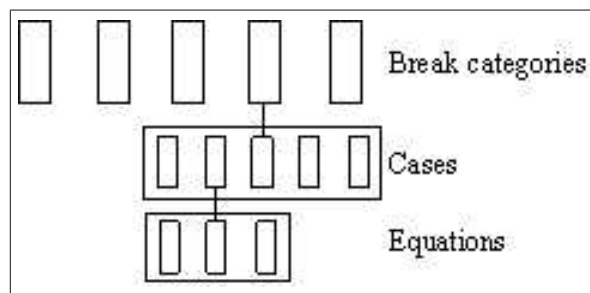


Figure 1: Structure of the Case Base [25].

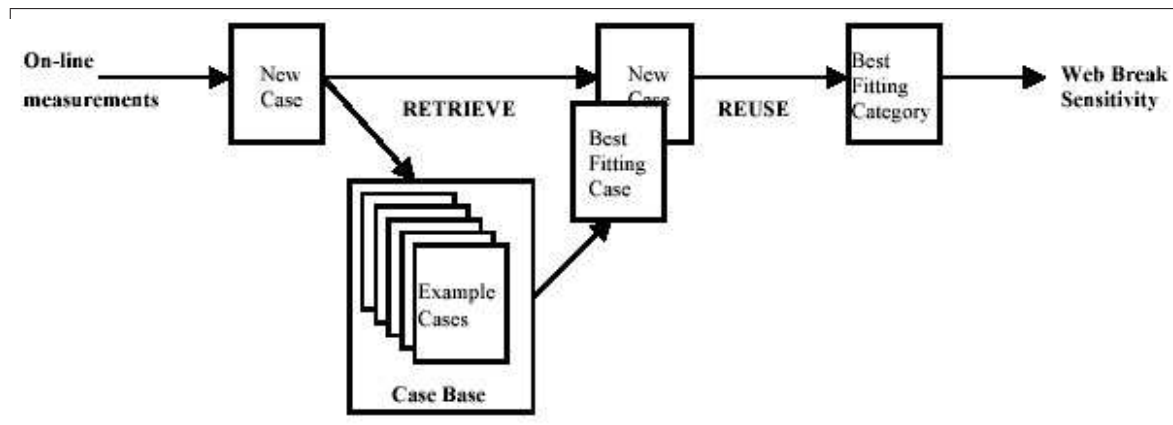


Figure 2: The structure of RETRIEVE and REUSE stages [25].

The Case Base of this application contains modelled example cases classified according to the number of related breaks. Models consist of equations that are stored as simple numerical matrices, which are indexed with break class information and number of examples in class. Equations itself describe the interactions between 3-5 variables. The variables in equations are found using a partly knowledge-based, partly model-based grouping technique.

For complex systems, a set of alternative variable groups are generated and models created with these groups. Process knowledge can be used in defining these groups. Another approach is to generate all

possible groups containing three, four or five variables and modelling them. Groups can also contain different number of variables.

Correlation analysis has also use in grouping. It should be noted that for prediction the input variables should have a high correlation with the output variable, but a low one with each other. For state detection, causality is not clear, and the group where all variables correlate with each other are acceptable. Here, however, the limitations given in Section 2 must be taken into account.

Groups with three, four or five variables can be generated automatically with FuzzEqu Toolbox [22]. The generation of the alternatives is based on groups with three variables: all groups with four variables have one variable in common, and all groups with five variables have two variables in common. The subsets of the variables and the common variables in the groups with four or five variables can be based on process knowledge.

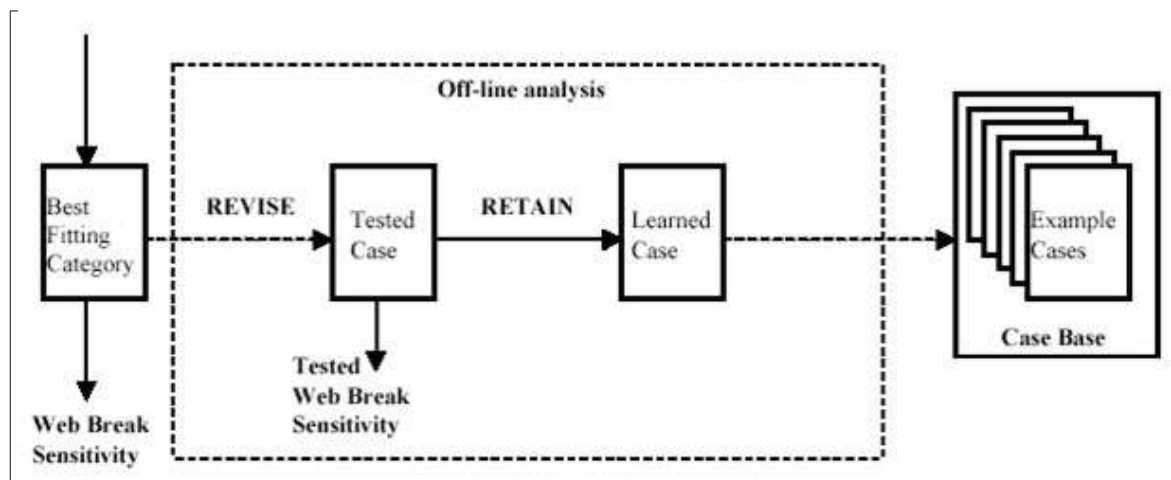


Figure 3: The structure of REVISE and RETAIN stages [25].

5.5 Some feedback \hat{U} importance of parameters

After modelling, the importance of variables was analysed based on the occurrence of variables in case models. This was also considered as a useful tool to reveal less important variables not included in the models. For the operating personnel, the list of importance might give some new information about which variables are responsible for different operating problems. There is also some difference in the collection of important variables between cases with different number of breaks.

The user interface presents continuously the six variables that best describe the current operating situation. These are simply the variables of the two best fitting equations of the best fitting case. In addition to these, also the most important variable of the best fitting case model is presented with a trend value of 8 hours. The variables are presented with their current measurement values marked with colours as normal, low or high and very low or very high. This information gives the process operators useful information of the current process conditions together with the reason for the current break sensitivity level.

The follow-up of the variable importance can lead to the need to update the whole system. The system updating is a straightforward task, all though time consuming when the whole case base is changed. A single model with 73 variables takes only few minutes to build. The same time is required for validation and tuning and it makes altogether 15 to 20 minutes. The updating of the system with 40 cases could take the working hours of two days. However, the automation of these tasks makes this time shorter.

6 Summary and Conclusions

This paper has considered possibilities to variable selection in large-scale industrial systems. It introduced knowledge-based, data-based and model-based methods for this purpose. As an example, Case-Based Reasoning application for the evaluation of the web break sensitivity in a paper machine was introduced. The application was build with Linguistic Equations approach and basic Fuzzy Logic.

The Case Base of the system contains models of example cases with different number of breaks. A new case is presented to the system as a collection of on-line measurements. The indicator compares the new case to the examples in the case base and uses the information of the best fitting case to evaluate the break sensitivity. The latest version of the indicator operates with a case base of 40 example models. Although the size of this case base is rather small, the results have been considerably good compared to the real break sensitivity.

The indicator combines the information of on-line measurements with expert knowledge and provides a continuous indication of the break sensitivity. The web break sensitivity defines the current operating situation at the paper mill and gives new information to the operators. The web break sensitivity is presented as a continuous signal with information of the actual web breaks as a trend of 8 hours. The trend shows how the situation has developed and the current value gives the prediction for next 24 hours if the situation stays as it is now. Together with information of the most important variables this prediction gives operators enough time to react to the changing operating situation.

The variable selection and grouping utilize knowledge-based and model-based approaches. Automatic group and model generation makes also the interactive variable selection possible.

References

- [1] C. Abrahamsson, J. Johansson, A. Sparén, and F. Lindgren, Comparison of different variable selection methods conducted on nir transmission measurements on intact tablets, *Chemometrics and Intelligent Laboratory Systems*, Vol. 69, pp. 3–12, 2003.
- [2] C.E.W. Gributs and D.H. Burns. Parsimonious calibration models for near-infrared spectroscopy using wavelets and scaling functions, *Chemometrics and Intelligent Laboratory Systems*, Vol. 83, pp. 44–53, 2006.
- [3] S. Gourvéneq, X. Capron, and D. L. Massart, Genetic algorithms (GA) applied to the orthogonal projection approach (OPA) for variable selection, *Analytica Chimica Acta*, Vol. 519, pp. 11–21, 2004.
- [4] L. Stordrange, T. Rajalahti, and F.O. Libnau, Multiway methods to explore and model NIR data from a batch process, *Chemometrics and Intelligent Laboratory Systems*, Vol. 70, pp. 137–145, 2004.
- [5] S.L.T. Lima, C. Mello, and R. J. Poppi, PLS pruning: a new approach to variable selection for multivariate calibration based on hessian matrix of errors, *Chemometrics and Intelligent Laboratory Systems*, Vol. 76, pp. 73–78, 2005.
- [6] M.J.C. Pontes, R. Kawakami, H. Galvão, M.C. Ugulino Araújo, P.N. Teles Moreira, O.D. Pessoa Neto, G.E. José, and T.C. Bezerra Saldanha, The successive projections algorithm for spectral variable selection in classification problems, *Chemometrics and Intelligent Laboratory Systems*, Vol. 78, pp. 11–18, 2005.
- [7] M. Arakawa, K. Hasegawa, and K. Funatsu, QSAR study of anti-HIV HEPT analogues based on multi-objective genetic programming and counter-propagation neural network, *Chemometrics and Intelligent Laboratory Systems*, Vol. 83, pp. 91-98, 2006.

- [8] Q. Shen, J.-H. Jiang, C.-X. Jiao, G. Shen, and R.-Q. Yu, Modified particle swarm optimization algorithm for variable selection in MLR and PLS modeling: QSAR studies of antagonism of angiotensin II antagonists, *European Journal of Pharmaceutical Sciences*, Vol. 22, pp. 145–152, 2004.
- [9] U. Norinder, Support vector machine models in drug design: applications to drug transport processes and qsar using simplex optimisations and variable selection, *Neurocomputing*, Vol. 55, pp. 337–346, 2003.
- [10] M. Smith, B. Pütz, D. Auer, and L. Fahrmeir, Assessing brain activity through spatial bayesian variable selection, *NeuroImage*, Vol. 20, pp. 802–815, 2003.
- [11] R. Narayanan and S.B. Gunturi, In silico ADME modelling: prediction models for blood-brain barrier permeation using a systematic variable selection method, *Bioorganic & Medicinal Chemistry*, Vol. 13, pp. 3017–3028, 2005.
- [12] E. Llobet, J. Brezmes, O. Gualdrón, X. Vilanova, and X. Correig, Building parsimonious fuzzy ARTMAP models by variable selection with a cascaded genetic algorithm: application to multisensor systems for gas analysis, *Sensors and Actuators B: Chemical*, Vol. 99, pp. 267–272, 2004.
- [13] O. Gualdrón, E. Llobet, J. Brezmes, X. Vilanova, and X. Correig, Coupling fast variable selection methods to neural network-based classifiers: Application to multisensor systems, *Sensors and Actuators B: Chemical*, Vol. 114, pp. 522–529, 2006.
- [14] M. Cocchi, J.L. Hidalgo-Hidalgo de Cisneros, I. Naranjo-Rodríguez, J.M. Palacios-Santander, R. Seeber, and A. Ulrici, Multicomponent analysis of electrochemical signals in the wavelet domain, *Talanta*, Vol. 59, pp. 735–749, 2003.
- [15] F. Westad, M. Hersleth, P. Lea, and H. Martens, Variable selection in pca in sensory descriptive and consumer data, *Food Quality and Preference*, Vol. 14, pp. 463–472, 2003.
- [16] J. Cadima, J. Orestes Cerdeira, and M. Minhoto, Computational aspects of algorithms for variable selection in the context of principal components. *Computational Statistics & Data Analysis*, 47(2):225–236, 2004.
- [17] M. Zarzo and A Ferrer, Batch process diagnosis: PLS with variable selection versus block-wise PCR, *Chemometrics and Intelligent Laboratory Systems*, Vol. 73, pp. 15–27, 2004.
- [18] L. H. Chiang and R.J. Pell, Genetic algorithms combined with discriminant analysis for key variable identification, *Journal of Process Control*, Vol. 14, pp. 143–155, 2004.
- [19] A. Alexandridis, P. Patrinos, H. Sarimveis, and G. Tsekouras, A two-stage evolutionary algorithm for variable selection in the development of RBF neural network models, *Chemometrics and Intelligent Laboratory Systems*, Vol. 75, pp. 149–162, 2005.
- [20] F. Dieterle, S. Busche, and G. Gauglitz, Growing neural networks for a multivariate calibration and variable selection of time-resolved measurements, *Analytica Chimica Acta*, Vol. 490, pp. 71–83, 2003.
- [21] I. Drezga and S. Rahman, Input variable selection for ann-based short-term load forecasting, *Power Systems, IEEE Transactions on*, Vol. 13, pp. 1238–1244, 1998.
- [22] E. K. Juuso, Integration of intelligent systems in development of smart adaptive systems, *International Journal of Approximate Reasoning*, Vol. 35, pp. 307–337, 2004.

- [23] A. Isokangas and M. Ruusunen, Systematic approach for data survey, in *Proceedings of the International Conference on Informatics in Control, Automation and Robotics. September 14 - 17, 2005, Barcelona, Spain*, pp. 60–65, 2005.
- [24] T. Ahola, Intelligent estimation of web break sensitivity in paper machines. Doctoral dissertation. University of Oulu, Department of Process and Environmental Engineering. Acta Universitatis Ouluensis, Technica C 232, 92 p., Oulu, 2005.
- [25] T. Ahola and K. Leiviskä, Case-based reasoning in web break sensitivity evaluation in a paper machine, *Journal of Advanced Computational Intelligence and Intelligence Informatics*, Vol. 9, pp. 555–561, 2005.

Timo Ahola, Esko Juuso, Kauko Leiviskä
University of Oulu, Control Engineering Laboratory
P.O. Box 4300, FI-90014 University of Oulu, Finland
E-mail: esko.juuso@oulu.fi
Received: March 21, 2007



Timo Ahola, born in Hämeenkyrö, Finland on December 18, 1965, received his M.Sc. (Eng.) in Process Engineering in 1992 and Lic. Tech. in Control Engineering in 2001 from the University of Oulu. He received D. Tech. in Control Engineering from the University of Oulu in 2006 with the thesis on intelligent estimation of web break sensitivity in paper machines. He worked as a research scientist in the Control Engineering laboratory at the University of Oulu 1993-2006. Since 2007 he has been belonging to the Outokumpu Stainless Oy, Tornio Research Centre, Finland. Presently he is a research engineer in process development and he is working with predictive maintenance issues.



Esko Juuso, born in Ylitornio, Finland on December 12, 1951, received M.Sc. (Eng.) in Technical Physics in 1979 from University of Oulu. He has worked as research engineer in Outokumpu Metallurgical Research Centre and computer analyst in Outokumpu Electronics. Since 1986, he has been belonging to University of Oulu, Oulu, Finland. Presently he is Senior Assistant in Control Engineering. He has been active in Finnish Simulation Forum (FinSim), Scandinavian Simulation Society (SIMS) and EUROSIM, currently he is chairman of FinSim. He has been member of Steering Committee and co-chairman of Technical Committee on Production Industry in EUNITE Network of Excellence, 2000-2004. His main research fields are intelligent systems and simulation in industrial applications, including control and fault diagnosis. In 1991 he introduced the linguistic equation (LE) methodology, and he has authored more than 200 publications on his research field.



Kauko Leiviskä, born in Pyhäntä, Finland, 1950, received M.Sc.(Eng.) in Process Engineering in 1975 and Lic. Tech. in Control Engineering in 1976 from the University of Oulu. He received the D. Tech. in Control Engineering from the University of Oulu in 1982 with the thesis on short term production scheduling of the pulp mill. Since 1975, he has been belonging to University of Oulu, Oulu, Finland. He has been Professor of Control Engineering and Head of Control Engineering Laboratory in the same University since 1988. He has been active in IFAC since 1988, currently he is member of IFAC TC on Large Scale Complex Systems, TC on Cognition and Control and TC on Mining, Mineral and Metal Processing. He has been member of Steering Committee and chairman of Technical Committee on Primary and Process Industries in ERUDIT Network of Excellence, 1997-2000, and the scientific director of EUNITE, the European Network of Excellence, 2000-2004. He is participating in EU/CA project NISIS (Nature-Inspired Smart Information Systems). A list of more than 200 publications of which he is (co)author is available. Recently his work concentrates on modelling and control of industrial processes, intelligent control methods, production scheduling and millwide control. He has also been consulting industry on control engineering and millwide control applications.

Modelling of Wastewater Treatment Plant for Monitoring and Control Purposes by State – Space Wavelet Networks

Adam Borowa, Mietek A. Brdys, Krzysztof Mazur

Abstract: Most of industrial processes are nonlinear, not stationary, and dynamical with at least few different time scales in their internal dynamics and hardly measured states. A biological wastewater treatment plant falls into this category. The paper considers modelling such processes for monitoring and control purposes by using State - Space Wavelet Neural Networks (SSWN). The modelling method is illustrated based on bioreactors of the wastewater treatment plant. The learning algorithms and basis function (multidimensional wavelets) are also proposed. The simulation results based on real data record are presented.

Keywords: neural network models, model approximation, learning algorithms, waste treatment.

1 Introduction

Biological wastewater treatment plants (WWTP) are very important due to their ability of neutralising results of human activity.

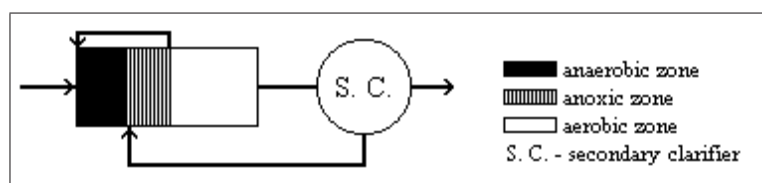


Figure 1: Activated sludge reactor (bioreactor) with secondary clarifier

Typical WWTP consists of three phases of treatment: mechanical, biological and chemical. Example of biological part of WWTP (bioreactor with secondary clarifier) is shown in figure 1. Before biological treatment the wastewater passes through mechanical treatment where coarse particle, inorganic solids and suspended particulate matter are removed. Chemical treatment may be implemented before, after or into biological treatment.

Activated sludge is responsible for nitrogen and phosphorus removing in bioreactor. Biological treatment, due nitrogen and phosphorous removing, consists of three phases: anaerobic, anoxic and aerobic. in figure 1 recirculation (from secondary clarifier to anoxic zone and from anoxic to anaerobic) are shown. Control of that recirculation, the air flow rate to aerobic zone and excessive sludge flow rate is very important for the process quality [8].

Wastewater treatment process is very complex due to its specific features such as: highly non-linear and multiple time scale dynamics, varying influent flow, high dimension of state vector with many states not accessible by hard sensors; see [4] for details.

Due to nationwide regulations, which force high standards on treated wastewater quality, a need for better treatment still exists. There are two solutions of this problem. The first is to enlarge the WWTP and the second, to implement an intelligent control system. It's obvious that the last proposition is better in an economic sense.

The main purpose of control of WWTP is to keep all quality parameters of treated waste under specific norm. Because of complexity of the process, the advanced control technologies are required. The multilevel-multilayer hierarchical control structure was recently derived [1] and further developed in [4].

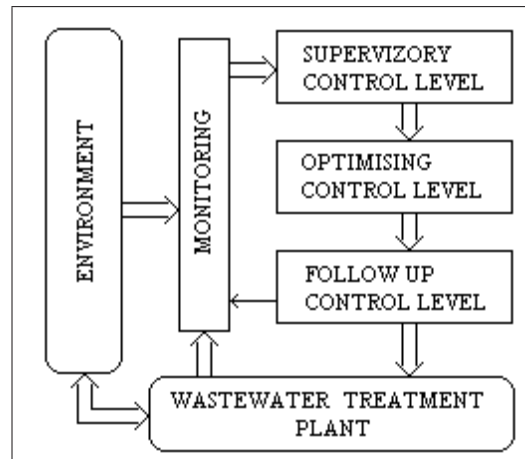


Figure 2: Multilevel-multilayer hierarchical control structure

This structure (figure 2) consists of 3 control levels: supervisory control level, optimising control level and follow up control level; each with different control objectives. Optimising Control Level (OCL) uses a Robust Model Predictive Control (RMPC) algorithm and it is decomposed into three layers each with different control horizon and time scale. This work assembles on modeling the WWTP for RMPC and supervisory control. A good candidate for modelling a dynamical, non-linear system with multiple time scales is a State Space Wavelet Network (SSWN). Using the wavelets as the basis functions solves the multiple time scale problem. The rest of article order goes as follow. Section 2 is a description of problem statement. Wavelet Networks are presented in section 3. A learning algorithm is proposed in section 4. Stability of SSWN is mentioned in section 5. Application and results are shown in section 6 and 7.

2 Problem statement

The paper aim is to verify the possibility of applying the SSWN modelling of a WWTP for monitoring and control purposes.

2.1 Monitoring of WWTP for supervisory control purposes

Monitoring the WWTP state is essential for the supervisory control purposes [4]. Most of the WWTP states are not measurable. Even though the SSWN for state monitoring can be made. First model of WWTP in SIMBA has to be calibrated. Methods for SIMBA model calibration are presented in [2]. When calibrated model in SIMBA is available then it can be used as a data generator for learning the SSWN.

2.2 Progress of the hierarchical intelligent control structure

The layer decomposition of the OCL raises a loss on optimality and certain problems with accommodating the constraints into the layer optimisation tasks. As the new SSWN model has an ability to combine the two time scales (fast and medium), the fast and medium layers can be combined into one layer, hence eliminating the above problems.

3 Wavelet network

Wavelet Network is a neural network with one hidden layer consisting wavelets as the basis function. Wavelets are specific mathematical functions and are described below. Neurons made up of wavelets are called wavelons.

3.1 Wavelets

Any function satisfying the conditions (1) and (2), where $\psi(\omega)$ is a Fourier transform of $\Psi(t)$, is called a mother wavelet. It is required that

$$C_{\psi} = \int_0^{\infty} \frac{|\psi(\omega)|^2}{\omega} d\omega < \infty \quad (1)$$

$$\int \Psi(t) dt = 0 \quad (2)$$

Examples of wavelets are Haar wavelet or Morlet wavelet shown in figure 3 and defined by the equations (3) and (4) respectively.

$$\Psi(x) = \begin{cases} 1, & 0 \leq x < 0.5 \\ -1, & 0.5 \leq x < 1 \\ 0, & \text{elsewhere} \end{cases} \quad (3)$$

$$\Psi(x) = \exp(-x^2/2) * \cos(5x) \quad (4)$$

A family of wavelets is made by dilating and translating mother wavelet. See equation (5), where d and

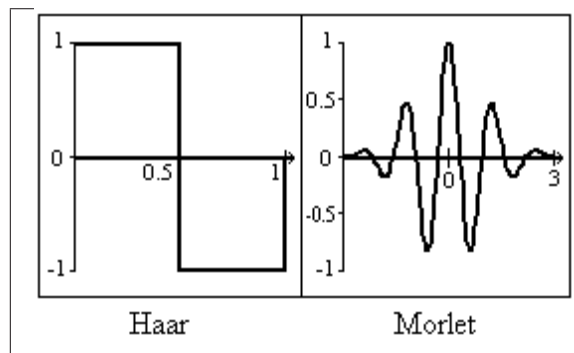


Figure 3: Haar and Morlet mother wavelets.

t represent the dilation and translation parameters respectively.

$$\Psi_{dt}(x) = \sqrt{d}\Psi(d(x-t)) \quad (5)$$

The translation and dilation parameter determine the position (time) and scale (frequency) domain. The wavelets represented by equation (5) may be design as orthonormal.

3.2 Multidimensional wavelets

The multidimensional wavelet (6) was presented first by Zhang and Beneveniste in [11]. In equation (6) $\Psi_S(x)$ is a one-dimensional wavelet, X is a vector and $\Psi(X)$ is a scalar

$$\Psi(X) = \Psi(x_1, \dots, x_n) = \prod_{j=1}^n \Psi_S(x_j) \quad (6)$$

Few months later Zhang [12] presented new multidimensional wavelet - radial wavelet (7) and (8) where $\Psi(X)$ is a one-dimensional wavelet.

$$\Psi(X) = \Psi_S(\|x\|) \quad (7)$$

$$\|x\| = (x^T x)^{\frac{1}{2}} \quad (8)$$

It is logical that for a multivariable function approximation the multidimensional wavelets are desirable. The multidimensional wavelet (6) was tested and some problems were encountered. First, often the optimisation problem to be solved by the learning algorithm was ill-conditioned. Second, a large n relatively easily led to overparameterization.

Only when radial wavelet (equations (7) and (8)) was used the network learning was successfully finalized. The radial multidimensional wavelet, used in the paper, is given by equations (9) - (13) and it is shown in figure 4.

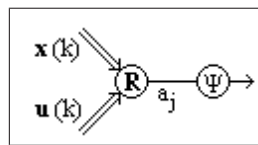


Figure 4: Multidimensional radial wavelon.

$$\mathbf{z} = [\mathbf{x}, \mathbf{u}] \quad (9)$$

$$\mathbf{d}_j = [d_{1,j}, \dots, d_{K+M+N,j}] \quad (10)$$

$$\mathbf{t}_j = [t_{1,j}, \dots, t_{K+M+N,j}] \quad (11)$$

$$\mathbf{A} = \text{diag}(\mathbf{d}_j) * (\mathbf{z} - \mathbf{t}_j)^T \quad (12)$$

$$a_j = R(\mathbf{A}(\mathbf{z}, \mathbf{d}, \mathbf{t})) = (\mathbf{A}^T \mathbf{A})^{\frac{1}{2}} \quad (13)$$

A one \tilde{U} dimension wavelet used in this application was a morlet wavelet (4). Proposed structure for a wavelon was described in [12].

3.3 Feed-forward and Input-Output Wavelet Networks.

It was presented in [5] that Feed-Forward Wavelet Network (FFWN) in the form (14) is a universal approximator for any function in Sobolev space.

Wavelet Networks have more freedom than other neural networks because of number of optimised parameters for each wavelon.

$$z = \sum_{i=1}^N w_i \Psi(d_i(x - t_i)) \quad (14)$$

The wavelet network for one-dimensional input network is described by equation (14), where d_i and t_i are dilation and translation parameters respectively, w_i are linear weights and N is a number of wavelons. A feed forward wavelet network (figure 5) was presented first by Zhang and Beneveniste [11]. A Dynamical Wavelet Network (DWN) was presented in [7]. The DWN structure is shown in figure 6. It is the Input-Output structure, which is commonly used but it has some drawbacks that do not permit this DWN to be used for WWTP modelling. Since 1992 many wavelet networks were presented and used in different places of human activity.

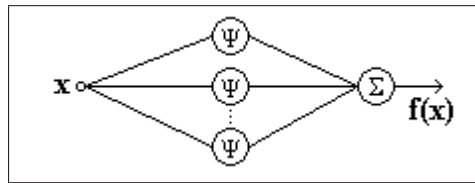


Figure 5: Feed-forward wavelet network.

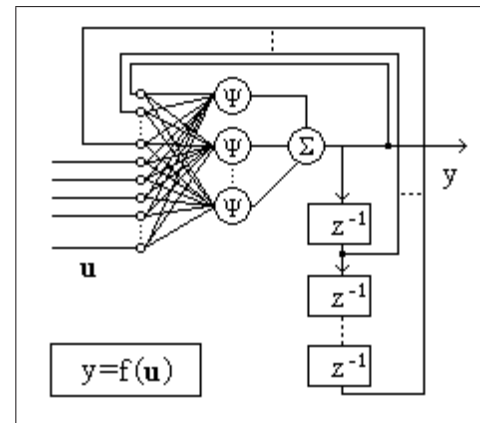


Figure 6: Dynamical wavelet network.

3.4 State Space Wavelet Network

State Space Wavelet Network (SSWN) is not as well known as FFWN or DWN but it has specific advantages over mentioned networks. Due to its of internal state space component the SSWN, better captures the modeled plant structure. Therefore the modeling error is smaller and the learning process is faster. Nevertheless, there is no proof for such a network to be the universal approximator.

SSWN with multidimensional wavelets is shown in figure 7 and described by the equations (15) and (16), where N is a number of outputs, $M + N$ is a number of state variables, K is a number of inputs (control and disturbance), L is a number of wavelons. A number of parameters to be estimated during learning process are then $(2 * L * (N + M + K) + L * (N + M))$. It combines the state space architecture of dynamic neural network [10] with the multidimensional wavelons as the processing nodes.

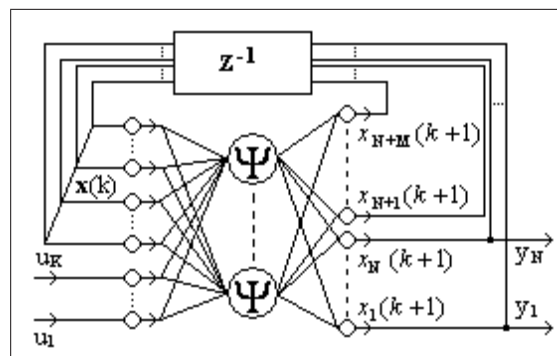


Figure 7: Structure of State Space Wavelet Network with multidimensional wavelons

$$\forall_{i=1}^{N+M} x_i(k+1) = \sum_{j=1}^L w_{i,j} * \Psi_j(x(k), u(k)) \quad (15)$$

$$\forall_{i=1}^N y_i(k) = \sum_{j=1}^L w_{i,j} * \Psi_j(x(k), u(k)) \quad (16)$$

4 Learning Algorithm

During searching for the right structure of wavelet network, selecting the learning algorithm was also investigated. Searching over a large set of combinatorial optimization algorithms included: Simple Genetic Algorithm (SGA), SGA with elitism, Evolutionary Algorithm (EA), EA with SBX crossover,

Simulated Annealing (SA) and parallel hybrid of SA and EA with the use of computer grid technology. The SA algorithm turned out to be the best for our problems.

4.1 Simulated annealing

Whilst inspiration for GA and EA was in biology (genetic), the SA was inspired by thermodynamics (Statistical Mechanics). The algorithm was motivated by the growing mechanism of a single crystal from a melt [6]. It was found that slow cooling (annealing) of melted metal goes to low state of energy while fast cooling does not.

A simple algorithm based on Monte Carlo search was proposed in (Metropolis 1953) which then became an important part of the SA algorithm [6]. This algorithm works on a chain of atoms (S). In each step one atom from the chain is disturbed and the new chain energy $E(S')$ is calculated. A difference between the chain S and S' energies is calculated by equation (17).

$$\Delta E = E(S') - E(S) \quad (17)$$

If $\Delta E < 0$ then new chain S' is accepted; otherwise the new chain is conditionally accepted with a probability given by Boltzmann probability factor (18), where k_B is a Boltzmann's constant and T is a temperature factor. Whilst T is not real temperature and k_B is a constant then product $k_B * T$ may be replaced for practical implementation by single factor T .

$$P(\Delta E) = \exp(-\Delta E / (k_B T)) \quad (18)$$

The annealing schedule was added in [6] in order to formalise the SA algorithm. The SA algorithm works iteratively as follow. For given temperature T_0 the Metropolis Monte Carlo (MMC) method is applied. When chain of atoms is said to be stable then new T is computed and MMC is applied; this procedure goes as long as the temperature reaches 0 or the energy obtains optimum. For the new T the computation-annealing schedule is given by equation (19), where T_0 is initial temperature and k is SA iteration counter. For practical matter chain of atoms is said to be stable when MMC iteration counter obtain an established number.

$$T(k) = \frac{T_0}{1 + \ln(k)} \quad (19)$$

4.2 A simple method for SSWN learning with simulated annealing algorithm

In the paper the learning procedure was implemented as follow:

- Initialize a chain of atoms,
- Set the initial temperature,
- Establish a number of MMC iterations needed to obtain a stable chain,
- Run iterative SA algorithm.

Initialization of chain of atoms is made using a random number generator with upper and lower constrains for weights and wavelet coefficients. The initial temperature is chosen by the user (between 0 and 1). Number of MMC iterations is proportional to length of the chain of atoms. In optimisation only one chain of atom should be used because convergence of SA is independent of initialization due to exploring nature of the method.

Before the SA is applied the energy function must be described. Energy function is given by equation (20), where Y is the plant output and Y^z is a SSWN output, N is a number of SSWN outputs, J is a number of samples taken for network learning.

$$E = \frac{1}{N * J} \sum_{i=1}^N \sum_{j=1}^J |Y_{i,j} - Y_{i,j}^z| \quad (20)$$

It is important to normalize the data before learning the network.

5 Stability of SSWN

It was proved that under certain conditions the State Space Neural Networks (SSNN) can be made stable [9]. The sufficient conditions for the stability suitable constrain the network weights.

It is not a subject of this work to proof the SSWN stability. However, we shall demonstrate by simulation that the stability is in place if the network parameters are constrained. A discharge of a randomly chosen initial state of the network is illustrated in figure 8. After the discharge has been finished the network output accurately follows the plant output.

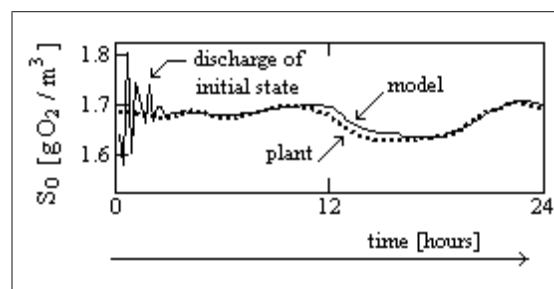


Figure 8: Discharge of randomly chosen initial state of SSWN

6 Application

After the SSWN structure and learning algorithm have been chosen the inputs, outputs, size of the state vector and the number of wavelons must be fixed.

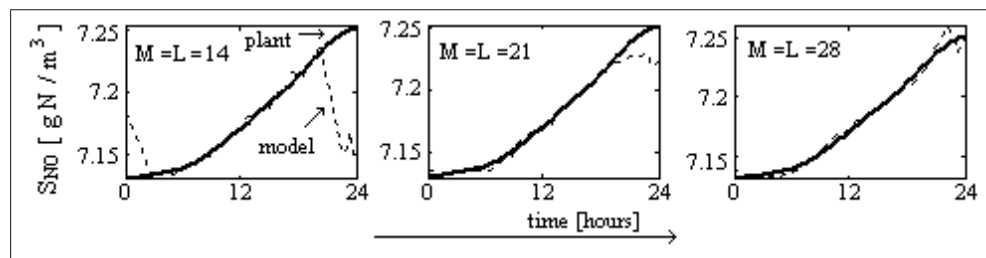


Figure 9: Learning results for different M and L

The presented SSWN was applied to wastewater treatment plant modelling. The modelled plant consists of 4 control inputs (flow rate to aerobic zone, 2 recirculation flow rates and excessive sludge flow rate), 4 measured or estimated disturbance inputs, including inflow and quality (COD , BOD , TSS), and 4 outputs: outflow (Q), concentrations of nitrate and nitrite nitrogen (S_{NO}) and $NH_4^+ + NH_3$ nitrogen (S_{NH}) in effluent and concentration of oxygen in aerobic zone (S_O).

Still the number of state space variables M and number of wavelons L are unknown. Fortunately we know the size of state of the modelled plant, which is 14 for each zone of biological treatment. Therefore M was searched in a set of values: 14, 21 and 28; the larger M implied a huge set of optimized parameters. In order to reduce the computational burden M wavelons were applied. The network was parameterized by 980, 1911 and 3136 parameters, respectively. Results of the learning (limited to single output) for these three parameterization examples are shown in figure 9. Finally M and L were selected as 28 to give small modelling error and acceptable learning time.

7 Results

The results of long term learning are shown in figure 10. The learning time was around 36 hours with simulated annealing algorithm. The bold line shows the modelled plant output while the dashed line illustrates the network output. The mean modelling error was about 1.84

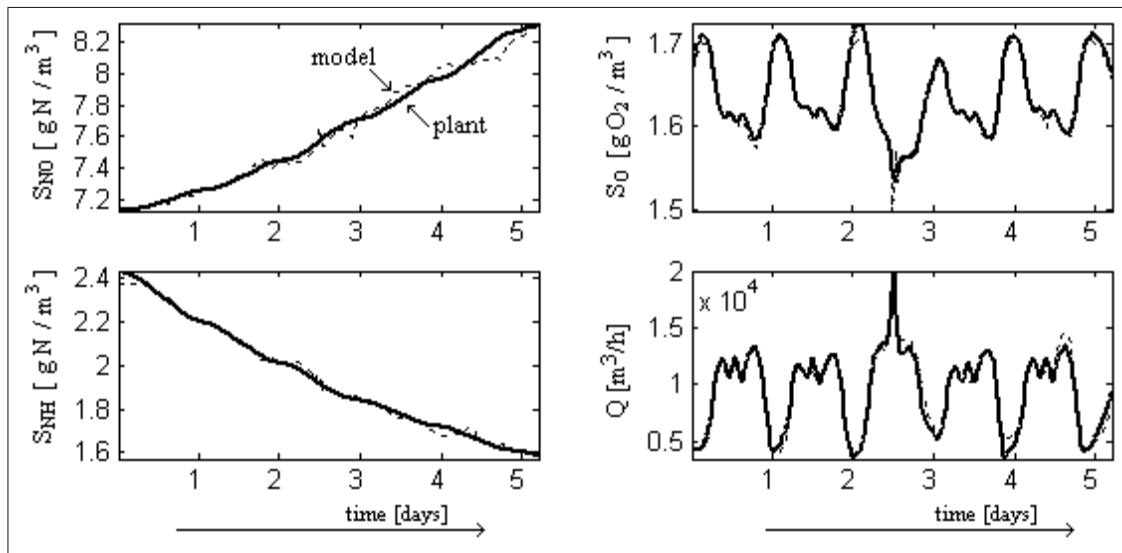


Figure 10: Comparison of model and plant outputs

8 Conclusions

A structure of a dynamical wavelet network called State Space Wavelet Network (SSWN) has been proposed. The Multidimensional Radial Wavelon has been proposed as the network processing nodes. The Simulated Annealing for the SSWN learning has been derived and validated by application to modelling the wastewater treatment plant.

It has been demonstrated that such a network is able to approximate this dynamic, nonlinear, not stationary process with several different time scales.

References

- [1] M. A. Brdys, M. Grochowski, K. Duzinkiewicz, W. Chotkowski, Y. Liu "Design of control structure for integrated wastewater treatment plant Ź- sewer systems", *International Conference on Technology, Automation and Control of Wastewater and Drinking Water Systems TiASWiK'02* Gdansk – Sobieszewo, Poland, June 19-21 2002.
- [2] W. Chotkowski, J. Mżkinia, M.A. Brdys, K. Duzinkiewicz, K. Konarczak "Mathematical modelling of the processes in integrated municipal wastewater systems", *Proc. of the 9th IFAC/IFORS/IMACS/IFIP Symposium on Large Scale Systems: Theory and Applications, Bucharest, July 18-20 2001.*
- [3] I. Daubechies *Ten Lectures on Wavelets* CBMS-NSF Regional Series in Applied Mathematics, SIAM, Philadelphia, 1992.

-
- [4] M. Grochowski, M.A. Brdys, T. Gmiński "Intelligent control structure for control of integrated wastewater systems" *IFAC 10th Symposium Large Scale Systems: Theory and Applications* Osaka Ű Japan July 26-28 2004.
- [5] A. Juditsky, Q. Zhang, B. Delyon, P-Y. Glorennec, A. Beneveniste "Wavelets in identification" *Rapport de recherche Nř2315* 1994.
- [6] S. Kirkpatrick, C.D. Gelatt, M.P. Vecchi "Optimization by Simulated Annealing" *Science* vol. 220 pp. 671-680, 1983.
- [7] Y. Oussar, I. Rivals, L. Personnaz, G. Dreyfus "Training Wavelet Networks for Nonlinear Dynamic Input-Output Modeling" *Neurocomputing*, vol. 20, pp. 173-188, 1998.
- [8] G. Olsson, R Newell *Wastewater Treatment Systems. Modelling, Diagnosis and Control*. IWA Publishing, London, 1999.
- [9] E.N. Sanchez, J.P. Perez "Input-to-State Stability (ISS) Analysis for Dynamic Neural Networks" *IEEE Transactions On Circuits And Systems Ű I: Fundamental Theory And Applications*, vol. 46, No. 11, pp 1395 Ű 1398, November 1999.
- [10] J.M. Zamarreno, V. Pastora "State space neural network. Properties and applications." *Neural Networks*, vol. 11, pp 1099-1112, 1998.
- [11] Q. Zhang, A. Beneveniste "Wavelet Networks" *IEEE Trans. on Neural Networks*, vol. 3, num. 6, pp 889-898, Nov. 1992.
- [12] Q. Zhang "Wavelet Network: the Radial Structure and an Efficient Initialization Procedure" *Technical Report of Linköping University, LiTH-ISY-I-1423*, October 1992.
- [13] J. Zhao, B. Chen, J. Shen "Multidimensional non-orthogonal wavelet basis function neural network for dynamic process fault diagnosis" *Computer and Chemical Engineering* vol. 23 pp. 83-92, 1998.

Adam Borowa, Krzysztof Mazur
Gdansk University of Technology,
Department of Automatic Control,
ul. G. Narutowicza 11/12, 80 952 Gdansk, Poland
E-mail: aborowa@ely.pg.gda.pl, kmazur@ely.pg.gda.pl

Mietek A. Brdys
The University of Birmingham, School of Engineering,
Department of Electronic, Electrical and Computer Engineering,
Birmingham B15 2TT, UK
E-mail: m.brdys@bham.ac.uk
Received: March 12, 2007



Adam Borowa received his M.Sc. degree in Control Engineering in 2002 from Electrical and Control Engineering Department at Gdansk University of Technology. Soon after he became a Ph.D. student in this Department. During the period of 2001 to 2002 he served one's apprenticeship on Wastewater Treatment Plant at Swarzewo. He published 6 publications. Mainly he focuses on modelling and monitoring of large scale systems, especially processes with many time scales.



Mietek A. Brdys received the M.Sc. degree in Electronic Engineering and the Ph.D. and the D.Sc. degrees in Control Systems from the Institute of Automatic Control at the Warsaw University of Technology in 1970, 1974 and 1980, respectively. From 1974 to 1983, he held the posts of Assistant Professor and Associate Professor at the Warsaw University of Technology. In 1992 he became Full Professor of Control Systems in Poland. Between 1978 and 1995, he held various visiting faculty positions at the University of Minnesota, City University, De Montfort University and University Polytechnic of Catalunya. Since January 1989, he has held the post of Senior Lecturer in the School of Electronic, Electrical and Computer Engineering at The University of Birmingham. Since February 2001 he has held the post of Full Professor of Control Systems in the Department of Automatic Control at Gdansk University of Technology. He has served as Consultant for Honeywell Systems and Research Center in Minneapolis, GEC Marconi and Water Authorities in UK, France, Spain, Germany and Poland. He is Head of Interdisciplinary Research Network on Decision Support and Control Systems at The University of Birmingham and Head of Intelligent Decision Support and Control System Group at Technical University of Gdansk. His research is supported by the UK and Polish Research Councils, and industry and European Commission. He is author and co-author of about 200 refereed papers and six books. His current research includes intelligent decision support and control of complex uncertain systems, robust monitoring and control, softly switched robustly feasible model predictive control. The applications include environmental systems, technological processes, autonomous intelligent vehicles and defence systems. He is a Chartered Engineer, a Member of the IEE, a Senior Member of IEEE, a Fellow of IMA and a Vice-Chair of IFAC Technical Committee on Large Scale Complex Systems. He the IPC Chair of the 11th IFAC Symposium on Large Scale Complex Systems, Gdansk, July 23-25, 2007.



Krzysztof Mazur received his M.Sc. degree in Control Engineering from Electrical and Control Engineering Department at Gdansk University of Technology in 2005. Currently a Ph.D. student in this Department. His research interests are in the areas of modelling, control and monitoring of large scale systems. Co-author of 4 publications.

Predictive Control of a Wastewater Treatment Process

Sergiu Caraman, Mihaela Sbarciog, Marian Barbu

Abstract: The paper deals with the design of a predictive controller for a wastewater treatment process. In the considered process, the wastewater is treated in order to obtain an effluent having the substrate concentration within the standard limits established by law (below 20 mg/l). This goal is achieved by controlling the concentration of dissolved oxygen to a certain value. The predictive controller uses a neural network as internal model of the process and alters the dilution rate in order to fulfill the control objective. This control strategy offers various possibilities for the control law adjustment by means of the following parameters: the prediction horizon, the control horizon, the weights of the error and the command. The predictive control structure has been tested in three functioning regimes, considered essential due to the frequency of their occurrence in current practice.

Keywords: predictive control, wastewater treatment, neural network, bioreactor

1 Introduction

The issue of wastewater treatment belongs to a larger area, namely the environment protection. The environment protection generally and biological wastewater treatments particularly is essential for the life of the human communities and received lately a lot of attention from specialized international organizations. In this context, European laws envisage a series of specific orientations for treating and maintaining the water quality within legal limits (eg. surface water directives, 75/440/EEC and 79/869/EEC, drinking water directives 80/778/EEC/15July1980 and 98/83/EEC/3 November 1998, urban wastewater treatment directive 91/271/EEC etc.).

Complementary to what has been already stated, the wastewater treatment processes are very complex, non-linear and characterized by many uncertainties w.r.t. the influent parameters, the structure and the coefficients of the model. Moreover, many wastewater treatment plants do not have measurement and control equipments. Therefore, there is a need in designing control strategies for the good operation of the process, strategies that may consider various types of models.

Process modelling: there has been a long transition between adopting the procedure of wastewater treatment using active sludge and setting up the theoretical framework to closely describe the procedure. The delay was mainly caused by the conflicting hypotheses related to the explanation of process mechanisms and their difficult translation into mathematical models [3, 7].

In 1983, International Water Association (IWA) formed a working group destined to promote and facilitate the practical methods of designing and operating the biological wastewater treatment systems. As a result, the Activated Sludge Model 1 (ASM1) has been presented in 1987 (see [8]). The model used 13 state variables and described the elimination of organic carbon and nitrogen. The same working group extended the model afterwards by adding the biological process of phosphorus elimination, and named this model the Activated Sludge Model 2 (ASM2) [9]. Two other improved versions of ASM2, named ASM2d and ASM3 appeared [10]. The major shortcoming of ASM1 is its complexity, which makes it difficult to be used in a control system. A simplified alternative of the ASM1 model was obtained by taking into consideration the significant variables on a medium time-scale (a few hours to several days). This is why the variables with a slow evolution were considered constant and those having a fast evolution were neglected [13]. These simplifications allowed the usage of ASM1 model in designing control laws. Lately, IWA established two major research areas:

- Modelling of different industrial wastewater treatment processes: cellulose and paper industry, agricultural farms, spun glass industry, etc. The research team tries to model each process, depending on the substances involved in the process. Contrary to the case of domestic wastewater treatment which is made naturally by the microorganisms, the industrial wastewater treatment is done by cultivation of microorganisms, sometimes genetically modified, which consume a certain organic substrate [17, 5].
- Conditioning the excess of active sludge in order to use it in other industrial activities, especially as a fertilizer in agriculture [16, 20].

Process control: the wastewater treatment systems are complex, non-linear processes, with multiple inputs and outputs (multivariable), which determine equivocal information about the influent's characteristics, the model's structure and parameters. Two approaches can be distinguished in choosing the control structure for such a process: the first one is process-driven and the second one is model-based. The first approach deals with the separate control of the most important variables. Within this category, the well-known problem of controlling the dissolved oxygen level is one of the most important issues for a good operation of the wastewater treatment plants. Thus, a good level of dissolved oxygen allows the optimal growth of microorganisms used in the process [12]. Recently, the control of nitrogen and phosphor level received also a lot of attention [21]. The second approach has been improved a number of times. These improvements are related to the type of the mathematical model used, as it is the case for state estimators. Using simplified models allowed the application of advanced control techniques (e.g. precise linearizing or adaptive control, robust control techniques etc.) [19]. However, when using more complex models, such as the ASM1 model, the issue of automatic control became very complicated and the established results were less numerous. For the ASM1 model classic control techniques are usually used (PI, PID controllers), arranged hierarchically, in a three-level structure [2]: at the higher level, a stable trajectory for the process is calculated for a certain period of time; the medium level deals with the trajectory optimization for the dissolved oxygen, the flow of the recycled active sludge and the recycled inflow for nitrogen removal; at the lower level, the control of dissolved oxygen concentration is achieved, based on the medium level reference. A well-suited approach for this type of process is the control based on artificial intelligence strategies. Thus the intelligent control exploits the knowledge and experience accumulated from managing the process and puts it across the control structures like expert, fuzzy, neuro-fuzzy systems [1, 15, 18].

The present study considers a simplified model of the biological wastewater treatment plant [19]. The process is controlled using a model-based predictive control (MPC) strategy. The predictive controller uses a neural network as internal model of the process. This offers various possibilities for the control law adjustment by means of the following parameters: the prediction horizon, the control horizon, the weights of the error and the command. The control purpose is to maintain the substrate concentration below an admissible limit, which is indirectly achieved by controlling the dissolved oxygen concentration, considering as control input the dilution rate D . The predictive control structure has been tested during several functioning regimes, which are essential due to their frequent occurrence in current practice.

The paper is structured as follows: the second section describes the process components and the mathematical model of the plant, the third section introduces theoretical considerations about the control structure used in the paper, while the fourth section refers to the neural network used as internal model of the predictive controller. The fifth section presents the simulation results of the proposed control structure and the last section is dedicated to conclusions.

2 The model of the wastewater treatment process

The mathematical model considered in this paper has been proposed in [19]. The model is based on the following assumptions:

- the system runs in steady-state regime ($F_{in} = F_{out} = F$, $D = F/V$);
- the recycled sludge is proportional to the process flow (F): $F_r = r \cdot F$, where r is the recycled sludge rate;
- the flow of the sludge removed from the bioreactor (sludge that is in excess) is considered proportional to the process flow (F): $F_\beta = \beta \cdot F$, where β is the removed sludge rate;
- there is no substrate or dissolved oxygen in the recycled sludge flow of the bioreactor;
- the output flow of the aerated tank is equal to the sum between the output flow of the clarifier tank (settler) and the recycled sludge flow.

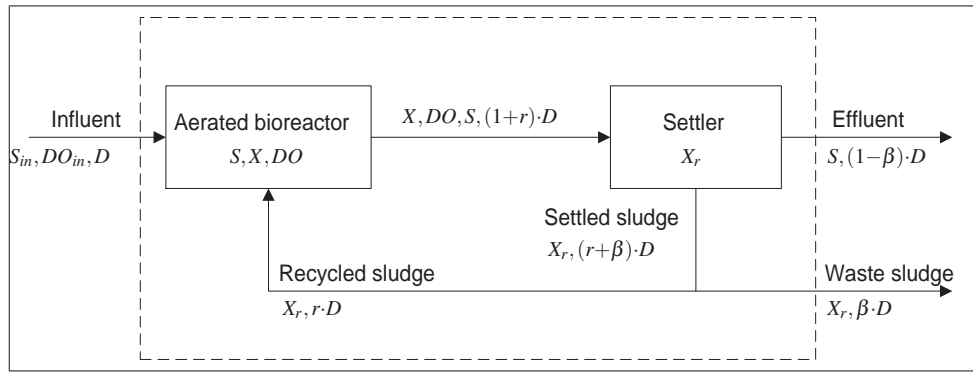


Figure 1: The schematic representation of the wastewater treatment process

Figure 1 presents the schematic representation of the wastewater treatment process. The *Aeration Tank* is a biological reactor containing a mixture of liquid and suspended solid, where a microorganism population is grown in order to remove the organic substrate from the mixture. The *Clarifier Tank* is a gravity settlement tank where the sludge and the clear effluent are separated. A part of the removed sludge is recycled back to the aeration tank and the other part removed [14].

Under these conditions the process model is given by the following mass balance equations:

$$\frac{dX}{dt} = \mu(t)X(t) - D(t)(1+r)X(t) + rD(t)X_r(t) \quad (1)$$

$$\frac{dS}{dt} = -\frac{\mu(t)}{Y}X(t) - D(t)(1+r)S(t) + D(t)S_{in} \quad (2)$$

$$\frac{dDO}{dt} = -K_0 \frac{\mu(t)}{Y}X(t) - D(t)(1+r)DO(t) + \alpha W(DO_{max} - DO(t)) + D(t)DO_{in} \quad (3)$$

$$\frac{dX_r}{dt} = D(t)(1+r)X(t) - D(t)(\beta+r)X_r(t) \quad (4)$$

$$\mu(t) = \mu_{max} \frac{S(t)}{k_S + S(t)} \frac{DO(t)}{K_{DO} + DO(t)} \quad (5)$$

where $X(t)$ - biomass, $S(t)$ - substrate, $DO(t)$ - dissolved oxygen, DO_{max} - maximum dissolved oxygen, $X_r(t)$ - recycled biomass, $D(t)$ - dilution rate, S_{in} and DO_{in} - substrate and dissolved oxygen concentrations in the influent, Y - biomass yield factor, μ - biomass growth rate, μ_{max} - maximum specific growth rate, k_S and K_{DO} - saturation constants, α - oxygen transfer rate, W - aeration rate, K_0 - model constant, r and β - ratio of recycled and waste flow to the influent. The model coefficients have the following values: $Y = 0.65$; $\beta = 0.2$; $\alpha = 0.018$; $K_{DO} = 2 \text{ mg/l}$; $K_0 = 0.5$ $\mu_{max} = 0.15 \text{ mg/l}$; $k_S = 100 \text{ mg/l}$; $DO_{max} = 10 \text{ mg/l}$; $r = 0.6$.

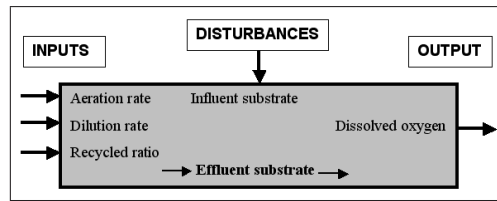


Figure 2: The systemic diagram of the wastewater treatment process

The systemic diagram of the process is given in figure 2.

Figure 3 illustrates the open loop response of the system for a step input $D = 0.1 \text{ h}^{-1}$ ($W = 80 \text{ h}^{-1}$). The initial conditions considered in this simulation are: $X(0) = 200 \text{ mg/l}$, $S(0) = 88 \text{ mg/l}$, $DO(0) = 5 \text{ mg/l}$, $Xr(0) = 320 \text{ mg/l}$, $DO_{in} = 0.5 \text{ mg/l}$, $S_{in} = 200 \text{ mg/l}$.

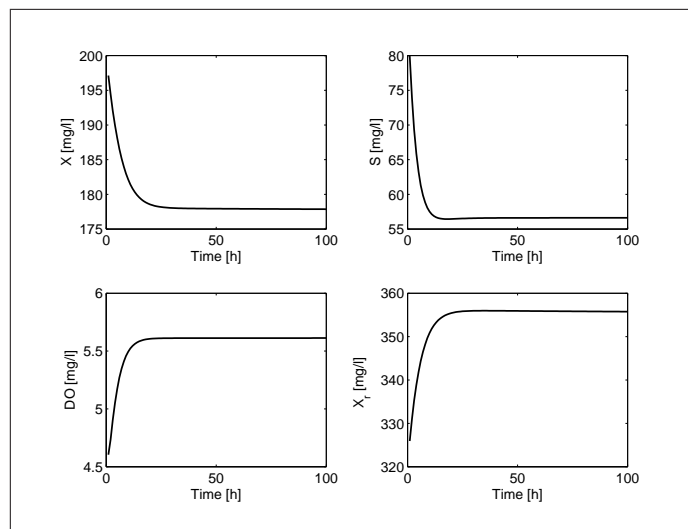


Figure 3: Open loop system response

During the normal functioning of the wastewater treatment process, three regimes have been identified: *rain* ($D = 1/20 \text{ h}^{-1}$, $W = 80 \text{ h}^{-1}$), *normal* ($D = 1/35 \text{ h}^{-1}$, $W = 60 \text{ h}^{-1}$) and *drought* ($D = 1/50 \text{ h}^{-1}$, $W = 20 \text{ h}^{-1}$). The first case is characterized by maximum values for the aeration and dilution rates, the second regime considers medium values for W and D and the third case is characterized by small values for the same parameters. In this study special attention has been paid to the predictive controller, such that it provides good performances for all the three functioning regimes.

3 Predictive control

Predictive control algorithms belong to the class of model-based control strategies, using a process model to incorporate the predicted future behavior of the process into the controller design procedure [6]. Independent of the type of model used and of the cost index minimized, the principle of MPC is the same. At each sampling instant t [4]:

- use the process model to predict the future output of the process over the *prediction horizon* N_2 , $\{y(t+k/t), k = 1 \dots N_2\}$, based on past inputs and outputs and *postulated* future inputs;
- minimize the cost index, taking into account possible constraints on input, output and states, in order to determine the optimal control sequence $\{u(t+k/t), k = 0 \dots N_u - 1\}$, where N_u is the *control horizon*;

- use the receding-horizon control mechanism that introduces feedback into the optimization problem by applying to the process the first optimal control action and discarding the consequent ones.

As it is straightforward from the description above (the model is the central element of the entire strategy), the success of MPC strategy is highly dependent on a reliable process model, that is a model which approximates well the process dynamics. A lot of research has been carried out up to now in the area of MPC based on linear models, however many of the real-life processes are characterized by complex non-linearities, the necessity of having a non-linear model of the process becoming straightforward. Therefore, emphasis is placed nowadays on using nonlinear models in the framework of predictive control that would lead to improved control performances.

The concept of the predictive control algorithm used in this work is illustrated in figure 4.

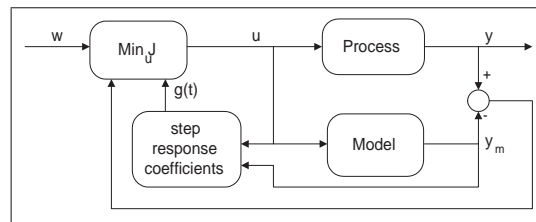


Figure 4: Predictive control structure

At each sampling instant t , the increment of the control input $\Delta u(t)$ is calculated by minimizing the cost function

$$J = \sum_{k=1}^{N_2} \delta^2(k) [w(t+k/t) - y(t+k/t)]^2 + \sum_{k=0}^{N_u-1} \lambda^2(k) [\Delta u(t+k/t)]^2 \quad (6)$$

where $w(t+k/t)$ is the setpoint prediction, $\delta(k)$ and $\lambda(k)$ are respectively the weighting coefficients of the prediction errors and of the control input increments. In order to calculate the output prediction, the step response of the model must be determined. To this end, it is necessary to admit that the model can be linearized around the current operating point. Then

$$y(t+k/t) = y_{free}(t+k/t) + y_{forced}(t+k/t) \quad (7)$$

where $\{y_{free}(t+k/t), k=1 \dots N_2\}$ is the model output produced by the control input sequence $\{u(t+k/t) = u(t-1), k=0 \dots N_2-1\}$ and

$$y_{forced}(t+k/t) = \sum_{i=1}^k g_i \Delta u(t+k-i/t) \quad (8)$$

with $\{g_i, i=1 \dots N_2\}$ the unit step response coefficients. In matrix notation, equation (7) becomes

$$Y = GU + Y_{free} \quad (9)$$

where

$$Y = [y(t+1/t) \quad \dots \quad y(t+N_2/t)]^T \quad (10)$$

$$U = [\Delta u(t/t) \quad \dots \quad \Delta u(t+N_u-1/t)]^T \quad (11)$$

$$Y_{free} = [y_{free}(t+1/t) \quad \dots \quad y_{free}(t+N_2/t)]^T \quad (12)$$

$$G = \begin{bmatrix} g_1 & 0 & \dots & 0 \\ g_2 & g_1 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ g_{N_2} & g_{N_2-1} & \dots & g_{N_2-N_u-1} \end{bmatrix} \quad (13)$$

Using the model (9) in (6), a quadratic relation with respect to U is obtained:

$$J = [\Delta \cdot W - \Delta(G \cdot U + Y_{free})]^T [\Delta \cdot W - \Delta(G \cdot U + Y_{free})] + (\Lambda \cdot U)^T (\Lambda \cdot U) \quad (14)$$

where

$$W = [w(t+1/t) \quad \dots \quad w(t+N_2/t)]^T \quad (15)$$

$$\Delta = \text{diag} [\delta(1) \quad \dots \quad \delta(N_2)] \quad (16)$$

$$\Lambda = \text{diag} [\lambda(0) \quad \dots \quad \lambda(N_u - 1)] \quad (17)$$

Only the first component of the vector U , $\Delta u(t/t) = \Delta u(t)$, is used. At the next sampling instant the whole procedure is repeated.

4 The internal model of the process

Artificial neural networks form an important class of nonlinear systems, with many applications in modeling and control. As mathematically proven (see [11]), any static continuous nonlinear function can be approximated arbitrary well over a compact interval by a multilayer neural network with one or more hidden layers.

In this contribution a feedforward neural network is used to model the behavior of the wastewater treatment process. The proposed neural network has three layers: the first one contains 15 neurons, the second one 7 neurons and the output layer 4 neurons. To appropriately capture the interconnections between all variables, up to four time-delayed values of the inputs and states were supplied to the network. The neural model predicts $X(t)$, $S(t)$, $X_r(t)$ and $DO(t)$ as functions of:

$$\begin{aligned} & D(t-1), \quad D(t-2), \quad D(t-3), \quad W(t-1), \quad W(t-2), \quad W(t-3) \\ & X(t-1), \quad X(t-2), \quad X(t-3), \quad X(t-4), \quad S(t-1), \quad S(t-2), \quad S(t-3) \\ & X_r(t-1), \quad X_r(t-2), \quad X_r(t-3), \quad X_r(t-4), \quad DO(t-1), \quad DO(t-2), \quad DO(t-3) \end{aligned}$$

The data used to train the neural network was obtained by integrating the differential equations (1)-(4), considering randomly varying dilution rates in the interval $[0, 0.1]$ and randomly varying aeration rates in the interval $[0, 100]$. Before training, the data was scaled to the interval $[0, 1]$. In the same manner, a second data set was generated and used to validate the accuracy of the model.

As figure 5 shows, there is hardly any difference between the measured values from the process and the ones predicted by the neural network for the dissolved oxygen and substrate concentrations. There is a noticeable shift between the biomass calculated based on the differential equations of the process and the biomass predicted by the neural network, but this is not going to affect the performance of the predictive controller since the neural network is used to predict the dissolved oxygen level.

5 Simulation results

Figure 6 illustrates the control principle. The predictive controller calculates the dilution rate, which forces the dissolved oxygen concentration to follow the setpoint. Controlling the concentration of dissolved oxygen has a benefic effect on the substrate concentration, which is brought within the limits imposed by the law (below 20mg/l).

Various configurations of the predictive controller parameters can be chosen in order to fulfill the control requirements. In these simulations a fast control was pursued, which can be generally achieved for a small prediction horizon, and less attention was paid to the magnitude of control input variations. The controller parameters were: $N_2 = 5$, $N_u = 1$, $\Delta = I_5$, $\lambda = 0$.

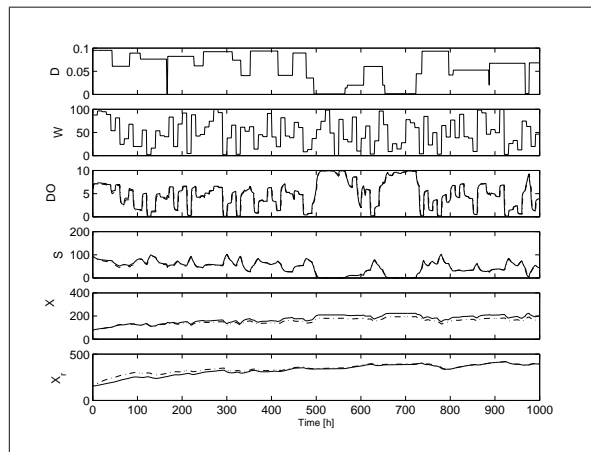


Figure 5: Validation of the neural network model: process - continuous line, model - dash-dotted line

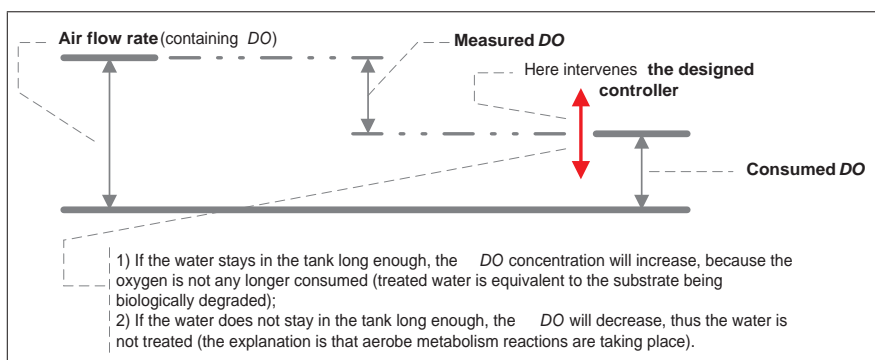


Figure 6: The principle of the dissolved oxygen concentration control

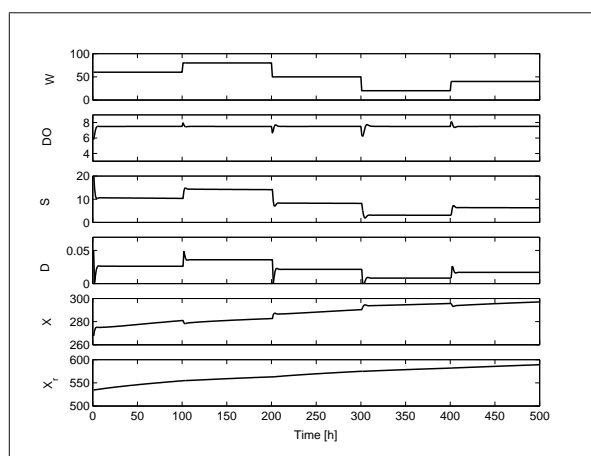


Figure 7: Simulation results for *DO* constant setpoint (7.5)

The results of the dissolved oxygen concentration control are presented in figures 7, 8 and 9. Figure 7 considers a constant dissolved oxygen setpoint (7.5), while figure 8 shows the case when the dissolved oxygen setpoint is variable. In both cases, the aeration rate W varies such that the process covers all three functioning regimes (rain, normal and drought).

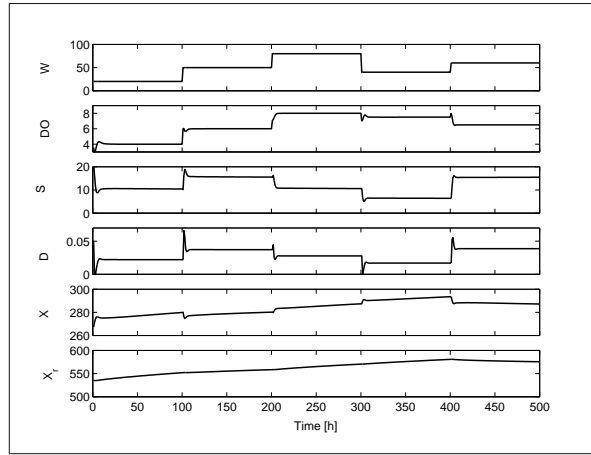


Figure 8: Simulation results for DO variable setpoint

Figure 9 considers a constant dissolved oxygen setpoint but a variable concentration of the substrate in the influent. At time $t = 150h$, S_{in} was changed from 200 mg/l to 300 mg/l and was kept constant to the new value until $t = 250h$, when it was changed to 150 mg/l. The controller adjusts the dilution rate and DO is brought back to the setpoint value.

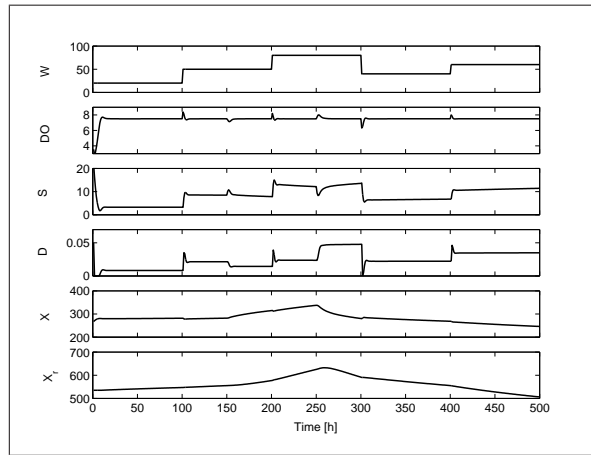


Figure 9: Variable substrate concentration in influent

6 Conclusions

Wastewater treatment is a complex process, which needs control for a good operation. This paper introduces a predictive controller for such a system and evaluates the control performances.

The success of the MPC strategy is highly dependent on a reliable process model, that is a model which approximates well the process dynamics. Taking into account the complexity of the wastewater treatment process, a neural network has been chosen as internal model for the predictive controller.

The simulation results show a good performance of the control loop. The controller manipulates the dilution rate and forces the dissolved oxygen concentration to follow the imposed setpoint. This has a benefic effect on the substrate concentration, which is maintained within the limits established by law. The control is effective for various operational regimes, defined by the aeration rate. Moreover, it is able to reject disturbances that might appear on the substrate concentration in the inflow.

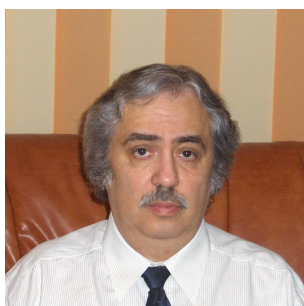
References

- [1] Z. Boger, Application of Neural Networks to Water and Wastewater Treatment Plant Operation, *ISA Trans.*, Vol. 31, pp. 25-33, 1992.
- [2] M. A. Brdys and Y. Zhang, Robust Hierarchical Optimising Control of Municipal Wastewater Treatment Plants, *Preprints of the 9th IFAC/IFORS/IMACS/IFIP Symposium Large Scale Systems: Theory & Applications Ū LSSŚ2001*, Bucharest, Romania, pp. 540-547, July 18-20, 2001.
- [3] A. M. Buswell and H. L. Long, Microbiology and Theory of Activated Sludge, *Journal of American Pollution Wks. Assn.*, Vol. 10, 1923.
- [4] E. F. Camacho and C. Bordons, *Model predictive control*, Springer-Verlag, London, 1999.
- [5] E. Choi et al., High strength nitrogen removal from nightsoil and piggery wastes, *Proc. of the 6th IWA Speciality Symposium on Strong Nitrogenous and Agro-Wastewater*, Seoul, Korea, 11-13 June, 2003.
- [6] R. De Keyser, Model Based Predictive Control, Invited Chapter in "*UNESCO Encyclopedia of Life Support Systems (EoLSS)*", EoLSS Publishers Co. Ltd., Oxford (www.eolss.net), 2003.
- [7] B. L. Goodman and A. J. Englande, A Unified Model of the Activated Sludge Process, *Journal of Water Pollution Control Fed.*, Vol. 46, pp. 312-332, 1974.
- [8] M. Henze et. al., Activated Sludge Model No. 1, *IAWQ Scientific and Technical Report No. 1*, IAWQ, London, Great Britain, 1987.
- [9] M. Henze et al., Activated Sludge Model No. 2, *IAWQ Scientific and Technical Report No. 3*, IAWQ, London, Great Britain., 1995.
- [10] M. Henze et al., *Activated Sludge Models ASM1, ASM2, ASM2d and ASM3*, IWA Publishing, London, Great Britain, 2000.
- [11] K. Hornik, M. Stinchcobe and H. White, Multilayer feedforward networks are universal approximators, *Neural Networks*, Vol. 2, pp. 359-366, 1989.
- [12] P. Ingildsen, *Realising Full-Scale Control in Wastewater Treatment Systems Using In Situ Nutrient Sensors*, Ph.D. thesis, Dept. of Industrial Electrical Engineering and Automation, Lund University, Sweden, 2002.
- [13] U. Jeppsson, *Modelling Aspects of wastewater treatment processes*, Ph.D. thesis, Dept. of Industrial Electrical Eng. and Automation, Lund University, Sweden, 1996.
- [14] M. R. Katebi, M.A. Johnson and J. Wilke, *Control and Instrumentation for Wastewater Treatment Plant*, Springer-Verlag, London, 1999.

- [15] R. E. King and A. Stathaki, A multi-layer perceptron for the control of a wastewater treatment plant, *12th Mediterranean Conference on Control and Automation - MED2004*, Kusadasi, Turkey, Proc. CD-ROM, June 6-9, 2004.
- [16] H. Kroiss, What is the potential for utilizing the resources in sludge, *Water Science and Technology*, Vol. 49, pp. 1-10, 2003.
- [17] G. Langergraber et al., Monitoring of a paper mill wastewater treatment plant using UV/VIS spectroscopy, *Water Science and Technology*, Vol. 49, pp. 9-14, 2003.
- [18] S. A. Manesis, D. J. Sapidis and R. E. King, Intelligent Control of Wastewater Treatment Plants, *Artificial Intelligence in Engineering*, Vol. 12, pp. 275-281, 1998.
- [19] F. Nejjari, A. Benhammou, B. Dahhou and G. Roux, Non-linear multivariable adaptive control of an activated sludge wastewater treatment process, *Int. J. Adapt. Control Signal Process.*, Vol. 13, pp. 347-365, 1999.
- [20] J. T. Novak and C. Park, Chemical Conditioning of Sludge, *Proc. of Int. Conf. on Wastewater Sludge as a Resource - Biosolids 2003*, Trondheim, Norway, 23-25 June, 2003.
- [21] S. Yagi, H. Kohara, Y. Nakamura and S. Shiba, Fuzzy Control of a Wastewater Treatment Plant for Nutrients Removal, *Proc. of the Int. Conf. on Artif. Intell. in Engineering & Technology*, Sabah, Malaysia, June 17-18, 2002.

Sergiu Caraman & Marian Barbu
"Dunarea de Jos" University of Galati, Romania
Department of Automatic Control and Electronics
Domneasca Street, no. 47, 800008-Galati, Romania
Sergiu.Caraman, Marian.Barbu@ugal.ro

Mihaela Sbarciog
Ghent University, Belgium
Department of Electrical Energy, Systems and Automation
Technologiepark 913, Zwijnaarde-Ghent, 9052, Belgium
mihaela@autoctrl.UGent.be
Received: February 19, 2007



Sergiu Caraman (born on September 30, 1955) received the License degree in 1980 from Polytechnical University of Bucharest and the PhD degree in 1996 from "Dunarea de Jos" University of Galati, Romania, both in Control Systems. Presently he is professor in the Faculty of Electrical Engineering and Electronics, "Dunarea de Jos" University. His research interest includes topics such as modelling and control of biotechnological processes, intelligent control.



Mihaela Sbarciog (born on May 21, 1976) graduated the Faculty of Naval and Electrical Engineering of “Dunarea de Jos” University of Galati in 1999. Presently she is a research assistant at Faculty of Engineering, Ghent University. Her main research field is the analysis and control of biochemical reaction systems.



Marian Barbu (born on March 22, 1978) received the License degree in 2001 and the PhD degree in 2006 from “Dunarea de Jos” University of Galati, both in Control Systems. Presently he is lecturer in the Faculty of Electrical Engineering and Electronics, “Dunarea de Jos” University. His research interest includes topics such as modelling and control of wastewater treatment processes, robust and intelligent control.

Self-Organizing Maps for Analysis of Expandable Polystyrene Batch Process

Mikko Heikkinen, Ville Nurminen, Yrjö Hiltunen

Abstract: Self-organizing maps (SOM) have been successfully applied in many fields of research. In this paper, we demonstrate the use of SOM-based method for the analysis of Expandable PolyStyrene (EPS) batch process. To this end, a data set of EPS-batch process was used for training a SOM. Reference vectors of the SOM were then classified by K-means algorithm into six clusters, which represent product types of the process. This SOM could also be used for estimating the optimal amounts of the stabilisation agent. The results of a validation data set showed a good agreement between the actual and estimated amounts of the stabilisation agent. Based on this model a Web application was made for test use at the plant. The results indicate that the SOM method can also be efficiently applied to the analysis of the batch process.

Keywords: Neural networks, self-organizing maps, process control, batch process

1 Introduction

Batch processes are typically based on predefined process recipes. If process circumstances, chemicals and recipes are constant, the product should basically be always the same. A batch process is also commonly used for producing Expandable PolyStyrene (EPS). However, in practice this polymerisation reaction is a very sensitive process and numerous variables affect it, which makes the process difficult to control. The EPS production has to be able to follow fast the aims and quality requirements of the market, which causes additional demands on the process control.

Archived process data is an important resource for the knowledge management of the process and it can be used for the optimization and improvement of productivity. Recent applications have demonstrated that artificial neural networks can provide an efficient and highly automated method for modelling industrial data [1], [2]. In particular, studies, which use standardized protocols, are most likely to benefit from automated ANN analysis [1], [2]. Self-organizing maps [1], [3]-[5] have also been successfully applied in many areas of research and are thus a tool for process optimization. The SOM method offers an efficient means of handling complex multidimensional data, which is typically the situation in industrial applications. In addition, the SOM method is robust for missing values of data.

Here, we apply self-organizing maps to the analysis of an EPS-batch process. The optimal amounts of the stabilisation agent can be estimated using the SOM model. In the study we have also included some features of the supervised approach in the designed unsupervised method.

2 Methods

2.1 The process and the data

The studied process was a typical suspension polymerisation batch process, which is commonly used for producing EPS (Expandable PolyStyrene). The polymerisation stage is executed in a pressure-temperature range below the boiling point of styrene-water suspension system. After the polymerisation stage the process continues into the impregnation stage, where the blowing agent is impregnated into the beads. The impregnation stage assumed to be negligible in the means of bead size distribution.

The biggest challenge in the suspension polymerisation process is to achieve the required bead size distribution. It is common knowledge that the basic variables in the term of the bead size are the mixing

properties and the amount and quality of the suspension stabilizers. However the suspension polymerisation of styrene is a very sensitive process and numerous variables affect it. Most of these variables cannot be measured or followed by in a reasonable way. For example to analyse all impurities from all raw materials is too heavy a task for any industrial laboratory. Some variables are quite easily measurable, but have not been traced due to the assumption that they would not have a significant contribution to the process. To be able to model the process the studied system required elimination of the variables, which were assumed to be inessential.

The data for the model were divided into three groups: recipe, results and process parameters. Process parameters, such as actual reactor temperature, were measured and stored automatically from each batch every one minute. Process parameter data was not used for modelling in the first part of this study and the target is to add it later to the model. Table I shows the recipe and result variables.

The data contained 15 production campaigns and about 450 batches in 4 reactors. The data had to be divided into two separate groups due to the process changes in the stabilisation system. These changes limited the size of useful data to 251 batches, which is a low amount for accurate modelling. Fortunately more production campaigns will be executed almost every month and more data will be available for further studies.

2.2 Computational methods

Self-organizing maps

Self-organizing maps (SOMs) are an artificial neural network methodology, which can transform an n-dimensional input vector into a one- or two-dimensional discrete map. The input vectors, which have common features, are projected to the same area of the map e.g. (in this case described as \check{S} neurons \check{T}). Each neuron is associated with an n-dimensional reference vector, which provides a link between the output and input spaces. During learning, the input data vector is mapped onto a particular neuron (best matching unit, BMU) based on the minimal n-dimensional distance between the input vector and the reference vectors of the neurons. Then the reference vectors of the activated neurons are updated. When the trained map is applied, the best matching units are calculated using these reference vectors. In this unsupervised methodology, the SOM can be constructed without previous a priori knowledge [1].

The data were coded into 11 inputs for the SOM. All input values were variance scaled. The SOM having 676 neurons in a 26x26 hexagonal arrangement was constructed. The linear initialization and

Table 1: Recipe and result variables of the particular StyroChem suspension polymerisation process.

<i>Recipe Variables</i>	<i>Unit</i>	<i>Result Variables</i>	<i>Unit</i>
Amount of stabilisation agent	% from MS	Under sized	%
Mixing speed	rpm	Product A	%
Polymerisation temperature	°C	Product B	%
Reactivity	min	Product C	%
Amount of styrene	kg	Product D	%
Additional stabilisation	kg	Product E	%
Product type	-	Product F	%
Batch number in campaign	-	Over sized	%
Reactor	-	Mean particle size	mm
		Delta	mm
		Narrowness	-

batch training algorithms were used in the training of the map. A Gaussian function was used as the neighbourhood function. The map was taught with 10 epochs and the initial neighbourhood had the value of 6. The SOM Toolbox [7] was used in the analysis under a Matlab-software platform (Mathworks, Natick, MA, USA).

The data set of an EPS-batch process (n= 251 batches) was divided into two subsets. The first subset (the first 190 batches) was the training set, which was used for training the map. The other subset (the last 61 batches) was the test set. Variables of each batch and amounts of the stabilisation agent were used as an input for the SOM in the training phase. However, the amounts of the stabilisation agent were missing values in the test phase, i.e. best matching neurons for the test set were sought using only those other batch variables specified above. Estimated values for the amounts of the stabilisation agent were obtained from the reference vectors of neurons.

K-means method

The K-means algorithm was applied to the clustering of the map. The K-means method is a well-known non-hierarchical cluster algorithm [8]. The basic version begins by randomly picking K cluster centers, assigning each point to the cluster whose mean is closest in a Euclidean distances sense, then computing the mean vectors of the points assigned to each cluster, and using these as new centers in an iterative approach.

3 Results and discussion

The map was obtained by training a self-organizing network with the training set of an EPS-batch process. The map and the six clusters calculated by the K-means method are shown in Figure 1. These clusters represent different types of the products. The brief descriptions of the clusters are also illustrated in Figure 1.

The method was validated by using the test set, i.e. the last 61 data vectors were in test. The results, when the SOM was trained to estimate the amounts of the stabilisation agent, are illustrated in Fig. 2. The correlation coefficient between the actual amount of the stabilisation agent and the estimated one was 0.851 (Fig 2 a). In Fig. 2 b) the actual and estimated amounts of the stabilisation agent are shown as a function of the batch number.

Table 2: Variables used for modelling.

<i>Variables</i>
Under sized
Product A
Product B
Product C
Product D
Product E
Mean particle size
Batch number in campaign
Amount of stabilisation agent in previous batch
Amount of stabilisation agent

A Web application based on this SOM model was made for estimating the optimal amount of stabilisation agent in practice at the plant. Figure 3 illustrates the interface of this application. The same

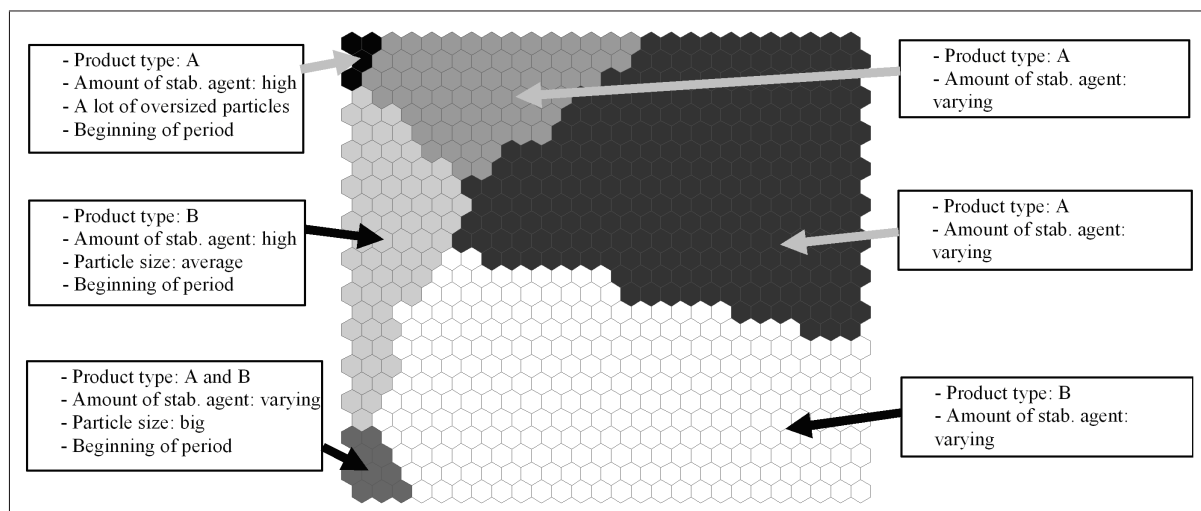


Figure 1: SOM using the data of an EPS-batch process. The background colours visualize the six clusters of the map. Short descriptions for each cluster are also shown.

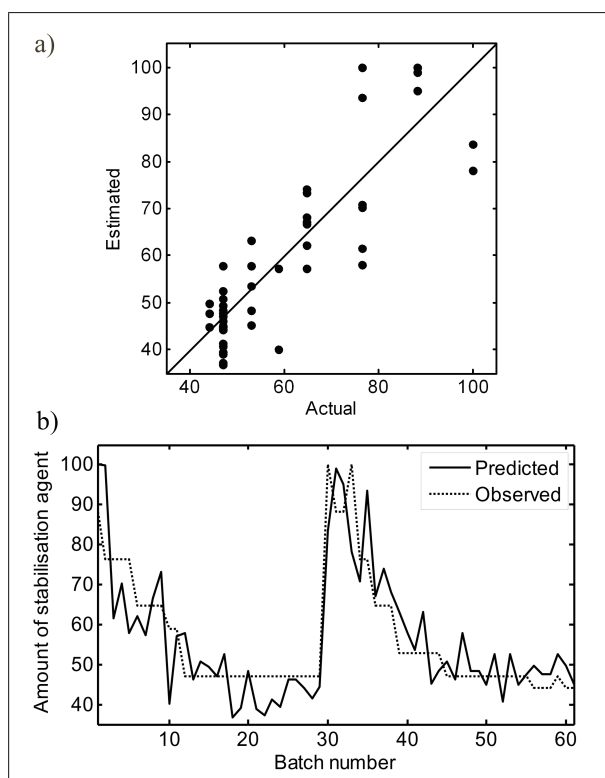
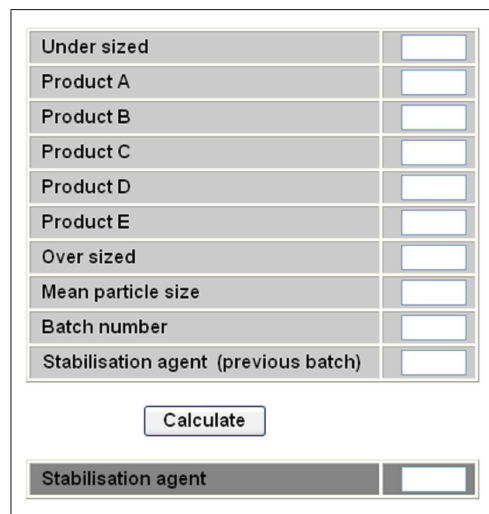


Figure 2: a) The correlation between the actual amount of the stabilisation agent and the one estimated by the SOM analysis. The results are from the test set (the last 61 batches) and the value for the correlation coefficient is 0.851. b) The actual and estimated amounts of the stabilisation agent as a function of the batch number.



Under sized	<input type="text"/>
Product A	<input type="text"/>
Product B	<input type="text"/>
Product C	<input type="text"/>
Product D	<input type="text"/>
Product E	<input type="text"/>
Over sized	<input type="text"/>
Mean particle size	<input type="text"/>
Batch number	<input type="text"/>
Stabilisation agent (previous batch)	<input type="text"/>

Stabilisation agent	<input type="text"/>
---------------------	----------------------

Figure 3: The interface of the Web application.

variables, which are shown in Table 2, have been used in this application. A user gives as many as possible out of the ten upmost variables and the application calculates the lowest one. The first seven variables define the target amounts of the product.

Our earlier results showed that the SOM method could be successfully applied to process state monitoring and optimization of NO_x emissions in the case of a continuous process [6]. The results presented here illustrate also the advantages of using SOM method in the analysis of a batch process. Figure 3 shows that the method can be used for estimation of optimal parameters of the process. It seems to reduce varieties of the process and so helps to get better products. Because our SOM method is also non-sensitive for the presence of missing values, it is feasible in the analysis of industrial data. Furthermore, SOM analysis does not require extensive knowledge of neural nets and it can easily be included in any kind of software. An attractive property of SOM is also that it can be retrained, if new product types of the process are to be analysed.

4 Conclusion

The SOM analysis provides an efficient and automated method for data analysis in the process industry. The present study shows that this kind of data-driven approach is a fruitful way of analysing a batch process.

5 Acknowledgement

This research was supported by StyroChem Ltd.

References

- [1] T. Kohonen, *Self-organizing Maps*, Springer-Verlag, Berlin Heidelberg New York, 2001.
- [2] S. Haykin, *Neural Networks: A Comprehensive Foundation*, Upper Saddle River, NJ: Prentice Hall, 1999.

- [3] J. Kaartinen, Y. Hiltunen, P. T. Kovanen, M. Ala-Korpela, Classification of Human Blood Plasma Lipid Abnormalities by 1H Magnetic Resonance Spectroscopy and Self-Organizing Maps, *NMR Biomed.*, Vol. 11, pp. 168-176, 1998.
- [4] M. T. Hyvönen, Y. Hiltunen, W. El-Deredy, T. Ojala, J. Vaara, P. T. Kovanen, M. Ala-Korpela, Application of Self-Organizing Maps in Conformational Analysis of Lipids, *Journal of the American Chemical Society*, Vol. 123, pp. 810-816, 2001.
- [5] M. Heikkinen, M. Kolehmainen, Y. Hiltunen, Classification of process phases using Self-Organizing Maps and Sammon's mapping for investigating activated sludge treatment plant in a pulp mill, *Proceedings of the Fourth European Symposium on Intelligent Technologies and their implementation on Smart Adaptive Systems*, pp. 281-297, 2004.
- [6] M. Heikkinen, A. Kettunen, E. Niemitalo, R. Kuivalainen, Y. Hiltunen, SOM-based method for process state monitoring and optimization in fluidized bed energy plant, ICANN 2005, *Lecture Notes in Computer Science 3696*, Eds. W. Duch, J. Kacprzyk, E. Oja, S. Zadrozny, Springer-Verlag Berlin Heidelberg, pp. 409-414, 2005.
- [7] Homepage of SOM toolbox, Helsinki University of Technology, Laboratory of Computer and Information Science (CIS), <http://www.cis.hut.fi/projects/somtoolbox/>.
- [8] J. MacQueen, Some methods for classification and analysis of multivariate observations, *In Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability*, Vol. I: Statistics, University of California Press, Berkeley and Los Angeles, pp. 281-297, 1967.

Mikko Heikkinen

University of Kuopio, Department of Environmental Sciences
P.O. Box 1627, FIN - 70211 Kuopio, Finland
E-mail: Mikko.Heikkinen@uku.fi

Ville Nurminen

StyroChem Ltd
P.O. Box 360, FIN - 06101 Porvoo, Finland
E-mail: Ville.Nurminen@styrochem.com

Yrjö Hiltunen

University of Kuopio, Department of Environmental Sciences
P.O. Box 1627, FIN - 70211 Kuopio, Finland
E-mail: Yrjo.Hiltunen@uku.fi

Received: March 6, 2007

Blind Steganalysis: Estimation of Hidden Message Length

Sanjay Kumar Jena, G.V.V. Krishna

Abstract: Steganography is used to hide the occurrence of communication. Discovering and rendering useless such covert message is an art of steganalysis. The importance of techniques that can reliably detect the presence of secret messages in images is increasing as images can hide a large amount of malicious code that could be activated by a small Trojan horse type of virus and also for tracking criminal activities over Internet. This paper presents an improved blind steganalysis technique. The proposed algorithm reduces the initial-bias, and estimates the LSB embedding message ratios by constructing equations with the statistics of difference image histogram. Experimental results show that this algorithm is more accurate and reliable than the conventional difference image histogram method. It outperforms other powerful steganalysis approaches for embedded ratio greater than 40% and comparable with RS steganalysis technique for shorter hidden message length.

Keywords: Steganography, steganalysis, hidden message extraction

1 Introduction

Steganography is the art of passing information through apparently innocent files in a manner that the very existence of the message is unknown. The term steganography in Greek literally means, "Covered Writing". The innocent files can be referred to as cover text, cover image, or cover audio as appropriate. After embedding the secret message it is referred to as stego-medium. A stego-key is used to control the hiding process so as to restrict detection and/or recovery of the embedded data. While cryptography is about protecting the content of messages (their meaning), steganography is about hiding the message so that intermediate persons cannot see the message.

Historically, steganography has been a form of security through obscurity where the security lies in that only sender and receiver know the method in which the message is hidden. This is in violation of Kirchoff's principle, which states that the security should lie in key alone. Steganography can be either "linguistic steganography" or "technical steganography" [1]. The ancient techniques that hide messages physically are called as technical steganographic systems. They include microdots, tattoos, invisible inks and semagrams. Recent techniques belong to the linguistic steganography. These techniques hide message in the cover images, which are of digital form.

Steganalysis is the process of detecting the existence of the steganography in a cover medium and rendering it useless. Current trend in steganalysis [4] seems to suggest two extreme approaches (a) little or no statistical assumptions about the image under investigation where statistics are learnt using a large database and (b) a parametric model is assumed for the image and its statistics are computed for steganalysis detection. The messages embedded into an image are often imperceptible to human eyes. But there exists some detectable artifacts in the images depending on the steganographic algorithm used [2,5]. The steganalyst uses these artifacts for the detection of the steganography.

By far the most popular and frequently used steganographic method is the Least Significant Bit embedding (LSB). It works by embedding message bits as the LSB's of randomly selected pixels. Several techniques for the steganalysis of the images for LSB embedding are present. Fridrich and Goljan [6,7] proposed LSB Steganography dual detection method, named RS method, based on probability statistics in the color or grayscale images. The basic idea is that LSB plane seems random in the typical cover images, but to some extent the other 7 bit planes could predict it. This method is suitable for detection of the non-sequential steganography reliably.

Pfitzmann and Westfeld [8] introduced a method based on statistical analysis of Pairs of Values (PoVs)

that are exchanged during message embedding. This method, which became known as the χ^2 attack, is quite general and can be applied to many embedding paradigms besides the LSB embedding. It provides very reliable results when the message placement for sequential embedding. Fridrich et al. [9] developed a steganographic method for detecting LSB embedding in 24-bit color images—the Raw Quick Pairs (RQP) method. The new method is based on analyzing close pairs of colors created by LSB embedding. On the condition that the number of unique colors in the cover image will be less than 30 percent that of the total pixels, it works reasonably well. When the number of unique colors exceeds about 50 percent that of total pixels, the results gradually become unreliable. This frequently happens for high resolution raw scans and images taken with digital cameras stored in an uncompressed format. Another disadvantage of the RQP method is that it can't be applied to grayscale images.

There are few papers in the field of the detecting pixels that contain the hidden message. Ian Davidson and Goutam Paul [10] proposed the hidden message location problem as outlier detection using probability/energy measures of images. Pixels contributing the most to the energy calculations of an image are deemed outliers. Though results for grayscale images are quite accurate; they are not as good as for color images. The algorithm can be defeated if the steganography algorithm has knowledge of probability/energy function or if the message is carefully embedded in the high-energy regions of an image.

The difference image histogram proposed by T.Zhang and X.Ping [10] consists of an initial-bias. The proposed algorithm constructs the embedding ratio-estimate equations using the difference image histogram and reduces the initial bias. Experimental results show that the novel algorithm is more accurate than the conventional difference image histogram method and other steganalysis techniques.

In the following section, we review principle of the difference image histogram method, and then in section 3, describe the improved difference image histogram method (IDIH) algorithm. Sections 4 show the experimental results and conclude this paper in section 5.

2 Principles Of Difference Image Histogram

Tao Zhang and Xijian Ping introduced the difference image histogram method, which uses the measure of weak correlation between successive bit planes to construct a classifier for discrimination between stego-images and cover images. Considering the property of LSB steganography, the difference image histogram is used as a statistical analysis tool. The difference image is defined as

$$D(i, j) = I(i + 1, j) - I(i, j). \quad (1)$$

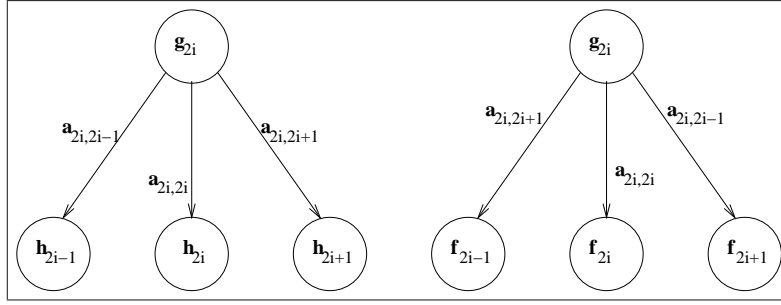
Where $I(i, j)$ denotes the value of the image I at the position (i, j) . T.Zhang and X.Ping found that there exists difference between the difference image histograms for normal image and the image obtained after flipping operation on the LSB plane. This fact is utilized to realize the steganalysis technique. To explain the details of difference image histogram method (DIH), we need to define some notions, Let I be the test image, which has $M \times N$ pixels. The embedding ratio p is defined as the percentage of the embedded message length to the maximum capacity.

If the difference image histogram of an image is represented by h_i , that of the image after flipping all bits in the LSB plane by f_i , and that of the image after setting all bits in the LSB plane to zero by g_i , there exist the following relationships between h_i , f_i and g_i .

$$\begin{aligned} h_{2i} &= f_{2i} = a_{2i,2i}g_{2i}, \\ h_{2i+1} &= a_{2i,2i+1}g_{2i} + a_{2i+2,2i+1}g_{2i+2}, \\ f_{2i+1} &= a_{2i,2i-1}g_{2i} + a_{2i+2,2i+3}g_{2i+2} \end{aligned} \quad (2)$$

in which $a_{2i,2i+j}$ is defined as the transition coefficient from the histogram g_i to h_i . When $j = 0, 1, -1$ then $0 < a_{2i,2i+j} < 1$, otherwise $a_{2i,2i+j} = 0$, and they satisfy

$$a_{2i,2i-1} + a_{2i,2i} + a_{2i,2i+1} = 1 \quad (3)$$

Figure 1: The transition diagram from g_i to h_i, f_i

Starting from the approximate symmetry of the difference histogram about $i = 0$, we first get $a_{0,1} \sim a_{0,-1}$. From the above Equations (2) we obtain the following iterative formula for calculating transition coefficients for all positive integers i :

$$\begin{aligned}
 a_{0,1} = a_{0,-1} &= \frac{g_0 - h_0}{2g_0}, \\
 a_{2i,2i} &= \frac{h_{2i}}{g_{2i}}, \\
 a_{2i,2i-1} &= \frac{h_{2i-1} - a_{2i-2,2i-1}g_{2i-2}}{g_{2i}}, \\
 a_{2i,2i+1} &= 1 - a_{2i,2i} - a_{2i,2i-1}.
 \end{aligned} \tag{4}$$

Assuming the embedded hidden message forms a random bit sequence, for the stego-image with the LSB plane fully embedded (i.e. $p = 100\%$) the LSB plane is independent of the neighboring bit planes. Therefore, for such stego-images we have $a_{2i,2i+1} \approx 0.25$, $a_{2i,2i} \approx 0.5$, $a_{2i,2i-1} \approx 0.25$.

The h_{2i+1} consists of two parts: $a_{2i,2i+1}g_{2i}$ and $a_{2i+2,2i+1}g_{2i+2}$. Statistical tests show that for natural images these two parts make an approximately equal contribution to h_{2i+1} , that is

$$a_{2i,2i+1}g_{2i} \approx a_{2i+2,2i+1}g_{2i+2} \tag{5}$$

If $\alpha_i = (a_{2i+2,2i+1})/(a_{2i,2i+1})$, $\beta_i = (a_{2i+2,2i+3})/(a_{2i,2i-1})$ and $\gamma_i = g_{2i}/g_{2i+2}$ then the statistical hypothesis of the steganalytic method is that for a natural image the following equation should be satisfied:

$$\alpha_i \approx \gamma_i \tag{6}$$

while for stego-images with the LSB plane fully embedded

$$\alpha_i \approx 1. \tag{7}$$

The physical quantity α_i , can be viewed as the measure of the weak correlation between the LSB plane and its neighboring bit planes. From further experiments they got that for a given i the value of α_i , decreases monotonically with the increasing length of embedded secret messages (p) and when the embedding ratio p increases to 100%, α_i decreases to 1 approximately. Figure-2 shows the functional relation between α_i and the embedding ration p when $i = 0$ for the ‘‘Lena’’ image.

The relationship between α , and the embedding ratio p will be modeled using a quadratic equation $y = ax^2 + bx + c$. The following four critical points $P_1 = (0, \gamma_i)$, $P_2 = (p, \alpha_i)$, $P_3 = (1, 1)$, and $P_4 =$

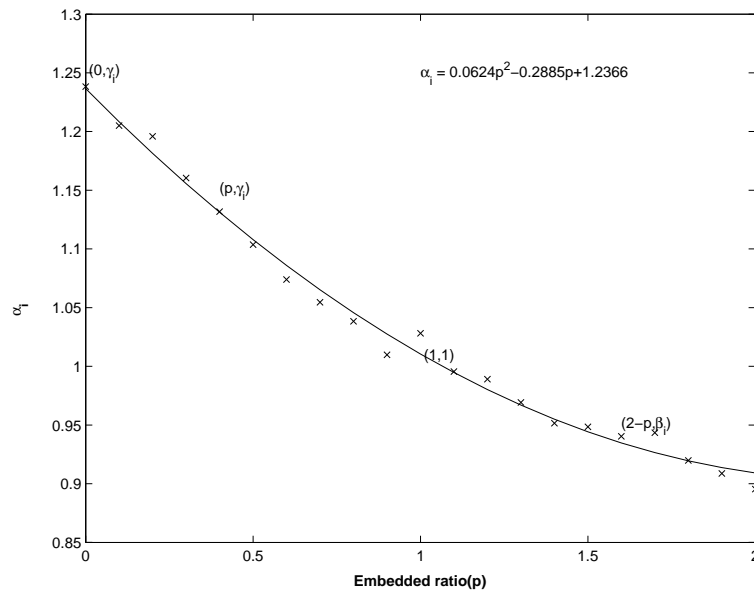


Figure 2: The functional relation between α_i and $p(i=0)$

$(2-p, \beta_i)$. Now the following equation set is obtained:

$$\begin{aligned}
 c &= \gamma_i, \\
 ap^2 + bp + c &= \alpha_i, \\
 a + b + c &= 1, \\
 a(2-p)^2 + b(2-p) + c &= \beta_i.
 \end{aligned} \tag{8}$$

Assume $d_1 = 1 - \gamma_i$, $d_2 = \alpha_i - \gamma_i$, $d_3 = \beta_i - \gamma_i$, then above equation set (8) can be simplified to

$$2d_1p^2 + (d_3 - 4d_1 - d_2)p + 2d_2 = 0 \tag{9}$$

The embedding ratio p can be obtained from the root of above whose absolute value is smaller. If the discriminant is smaller than zero, then $p \approx 1$.

3 Principles of Improved Difference Image Histogram steganalysis

The Difference image histogram algorithm was primarily based on the statistical hypothesis that for the natural images

$$\alpha_i \approx \gamma_i \tag{10}$$

and for a stego-images with the LSB plane fully embedded

$$\alpha_i \approx 1. \tag{11}$$

Obviously, the hypotheses given in Equations(10) and (11) will affect the precision of the Difference image histogram method. Once in these hypotheses there exists some initial bias, the estimate value via the Equation(9) will not be reliable. When the embedding ratio is low, the bias of these hypotheses will lead the incorrect decision, and if there are no embedding messages in images, the false alarm rate is high. Table 1 will show the mean and variance of the γ_i to α_i value. With the increase in i the variance increases and the mean begins to deviate from 1. In some cases the detection lead to an incorrect decision

	i=0	i=1	i=2	i=3
mean	1.0013	1.0034	1.0079	1.0443
variance	7.6E-04	1.4E-03	2.8 e-03	4.9E-03

Table 1: Statistical data on the ratio of γ_i to α_i for natural images

of estimating more than 1% embedding, for the normal images.

The Figure "Proposed" shows the initial value of difference between α_i and γ_i for a lena image and Figure "Proposed (b)" shows a close look at the α_i and γ_i at $p = 0$ values. This initial deviation may lead a serious estimate error. The initial-bias in detection of the message in the normal image is affecting the detection of stego-images, as the error present in the normal image will effect the estimation of the hidden message length for stego-image.

If the stego-image created with embedded embedding ratio p is denoted as S_p , and the image created by flipping all bits in the LSB plane of S_p as R_p , the value of α_i can be calculated for the images S_p and R_p (note that the value of α_i for the images R_p is equal to the value of β_i for the image S_p). Moreover, we note that in S_p only $p/2$ of the pixels are flipped by message embedding, while in R_p about $1 - (p/2)$ of pixels are flipped. Therefore, R_p is equivalent to a "stego-image" with "embedding ratio" $2 - p$. So given a stego-image we can calculate the values of α_i at p and $2 - p/2$, as the value of β_i at p is equal to the value of α_i at value $2 - (p/2)$.

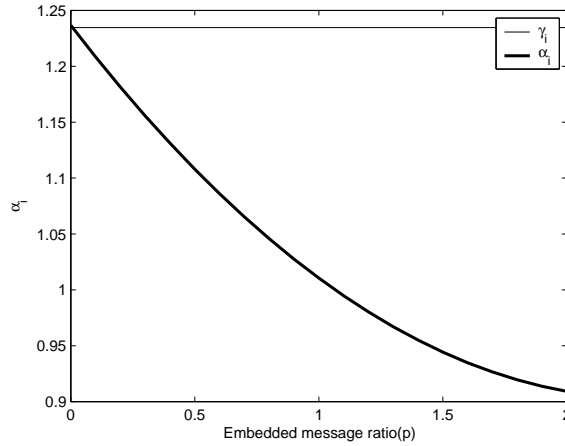


Figure 3: Proposed

Let $\alpha_i(0)$ be the initial value of α_i and $\gamma_i(0)$ be the initial value of γ_i (i.e. when embedding ratio is zero). The error ' ϵ ' be the initial-bias between the γ_i and α_i . So we have

$$\epsilon = \gamma_i(0) - \alpha_i(0). \quad (12)$$

From the difference image histogram method, $\gamma_i = g_{2i}/g_{2i+1}$ where g is the difference image histogram after setting all bits in LSB plane to zero. The grayscale value of pixels in the image will be even as the LSB are set to zero. When the image is embedded with hidden message using the LSB insertion and then performing the operation of setting all LSB plane bits to zero for the stego-image will result in the values of g_{2i} and g_{2i+1} unmodified. Hence

$$\gamma_i(0) = \gamma_i \quad \forall p \quad (13)$$

so the value of error ϵ will become

$$\epsilon = \gamma_i - \alpha_i(0). \quad (14)$$

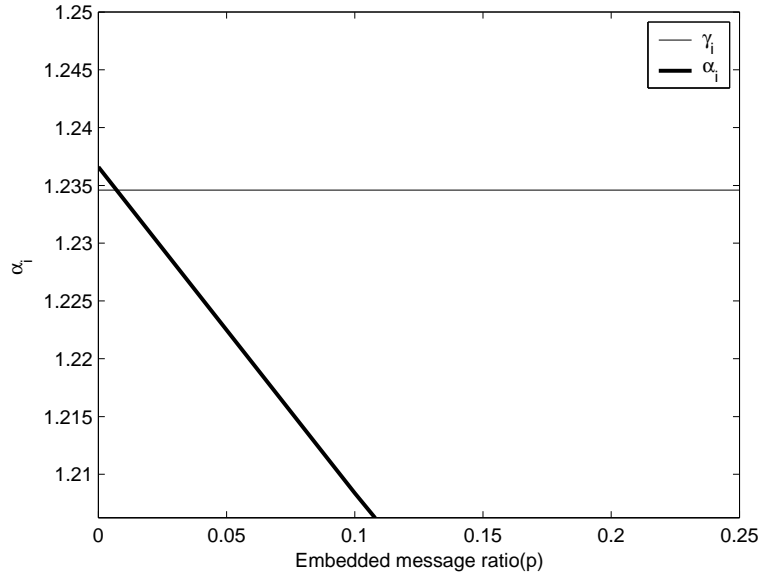


Figure 4: Proposed-(b)

The Difference image histogram models the relationship between α_i and the embedded ratio (p) using a quadratic equation $y = ax^2 + bx + c$. Considering the statistical hypotheses given in Equation (15) to be correct initially, we can find (p, α_i) , $(1, 1)$, $(2 - p, \beta_i)$ are the three points on the curve $y = ax^2 + bx + c$. Now we obtain the following equation set:

$$\begin{aligned} ap^2 + bp + c &= \alpha_i, \\ a + b + c &= 1, \\ a(2 - p)^2 + b(2 - p) + c &= \beta_i. \end{aligned} \quad (15)$$

Assume $e_1 = 1 - p$, $e_2 = 1 - \alpha_i$, $e_3 = 1 - \beta_i$ and then constant value “ c ” in equation 4.21 can be simplified to

$$c = \frac{2e_1^2 - 2e_1e_2 - (e_2 + e_3)(1 - e_1)}{2e_1^2}. \quad (16)$$

The value of ‘ c ’ in Equation (15) will give the $\alpha_i(0)$ for an image. Hence subtracting the error from the estimated ratio p will remove the initial bias in the image. Hence the new estimated ratio “ $p_{modified}$ ” will be

$$p_{modified} = p - \varepsilon \quad (17)$$

4 Description of IDIH algorithm

We now describe our detection algorithm

Input: A set of BMP images for detecting.

Output: The embedded ratio estimate $p_{modified}$ for each image.

Step 1. Select one image in the image set;

Step 2. Obtain difference image histogram of the image before (h_i) and after flipping the LSB bit planes to “zero” (g_i);

Step 3. Do from the step 4 to 8 for each value of $i = 0, 1, 2$;

Step 4. Calculate the statistical values for the image i.e $\alpha_i = (a_{2i+2, 2i+1}) / (a_{2i, 2i+1})$, $\beta_i = (a_{2i+2, 2i+3}) / (a_{2i, 2i-1})$

and $\gamma_i = g_{2i}/g_{2i+2}$, where the transition co-efficient can be estimated using the following equation

$$\begin{aligned} a_{0,1} &= a_{0,-1} = \frac{g_0 - h_0}{2g_0}, \\ a_{2i,2i} &= \frac{h_{2i}}{g_{2i}}, \\ a_{2i,2i-1} &= \frac{h_{2i-1} - a_{2i-2,2i-1}g_{2i-2}}{g_{2i}}, \\ a_{2i,2i+1} &= 1 - a_{2i,2i} - a_{2i,2i-1}. \end{aligned}$$

Step 5. Obtain the value of ‘p’ from the root of the below equation whose absolute value is smaller.

$$2d_1p^2 + (d_3 - 4d_1 - d_2)p + 2d_2 = 0$$

where $d_1 = 1 - \gamma_i$, $d_2 = \alpha_i - \gamma_i$, $d_3 = \beta_i - \gamma_i$;

Step 6. Calculate the value $\alpha_i(0)$ which represents the estimation of α_i for zero embedded message length using the following equation

$$\alpha_i(0) = \frac{2e_1^2 - 2e_1e_2 - (e_2 + e_3)(1 - e_1)}{2e_1^2}$$

where $e_1 = 1 - p$, $e_2 = 1 - \alpha_i$ and $e_3 = 1 - \beta_i$;

Step 7. Calculate the initial bias ‘ ε ’ as

$$\varepsilon = \gamma_i - \alpha_i(0).$$

Step 8. Subtract the error ‘ ε ’ from the p to obtain the modified estimation ratio $p_{modified}(i)$.

$$p_{modified}(i) = p - \varepsilon$$

Step 9. The average of $p_{modified}(i)$ for $i=0,1,2$ will give the final embedded ratio $p_{modified}$.

5 Experimental Results

We select 150 standard 512×512 test images (such as Lena, Peppers and so on). Applying random and sequential LSB replacement to embed the images with the ratio of $p=0, 10\%, 20\%, \dots, 90\%, 100\%$ respectively with 10% increments we created two databases. Then we have use the RS method [7], DIH method [11] and GEFR method [12] to estimate the embedding ratio of secret information respectively. The mask used in the RS method is [1,0; 0,1].

The testing results of the test images got by DIH method and the proposed method (IDIH) are shown in Table 2. The leftmost column in Table 2 is the real embedding ratio, and column ‘‘IDIH’’, ‘‘DIH’’ represent the estimate embedding ratio got by Improve Difference Image Histogram method (proposed method) and Difference Image Histogram Method (DIH) respectively. It can be seen in Table 2 that the estimate precision of IDIH is higher than DIH obviously.

Figure [4] and [5] shows the corresponding plot of the embedded message length to the mean absolute error of the estimated values for random embedding and sequential embedding. Figure [4] indicates that the proposed algorithm (IDIH) algorithm outperforms the other three steganalysis techniques for embedded ratios greater than 40%. Improved Difference Image Histogram algorithm has performance comparable to the RS steganalysis for short messages (when p is smaller than 40%). However, because it is harder to detect smaller messages than large messages, the accuracy of the estimate is far more important for smaller message embedding. The proposed algorithm proves to be effective and reliable when complete range of embedding lengths is considered and compared to the existing algorithms.

Embedding ratio(%)	Random		Sequential	
	IDIH	DIH	IDIH	DIH
0%	0.3052	1.6855	0.3052	1.6855
10%	14.7804	15.3881	15.6703	16.0368
20%	20.38	20.80	27.98	28.11
30%	20.3764	20.8017	27.9818	28.1124
40%	40.1524	42.9062	44.3258	44.922
50%	48.6793	52.2864	49.7154	48.5228
60%	62.245	63.8	60.5394	56.5979
70%	72.7311	66.67118	69.7919	68.726
80%	84.6388	73.4632	80.8796	72.2582
90%	90.9915	85.8664	84.8516	81.955
100%	98.6088	92.5193	98.6088	92.5193

Table 2: Comparison between IDIH and DIH

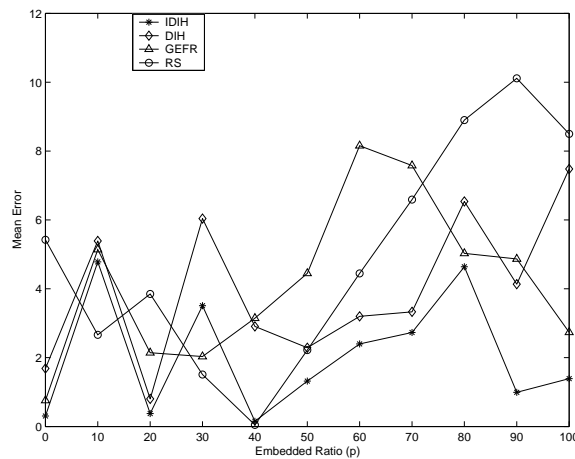


Figure 5: Comparison with other steganalytic techniques for random LSB embedding

In case of sequential embedding (as shown in Figure ??), the accuracy is much higher than the case of random embedding for the embedded ratios of greater than 40%. It is having a higher performance to all the other steganalytic techniques for entire range of possible embedding lengths.

6 Summary and Conclusions

This paper proposes a new detection algorithm, which is an improved algorithm to the difference image histogram algorithm and performed tests on a group of raw lossless images. Experimental results show that the improved difference image histogram steganalysis method is more accurate and reliable than the conventional difference image histogram method. The proposed algorithm reduces the mean error by 50% for embedding ratios greater than 40% when compared to the DIH algorithm.

References

- [1] F.A.P.Petitcolas, R.J.Anderson and M.G.Kuhn, "Information Hiding-A Survey", *Proceedings of the IEEE*, vol.87(7), *special issue on protection of multimedia content*, June 1999, pp.1062-1018.

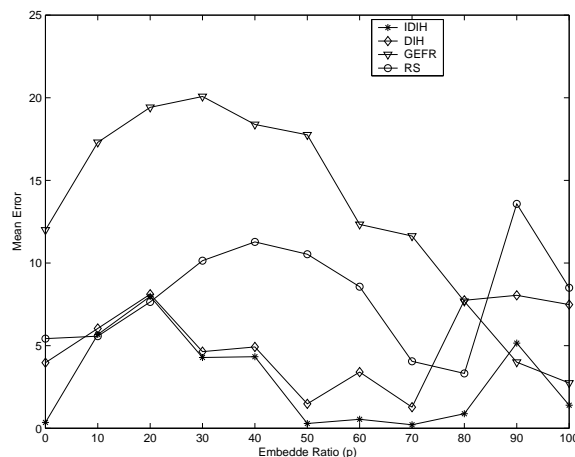


Figure 6: Comp. with other steganalytic techniques for sequential LSB embedding

- [2] N.F.Johnson and S.Jajodia, "Steganalysis of images created using current steganography software", *Lecture Notes in Computer Science*, vol.1525, Springer, Berlin, April 1998, pp.273-289.
- [3] N.F.Johnson, and S.Jajodia, "Exploring steganography: Seeing the unseen ", *IEEE, Computer*, February 1998, pp. 26-34.
- [4] R.Chandramouli and K.P.Subbalakshmi, "Current trends in steganalysis: a critical survey ", *International Control, Automation, Robotics and Vision Conference 2004*, Volume 2, December 2004, pp.964 - 967
- [5] Neil F. Johnson and Sushil Jajodia, "Steganalysis: The Investigation of Hidden Information", *IEEE Information Technology Conference*, September 1998, pp.113-116.
- [6] J.Fridrich, M.Goljan and R.Du, "Detecting LSB Steganography in Color and Grayscale Images", *IEEE*, vol.8(4) *Multimedia*, October-December 2001, pp.22-28.
- [7] J. Fridrich, M. Goljan, R. Du, "Reliable detection of LSB Steganography in grayscale and color images", *Proceeding of ACM, Special Session on Multimedia Security and Watermarking*, Ottawa, Canada, 2001, pp. 27-30.
- [8] A.Westfeld, A. Pitzmann, "Attacks on steganographic systems", *Information hiding, Third International Workshop, IH'99*, Dresden, Germany, 29 September-1 October, 1999.
- [9] J.Fridrich, R.Du and M.Long, "Steganalysis of Lsb Encoding in Color Images', *In Proceedings of ICME 2000*, July-August 2000.
- [10] T.Zhang and X.Ping, "Reliable detection of LSB steganography based on the difference image histogram", *IEEE International Conference on Acoustics, Speech, and Signal Processing*, Volume 3, April 2003, pp.545-548.
- [11] Li Zhi and Sui Ai Fen, "Detection of random LSB image steganography", *Vehicular Technology Conference IEEE 60th*, Vol.3, September 2004, pp.2113 - 2117.
- [12] A.Pinnet, *SecurEngine professional v1.0*, 2004, <http://securengine.isecuqrelabs.com>.
- [13] CBIR image database, University of washington, available at <http://www.cs.washington.edu/research/imagetadbse/groundtruth/>.

[14] The USC-SIPI image database, available at: <http://sipi.usc.edu/services/database/Database.html>.

Sanjay Kumar Jena
National Institute of Technology
Department of Computer Science and Engineering
Rourkela, Orissa
India, 769008
E-mail: skjena@nitrkl.ac.in

G.V.V. Krishna
National Institute of Technology
Department of Computer Science and Engineering
Rourkela, Orissa
India, 769008
E-mail: gvvkrishna@yahoo.co.in
Received: August 18, 2006



Dr. S.K. Jena was born in 28 April, 1954. He received his Ph.D. from Indian Institute of Technology, Bombay and M.Tech from Indian Institute of Technology, Kharagpur. He has joined National Institute of Technology as Professor in the Department of Computer Science and Engineering in 2002. Currently he is working as Professor and Head of Computer Science and Engineering department. He has more than 35 publications in International Journals and conferences. His research areas of interest are Database Engineering, Distributed Computing, Parallel algorithm, Information Security and Data Compression.



Mr. G.V.Vamsi.Krishna was born in 17 June, 1985. He received his M.Tech in Computer Science & Engineering from National Institute of Technology, Rourkela in 2006 and B.Tech in Computer Science & Engineering from Sarada Institute of Science, Technology and Management, Hyderabad in 2004. Currently he is working as Software Engineer in IBM India Pvt. Ltd., Bangalore.

An Enhancement of the Random Sequence 3-Level Obfuscated Algorithm for Protecting Agents Against Malicious Hosts

Kamalrulnizam Abu Bakar, Bernard S. Doherty

Abstract: With the advent of agent technology, the need for security of applications becomes a concern that is beyond classical data security capability and considerations. Without proper security protection especially against a malicious host attack, the widespread use of agent technology can be severely impeded. The Random Sequence 3-level obfuscated algorithm has been proposed by the authors to improve agent security: in this paper, an enhancement to the protection level of this algorithm is proposed. The effectiveness of the obfuscation algorithm is enhanced by addition of noise, which surrounds the true value carried by the agent with false values. A malicious host can thus at best guess the true value carried by the agent.

Keywords: Agent security, Malicious host, Spying attack, Noise code.

1 Introduction

Security is considered as an important factor for an agent used in real world applications. The problems in agent security arise when an agent is used in open and unsecured environments [8, 3]. One kind of attack by an execution host (malicious host) is spying on the agent's code, data and state [6, 11]. Spying attack by the malicious host may invade the agent's privacy, especially an agent's critical data, for example a user's maximum budget carried by the agent. Knowledge of an agent's critical data gives a malicious host an advantage in any competition over other hosts because the malicious host knows what is expected by the agent. For example, a customised agent is sent out (in an open and unsecured environment) to find a suitable flight with the fare price under or equal to 500 pounds. Malicious host attack based on a spying attack is to raise the offered price until it meets the maximum price that has been set by the agent's owner, even though the normal price is much lower. Spying attack by the malicious host on an agent's data or state is difficult to detect, because the attack doesn't leave any detectable trace [6, 11, 5]. The executing host has to read the agent's code, must have access to the agent's data, and must be able to manipulate the agent's variable data in order to execute the agent [6, 5, 8]. Therefore, the executing host can see and access all of the agent's code including data and state, and is thus able to hide any traces of an attack, which makes any attempt to address spying attack difficult.

To reduce the likelihood of this attack succeeding, the authors have proposed the Random Sequence 3-level (RS3) obfuscated algorithm [1], which obfuscates the actual value of an agent's critical data to prevent a malicious host spying attack. The RS3 obfuscated algorithm consists of multiple polynomial functions which obfuscating the actual value of the agent's critical data to produce an obfuscated value that is meaningless to the attacker. Only selected polynomial functions are used in the conversion process and in each selected function, multiple random inputs are implied. The main objective of using the RS3 obfuscated algorithm is to enable the comparing of confidential values within an unsecured remote host environment without exposing the actual confidential value to an unauthorised party, using comparison of obfuscated values rather than actual values. Unfortunately, implementing the RS3 obfuscated algorithm alone may expose the algorithm to an attack, which could execute multiple copies of the algorithm many times in order to analyse it, and propose actual values that will obfuscate to give the comparison outcome sought by the malicious host.

This paper proposes an enhancement of the protection level of the RS3 obfuscated algorithm by adding noise code in the agent application that executes in the remote host environment, to hide the actual obfuscated value among a number of false values. The enhancement makes it more difficult for

the attacker to analyse the actual obfuscated value in order to discover the actual value of an agent's critical data: the attacker can deobfuscate the set of values carried by the agent, but can at best guess which is the correct value.

This paper is organized as follows: Section 2 present an overview of the Random Sequence 3-level obfuscated algorithm. Section 3 describes the random sequence 3-level obfuscated algorithm with noise code, the implementation of noise code to enhance the obfuscated algorithm protection level and the deobfuscation time of the RS3 obfuscated algorithm. Section 4 presents the experimental results on the overhead of implementing noise code with the RS3 obfuscated algorithm. Section 5 analyse the strength of the RS3 obfuscated algorithm with noise code. Section 6 presents a discussion and the conclusion is presented in section 7.

2 An Overview of The Random Sequence 3-level Obfuscated Algorithm

The Random Sequence 3-level obfuscated algorithm [1] is an algorithm that is designed to protect the confidentiality of an agent against malicious host spying attack. This algorithm uses three polynomial functions for obfuscating the actual value of the agent's critical data to an obfuscated value that is meaningless to the malicious host, in order to prevent the malicious host from spying on the agent critical data (Figure 1 show the obfuscation process of the RS3 obfuscation algorithm). The obfuscated method used in the Random Sequence 3-level obfuscated algorithm enables the execution host to execute the process of comparing its offer with the agent budget with both values in an obfuscated format without the execution host having any knowledge of the actual value of the agent's critical data. This comparing process can be done without needing deobfuscation of the data, unlike cryptographic methods, which require decryption of the data, thus revealing its value, before a comparison can be made.

Although the RS3 obfuscated algorithm is able to obfuscate an agent's critical data to make it more difficult for the malicious host to spy, the malicious host can execute multiple copies of the obfuscated algorithm in parallel in order to analyse the algorithm quickly, making this obfuscated algorithm vulnerable. This attack can be addressed by limiting the processing time available to the host before the agent is discarded [6]. However, the problem in determining an effective protection interval that can prevent the malicious host having enough time to execute multiple copies of the obfuscated algorithm also makes it difficult for this obfuscated algorithm to be implemented in real applications, where sufficient time must be allowed for legitimate processing. In order to overcome the problem of a malicious host executing multiple copies of the RS3 obfuscated algorithm and of determining an effective protection interval to protect the algorithm, noise code [9, 10] is introduced for an agent that executes in the remote host environment. The RS3 obfuscated algorithm with noise code is described in the next section.

3 The Random Sequence 3-level Obfuscated Algorithm with Noise Code

The objective for implementing the noise code in the agent application is to hide the actual obfuscated value among a numbers of fake obfuscated values so the malicious host can at best guess at the true value of the agent's critical data [9, 10]. The difference between ordinary RS3 algorithm and RS3 with noise code algorithm is in the number of obfuscated values generated and added by the master agent into the slave agent application before the slave agent is dispatched to the remote host execution environment to execute its tasks.

In order to discover the true value of the agent's critical data, the malicious host must first guess the actual obfuscated value among a number of fake obfuscated values. Any wrong decision in choosing the obfuscated value will result in using a wrong true value of the agent's critical data. For instance, if the agent is equipped only with the actual obfuscated value, X without adding any noise code, the probability that the malicious host could discover the actual obfuscated value by searching a range of values is one,

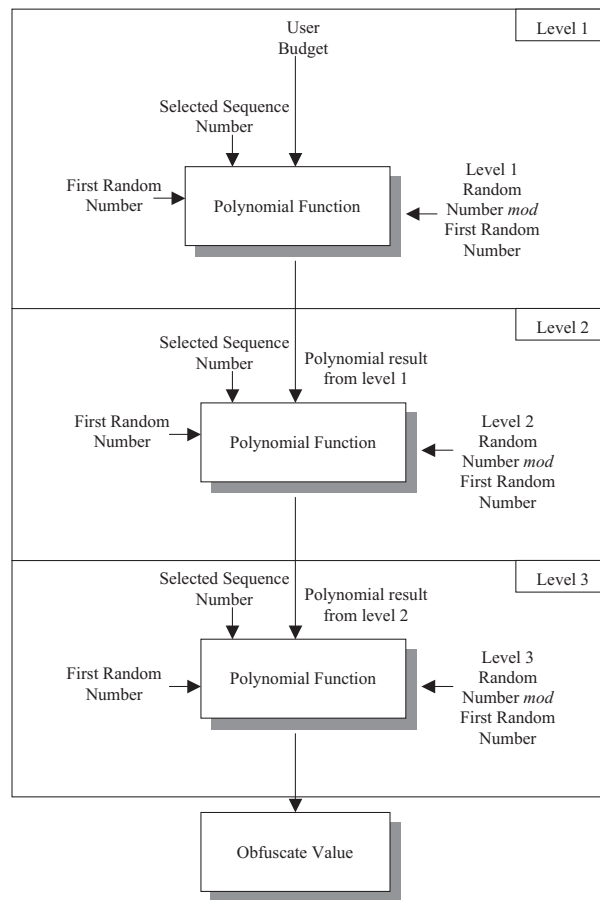


Figure 1: The RS3 Obfuscation Algorithm Obfuscation Process

i.e. $P(X) = 1$. However, if noise codes E_i (fake obfuscated values) are added to the agent, where $i = 1, 2, \dots, 100 - 1$, the probability of discovering the actual obfuscated value is $\frac{1}{100}$. The probability of guessing the actual obfuscated value becomes smaller as more noise codes are added. Figure 2 illustrates the effect of introducing noise codes into the agent application. In addition, the time needed to guess the actual obfuscated value will delay the malicious host in analysing the obfuscated algorithm. Therefore, the use of an effective protection interval in enhancing the obfuscated algorithm protection will be less important.

3.1 Implementing RS3 Obfuscated Algorithm with Noise Code

The operation of the RS3 obfuscated algorithm with noise code is almost the same as the operation of the RS3 obfuscated algorithm without noise code (refer to [1]). The only difference between these two obfuscated algorithms is in the number of obfuscated values generated and added by the master agent into the slave agent application before the slave agent is dispatched to the remote host execution environment to execute its tasks.

In the operation of RS3 obfuscated algorithm without noise code, the master agent only has to obfuscate the value of user's budget and add the obfuscated value into the slave agent application before dispatching the slave agent to execute its tasks in the remote host execution environment. However, in the operation of RS3 obfuscated algorithm with noise code, the master agent has to generate more than one obfuscated value (the extra obfuscated values serve as noise codes) and add these obfuscated values

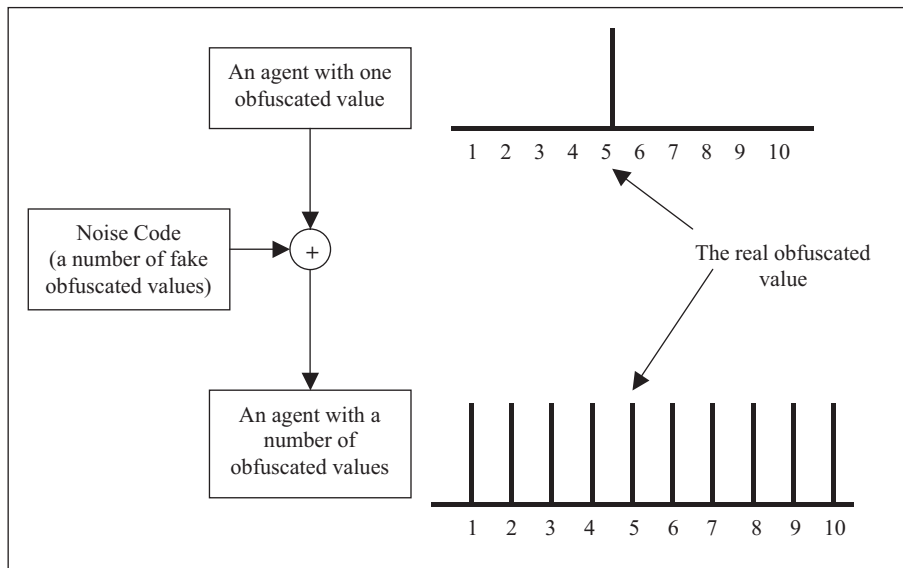


Figure 2: The Effect of Adding Noise Codes Into The Agent Application

into the slave agent application before dispatching the slave agent to execute in the remote host execution environment (see figure 3).

```

Vector hostAddress = new Vector();
double newOffer=0;
double bestOffer1, bestOffer2, bestOffer3;
URL bestShop1, bestShop2, bestShop3;

      ⋮

if(NewObfuscationValue <= ObfuscationValue1)) //fake obfuscated value
{
    bestOffer1 = newOffer;
    bestShop1 = hostAddress;
} else
if(NewObfuscationValue <= ObfuscationValue2)) //true obfuscated value
{
    bestOffer2 = newOffer;
    bestShop2 = hostAddress;
} else
if(NewObfuscationValue <= ObfuscationValue3)) //fake obfuscated value
{
    bestOffer3 = newOffer;
    bestShop3 = hostAddress;
}

```

Figure 3: A Slave Agent Program added with Noise Code (data block)

To illustrate, the noise code is a fake obfuscated value that is generated by the master agent from a fake user budget value. This fake user budget value is created by the master agent by adding or subtracting a random number to or from the actual user budget value. For example, say the actual user budget value is €500. To create a fake user budget value, the master agent needs to generate a random

number, e.g. 176. If the master agent chooses to add the random number to the actual user budget value, the fake user budget value becomes 676. This value is then converted to the obfuscated value to represent the fake obfuscated value. On the other hand, if the master agent chooses to subtract the random number with the actual user budget value, the fake user budget value becomes 324. The same conversion process will be applied to convert the fake user budget value to a fake obfuscated value. To generate more fake obfuscated values, the master agent has to generate more random numbers.

Once the obfuscation process in the home host is completed, the master agent dispatches the slave agent together with all the obfuscated values generated (including fake obfuscated values) to the remote host to execute its given tasks. In the remote host execution environment, the slave agent starts its execution process by converting any offer that was gathered from the remote host into an obfuscated value to be used in the comparing process. For example, if the slave agent has 4 obfuscated values (one actual and three fake values), the slave agent has to execute 4 comparing processes. If any of the obfuscated user budget value matches the obfuscated offer value, the corresponding obfuscated user budget value will be excluded from the obfuscation stage (see figure 4). The next obfuscation stage starts by obfuscating a new remote host offer. In the authors' work, a maximum of three obfuscation stages are used to searches for a flight offer (the obfuscated user budget value is compared with the first class obfuscated fare value in the obfuscation stage 1, business class obfuscated fare value in the obfuscation stage 2 and economy class obfuscated fare value in the obfuscation stage 3) or less than three times if all the obfuscated user budget values have been excluded.

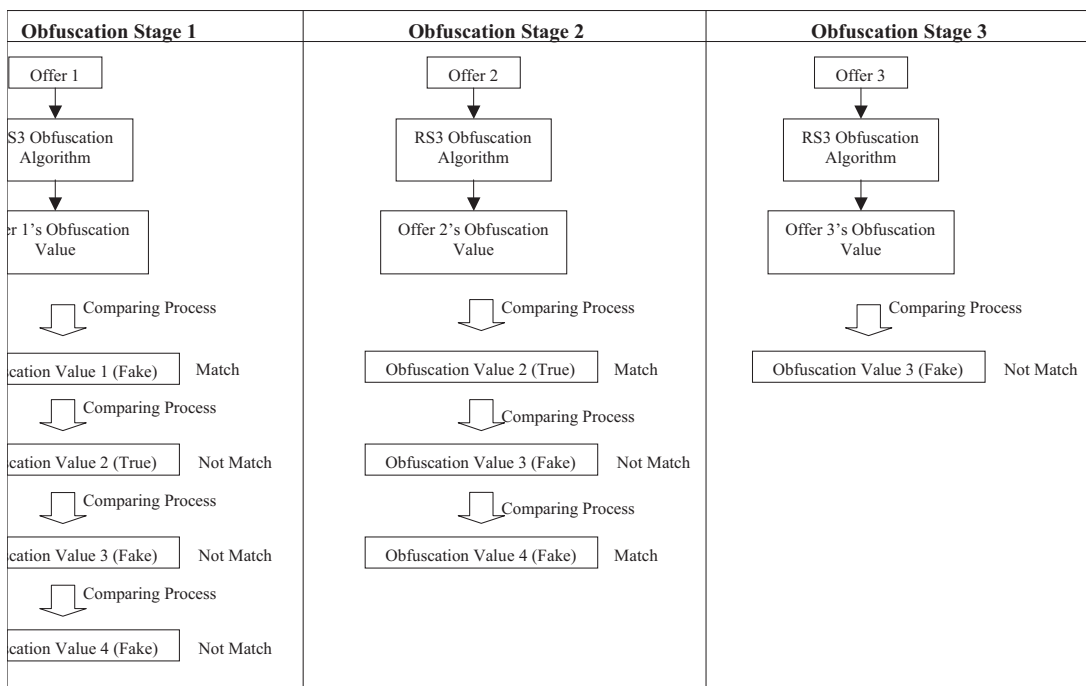


Figure 4: The illustration of the Obfuscation Stages and the Comparing Process

After completing the execution process in the remote host, the slave agent returns to its home host together with the remote host offer. The home host then extracts only the offer that fulfils the requirement of the true obfuscated value for further actions.

3.2 The Deobfuscation Time of the RS3 Obfuscated Algorithm with Noise Code

The experiment on the deobfuscation time of the RS3 obfuscated algorithm with noise code is conducted to examine the time taken by the execution host (assumed to be the malicious host) to deobfuscate a full set of obfuscated value. The experiment is conducted using one 700 MHz personal computer with 128 MB of main memory, which running Windows 98 operating system.

In this experiment, the deobfuscation time is taken starting from the start of the deobfuscation process for the first obfuscated value and ending when the last obfuscated value from a set of obfuscated value have been deobfuscated. The experiment is performed by the execution host by executing the RS3 obfuscated algorithm using different input value many times until the matching obfuscated value is produced and continue until a full set of the obfuscated value are deobfuscated (see Figure 5). Twenty sets of obfuscated values are examine with the number of values in each set ranging from five to one thousand obfuscated values. Each set of obfuscated values contained one correct value and the remainder false values.

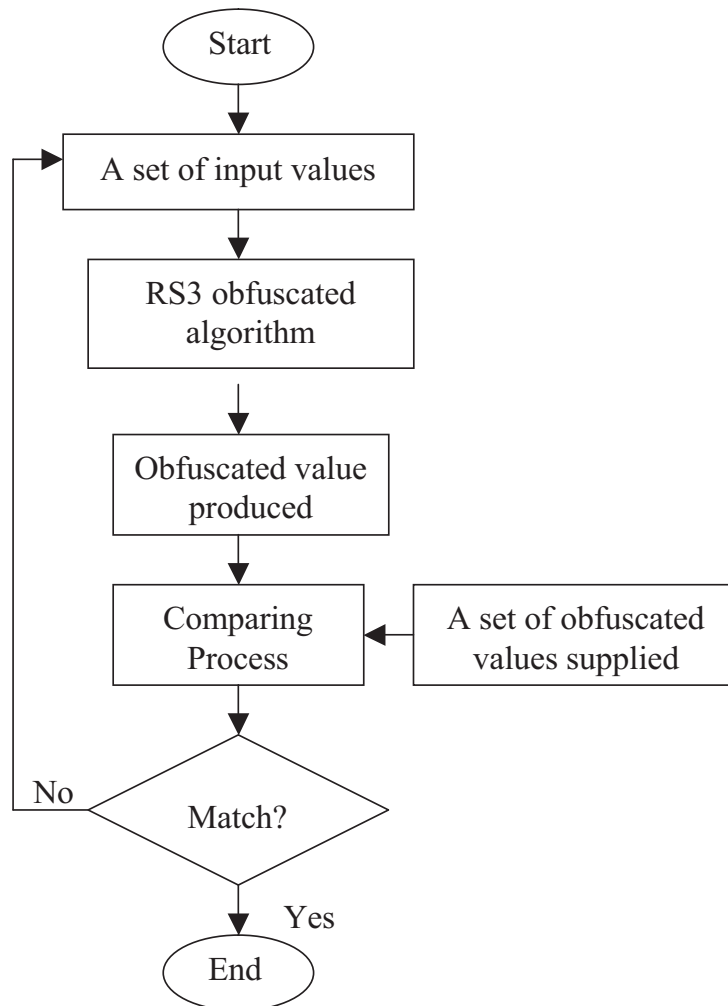


Figure 5: The Deobfuscation Process

The deobfuscation time results (which were gathered in milliseconds and then converted into seconds) of the RS3 obfuscated algorithm with noise code are shown in Table 1 and 2, and illustrated in Figure 6 and 7 respectively.

Obfuscated Values	5	10	15	20	25	30	35	40	45	50
Deobfuscation Time	8.95	21.57	29.42	37.69	49.96	60.71	62.53	72.46	81.66	85.58

Table 1: Summary Statistics of the Random Sequence 3-level Obfuscated Algorithm Deobfuscation Time for Small Number of Obfuscated Values

Obfuscated Values	100	200	300	400	500	600	700	800
Deobfuscation Time	180.15	376.33	581.28	779.12	982.06	1142.27	1376.64	1568.75

Obfuscated Values	900	1000
Deobfuscation Time	1765.94	1887.83

Table 2: Summary Statistics of the Random Sequence 3-level Obfuscated Algorithm Deobfuscation Time for Large Number of Obfuscated Values

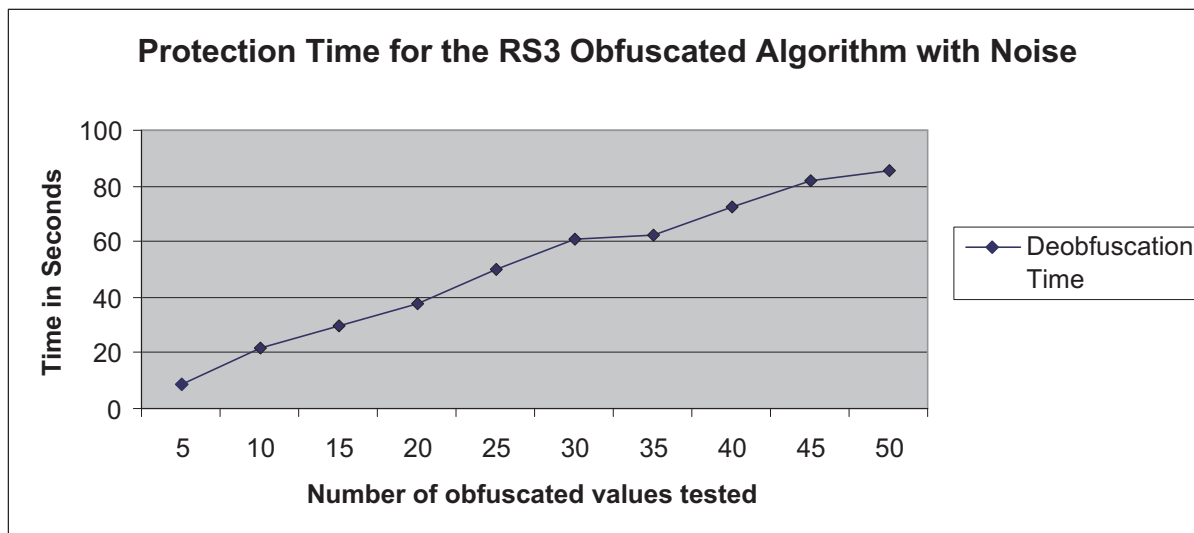


Figure 6: The Deobfuscation Time of the Random Sequence 3-level Obfuscated Algorithm for Small Number of Obfuscated Values

From the results given in Table 1 and 2, and illustrated in Figure 6 and 7 respectively, it can be seen that the deobfuscation time of the RS3 obfuscated algorithm with noise code increases linearly with the number of obfuscated values in the test set. Note that even after deobfuscation, the malicious host is left with a set of values of which only one is the correct value, and has no way of knowing which one is the correct value.

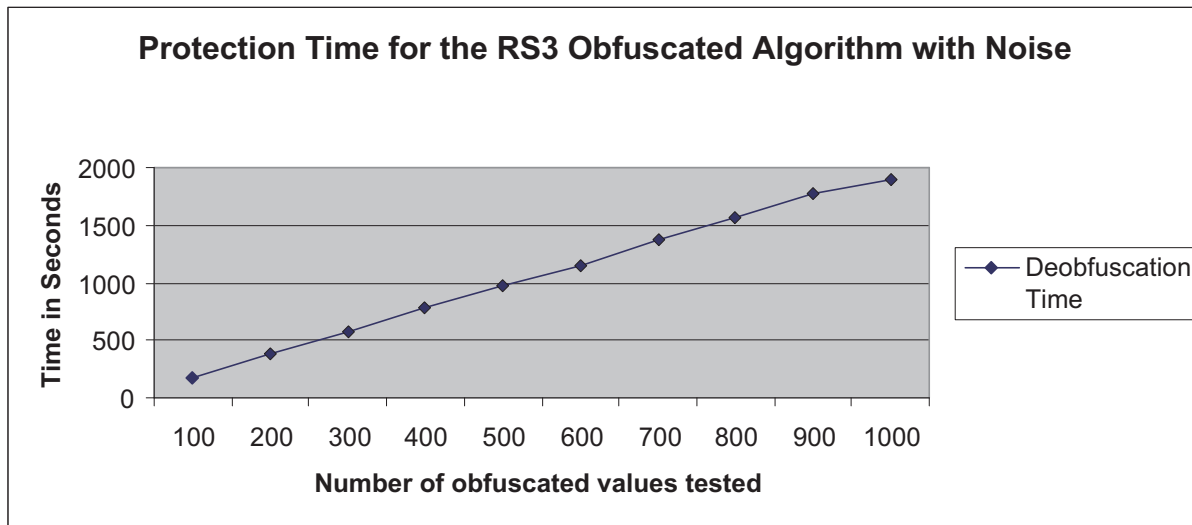


Figure 7: The Deobfuscation Time of the Random Sequence 3-level Obfuscated Algorithm for Large Number of Obfuscated Values

4 The Overhead of Implementing Noise Code with the RS3 Obfuscated Algorithm

The experiments to measure the overhead of implementing the RS3 obfuscated algorithm with noise code were conducted using six 400 MHz Sun Ultra Sparc 5 workstations with 128 MB of main memory. Each of the workstations is running the Solaris 8 operating system and is connected to the others using 100 Mbit/s UTP¹ cable. All of the workstations involved in this experiment were situated in the same room.

In this configuration, one workstation was chosen among the six workstations to be the home host for the agent, and only this host had permission to manage and dispatch the agent. The rest of the workstations were assumed to be remote hosts, having only the capability to receive the agent and to dispatch the agent back to its home host.

To examine the security overhead in implementing noise code with the Random Sequence 3-level obfuscated algorithm in an agent-based application, times are measured starting from sending of the agents to the remote hosts and ending with the home host receiving the agents back from the remote hosts. Four different experiments are conducted starting with one remote host, two remote hosts, three remote hosts and five remote hosts using three different kinds of agent: a plain agent², an agent with the Random Sequence 3-level obfuscated algorithm (RS3) and an agent with the RS3 obfuscated algorithm with noise code (RS3N). The times were measured using the “System.currentTimeMillis()” method in the Java language. This method produces a specific instant in time with millisecond precision [7].

The experiment is performed for 20 runs³ and the time for each run gathered in milliseconds. The average result of the 20 runs is taken and converted into seconds. The result is then rounded and presented in two decimal places as shown and illustrated in Tables 4 to 7 and Figure 8 to 11 respectively.

The results of the security overhead of the Random Sequence 3-level obfuscated algorithm without noise (RS3) and the Random Sequence 3-level obfuscated algorithm with noise are compared to the plain agent as shown in Table 3. From these results, it can be seen that there is not much difference in

¹Unshielded Twisted Pair Category 5e

²agents without security mechanisms

³The results were observed to be consistent after twenty runs

Number of Remote Hosts	Mean		
	Plain	RS3	RS3N
1	1.49	1.5	1.69
2	2.45	2.52	2.67
3	3.36	3.56	3.63
5	5.42	5.43	5.66

Table 3: Summary Statistics of The Random Sequence 3-Level Obfuscated Algorithm (1 Cycle, 1 Obfuscation Value Experiment(without noise code) and 1000 Obfuscation Value Experiment(with noise code))

the security overhead for the RS3 without noise and RS3 with noise for all numbers of remote hosts. The highest difference is for one remote host, where the security overhead for RS3 with noise is just 12.67 % higher than the security overhead of RS3 without noise. Therefore, it is concluded that RS3 with noise will add a small and acceptable overhead. More detail of the results now follows:

Firstly, consider plain and RS3, then plain and RS3N experimental results.

Number of Remote Hosts	Mean		Standard Error		Standard Deviation	
	Plain	RS3	Plain	RS3	Plain	RS3
1	1.49	1.5	0.001	0.001	0.003	0.004
2	2.45	2.52	0.003	0.005	0.011	0.022
3	3.36	3.56	0.003	0.004	0.011	0.02
5	5.42	5.43	0.007	0.006	0.031	0.025

Table 4: Summary Statistics of The Random Sequence 3-Level Obfuscated Algorithm Overhead (1 Cycle and 1 Obfuscation Value Experiment(without noise code))

Number of Remote Hosts	Mean		Standard Error		Standard Deviation	
	Plain	RS3N	Plain	RS3N	Plain	RS3N
1	1.49	1.54	0.001	0.001	0.003	0.003
2	2.45	2.52	0.002	0.007	0.011	0.031
3	3.36	3.51	0.003	0.005	0.011	0.022
5	5.42	5.44	0.007	0.007	0.031	0.031

Table 5: Summary Statistics of The Random Sequence 3-Level Obfuscated Algorithm (1 Cycle and 100 Obfuscation Value Experiment(with noise code))

From the results given in Tables 4 and 5 and illustrated in Figure 8 and 9 respectively, it can be seen that the mean of the security overhead for a plain agent is almost the same as the security overhead for RS3 and RS3N, where when comparing with RS3, the highest difference is just 5.95 % and 4.46 % when comparing with RS3N. However, from Table 6 and illustrated in Figure 10, it can be seen that as the number of noise codes is increased, the difference becomes larger, where the highest difference is 13.42 %, which is still considered acceptable.

From Table 7 and Figure 11, the security overhead for both the agents is almost the same as the

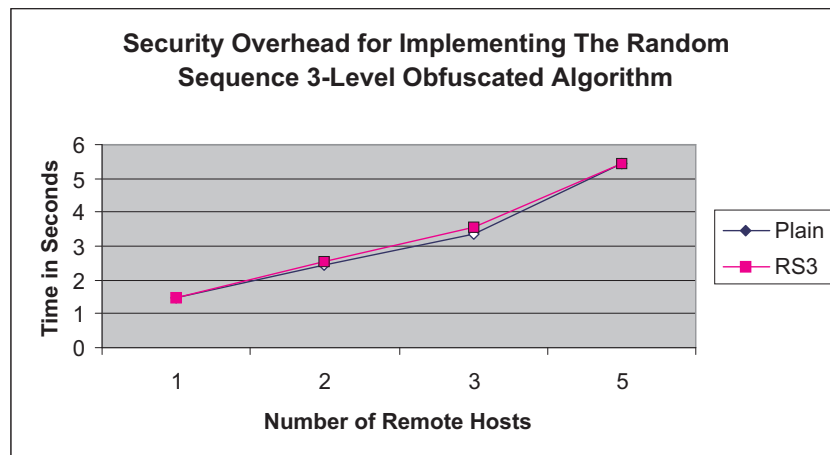


Figure 8: Security Overhead of The Random Sequence 3-Level Obfuscated Algorithm (1 Cycle and 1 Obfuscation Value Experiment(without noise code))

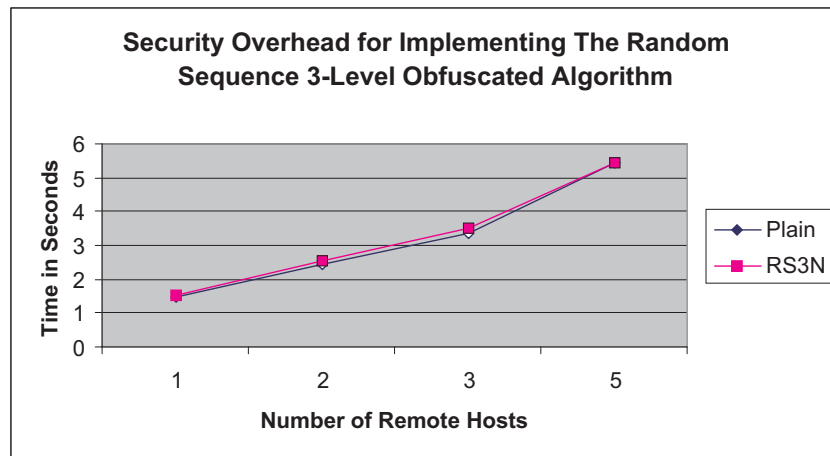


Figure 9: Security Overhead of The Random Sequence 3-Level Obfuscated Algorithm (1 Cycle and 100 Obfuscation Value Experiment(with noise code))

security overhead given in Tables 4 to 6, even though now the number of cycles⁴ has been increased to 100.

These results show that implementation of the Random Sequence 3-level obfuscated algorithm with noise code does increase the overhead by up to 13.42 % compared to the plain agent, but the noise code adds little to the overhead.

5 The analysis of RS3 obfuscated algorithm strength with noise code implementation

To analyse the strength of the RS3 obfuscated algorithm after the implementation of noise code, the authors have listed vulnerabilities and the way to address them as follows:

⁴a loopings that simulate an agent tasks

Number of Remote Hosts	Mean		Standard Error		Standard Deviation	
	Plain	RS3N	Plain	RS3N	Plain	RS3N
1	1.49	1.69	0.001	0.004	0.003	0.016
2	2.45	2.67	0.003	0.006	0.011	0.028
3	3.36	3.63	0.003	0.003	0.011	0.015
5	5.42	5.66	0.007	0.006	0.031	0.028

Table 6: Summary Statistics of The Random Sequence 3-Level Obfuscated Algorithm (1 Cycle and 1000 Obfuscation Value Experiment(with noise code))

Number of Remote Hosts	Mean		Standard Error		Standard Deviation	
	Plain	RS3N	Plain	RS3N	Plain	RS3N
1	1.48	1.65	0.0003	0.002	0.001	0.008
2	2.46	2.65	0.002	0.003	0.01	0.014
3	3.45	3.67	0.007	0.005	0.03	0.021
5	5.46	5.67	0.013	0.004	0.06	0.02

Table 7: Summary Statistics of The Random Sequence 3-Level Obfuscated Algorithm (100 Cycle and 1000 Obfuscation Value Experiment(with noise code))

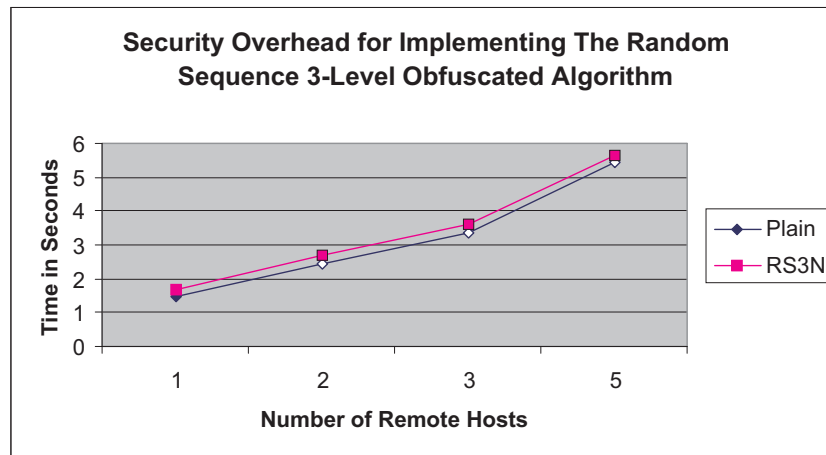


Figure 10: Security Overhead of The Random Sequence 3-Level Obfuscated Algorithm (1 Cycle and 1000 Obfuscation Value Experiment(with noise code))

5.1 The vulnerabilities of RS3 obfuscated algorithm with noise code

There are two main weaknesses of the RS3 obfuscated algorithm with noise code that have been identified.

- If the attacker (malicious host) takes a guess at the correct obfuscated value from many obfuscated values (including noise code) carried by the visiting agent and the attacker is given enough time to execute, the attacker can execute the RS3 obfuscated algorithm many times using different result values and watch the pattern of the RS3 obfuscated algorithm outcomes (which result value the

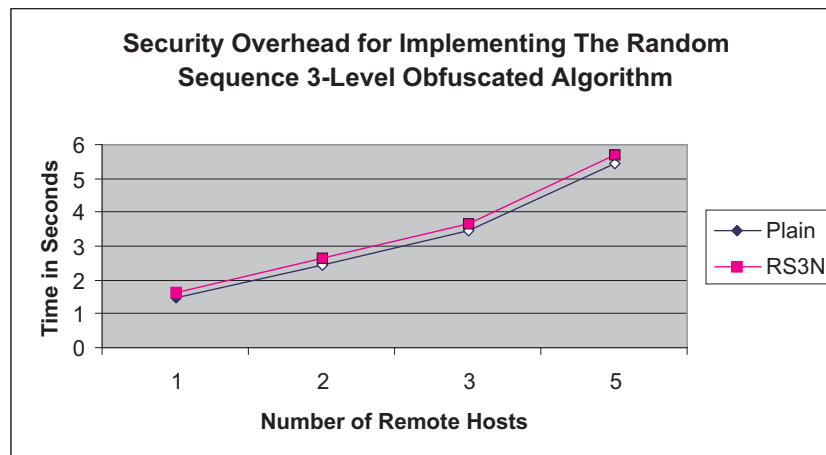


Figure 11: Security Overhead of The Random Sequence 3-Level Obfuscated Algorithm (100 Cycle and 1000 Obfuscation Value Experiment(with noise code))

agent accepts and which it rejects) to analyse the chosen obfuscated value in order to discover the actual value of agent's critical data.

- If the noise code generated from the actual value are very much out of range from the reasonable value, the malicious host could easily omit that values and take a guess from fewer values.

5.2 Addressing vulnerabilities of RS3 obfuscated algorithm with noise code

In order to overcome the weaknesses of the RS3 obfuscated algorithm with noise code, two main points are suggested:

- To prevent the attacker (malicious host) from being able to guess a correct obfuscated value among many obfuscated values in a short time, and running many tests on RS3 obfuscated algorithm by analysing which result value the agent accepts and which it rejects, the number of noise codes added in the agent application, could be made bigger. This is due to the fact that the probability to guess the correct obfuscated value becomes smaller as the number of noise code added become bigger, i.e. $P(X) \rightarrow 0$ as $N \rightarrow \infty$. The malicious host can at best guess the actual value from among the noise code.
- To prevent the malicious host able to omit any of the noise code values, the value of the noise codes must be within a reasonable range of the actual value and this can be done by limiting the range of the random numbers.

6 Discussion

In this paper, the problem of a malicious host spying on the actual value of an agent's critical data, such as the user maximum budget, has been discussed. The Random Sequence 3-level obfuscated algorithm, which is able to obfuscate the actual value of an agent's critical data in order to make it more difficult for the malicious host to spy on the actual value of the critical data, has been previously proposed by the authors. To address weaknesses in the RS3 obfuscated algorithm, an enhancement has been proposed in this paper.

The implementation of the RS3 obfuscated algorithm alone exposes the obfuscated algorithm to the attacker that could execute multiple copies of the obfuscated algorithm many times to analyse the algorithm. The agent owner could also face the problem of determining an effective protection interval for the obfuscated value that is carried by the agent in order to prevent the value from being analysed and discovered by the malicious host. These problems can be overcome by introducing noise codes carried by the agent application to force a malicious host to guess the actual obfuscated value, thus leaving a malicious host with at best a guess at the actual value of the user maximum budget. The noise code implementation also delays the malicious host in analysing the obfuscated algorithm, which has been shown in Section 3.2. Therefore, the use of an effective protection interval to enhance the level of obfuscated algorithm protection is less important.

On the other hand, based on the experimental results on the overhead of implementing the RS3 obfuscated algorithm with noise code, the implementation of the RS3 obfuscated algorithm does increase the overhead by up to 13.42 % compared to the plain agent, but this is considered acceptable. This suggests the Random Sequence 3-level obfuscated algorithm can be implemented in “real world” agent applications to protect the agent application from the spying attack by the malicious host. Experiment shown that adding noise codes to the RS3 obfuscated algorithm gives little (12.67 %) increase in security overhead.

7 Conclusion

The Random Sequence 3-level obfuscated algorithm is an algorithm that improves the level of protection of an agent against malicious host spying attack. This obfuscated algorithm does not protect against all spying attacks by the malicious host, only an attack to the agent’s critical data. However, the implementation of noise code in the agent application prevents the malicious host discovering the actual value of critical data carried by the agent; the malicious host can at best guess the actual value from among a number of noise values. The addition of noise code has strengthened the protection of the obfuscated algorithm and has reduced the likelihood of successful attack on the RS3 obfuscated algorithm, with very small increase in execution time.

References

- [1] Abu Bakar, K. and Doherty, B. S.: A Random Sequence 3-level Obfuscated Algorithm for Protecting Mobile Agents Against Malicious Hosts. Proceedings of the 2003 International Conference on Informatics, Cybernetics and Systems. I-Shou University(2003) 525 – 530
- [2] DiVincenzo, D. P., Leung, D. W. and Terhal, B. M.: Quantum Data Hiding. IEEE Transactions on Information Theory, Vol. 48, No. 3. IEEE(2002) 580–598
- [3] Farmer, W.M. and Guttman, J.D. and Swarup, V.: Security for Mobile Agents: Issues and Requirements. Proceedings of the 19th National Information System Security Conference. Baltimore (1996) 591–597
- [4] Harmsen, J. J. and Pearlman, W. A.: Steganalysis of Additive Noise Modelable Information Hiding. Proceedings of SPIE Electronic Imaging 5022. SPIE (2003) 21–24
- [5] Hohl, F.: A Framework to Protect Mobile Agents by Using Reference States. In: Proceedings of the 20th international conference on distributed computing systems (ICDCS 2000). IEEE Computer Society (2000) 410-417

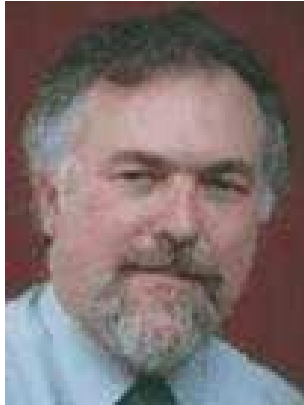
- [6] Hohl, F.: Time Limited Blackbox Security: Protecting Mobile Agents from Malicious Hosts. In: G. Vigna (Ed.). *Mobile Agent and Security. Lecture Notes in Computer Science*, Vol. 1419. Springer-Verlag, Berlin(1998) 92–113
- [7] Sun Microsystems, Inc. Java 2 Platform Std. Ed. V1.3.1 <http://java.sun.com/j2se/1.3/docs/api/index.html> (2004)
- [8] Mandry, T., Pernul, G. and Rohm, A.: Mobile Agents in Electronic Markets: Opportunities, Risks, Agent Protection. *International Journal of Electronic Commerce*. M.E. Sharpe (2001) 47–60
- [9] Ng, S. K. and Cheung, K. W.: Protecting Mobile Agents against Malicious Hosts by Intention Spreading. In H. Arabnia (ed.), *Proc. International Conference on Parallel and Distributed Processing Techniques and Applications (PDPTA'99)*. CSREA (1999) 725–729
- [10] Ng, S. K. and Cheung, K. W.: Intention Spreading: An extensible theme to protect mobile agents from read attack hoisted by malicious hosts. In Jimming Liu, Ning Zhong(ed.), *Intelligent Agent Technology: Systems, Methodologies and Tools*. World Scientific (1999) 406–415
- [11] Sander, T. and Tschudin, C.: Protecting Mobile Agent Against Malicious Hosts. In: G. Vigna (Ed.). *Mobile Agent and Security. Lecture Notes in Computer Science*, Vol. 1419. Springer-Verlag, Berlin(1998) 44–60
- [12] Yeh, W. H. and Hwang, J. J.: Hiding Digital Information Using a Novel System Scheme. *Computers and Security*, Vol. 20, No. 6. Elsevier Science (2001) 533–538

Kamalrulnizam Abu Bakar
Faculty of Computer Science and Information System
Universiti Teknologi Malaysia
81310 UTM Skudai
Johor D. T.
Malaysia
E-mail: kamarul@fsksm.utm.my

Bernard S. Doherty
School of Engineering and Applied Science
Aston University
Aston Triangle, Birmingham B4 7ET
United Kingdom
E-mail: b.s.doherty@aston.ac.uk
Received: July 22, 2006



Kamalrulnizam Abu Bakar is a lecturer at Universiti Teknologi Malaysia, Malaysia. He received the diploma and degree of Computer Science in 1994 and 1996 respectively from Universiti Teknologi Malaysia, Malaysia. He then received Masters in Computer Communication and Networks degree from Leeds Metropolitan University, United Kingdom in 1998 and PhD in Network Security from Aston University, United Kingdom in 2004. His current research interests include computer and network security, distributed systems and parallel processing, grid computing, wireless and cellular network.



Bernard S. Doherty (born October 2nd, 1945) obtained the degrees of Bachelor of Engineering (Electrical), Bachelor of Arts and Master of Engineering Science from the University of Melbourne in 1967, 1971 and 1981 respectively. He has held positions with the State Electricity Commission of Victoria, LM Ericsson Pty Ltd, Swinburne College of Technology (all in Melbourne) and, since 1980, at Aston University (Birmingham, UK), where is presently Lecturer in Computer Science. His main fields of teaching and research are Distributed and Networked applications and Information Security. In addition to supervising a number of Doctoral students, he has developed computer-based administration and teaching software, written a number of papers and presented papers at international conferences.

Ant systems & Local Search Optimization for flexible Job Shop Scheduling Production

Noureddine Liouane, Ihsen Saad, Slim Hammadi, Pierre Borne

Abstract: The problem of efficiently scheduling production jobs on several machines is an important consideration when attempting to make effective use of a multi-machines system such as a flexible job shop scheduling production system (FJSP). In most of its practical formulations, the FJSP is known to be NP-hard [8][9], so exact solution methods are unfeasible for most problem instances and heuristic approaches must therefore be employed to find good solutions with reasonable search time.

In this paper, two closely related approaches to the resolution of the flexible job shop scheduling production system are described. These approaches combine the Ant system optimisation meta-heuristic (AS) with local search methods, including tabu search. The efficiency of the developed method is compared with others.

Keywords: Flexible production, Ant colony, Tabu search, job shop scheduling, makespan, optimisation.

1 Introduction

Modern hybrid heuristics are by their nature non-exhaustive, and so there is often scope for different approaches to better previous solution methods according to the execution speed or the quality of feasible solutions. Traditional approaches to resolve the FJSP are as varied as the different formulations of the problem, but include fast, simple heuristics [2][12], tabu search [15], evolutionary approaches [5] and modern hybrid meta-heuristics that consolidate the advantages of various different approaches [1][13].

The ant colony optimisation (ACO) was described by Dorigo in his PhD thesis [6] and was inspired by the ability and the organisation of real ant colony using external chemical *pheromone trails* acting as a means of communication.

Ant system algorithms have since been widely employed on the NP-hard combinatorial optimisation problems including problems related to Continuous Design Spaces research [4], and job shop scheduling [16]. However, they have not previously been applied to the FJSP described in what follows.

Local search methods encompass many optimisation approaches and have been shown that the efficiency of their use with an ant system approach [7].

The approach described in this paper for the FJSP shows the quality of solutions found, using benchmark problems. The performances of the proposed approach are evaluated and compared with the results obtained from other methods.

In this paper, an application of the ant system algorithms combined by the tabu search heuristic is proposed for solving the FJSP. Thus, The FJSP is described and formulated in section 2. Then, in section 3, The suggested approach by ACO with the tabu search is described. An illustrative example is given in section 4. The last section will be devoted to the presentation of some results and some conclusions relating to this research work.

2 Problem formulation

The FJSP may be formulated as follows. Consider a set of n independent jobs, noted $\mathfrak{J} = \{J_1, J_2, \dots, J_n, 1 \leq j \leq J\}$, which are carried out by K machines $M_k, M = \{M_1, M_2, \dots, M_k, 1 \leq k \leq K\}$. Each job J_j consists of a sequence of n_j operations $O_{i,j}$,

$i = 1, 2, \dots, n_j$. Each routing has to be performed to achieve a job. The execution of each operation i of a job J_j requires one resource selected from a set of available machines. The assignment of the operation $O_{i,j}$ to the machine $M_k \subseteq M$ entails the occupation of the latter one during a processing time, noted $p_{i,j,k}$. The problem is thus to both determine an assignment scheme and a sequence of the operations on all machines that minimize some criteria.

- A set of J independent jobs.
- Each job is characterized by the earliest starting time r_j and the latest finishing time d_j .
- Denote by $pt_{i,j}$ and $r_{i,j}$ respectively the processing time and the ready date of the operation $O_{i,j}$. The $p_{i,j,k}$ represent the processing time $pt_{i,j}$ with the machine M_k .
- A started operation can not be interrupted.
- Each machine can not perform more than one operation at the same time.
- The objective is to find an operation ordering set satisfying a cost function under problem constraints. In this paper, the considered objective is to minimize the makespan C_{max} .

3 ACO and Tabu search for FJSP Scheduling

In this stage, the application of the combined ant systems with tabu search techniques in the resolution of FJSP problem are described.

3.1 Construction Graph and Constraints

Generally, the FJSP can be represented by a bipartite graph with two categories of nodes: $O_{i,j}$ and M_k . A task is mapped to a $O_{i,j}$ node; a machine is mapped to a M_k . There is an edge between the $O_{i,j}$ node and the M_k node if and only if the corresponding task can be assigned to the corresponding machine while respecting the availability of the machine and the precedence constraints among the operations of different jobs. The cost of assignment is directly related to the processing time of the task upon the machine.

To model the process in a more straightforward manner, we use the construction graph that is derived from the utilization matrix. Below is a sample construction graph.

Table 1: Construction graph of 4 machines and 7 tasks.

	M_1	M_2	M_3	M_4
$O_{1,1}$	10	7	6	13
$O_{2,1}$	4	5	8	12
$O_{3,1}$	9	5	6	12
$O_{1,2}$	15	12	8	6
$O_{2,2}$	9	5	7	13
$O_{1,3}$	7	16	5	11
$O_{2,3}$	9	16	8	11

With this construction graph, we can transform the FJSP into a traveling ant problem. Specifically, given the representative table of n rows and m columns, and each of its cells is associated with $p_{i,j,k}$, representing this one distance among $O_{i,j}$ and M_k . An ant seeks to travel across the table in such a way

that all of the following constraints will be satisfied: one and only one cell is visited for each of the rows. In the rest of this paper, "tour" and "solution" are used interchangeably; a pair of (operation, machine) means: operation is assigned to machine, table 2.

Table 2: Solution of Construction graph table 1

	M_1	M_2	M_3	M_4
$O_{1,1}$			6	
$O_{2,1}$	4			
$O_{3,1}$		5		
$O_{1,2}$				6
$O_{2,2}$		5		
$O_{1,3}$			5	
$O_{2,3}$			8	

3.2 Ant systems scheduling

The Ant system approach was inspired by the behaviour of the real ants. The ants deposit the chemical *pheromone* when they move in their environment, they are also able to detect and to follow pheromone trails.

In our case, the pheromone trail describes how the ant systems build the solution of the FJSP problem. The probability of choosing a branch at a certain time depends on the total amount of pheromone on the branch, which in turn is proportional to the number of ants that used the branch until that time. The probability P_{ijk}^f that an ant will assign an operation $O_{i,j}$ of job J_j to an available machine M_k . Each of the ants builds a solution using a combination of the information provided by the pheromone trail τ_{ijk} and by the heuristic function defined by $\eta_{ijk} = p_{i,j,k}$.

Formally, the probability of picking that an ant f^{th} will assign an operation $O_{i,j}$ of job J_j to the machine M_k is given in equation 1.

$$P_{ijk}^f = \begin{cases} \frac{(\tau_{ijk})^\alpha * (\eta_{ijk})^{-\beta}}{\sum_{l \in D} (\tau_{ijk})^\alpha * (\eta_{ijk})^{-\beta}} & \text{if } j \in D \\ 0 & \text{if } j \notin D \end{cases} \quad (1)$$

In this equation, D denotes the set of available non-executed operations set and where α and β are parameters that control the relative importance of trail versus visibility. Therefore the transition probability is a trade-off between visibility and trail intensity at the given time.

3.3 Updating the pheromone trail

To allow the ants to share information about good solutions, the updating of the pheromone trail must be established. After each iteration of the ant systems algorithm, equation 2 describes in detail the pheromone update used when all ants have completed an own scheduling solution denote L^{ants} , that represent the length of ant tour. In order to guide the ant systems towards good solutions, a mechanism is required to assess the quality of the best solution. The obvious choice would be to use the best makespan $L^{min} = C_{max}$ of all solutions given by a set of ant.

$$\Delta\tau_{ijk}^f = \begin{cases} \frac{L^{min}}{L^{Ants}} & \text{if } i, j, k \in L^{Ants} \\ 0 & \text{otherwise} \end{cases} \quad (2)$$

After all of the ants have completed their tours, the trail levels on all of the arcs need to be updated. The evaporation factor ρ ensures that pheromone is not accumulated infinitely and denotes the proportion of τ pheromone that is carried over to the next iteration of the algorithm. Then for each edge the pheromone deposited by each ant that used this edge are added up, resulting in the following pheromone-level-update equation:

$$\tau_{ijk} = \rho \cdot \tau_{ijk} + \sum_{f=1}^{N^{BA}} \Delta\tau_{ijk}^f \quad (3)$$

where N^{BA} defines the number of ants to use in the colony.

3.4 Tabu search optimisation

A simple tabu search was also implemented for this optimisation FJSP problem. The proposed approach is to allow the ants to build their solutions as described in section 3.2 and then the resulting solutions are taken to a local optimum by the local search mechanism.

Each of these ant solutions is then used in the pheromone update stage. The local search is performed on every ant solution, every iteration, so it needs to be fairly fast. In the case of the FJSP problem, the method is to pick the machine responsible to the C_{max} and check if any operations $O_{i,j}$ could be swapped between other machines which would result in a lower makespan.

Following their concept, the local search considers one problem machine at a time and attempts to swap one operation from the problem machine with any other (non-problem) machine in the solution (non-problem operations). Then the ants are used to generate promising scheduling production solutions and the tabu search algorithm is used to try to improve these solutions.

The tabu search is performed on each problem machine and continues until there is no further improvement in the makspean value of the solution. An example of this algorithm is shown in section 4.

3.5 The set up parameter values

The set up parameter values used in the ant system scheduling algorithms are often very important in getting good results, however the appropriate values are very often entirely problem dependent [7], and cannot always be derived from features of the problem itself:

- α determines the degree to which pheromone trail is used as the ants build their solution. The lower the value, the less ‘attention’ the ants pay to the pheromone trail, but the higher values implicate the ants then perform too little exploration, after testing values in the range 0.1-0.75 this algorithm works well with relatively high values (around 0.5-0.75).
- β determines the extent to which heuristic information is used by the ants. Again, values between 0.1-0.75 were tested, and a value around 0.5 appeared to offer the best trade-off between following the heuristic and allowing the ants to explore the research space.
- τ is the value to which the pheromone trail values are initialized. Initially the value of the parameter should be moderately high to encourage initial exploration, while the pheromone evaporation procedure will gradually stabilise the pheromone trail.
- ρ is the pheromone evaporation parameter and is always set to be in the range $[0 < \rho < x]$. It defines how quickly the ants ‘forget’ past solutions. A higher value makes for a more aggressive search; it tests a value of around 0.5-0.75 to find good solutions.

- N^{BA} defines the number of ants to use in the colony, a low value speeds the algorithm up because less search is done, a high value slows the search down, as more ants run before each pheromone update is performed. A value of 10 appeared to be a good compromise between execution speed and the quality of the solution achieved.

It is interesting to note that for each value of parameters the ant systems scheduling meta-heuristics yields a good solution. Moreover, its convergence speed depends essentially on the number of used ants N^{BA} .

3.6 Building a solution steps

The main steps in the strategy of the FJSP system by ant systems and tabu search algorithm are given below.

- Initialize parameters N^{BA} , α , β , τ_0 , ρ .

- Create an initial *solution* and an empty tabu list of a given size.

In order to generate feasible and diverse solutions, initial ants are represented by solutions issued from heuristic rules [12] (SPT, DL, FIFO, etc) and a random method. Heuristics are used to approximate an optima solution as near as possible.

- **Repeat the following steps until the termination criteria are met:**

- Find *new solution* by ant systems procedure scheduling given in section 3.2.
- Evaluate the quality of the new solution.
- If a *new solution* is improved then the current *best solution* becomes *new solution*
- else If no *new solution* was improved then apply the tabu search optimisation given in section 3.4.
- Add *solution* to the tabu list, if the tabu list is full then delete the oldest entry in the list.
- Apply the updating pheromone trail procedure given in section 3.3.

- **END Repeat**

4 Illustration example

Let us consider a flexible job shop scheduling problem, this example is to execute three jobs J_j ($j=1,2,3$) and six machines M_k ($k = 1, \dots, 6$) described in table 1.

Applying the ant systems meta-heuristic, the simulation propose four different scheduling with $C_{max} = 19$ ut (unit of time), shown in table 2 to 7.

The solution given in the table 7 has a makespan equal to 19 ut. The machine M_5 is the cause of this value of makespan. To solve this problem, the tabu search optimisation is applied for this solution. Indeed, this method finds the operation $O_{2,2}$ for job J_2 on M_2 that can be swapped with other machines which will reduce makespan to 18 ut. And this method finds that the operation $O_{1,3}$ for the job J_1 executed by M_2 and can be swapped with M_5 who will execute the operation $O_{2,2}$ for the job J_2 . Finally, the obtained solution by the tabu search is better than before, table 8.

Table 3: Example benchmark 3 jobs - 6 machines : processing time and ordering operation.

		M_1	M_2	M_3	M_4	M_5	M_6
J_1	$O_{1,1}$	10	7	6	13	5	1
	$O_{2,1}$	4	5	8	12	7	11
	$O_{3,1}$	9	5	6	12	6	17
	$O_{4,1}$	7	8	4	10	15	3
J_2	$O_{1,2}$	15	12	8	6	10	9
	$O_{2,2}$	9	5	7	13	4	7
	$O_{3,2}$	14	13	14	20	8	17
J_3	$O_{1,3}$	7	16	5	11	17	9
	$O_{2,3}$	9	16	8	11	6	3
	$O_{3,3}$	6	14	8	18	21	14

Table 4: Solution 1 for benchmark 3 jobs - 6 machines.

$$N^{BA} = 10; \alpha = 0.5; \beta = 0.5; \tau_0 = 0.01; \rho = 0.5$$

S_1	O_1	O_2	O_3	O_4
J_1	$M_6: [0,1]$	$M_1: [1,5]$	$M_2: [11,16]$	$M_6: [16,19]$
J_2	$M_4: [0,6]$	$M_2: [6, 11]$	$M_5: [11,19]$	***
J_3	$M_3: [0,5]$	$M_3: [5,13]$	$M_1: [13; 19]$	***

Table 5: Solution 2 for benchmark 3 jobs - 6 machines.

$$N^{BA} = 10; \alpha = 0.75; \beta = 0.25; \tau_0 = 0.01; \rho = 0.5$$

S_2	O_1	O_2	O_3	O_4
J_1	$M_6: [0,1]$	$M_1: [1,5]$	$M_2: [5,10]$	$M_6: [10,13]$
J_2	$M_4: [0,6]$	$M_5: [6,10]$	$M_5: [10,18]$	***
J_3	$M_3: [0,5]$	$M_3: [5,13]$	$M_1: [13,19]$	***

Table 6: Solution 3 for benchmark 3 jobs - 6 machines.

$$N^{BA} = 10; \alpha = 0.25; \beta = 0.75; \tau_0 = 0.01; \rho = 0.5$$

S_3	O_1	O_2	O_3	O_4
J_1	$M_6: [0,1]$	$M_1: [1,5]$	$M_5: [5,11]$	$M_6: [11,14]$
J_2	$M_4: [0,6]$	$M_2: [6,11]$	$M_5: [11; 19]$	***
J_3	$M_3: [0,5]$	$M_6: [5,8]$	$M_1: [8,14]$	***

Table 7: Solution 4 for benchmark 3 jobs - 6 machines.

$$N^{BA} = 10; \alpha = 0.3; \beta = 0.7; \tau_0 = 0.01; \rho = 0.5$$

S_4	O_1	O_2	O_3	O_4
J_1	$M_6: [0,1]$	$M_1: [1,5]$	$M_5: [5,11]$	$M_6: [10,13]$
J_2	$M_4: [0,6]$	$M_2: [6,11]$	$M_5: [11,19]$	***
J_3	$M_3: [0,5]$	$M_6: [5,8]$	$M_1: [8,14]$	***

Table 8: Tabu search optimisation solution.

	O_1	O_2	O_3	O_4
J_1	$M_6; 0; 1$	$M_1; 1; 5$	$M_2; 5; 10$	$M_6; 10; 13$
J_2	$M_4; 0; 6$	$M_5; 6; 10$	$M_5; 11; 18$	***
J_3	$M_3; 0; 5$	$M_6; 5; 8$	$M_1; 8; 14$	***

Table 9: Results of problem sets solution.

Problem Set	Results problem sets from [11]		
	Kacem et al	GENACE	Ant systems and tabu search optimisation
FJSP 4-5	16	11	11
FJSP 10-7	15	12	11
FJSP 10-10	7	7	7
FJSP 15-10	23	12	12
Results problem sets from [3]			
FJSP 10-6	32	29	28
FJSP 10-15	86	68	68

Table 10: Result example : FJSP 10-10 from [11]

$N^{BA} = 10; \alpha = 0.1; \beta = 0.9; \tau_0 = 0.1; \rho = 0.25$

*	O_1	O_2	O_3
J_1	$M_7 : [0,2]$	$M_3 : [2,3]$	$M_4 : [3,4]$
J_2	$M_1 : [1,3]$	$M_{10} : [3,4]$	$M_{10} : [4,6]$
J_3	$M_{10} : [0,1]$	$M_8 : [1,2]$	$M_8 : [2,4]$
J_4	$M_9 : [0,1]$	$M_3 : [3,6]$	$M_4 : [6,7]$
J_5	$M_9 : [1,3]$	$M_9 : [3,4]$	$M_4 : [4,5]$
J_6	$M_6 : [1,3]$	$M_9 : [4,6]$	$M_9 : [6,7]$
J_7	$M_1 : [0,1]$	$M_3 : [1,2]$	$M_4 : [2,3]$
J_8	$M_5 : [0,2]$	$M_2 : [2,5]$	$M_2 : [5,7]$
J_9	$M_3 : [0,1]$	$M_7 : [2,3]$	$M_6 : [3,4]$
J_{10}	$M_6 : [0,1]$	$M_4 : [1,2]$	$M_7 : [3,5]$

Table 11: Result example : FJSP 10-6 from [3]
 $N^{BA} = 10; \alpha = 0.25; \beta = 0.75; \tau_0 = 0.2; \rho = 0.5$

*	O_1	O_2	O_3	O_4	O_5	O_6
J_1	$M_6 : [0,1]$	$M_6 : [4,9]$	$M_1 : [9,10]$	$M_2 : [15,21]$	$M_6 : [23,26]$	$M_2 : [27,28]$
J_2	$M_4 : [1,3]$	$M_6 : [9,12]$	$M_5 : [13,15]$	$M_3 : [16,19]$	$M_3 : [19,21]$	$M_1 : [21,24]$
J_3	$M_1 : [2,3]$	$M_1 : [5,7]$	$M_2 : [14,15]$	$M_5 : [15,17]$	$M_4 : [21,22]$	$M_5 : [24,27]$
J_4	$M_5 : [0,2]$	$M_6 : [3,4]$	$M_2 : [8,14]$	$M_1 : [14,15]$	$M_4 : [15,20]$	$M_4 : [20,21]$
J_5	$M_1 : [0,2]$	$M_2 : [2,5]$	$M_4 : [5,11]$	$M_1 : [11,13]$	$M_4 : [13,15]$	$M_3 : [24,27]$
J_6	$M_5 : [2,5]$	$M_3 : [8,12]$	$M_6 : [12,14]$	$M_1 : [15,17]$	$M_2 : [21,27]$	$M_6 : [27,28]$
J_7	$M_4 : [0,1]$	$M_5 : [5,7]$	$M_1 : [7,9]$	$M_5 : [10,13]$	$M_6 : [20,23]$	$M_4 : [24,27]$
J_8	$M_3 : [1,4]$	$M_2 : [5,8]$	$M_1 : [10,12]$	$M_5 : [17,20]$	$M_5 : [20,24]$	$M_1 : [24,27]$
J_9	$M_3 : [4,8]$	$M_3 : [12,16]$	$M_6 : [16,17]$	$M_1 : [17,21]$	$M_3 : [21,24]$	$M_4 : [27,28]$
J_{10}	$M_6 : [1,3]$	$M_1 : [3,5]$	$M_5 : [7,10]$	$M_4 : [11,12]$	$M_6 : [17,20]$	$M_4 : [22,24]$

5 Validation and comparison

All ant systems and tabu search optimisation results presented are for 1000 iterations with 10 the number of ants, and each run was performed 10 times. The algorithms have been coded in VB language and tested using a P4 Pentium processor 2.4 GHz and Windows XP system.

To illustrate the effectiveness and performance of the algorithm proposed in this paper, six representative benchmark FJSP instances (represented by problem $n \times m$) based on practical data have been selected to compute. These benchmark instances are all taken from of Brandimarte [3] and from Kacem [11] as well as those from GENACE [14]. The different results obtained by proposed approach is presented and compared with the other methods in table 9.

Concerning the FJSP instances, the different results show that the solutions obtained are generally acceptable and satisfactory. The values of the different objective functions show the efficiency of the suggested approach, table 9. Moreover, the proposed method enables us to obtain good results in a polynomial computation time. In fact, the efficiency of this approach can be explained by the quality of the ant system algorithms combined by the tabu search heuristic to the optimization of solutions.

6 Conclusion

In this paper, a new approach based on the combination of the ant system with tabu search algorithm for solving flexible job-shop scheduling problems, is presented. The results for the reformulated problems show that the ant systems with local search meta-heuristic can find optimal solutions for different problems that can be adapted to deal with the FJSP problem. The performances of the new approach are evaluated and compared with the results obtained from other methods. The obtained results show the effectiveness of the proposed method. Ant system algorithms and the tabu search techniques described are very effective and they alone can outperform all the alternative techniques.

References

- [1] A. C. F. Alvim, F. Glover, C. C. Ribeiro and D. J. Aloise. A hybrid improvement heuristic for the bin packing problem, 2002. Available from: <http://citeseer.nj.nec.com/557429.html>.
- [2] T. D. Braun, H. J. Siegel, N. Beck, L. L. Bölöni, M. Maheswaran, A. I. Reuther, J. P. Robertson, M. D. Theys, B. Yao, D. Hensgen and R. F. Freund. A comparison of eleven static heuristics for

- mapping a class of independent tasks onto heterogeneous distributed computing systems. *Journal of Parallel and Distributed Computing*, 61(6):810-837, 2001.
- [3] Brandimarte P., Routing and Scheduling in a Flexible Job-Shop by Tabu Search, *Annals of Operations Research*, vol. 2, pp. 158-183, 1993.
- [4] Bilchev, G., Parmee, I.C. : The Ant Colony Metaphor for Searching Continuous Design Spaces. *Lecture Notes in Computer Science*, 993, pp. 25-39, 1995.
- [5] A. L. Corcoran and R. L. Wainwright. A parallel island model genetic algorithm for the multiprocessor scheduling problem. In *Selected Areas in Cryptography*, pp. 483-487, 1994.
- [6] M. Dorigo. Optimization, Learning and Natural Algorithms. PhD thesis, DEI, Polytecnico di Milano, Milan, 1992.
- [7] M. Dorigo and T. Stützle. The ant colony optimization metaheuristic: Algorithms, applications, and advances. In Glover, F. and Kochenberger, G., editors, *Handbook of Metaheuristics*, vol. 57 of *International Series in Operations Research and Management Science*, pp. 251-285. Kluwer Academic Publishers, 2002.
- [8] M. Garey, D. Johnson, R. Sethi. The Complexity of Flow Shop and Job-shop Schedules. *Mathematics of Operations Research*, vol. 1(2), pp. 117-129, 1976.
- [9] M. Garey and D. Johnson. *Computers and Intractability: A Guide to the theory of NP-Completeness*. *Freeman and Company*, San Francisco, 1979.
- [10] I. Kacem, S. Hammadi and P. Borne. Approach by Localization and Multiobjective Evolutionary Optimization for Flexible Job-shop Scheduling Problems. *IEEE Transactions on Systems, Man and Cybernetics*, vol. 32(1), pp. 1-13, 2002.
- [11] I. Kacem, S. Hammadi and P. Borne. Pareto-optimality Approach for Flexible Job-Shop Scheduling Problems: Hybridization of Evolutionary Algorithms and Fuzzy Logic. *Mathematics and Computer in Simulation*, vol. 60, pp. 245-276, 2002.
- [12] N. Liouane, S. Hammadi and P. Borne. Robust methodology for scheduling production in uncertain environment. *IMACS Multi-Conference on Computational Engineering in Systems Applications, CESA'98*, Hammamet, 1998.
- [13] K. Mesghouni. Application des algorithmes évolutionnistes dans les problèmes d'optimisation en ordonnancement de production. Thèse de Doctorat, Université de Lille1, Lille, 1998.
- [14] J. C. Tay and N. B. Ho. GENACE: An Efficient Cultural Algorithm for Solving the Flexible Job-Shop Problem. *Proceedings of the IEEE Congress of Evolutionary Computation*, pp. 1759-1766, 2004.
- [15] A. Thiesen., Design and evaluation of tabu search algorithms for multiprocessor scheduling. *Journal of Heuristics*, Vol. 4, pp.141-160, 1998.
- [16] S. van der Zwaan and C. Marques. Ant colony optimisation for job shop scheduling. In *Proceedings of the Third Workshop on Genetic Algorithms and Artificial Life, GAAL 99*, 1999.

Nouredine Liouane¹, Ihsen Saad^{2,3}, Slim Hammadi², Pierre Borne²

¹ATSI : Ecole Nationale d'Ingénieurs de Monastir,
rue Ibn El Jazzar, 5019 Monastir, Tunisie

²Ecole Centrale de Lille, Cité scientifique
Laboratoire d'Automatique, Genie Informatique et Signal
BP 48, 59651 Villeneuve d'Ascq Cedex, France

³Ecole Nationale d'Ingénieurs de Tunis
Unité de recherche LARA-Automatique
BP 37, Le Belvédère, 1002 Tunis, Tunisie

E-mail: noureddine.liouane@enim.rnu.tn, ihsen.saad@enit.rnu.tn,
slim.hammadi@ec-lille.fr, pierre.borne@ec-lille.fr



Nouredine Liouane was born in Kairouan, Tunisia, in 1963. He received the Master degree of science electrical genius in 1988 from the “Ecole Normale Supérieure de l’Enseignement Technique de Tunis”. He received the Ph.D. degree from Ecole Centrale de Lille, France, in 1998. He is currently “Maitre Assistant” in the “Ecole Nationale d’Ingénieurs de Monastir” and Director of the “Institut Supérieure des Sciences Appliquées et de la Technologie de Kairouan”. His research is related to the evolutionary optimization methods for discrete events systems, computer science and operational research.



Ihsen Saad was born in Monastir, Tunisia, in 1977. He received the Engineer Diploma degree in electrical automatic control engineering from the “Ecole Nationale d’Ingénieur de Gabès”, Tunisia, in 2002. He obtained the Master of automatic and signal treatment in 2004 at the “Ecole Nationale d’Ingénieur de Tunis”. He is currently preparing the Ph.D. degree in automatic and computer science within the framework of LAGIS-EC-Lille and LARA-ENIT cooperation. His research is related to the evolutionary optimization methods for discrete events systems, computer science and operational research.



Slim Hammadi is a full Professor of production planning and control at the Ecole Centrale de Lille (French “Grande Ecole”). Born in Gafsa (Tunisia) in 1962, he has obtained by 1988 the Master degree in Computer science from the University of Lille (France). Pr Hammadi obtained a Ph.D degree in job-shop scheduling and control in 1991 at Ecole Centrale de Lille. He is a senior member of IEEE/SMC and has served as a referee for numerous journals including the IEEE Transactions on SMC. Pr. S. Hammadi was Co-Organizer of a Symposium (IMS) of the IMACS/IEEE SMC Multiconference CESA’98 held in Hammamet (Tunisia) in April 1998. He has organized several invited sessions in different SMC conferences where he was session chairman. He was chairman of the International congress on “Logistic and Transport” LT’2004, MHOSI’2005 and LT’2006. His teaching and research interests focus on the areas of production control, production planning, computer science, discrete and dynamic programming and computer integrated manufacturing.



Pierre Borne was born in Corbeil, France in 1944, he received the Master degree of Physics in 1967, the Masters of Electronics, of Mechanics and of Applied Mathematics in 1968. The same year he obtained the Diploma of "Ingénieur IDN" (French "Grande Ecole"). He obtained the PhD in Automatic Control of the University of Lille in 1970 and the DSc of the same University in 1976. He became Doctor Honoris Causa of the Moscow Institute of Electronics and Mathematics (Russia) in 1999 and of the University of Waterloo (Canada) in 2006. He is author or co-author of about 200 Journal articles and book chapters, and of 30 plenary lectures and of more than 250 communications in international conferences. He has been the supervisor of 64 PhD thesis and is author of 20 books. He has been President of the IEEE/SMC society in 2000 and 2001. He is presently Professor "de classe exceptionnelle" at the Ecole Centrale de Lille and director of the french pluriformations national group of research in Automatic Control.

Fault Detection for Large Scale Systems Using Dynamic Principal Components Analysis with Adaptation

Jesús Mina, Cristina Verde

Abstract: The Dynamic Principal Component Analysis is an adequate tool for the monitoring of large scale systems based on the model of multivariate historical data under the assumption of stationarity, however, false alarms occur for non-stationary new observations during the monitoring phase. In order to reduce the false alarms rate, this paper extends the DPCA based monitoring for non-stationary data of linear dynamic systems, including an on-line means estimator to standardize new observations according to the estimated means. The effectiveness of the proposed methodology is evaluated for fault detection in a interconnected tanks system.

Keywords: Fault Detection, Statistical Analysis, Dynamic Principal Component Analysis, Time Series Analysis, Non-Stationary Signals.

1 Introduction

The on-line process monitoring for fault detection and isolation, FDI, is an important task to ensure plant safety and product quality. One of the most consolidated FDI techniques of the last twenty years is the analytical approach, which is based on explicit modeling, this is, models obtained from primary physical principles, some of the analytical approaches are very well revised e.g. [1]. In the case of large scale processes whose analytical models are not available or are difficult to obtain the FDI techniques based on data-driven can help to overcome the problem of modeling, these techniques are based on implicit modeling through multivariate statistical methods, some of these methods are resumed in [2].

Principal Component Analysis (PCA) is a multivariate statistical method which models the linear correlation structure of a multivariate process from nominal historical data. PCA transforms a set of multivariate observations to a lower dimension orthogonal space, retaining the most variability of the original data [3]. Because of the simplification and the orthogonal property obtained with PCA, this has been used with success for fault diagnosis issues as in [4] and [5].

It is important to note that like other multivariate statistical methods, PCA works under three assumptions: the data follows a multivariate normal distribution; there exist no auto-correlation among observations; and the variables are stationary, this is, the variables should keep constant mean and standard deviation over time [6], [7]. In the case of data with non normal distribution it is possible to carry out an appropriate transformation like square root or logarithm [8], in order to improve the distribution of data. In the case of dynamic systems the auto-correlation in variables is taking into account incorporating time lags of the time series during the modeling stage, this extension is called dynamic principal component analysis, DPCA [9].

Since PCA and DPCA assume stationarity during modeling process, high rate of false alarms are generated in the diagnosis stage if the test data are non-stationary. The non stationary problem has been tackled with adaptive versions of PCA like in [10] and [11]. Although these algorithms adapt means, covariance, and the PCA model, however, they can not be used for FDI tasks since the adaptation is based in the variations of actual multivariate observations without distinguish the real causes of changes in the variables.

A non-stationary condition has many possible causes e.g. due to components aging, to faults, even to normal changes in the operating point of the plant, this problem motivated the development of a fault detection algorithm which is robust to changes in the operating point but sensitive to faults.

The proposal is based in the fact that the correlation structure between system signals is invariant under nominal conditions, i.e. the relations between variables is the same despite the nominal changes in the mean of the signals, this is a result of the affinity property of nominal signals in linear systems [12]; therefore this work proposes: in the modeling stage, obtain a nominal DPCA model from nominal historical data and identify a set of nominal inputs-output relations; in the diagnosis stage, keep the nominal DPCA model but estimate actual means of input variables through exponentially weighted moving average (EWMA) and estimate means of the output variables from input means through the identified nominal inputs-output relations in order to carry out an appropriate standardization with respect to the estimated means.

In the following sections the recursive means estimation process is summarized; next, the proposed extension of DPCA based fault detection for changes in the operating point will be described. Finally the methodology will be evaluated for faults detection in a three interconnected tanks system.

2 Backgrounds

2.1 Identification of the Inputs-Output Relations

It is proposed the recursive mean estimation of the input signals using EWMA and the outputs mean estimation from inputs estimated means and through inputs-output nominal relations. Here is proposed to identify the inputs-output relations with moving average models, MA(), for each one of the output variables.

Lets consider the case of a MIMO linear system with r inputs and s outputs

$$\mathbf{y} = \mathbf{A}\mathbf{u} \quad (1)$$

$$\begin{aligned} \mathbf{u} &= \mathbf{u}_n + \boldsymbol{\eta} \\ \mathbf{y} &= \mathbf{y}_n + \mathbf{v} \end{aligned} \quad (2)$$

where $\boldsymbol{\eta}$ and \mathbf{v} are stationary white noise vectors added to the inputs and to the outputs, with zero mean and variances $\sigma_{\boldsymbol{\eta}}^2$ and $\sigma_{\mathbf{v}}^2$, respectively.

Each one of the output variables can be expressed as a linear combination of the inputs with corresponding time lag orders $q_{1i}, q_{2i}, \dots, q_{ri}$

$$y_{n_i}(t) = \sum_{k=0}^{q_{1i}} a_{k1i} u_1(t-k) + \dots + \sum_{k=0}^{q_{ri}} a_{kri} u_r(t-k) \quad (3)$$

for $i = 1, \dots, s$. The a_{k1i}, \dots, a_{kri} parameters are obtained through correlation analysis [13].

The identified input-output relations can be expressed in compact form as follows

$$\hat{y}_{n_i}(t) = \hat{\mathbf{a}}_i \vec{\mathbf{u}}_i(t) \quad (4)$$

where

$$\hat{\mathbf{a}}_i = \left[\hat{a}_{01i} \quad \dots \quad \hat{a}_{q_{1i}1i} \quad \dots \quad \hat{a}_{0ri} \quad \dots \quad \hat{a}_{q_{ri}ri} \right]$$

and

$$\vec{\mathbf{u}}_i(t) = [u_1(t) \quad \cdots \quad u_1(t - q_{1i}) \quad \cdots \quad u_r(t) \quad \cdots \quad u_r(t - q_{ri})]^T$$

The orders q_{ji} of the model (4) is selected through a validation procedure taking the minimal value of the sum of square error as a function of q_{ji} .

Finally, from (4) for all of the outputs, the relations in matrix notation are given by

$$\hat{\mathbf{y}}_n(t) = \hat{\mathbf{A}} \vec{\mathbf{u}}(t) \quad (5)$$

where

$$\hat{\mathbf{y}}_n(t) = [\hat{y}_{n_1}(t) \quad \cdots \quad \hat{y}_{n_s}(t)]_{(s \times 1)}^T$$

$$\vec{\mathbf{u}}(t) = [u_1(t) \quad \cdots \quad u_1(t - L) \quad \cdots \quad u_r(t) \quad \cdots \quad u_r(t - L)]_{(o \times 1)}^T$$

and $\hat{\mathbf{A}} \in \mathfrak{R}^{s \times o}$ is made up by the coefficients

$$\hat{a}_{kji} = \begin{cases} \hat{a}_{kji} & , \quad k \leq q_{ji} \\ 0 & , \quad q_{ji} < k \leq L \end{cases} \quad \text{for } \begin{matrix} j = 1, \dots, r \\ i = 1, \dots, s \end{matrix}$$

$L = \text{Max} \{ q_{11}, \dots, q_{r1}, \dots, q_{1s}, \dots, q_{rs} \}$ and $o = r(L + 1)$.

2.2 Means Recursive Estimation

Once the identification of MA models (5), for the system (1), are obtained, the recursive means estimation is carried out in the following way.

The inputs mean recursive estimation can be computed as follows

$$\hat{\mu}_{\vec{\mathbf{u}}}(t) = \beta \hat{\mu}_{\vec{\mathbf{u}}}(t-1) + (1 - \beta) \vec{\mathbf{u}}(t) \quad (6)$$

with $(0 < \beta \leq 1)$ the forgetting factor of EWMA.

By the other side, the outputs means at time instant t are estimated from the input means given in (6) and through (5) according with

$$\hat{\mu}_{\mathbf{y}(t)} = \hat{\mu}_{\mathbf{y}_n(t)} = \hat{\mathbf{A}} \hat{\mu}_{\vec{\mathbf{u}}}(t) \quad (7)$$

3 DPCA Based Fault Detection with Mean Parameter Estimation

The DPCA statistical tool is used to obtain an implicit model of a dynamic system from nominal historical data, and use this implicit model to carry out fault detection tasks. The proposed algorithm is illustrated in Fig. 1. The idea is not only to obtain a DPCA based statistical model but also identify nominal inputs-output relations. So, estimating actual means of input variables through EWMA and estimating means of the output variables from input means using the nominal inputs-output relations, an appropriate standardization can be carried out.

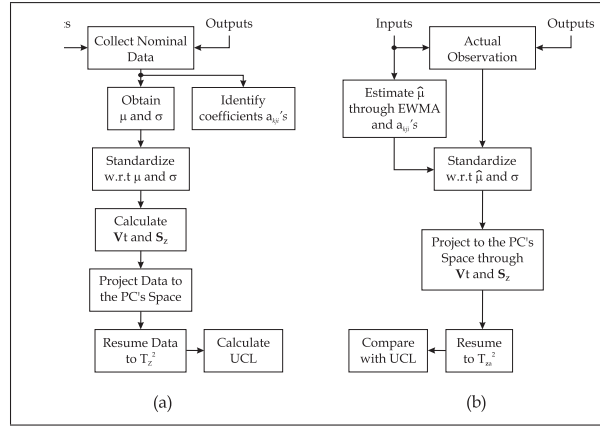


Figure 1: Proposed fault detection algorithm. (a) Modeling stage carried out off-line, (b) Detection stage carried out on-line

3.1 DPCA based Statistical Modeling

Let matrix \mathbf{X} be a set of historical data made up of n_t observations of r input variables and s output variables, taken from a dynamic system working under nominal conditions and around an operating point

$$\mathbf{X} = \begin{bmatrix} \mathbf{u}_1 & \cdots & \mathbf{u}_r & \mathbf{y}_1 & \cdots & \mathbf{y}_s \end{bmatrix}_{(n_t \times p)} \quad (8)$$

Each column in \mathbf{X} represents an auto-correlated time series. In DPCA the serial correlation is included constructing the named *trajectory matrix* applying w time lags on each time series, this is

$$\vec{\mathbf{X}} = \begin{bmatrix} \vec{\mathbf{U}}_1 & \cdots & \vec{\mathbf{U}}_r & \vec{\mathbf{Y}}_1 & \cdots & \vec{\mathbf{Y}}_s \end{bmatrix}_{(n \times m)} \quad (9)$$

where e.g. $\vec{\mathbf{U}}_1 = [\mathbf{u}_1(t) \ \mathbf{u}_1(t-1) \ \dots \ \mathbf{u}_1(t-w)]$; $n = n_t - w$ and $m = p(w+1)$.

To avoid that some particular variables dominate the modeling process, it is convenient to carry out a data standardization in matrix $\vec{\mathbf{X}}$ in relation to its means and standard deviations. Thus, the means of $\vec{\mathbf{X}}$ are given by

$$\hat{\boldsymbol{\mu}}_{\vec{\mathbf{X}}} = \left[\frac{1}{n} \vec{\mathbf{X}}^T \mathbf{1} \right]^T = \begin{bmatrix} \hat{\boldsymbol{\mu}}_{\vec{\mathbf{U}}} & \hat{\boldsymbol{\mu}}_{\vec{\mathbf{Y}}} \end{bmatrix}_{(1 \times m)} \quad (10)$$

where $\mathbf{1} = [1, 1, \dots, 1]^T \in \mathfrak{R}^n$. By the other side, the covariance matrix of $\vec{\mathbf{X}}$ is given by

$$\mathbf{S} = \frac{1}{n-1} \left(\vec{\mathbf{X}} - \mathbf{1} \hat{\boldsymbol{\mu}}_{\vec{\mathbf{X}}} \right)^T \left(\vec{\mathbf{X}} - \mathbf{1} \hat{\boldsymbol{\mu}}_{\vec{\mathbf{X}}} \right)$$

from which the standard deviations can be obtained

$$\hat{\boldsymbol{\sigma}}_{\vec{\mathbf{X}}} = \sqrt{\text{diag}(\mathbf{S})} = \begin{bmatrix} \hat{\boldsymbol{\sigma}}_{\vec{\mathbf{U}}} & \hat{\boldsymbol{\sigma}}_{\vec{\mathbf{Y}}} \end{bmatrix}_{(1 \times m)} \quad (11)$$

Thus, the data standardization is computed in the following way

$$\tilde{\mathbf{X}}(i, j) = \frac{\vec{\mathbf{X}}(i, j) - \hat{\boldsymbol{\mu}}_{\vec{\mathbf{X}}}(j)}{\hat{\boldsymbol{\sigma}}_{\vec{\mathbf{X}}}(j)}$$

for $i = 1, \dots, n$ and $j = 1, \dots, m$.

The uncorrelated principal components \mathbf{Z} of dimension $n \times l$ are obtained through the following transformation

$$\mathbf{Z} = \tilde{\mathbf{X}}\mathbf{V}_t \quad (12)$$

where the orthonormal transformation matrix $\mathbf{V}_t \in \mathfrak{R}^{m \times l}$ is composed of an appropriate selection of l eigenvectors, called loading vectors, associated to the correlation matrix $\mathbf{R} = \frac{1}{n-1} \tilde{\mathbf{X}}^T \tilde{\mathbf{X}}$.

The data matrix $\tilde{\mathbf{X}}$ can be expressed as

$$\tilde{\mathbf{X}} = \hat{\tilde{\mathbf{X}}} + E = \mathbf{Z}\mathbf{V}_t^T + E$$

where $\hat{\tilde{\mathbf{X}}}$ is the information captured by the l -principal components and E is the information in the neglected $m - l$ -components. So, for the detection purpose it is possible to use the Hotelling statistic, $T_{Z_i}^2$ from \mathbf{Z} and/or the squared prediction error (SPE) from E . In this paper the Hotelling statistic is used for fault detection because our interest is not the evaluation of these two statistical parameters but the illustration of reduction of false alarms.

For each l -variate observation in \mathbf{Z} the Hotelling univariate statistic $T_{Z_i}^2$ is given by

$$T_{Z_i}^2 = \mathbf{Z}_i \mathbf{S}_{\mathbf{Z}}^{-1} \mathbf{Z}_i^T \quad (13)$$

where $\mathbf{S}_{\mathbf{Z}}$ is the covariance matrix of \mathbf{Z} . Finally, a threshold of normal condition from the probability density function of the set of parameters $T_{Z_i}^2$ is calculated. [6] propose, among others, for a beta distribution of the data set $T_{Z_i}^2$ the threshold UCL as

$$UCL = \frac{(n-1)^2 \left(\frac{l}{n-l-1}\right) F\left(\frac{\alpha}{2}; l, n-l-1\right)}{n \left(1 + \left(\frac{l}{n-l-1}\right) F\left(\frac{\alpha}{2}; l, n-l-1\right)\right)} \quad (14)$$

where n and l are the dimensions of \mathbf{Z} and α is a level of significance.

In a DPCA based modeling conventional approach the implicit model consist of the means and standard deviations vectors (10), (11); the loading vectors in \mathbf{V}_t ; the variance in the principal components given by $\mathbf{S}_{\mathbf{Z}}$; and the nominal threshold UCL. However, according with the proposal it is just considered $\hat{\sigma}_{\tilde{\mathbf{X}}}$, \mathbf{V}_t , $\mathbf{S}_{\mathbf{Z}}$ and UCL, since $\hat{\mu}_{\tilde{\mathbf{X}}}$ will be estimated recursively. Additionally to the statistical modeling it is carried out the identification of the inputs-output relations (5).

3.2 Fault Detection

DPCA detects a deviation of vector of actual observation $\vec{\mathbf{x}}_a$ from the nominal reference in terms of its mean and its standard deviation. However, it is important to note that the modeling process is based in the data set $\vec{\mathbf{X}}$ which was obtained in a particular operating point of the system, so any change in the nominal values of the signals is interpreted by DPCA as a fault, even when the process is healthy, this misinterpretation is because of the time variant behavior of components in (10).

For linear systems, a change in the operating point means a new assignment in the input variables with consequent variations in the output variables, this is, changes in the mean values of input and output variables but no changes in their correlation structure. However, faults in the system produce changes in the mean values and in the correlation structure between variables. Thus, here it is proposed during the detection stage, the on line estimation of the statistical set (10) using nominal linear inputs-output relations in order to adapt the standardization procedure.

So, according to the proposed extension to DPCA based fault detection algorithm, the procedure to evaluate and classify an actual observation $\vec{\mathbf{x}}_a \in \mathfrak{R}^{1 \times m}$ is summarized as follows. Be the actual observation vector, with input and output variables, expressed in w time lags

$$\vec{\mathbf{x}}_a = [\vec{\mathbf{u}}_{a1} \cdots \vec{\mathbf{u}}_{ar} \quad \vec{\mathbf{y}}_{a1} \cdots \vec{\mathbf{y}}_{as}] \quad (15)$$

1. Estimate through (6) the means of the actual input data, $\hat{\mu}_{\vec{\mathbf{u}}_a}$, and through (7) the nominal means of the output variables, $\hat{\mu}_{\vec{\mathbf{y}}}$; next construct the vector

$$\hat{\mu}_{\vec{\mathbf{x}}_a} = [\hat{\mu}_{\vec{\mathbf{u}}_a} \quad \hat{\mu}_{\vec{\mathbf{y}}}]_{(1 \times m)} \quad (16)$$

2. Standardize the m terms in (15) using the means estimated given in (16) and the historical standard deviations (11), this is

$$\tilde{\mathbf{x}}_a(j) = \frac{\vec{\mathbf{x}}_a(j) - \hat{\mu}_{\vec{\mathbf{x}}_a}(j)}{\hat{\sigma}_{\vec{\mathbf{x}}}(j)}$$

for $j = 1, \dots, m$.

3. Transform the $\tilde{\mathbf{x}}_a$ vector to the principal components subspace \mathbf{z}_a through \mathbf{V}_t

$$\mathbf{z}_a = \tilde{\mathbf{x}}_a \mathbf{V}_t$$

4. Map \mathbf{z}_a in the behaviour symptom $T_{z_a}^2$ through

$$T_{z_a}^2 = \mathbf{z}_a \mathbf{S}_Z^{-1} \mathbf{z}_a^T$$

5. If the resulting value deviates from the normal condition threshold UCL then a fault is present in the system.

The key of the proposed methodology is in the continuous estimation of nominal means (16) using the nominal linear inputs-output relations (5) in order to carry out an appropriate standardization.

In the following section the proposed fault detection algorithm is applied to detect faults in a three interconnected tanks system considering simple relations for the means estimation.

4 Three Tanks System

The tanks system is composed of three cylindrical tanks, interconnected at the bottom by pipes and with valves V_1 in the link between tanks 2 and 3, and V_2 in the link between tank 2 and the outside, which aperture can be manipulated in order to emulate faults (e.g. pipe blockage), see Fig. 2. The tank dimensions are: $h_T = 0.63m$, $A_T = 0.01539m^2$. The system is feed by two inputs Q_1 to the tank 1 and Q_2 to the tank 2 which are measured just as the output variables h_1 , h_2 and h_3 which correspond to tanks levels. The mathematical model is the following

$$\begin{aligned} A_T \frac{dh_1}{dt} &= Q_1 + Q_{31} - Q_{10} & Q_{10} &= K_1 \sqrt{h_1} \\ A_T \frac{dh_3}{dt} &= Q_{23} - Q_{31} & Q_{31} &= K_{31} \rho(h_3 - h_1) \\ A_T \frac{dh_2}{dt} &= Q_2 - Q_{23} - Q_{20} & Q_{23} &= K_{23} \rho(h_2 - h_3) \\ & & Q_{20} &= K_2 \sqrt{h_2} \end{aligned} \quad (17)$$

where $\rho(x) \triangleq \text{sgn}(x) \sqrt{|x|}$.

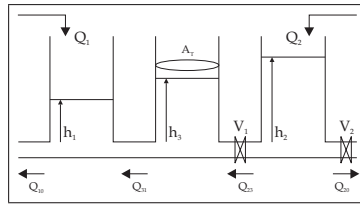


Figure 2: Three Tanks System

For the experiments the system was simulated under the following operation point: $Q_1^0 = 4.75e - 5 m^3/s$, $(\sigma_{Q_1}^2 = 1.07e - 10)$; $Q_2^0 = 7.35e - 5 m^3/s$, $(\sigma_{Q_2}^2 = 1.05e - 10)$; $h_1^0 = 0.147m$, $(\sigma_{h_1}^2 = 1.96e - 4)$; $h_2^0 = 0.276m$, $(\sigma_{h_2}^2 = 4.81e - 4)$; $h_3^0 = 0.195m$, $(\sigma_{h_3}^2 = 2.65e - 4)$; $K_1 = 1.816e - 4$, $K_{31} = 1.005e - 4$, $K_2^0 = 9.804e - 5$ and $K_{23}^0 = 7.804e - 5$.

Taking a set of 400 nominal observations measured every 10s it was obtained a DPCA based principal components space of dimensions 301×68 , so, for an $\alpha = 0.01$ the resulting threshold is $UCL = 95.886$. By the other side, the inputs-output relations identified were $h_1 = f(Q_1, Q_2, q_1)$, $h_2 = f(Q_1, Q_2, q_2)$ and $h_3 = f(Q_1, Q_2, q_3)$ with time lags of order $q_1 = 61$, $q_2 = 61$ and $q_3 = 60$, respectively.

4.1 Detection Results

Using a forgetting factor of $\beta = 0.95$ for the recursive inputs means estimation (6) the fault detection algorithm is evaluated considering the following cases:

1. Fault condition, blockage in the pipe which links tanks 2 and 3, the fault occurrence is at $8000s$.
2. Normal operation of the system during $15000s$, with changes in the means of U_1 of $+20\%$ in $3000s < t < 6000s$; -20% in $9000s < t < 12000s$ and in U_2 of $+20\%$ in $4500s < t < 7500s$; -20% in $10500s < t < 13500s$.
3. Change in the mean of U_1 of $+20\%$ from $4000s$ and fault condition, blockage in the pipe between tank 2 and 3, at $8000s$.

The first test is to compare the performance of the DPCA based conventional fault detection and the fault detection based in the proposed algorithm, under fault conditions. The monitoring results are given in Fig. 3 which shows that both algorithms are able to detect the fault.

The second test evaluate the performance of both algorithms before changes in the operation point, the monitoring results are given in Fig. 4, where it is cleared observed that the traditional DPCA-based fault detection (Mon1) interprets the normal changes in the operation point as faults, however, the proposed algorithm (Mon2) is robust before these changes, which reduces the false alarm rate.

Finally, the third test shows the capability of the proposed fault detection algorithm to distinguish between normal variations in the operating point and the presence of faults, see Fig. 5.

5 Conclusions

Here, a modification to the DPCA algorithm for fault detection has been proposed, in which an appropriate standardization with respect to on-line estimated statistical parameters is carried out if simple

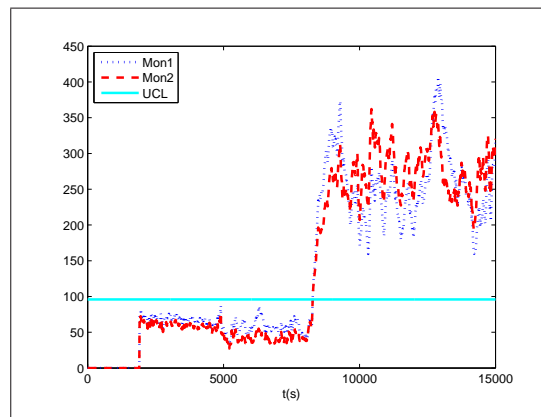


Figure 3: Fault condition: UCL - Threshold of normal condition; Mon1 - DPCA-based monitoring; Mon2 - DPCA with adaptation

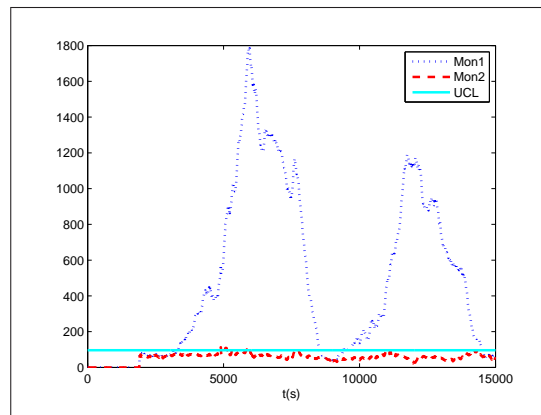


Figure 4: Normal condition: UCL - Threshold of normal condition; Mon1 - DPCA-based monitoring; Mon2 - DPCA with adaptation

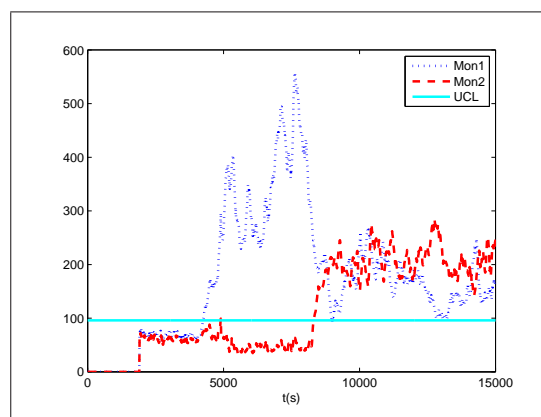


Figure 5: Normal and Fault condition: UCL - Threshold of normal condition; Mon1 - DPCA-based monitoring; Mon2 - DPCA with adaptation

healthy relations between variables can be obtained. This idea allows to deal with non-stationary signals and to reduce significantly the rate of false alarms. It was shown through a series of tests the effectiveness of the proposed fault detection algorithm to distinguish between normal changes in signals and the variations due to the presence of faults.

References

- [1] R. J. Patton, P. M. Frank, R. N. Clark, *Issues of Fault Diagnosis for Dynamic Systems*, Springer-Verlag, 1989, London.
- [2] L. H. Chiang, E. L. Russell, R. D. Braatz, *Fault Detection and Diagnosis in Industrial Systems*, Advanced Textbooks in Control and Signal Processing, Springer-Verlag, 2001, London.
- [3] J. E. Jackson, *A Users Guide to Principal Components*, John Wiley, 1991, New York.
- [4] J. V. Kresta, J. F. MacGregor, T. E. Marlin, Multivariate Statistical Monitoring of Process Operating Performance, *The Canadian Journal of Chemical Engineering*, Vol. 69, pp. 35-47, February, 1991.
- [5] A. Raich, A. Çinar, Statistical Process Monitoring and Disturbance Diagnosis in Multivariable Continuous Processes, *AIChE Journal*, Vol. 42, No. 4, pp. 995-1009, April, 1996.
- [6] N. D. Tracy, J. C. Young, R. L. Mason, Multivariate Control Charts for Individual Observations, *Journal of Quality Technology*, Vol. 24, No. 2, pp. 88-95, April, 1992.
- [7] A. Norvilas, A. Negiz, J. DeCicco, A. Çinar, Intelligent Process Monitoring by Interfacing Knowledge-based Systems and Multivariate Statistical Monitoring, *Journal of Process Control*, Vol. 10, No. 4, pp. 341-350, August, 2000.
- [8] D. C. Montgomery, *Introduction to Statistical Quality Control*, John Wiley, 2001, New York.
- [9] W. Ku, R. H. Storer, Ch. Georgakis, Disturbance Detection and Isolation by Dynamic Principal Component Analysis, *Chemometrics and Intelligent Laboratory Systems*, Vol. 30, No. 1, pp. 179-196, November, 1995.
- [10] N. B. Gallagher, B. M. Wise, S. W. Butler, D. D. White, G. G. Barna, Development and Benchmarking of Multivariate Statistical Process Control Tools for a Semiconductor Etch Process: Improving Robustness Through Model Updating, *ADCHEM'97*, Banff, Canada, pp. 78-83, 9-11 June, 1997.
- [11] W. Li, H. H. Yue, S. Valle-Cervantes, S. J. Qin, Recursive PCA for Adaptive Process Monitoring, *Journal of Process Control*, Vol. 10, No. 5, pp. 471-486, October, 2000.
- [12] T. Kailath, A. H. Sayed, B. Hassibi, *Linear Estimation*, Prentice Hall, 2000, New Jersey.
- [13] G. E. P. Box, G. M. Jenkins, G. C. Reinsei, *Time Series Analysis: Forecasting and Control*, Prentice Hall, 1994, New Jersey.

Jesús Mina, Cristina Verde
Instituto de Ingeniería-UNAM
Automatización
Coyoacán, DF, 04510, México, Fax: (52)-55-56233600 ext 8052
E-mail: jminaa@iingen.unam.mx, verde@servidor.unam.mx
Received: January 15, 2007



Jesús Mina received the BS degree in Electric Eng. from Tuxtla Gutiérrez Technological Institute, México in 1999; the MS degree in Electronic Eng. from the Research and Technological Development National Center, México in 2002; and currently is a student in the PhD program in Electrical Eng. of the National University from México. He was professor from 2002 to 2003 in the Zacatepec Technological Institute. He has carried out research in non-linear control for power active filters and currently is interested in fault diagnosis based in multivariate statistical analysis. Jesús Mina is a member of the International Society of Automation.



Cristina Verde received the BS degree in Electronic and Communication Eng. from National Polytechnic, México in 1973, the MS degree in Electrical Eng. from the National Polytechnic, in México 1974 and the PhD Degree in Electrical Eng. from the Duisburg University in Germany in 1983. In 1984, she joined the National University from México (UNAM) and became the head of the automatic control department in 1988. She has been the Coordinator of the Postgraduate School in Computer Science and Engineering from the National University. She has used the control theory to improve distribution, regulation and quality of water in México and her main topics of research interests include automatic fault detection and diagnosis for dynamic systems and integrity of industrial process. She got the Prize Sor Juana Ines de la Cruz given by the National University to the outstanding women in the engineering field in 2005.

Robust PID Decentralized Controller Design Using LMI

Danica Rosinová, Vojtech Veselý

Abstract: The new LMI based method for robust stability analysis for linear uncertain system with PID controller is proposed. The general constrained structure of controller matrix is considered appropriate for both output feedback and decentralized control and the respective guaranteed cost control design scheme is presented. The sufficient robust stability condition is developed for extended quadratic performance index including first derivative of the state vector to damp oscillations. The obtained stability condition is formulated for parameter-dependent Lyapunov function.

Keywords: Uncertain systems, Robust stability, Decentralized control, Linear matrix inequalities (LMI), Lyapunov function

1 Introduction

Robust stability and robust control belong to fundamental problems in control theory and practice; various approaches in this field have been proposed to cope with uncertainties that always appear in real plant ([2];[8];[7];[5];[4]). The development of Linear Matrix Inequality (LMI) computational techniques has brought an efficient tool to solve a large set of convex problems in polynomial time (e.g. [2]). Significant effort has been therefore made to formulate crucial control problems in algebraic way ([12]), so that the numerical LMI solution can be adopted. This approach is advantageously used in solving control problems for linear systems with convex (affine or polytopic) uncertainty domain. However, many important problems in linear control design, such as decentralized control, simultaneous SOF or more generally - structured linear control problems have been proven as NP hard ([1]). Intensive research has been devoted to overcome nonconvexity and transform the nonconvex or NP-hard problem into convex optimisation problem in LMI framework. Various techniques have been developed using inner or outer convex approximation of the respective nonconvex domain. The common tool in both inner and outer approximation is the use of linearization or convexification. In ([6]; [3]) the general convexifying algorithm for the nonconvex function together with potential convexifying functions for both continuous and discrete-time case have been proposed. Linearization approach for continuous and discrete-time system design was independently used in ([13];[11]).

Proportional-integral-derivative (PID) controllers belong to the most popular ones in the industrial world. The derivative part of the controller, however, causes difficulties when uncertainties are considered. In multivariable PID control schemes using LMI developed recently ([14]) the incorporation of the derivative part requires inversion of the respective matrix, which does not allow including uncertainties. The other way to cope with the derivative part is to assume the special case when output and its derivative are state variables, robust PID controller for first and second order SISO systems are proposed for this case in ([7]).

In this paper a state space approach to designing decentralized (multi-loop) PID robust controllers is proposed for linear uncertain system with guaranteed cost with a new quadratic cost function. The major contribution is in considering the derivative part in robust control framework. We adopt the new PID control problem formulation using LMI that is appropriate for polytopic uncertain systems. The robust PID control scheme is proposed for structured control gain matrix, thus enabling decentralized PID control design. In Section 2 the robust control design problem with structured control gain matrix is formulated in general. The robust optimal control design procedure in state space with the extended cost function is proposed in Section 3. The main result-robust PID controller design approach is provided in Section

4, the developed approach is appropriate for decentralized control structure. In Section 5 the results are illustrated on the examples.

2 Problem Formulation And Preliminaries

Consider a linear affine uncertain system:

$$\begin{aligned}\delta x(t) &= (A + \delta A)x(t) + (B + \delta B)u(t) \\ y(t) &= Cx(t)\end{aligned}\quad (1)$$

where

$$\begin{aligned}\delta x(t) &= \dot{x}(t) \text{ for continuous - time system} \\ \delta x(t) &= x(t+1) \text{ for discrete - time system}\end{aligned}$$

$x(t) \in R^n$, $u(t) \in R^m$, $y(t) \in R^l$ are state, control and output vectors respectively; A, B, C are known constant matrices of appropriate dimensions corresponding to the nominal system, $\delta A, \delta B$ are matrices of uncertainties of the respective dimensions. The uncertainties are considered to be affine of the form

$$\delta A = \sum_{j=1}^p \varepsilon_j \tilde{A}_j, \quad \delta B = \sum_{j=1}^p \varepsilon_j \tilde{B}_j \quad (2)$$

where $\underline{\varepsilon}_j \leq \varepsilon_j \leq \bar{\varepsilon}_j$ are unknown uncertainty parameters; $\tilde{A}_j, \tilde{B}_j, j = 1, 2, \dots, p$ are constant matrices of uncertainties of the respective dimensions and structure. The uncertain system (1), (2) can be equivalently described by a polytopic model given by its vertices

$$\{(A_1, B_1, C), (A_2, B_2, C), \dots, (A_N, B_N, C)\}, N = 2^p.$$

The decentralized feedback control law is considered in the form

$$u(t) = FCx(t) \quad (3)$$

where F is a matrix corresponding to decentralized controller. The uncertain closed-loop polytopic system is then

$$\delta x(t) = A_C(\alpha)x(t) \quad (4)$$

where

$$\begin{aligned}A_C(\alpha) &\in \left\{ \sum_{i=1}^N \alpha_i A_{Ci}, \sum_{i=1}^N \alpha_i = 1, \alpha_i \geq 0 \right\}, \\ A_{Ci} &= A_i + B_i FC.\end{aligned}\quad (5)$$

To assess the performance quality a quadratic cost function known from LQ theory is often used. However, in practice the response rate or overshoot are often limited. Therefore we include into the cost function the additional derivative term for state variable to open the possibility to damp the oscillations and limit the response rate.

$$J_c = \int_0^{\infty} [x(t)^T Qx(t) + u(t)^T Ru(t) + \delta x(t)^T S \delta x(t)] dt$$

for a continuous-time and

$$J_d = \sum_{k=0}^{\infty} [x(k)^T Qx(k) + u(k)^T Ru(k) + \delta x(k)^T S \delta x(k)] \quad (6)$$

for a discrete-time system,

where $Q, S \in R^{n \times n}, R \in R^{m \times m}$ are symmetric positive definite matrices. The concept of guaranteed cost control is used in a standard way: Let there exist a feedback gain matrix F_0 and a constant J_0 such that

$$J \leq J_0 \quad (7)$$

holds for the closed loop system (4), (5). Then the respective control (3) is called the guaranteed cost control and the value of J_0 is the guaranteed cost. The main aim of this paper is to develop a decentralized PID control algorithm that stabilizes the uncertain system (1), with *guaranteed cost* with respect to the cost function (6).

We start with basic notions concerning Lyapunov stability and convexifying functions. In the following we use D -stability concept ([4]) to receive the respective stability conditions in more general form.

Definition 1. (D -stability) Consider the D -domain in the complex plain defined as

$$D = \{s \text{ is complex number} : \begin{bmatrix} 1 \\ s \end{bmatrix}^* \begin{bmatrix} r_{11} & r_{12} \\ r_{12}^* & r_{22} \end{bmatrix} \begin{bmatrix} 1 \\ s \end{bmatrix} < 0\}.$$

The considered linear system (1) is D -stable if all its poles lie in the D -domain.

(To simplify the reading of formulas we use in the Definition 1 scalar values of the parameters r_{ij} , in general the stability domain can be defined using matrix values of parameters r_{ij} with the respective dimensions.) The standard choice of r_{ij} is $r_{11} = 0, r_{12} = 1, r_{22} = 0$ for a continuous-time system; $r_{11} = -1, r_{12} = 0, r_{22} = 1$ for a discrete-time system. The *quadratic D -stability* is equivalent to the existence of one Lyapunov function for the whole set that describes the uncertain system model.

Definition 2. (Quadratic D - stability)

The uncertain system (4) is quadratically D -stable if and only if there exists a symmetric positive definite matrix P such that

$$r_{12}PA_C(\alpha) + r_{12}^*A_C^T(\alpha)P + r_{11}P + r_{22}A_C^T(\alpha)PA_C(\alpha) < 0 \quad (8)$$

Instead of quadratic stability, a robust stability notion is considered based on the parameter dependent Lyapunov function (PDLF) defined as

$$P(\alpha) = \sum_{i=1}^N a_i P_i \text{ where } P_i = P_i^T > 0 \quad (9)$$

to obtain less conservative results than using quadratic stability with unique Lyapunov function.

Definition 3. ([5]) System (4) is *robustly D -stable* in the convex uncertainty domain (5) with parameter-dependent Lyapunov function (9) if and only if there exists a matrix $P(\alpha) = P(\alpha)^T > 0$ such that

$$r_{12}P(\alpha)A_C(\alpha) + r_{12}^*A_C^T(\alpha)P(\alpha) + r_{11}P(\alpha) + r_{22}A_C^T(\alpha)P(\alpha)A_C(\alpha) < 0 \quad (10)$$

for all α such that $A_C(\alpha)$ is given by (5).

The sufficient robust D -stability condition which can be considered as not too conservative has been proposed in ([9]), recalled in the following lemma.

Lemma 4. *If there exist matrices $E \in R^{n \times n}, G \in R^{n \times n}$ and N symmetric positive definite matrices $P_i \in R^{n \times n}$ such that for all $i = 1, \dots, N$:*

$$\begin{bmatrix} r_{11}P_i + A_{Ci}^T E^T + E A_{Ci} & r_{12}P_i - E + A_{Ci}^T G \\ r_{12}^*P_i - E^T + G^T A_{Ci} & r_{22}P_i - (G + G^T) \end{bmatrix} < 0 \quad (11)$$

then system (4) is robustly D -stable.

Note that matrices E and G are not restricted to any special form; they were included to relax the conservatism of the sufficient condition. To transform nonconvex problem of structured control (decentralized control in our case) into convex form, the convexifying (linearizing) function can be used ([6]; [3];[11];[13]). The respective potential convexifying function for X^{-1} and XWX has been proposed in the linearizing form:

- The linearization of $X^{-1} \in R^{n \times n}$ about the value $X_k > 0$ is

$$\Phi(X^{-1}, X_k) = X_k^{-1} - X_k^{-1}(X - X_k)X_k^{-1} \quad (12)$$

- The linearization of $XWX \in R^{n \times n}$ about X_k is

$$\Psi(XWX, X_k) = -X_k W X_k + X W X_k + X_k W X \quad (13)$$

Both functions defined in (12) and (13) meets one of the basic requirements on convexifying function: to be equal to the original nonconvex term if and only if $X_k = X$. However, the question how to choose the appropriate nice convexifying function remains still open.

In the sequel, $X > 0$ denotes positive definite matrix; $*$ in matrices denotes the respective transposed term to make the matrix symmetric; I denotes identity matrix and 0 denotes zero matrix of the respective dimensions.

3 Robust Optimal Controller Design

In this section the new design algorithm for optimal control with guaranteed cost is developed using parameter dependent Lyapunov function and convexifying approach employing iterative procedure. The proposed control design approach uses sufficient stability condition inspired by the result of ([9]). The next theorem provides the new form of robust stability condition for linear uncertain system with guaranteed cost.

Theorem 5. Consider uncertain linear system (1), (2) with static output feedback (3) and cost function (6). The following statements are equivalent:

i) Closed loop system (4) is robustly D -stable with PDLF (9) and guaranteed cost with respect to cost function (6): $J \leq J_0 = x^T(0)P(\alpha)x(0)$.

ii) There exist matrices $P(\alpha) > 0$ defined by (9) such that

$$\begin{aligned} r_{12}P(\alpha)A_C(\alpha) + r_{12}^*A_C^T(\alpha)P(\alpha) + r_{22}A_C^T(\alpha)P(\alpha)A_C(\alpha) + \\ + r_{11}P(\alpha) + Q + C^T F^T R F C + A_C^T(\alpha)S A_C(\alpha) < 0 \end{aligned} \quad (14)$$

iii) There exist matrices $P(\alpha) > 0$ defined by (9) and H , G and F of the respective dimensions such that

$$\begin{aligned} \left[\begin{array}{c} r_{11}P(\alpha) + A_C^T(\alpha)H^T + H A_C(\alpha) + Q + C^T F^T R F C \\ r_{12}^*P(\alpha) - H^T + G^T A_C(\alpha) \end{array} \right] \cdot \\ \left[\begin{array}{c} * \\ r_{22}P(\alpha) - (G + G^T) + S \end{array} \right] < 0 \end{aligned} \quad (15)$$

$A_{Ci} = (A_i + B_i F C)$ denotes the i -th closed loop system vertex. Matrix F is the guaranteed cost decentralized control gain for the uncertain system (4), (5).

Proof. For brevity the detail steps of the proof are omitted where standard tools are applied. (i) \Leftrightarrow (ii): the proof is analogous to that in ([10]). The (ii) \Rightarrow (i) is shown by taking $V(t) = x(t)P(\alpha)x(t)$ as a candidate Lyapunov function for (4) and writing $\delta V(t)$, where

$$\begin{aligned}\delta V(t) &= \dot{V}(t) \text{ for continuous - time system} \\ \delta V(t) &= V(t+1) - V(t) \text{ for discrete - time system}\end{aligned}$$

$$\delta V(t) = r_{12}^* \delta x(t)^T P(\alpha)x(t) + r_{12}x(t)^T P(\alpha)\delta x(t) + r_{11}x(t)^T P(\alpha)x(t) + r_{22}\delta x(t)^T P(\alpha)\delta x(t) \quad (16)$$

Substituting for δx from (4) to (16) and comparing with (14) provides D -stability of the considered system when the latter inequality holds. The guaranteed cost can be proved by summing or integrating both sides of the following inequality for t from 0 to ∞ :

$$\delta V(t) < -x(t)^T [Q + C^T F^T R F C + A_C^T(\alpha) S A_C(\alpha)] x(t)$$

The (i) \Rightarrow (ii) can be proved by contradiction. (ii) \Leftrightarrow (iii): The proof follows the same steps to the proof of Lemma 1: (iii) \Rightarrow (ii) is proved in standard way multiplying both sides of (15) by the full rank matrix:

$$\begin{bmatrix} I & A_C^T(\alpha) \end{bmatrix} \{l.h.s.(15)\} \begin{bmatrix} I \\ A_C(\alpha) \end{bmatrix} < 0.$$

(ii) \Rightarrow (iii) follows from applying a Schur complement to (14) rewritten as

$$\begin{aligned}r_{12}P(\alpha)A_C(\alpha) + r_{12}^*A_C^T(\alpha)P(\alpha) + Q + C^T F^T R F C + \\ + r_{11}P(\alpha) + A_C^T(\alpha)[r_{22}P(\alpha) + S]A_C(\alpha) < 0\end{aligned}$$

Therefore $\begin{bmatrix} X_{11} & X_{12} \\ X_{12}^T & X_{22} \end{bmatrix} < 0$ where

$$\begin{aligned}X_{11} &= r_{11}P(\alpha) + r_{12}P(\alpha)A_C(\alpha) + r_{12}^*A_C^T(\alpha)P(\alpha) + Q + \\ &\quad + C^T F^T R F C \\ X_{12} &= A_C^T(\alpha)[r_{22}P(\alpha) + S] \\ X_{22} &= -[r_{22}P(\alpha) + S]\end{aligned}$$

which for $H = r_{12}P(\alpha)$, $G = [r_{22}P(\alpha) + S]$ gives (15). \square

The guaranteed cost control design is based on the robust stability condition (15). Since the matrix inequality (15) is not LMI we use the inner approximation for the continuous time system applying linearization formula (13) together with using the respective quadratic forms to obtain LMI formulation, which is then solved by iterative procedure.

4 PID Robust Controller Design For Continuous-Time Systems

Control algorithm for PID is considered as

$$u(t) = K_P y(t) + K_I \int_0^t y(t) dt + F_d C_d \dot{x}(t) \quad (17)$$

The proportional and integral term can be included into the state vector in the common way defining the auxiliary state $z = \int_0^t y(t) dt$, i.e. $\dot{z}(t) = y(t) = Cx(t)$. Then the closed-loop system for PI part of the controller is

$$\dot{x}_n = \begin{bmatrix} \dot{x} \\ \dot{z} \end{bmatrix} = \begin{bmatrix} A + \delta A & 0 \\ C & 0 \end{bmatrix} \begin{bmatrix} x \\ z \end{bmatrix} + \begin{bmatrix} B + \delta B \\ 0 \end{bmatrix} u(t)$$

and

$$u(t) = FCx(t) + F_d C_d \dot{x}(t) \quad (18)$$

where $FCx(t)$ and $F_d C_d \dot{x}(t)$ correspond respectively to the PI and D term of PID controller. The resulting closed loop system with PID controller (17) is then

$$\dot{x}_n(t) = A_C(\alpha)x_n(t) + B(\alpha) \begin{bmatrix} F_d C_d & 0 \end{bmatrix} \dot{x}_n(t) \quad (19)$$

where the PI controller term is included in $A_C(\alpha)$. (For brevity we omit the argument t .) To simplify the denotation, in the following we consider PD controller (which is equivalent to the assumption, that the I term of PID controller has been already included into the system dynamics in the above outlined way) and the closed loop is described by

$$\dot{x}(t) = A_C(\alpha)x(t) + B(\alpha)F_d C_d \dot{x}(t) \quad (20)$$

Let us consider the following performance index

$$J_s = \int_0^{\infty} \begin{bmatrix} x & \dot{x} \end{bmatrix}^T \begin{bmatrix} Q + C^T F^T R F C & 0 \\ 0 & S \end{bmatrix} \begin{bmatrix} x \\ \dot{x} \end{bmatrix} dt \quad (21)$$

which formally corresponds to (6). Then for Lyapunov function (9) we have the necessary and sufficient condition for robust stability with guaranteed cost in the form (14), i.e. for continuous time system:

$$\begin{bmatrix} x & \dot{x} \end{bmatrix}^T \begin{bmatrix} Q + C^T F^T R F C & P(\alpha) \\ P(\alpha) & S \end{bmatrix} \begin{bmatrix} x \\ \dot{x} \end{bmatrix} < 0. \quad (22)$$

The main result on robust PID control stabilization is summarized in the next theorem

Theorem 6. Consider a continuous uncertain linear system (1), (2) with PID controller (17) and cost function (21). The following statements are equivalent:

- Closed loop system (19) is robustly D-stable with PDLF (9) and guaranteed cost with respect to cost function (21):

$$J \leq J_0 = x^T(0)P(\alpha)x(0).$$

- There exist matrices $P(\alpha) > 0$ defined by (9), and H , G , F and F_d of the respective dimensions such that

$$\begin{bmatrix} A_{Ci}^T H^T + H A_{Ci} + Q + C^T F^T R F C \\ P_i - M_{di}^T H + G^T A_{Ci} \\ -M_{di}^T G - G^T M_{di} + S \end{bmatrix} < 0 \quad (23)$$

$A_{Ci} = (A_i + B_i F C)$ denotes the i -th closed loop system vertex, M_{di} includes the derivative part of the PID controller: $M_{di} = I - B_i F_d C_d$.

Proof. Owing to (20) for any matrices H and G :

$$\begin{aligned} & (-x^T H - \dot{x}^T G^T) (\dot{x} - A_C(\alpha)x - B(\alpha)F_d C_d \dot{x}) + \\ & + (\dot{x} - A_C(\alpha)x - B(\alpha)F_d C_d \dot{x})^T (H^T x - G \dot{x}) = 0 \end{aligned} \quad (24)$$

Summing up the l.h.s of (24) and (22) and taking into consideration linearity w.r.t. α we get condition (23). *Theorem 6* provides the robust stability condition for the linear uncertain system with PID controller. Notice that the derivative term does not appear in the matrix inversion and allows including the uncertainty in control matrix B into the stability condition.

Considering PID control design, there are unknown matrices H , G , F and F_d to be solved from (23). (Recall that $A_{Ci} = (A_i + B_iFC)$, $M_{di} = I - B_iF_dC_d$.) Then the inequality (23) is not LMI, to cope with the respective unknown matrices products the linearizing approach using (13) has been adopted and the PID iterative control design algorithm based on LMI (4x4 matrix) has been proposed. The resulting closed loop system with PD controller is

$$\dot{x}(t) = (I - B_iF_dC_d)^{-1}(A_i + B_iFC)x(t), i = 1, \dots, N \quad (25)$$

□

The extension of the proposed algorithm to decentralized control design is straightforward since the respective F and F_d matrices are assumed as being of the prescribed structure, therefore it is enough to prescribe the decentralized structure for both matrices.

5 Examples

In this section the major contribution of the proposed approach: design of robust derivative feedback is illustrated on the examples. The results obtained using the proposed new iterative algorithm based on (23) to design the PD controller are provided and discussed. The impact of matrix S choice is studied as well.

We consider affine models of uncertain system (1), (2) with symmetric uncertainty domain:

$$\underline{\varepsilon}_j = -q, \bar{\varepsilon}_j = q \quad (26)$$

Example 1.

Consider the uncertain system (1), (2) where:

$$A = \begin{bmatrix} -4.365 & -0.6723 & -0.3363 \\ 7.0880 & -6.5570 & -4.6010 \\ -2.4100 & 7.5840 & -14.3100 \end{bmatrix} \quad B = \begin{bmatrix} 2.3740 & 0.7485 \\ 1.3660 & 3.4440 \\ 0.9461 & -9.6190 \end{bmatrix}$$

$$C = C_d = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

uncertainty parameter $q=1$; uncertainty matrices:

$$\tilde{A}_1 = \begin{bmatrix} -0.5608 & 0.8553 & 0.5892 \\ 0.6698 & -1.3750 & -0.9909 \\ 3.1917 & 1.7971 & -2.5887 \end{bmatrix} \quad \tilde{B}_1 = \begin{bmatrix} -0.1602 & -0.3521 \\ 0.1162 & -2.4839 \\ -0.1106 & -4.6057 \end{bmatrix}$$

$$\tilde{A}_2 = \begin{bmatrix} 0.6698 & -1.3750 & -0.9909 \\ -2.8963 & -1.5292 & 10.5160 \\ -3.5777 & 2.8389 & 1.9087 \end{bmatrix} \quad \tilde{B}_2 = \begin{bmatrix} 0.1562 & 0.1306 \\ -0.4958 & 4.0379 \\ -0.0306 & 0.8947 \end{bmatrix}$$

The uncertain system can be described by four vertices; the corresponding maximal eigenvalues in the vertices of open loop system are respectively:

$-4.0896 \pm 2.1956i$; -3.9243 ; **1.5014**; -4.9595

Notice that the open loop uncertain system is unstable (see third vertex). The stabilizing optimal PD controller has been designed. Optimality is considered in the sense of guaranteed cost w.r.t. cost function (21) with matrices $R = I_{2 \times 2}$, $Q = 0.001 * I_{3 \times 3}$. The results summarized in the Tab. 1 indicate the differences between results obtained for different values of S

S	F (proportional part) Fd (derivative part)	Max Eig in vertices
1e-6 *I	-1.0567 -0.5643 -2.1825 -1.4969 -0.3126 -0.2243 -0.0967 0.0330	-4.8644 -2.4074 $-3.8368 \pm 1.1165i$ -4.7436
0.1*I	-1.0724 -0.5818 -2.1941 -1.4642 -0.3227 -0.2186 -0.0969 0.0340	-4.9546 -2.2211 $-3.7823 \pm 1.4723i$ -4.7751

Table 1: Example 1, Example 2

Consider the uncertain system (1), (2) where:

$$A = \begin{bmatrix} -2.9800 & 0.9300 & 0 & -0.0340 \\ -0.9900 & -0.2100 & 0.0350 & -0.0011 \\ 0 & 0 & 0 & 1 \\ 0.3900 & -5.5550 & 0 & -1.89 \end{bmatrix},$$

$$\tilde{A}_1 = \begin{bmatrix} 0 & 1.5 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix},$$

$$B = \begin{bmatrix} -0.0320 \\ 0 \\ 0 \\ -1.6000 \end{bmatrix}, \tilde{B}_1 = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix},$$

$$C = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

The results are summarized in Tab. 2 for $R = 1$, $Q = 0.0005 * I_{4 \times 4}$ for various values of matrix S in cost function. As indicated in Tab.2, increasing values of S slow down the response (*max.eig.CL* shifted to zero) as assumed.

6 Conclusion

The new robust PID controller design method is proposed for uncertain linear system based on LMI. The important feature of this PID design approach is that the derivative term appears in such form that enables to consider the model uncertainties. Since the structured feedback matrix is assumed, this approach is appropriate for decentralized PID control design. The guaranteed cost control is proposed with a new quadratic cost function including the derivative term for state vector as a tool to influence the overshoot and response rate. The obtained results are illustrated on the examples: to show the robust PID control design and the influence of the choice of matrix S in the extended cost function.

Acknowledgment The work has been supported by Slovak Scientific Grant Agency, Grant N 1/3841/06.

$S=10e-8*I_{4x4}$	
q_{max}	1.1
$max.eig.CL$	-0.189
$S=0.1*I_{4x4}$	
q_{max}	1.1
$max.eig.CL$	-0.1101
$S=0.2*I_{4x4}$	
q_{max}	1.1
$max.eig.CL$	-0.0863
$S=0.29*I_{4x4}$	
q_{max}	1.02
$max.eig.CL$	-0.0590

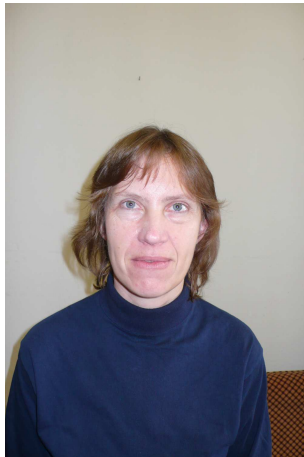
Table 2: Comparison for various S - Example 2

References

- [1] Blondel, V. and J.N. Tsitsiklis (1997). NP-hardness of some linear control design problems. *SIAM J. Control Optim.*, **35**, pp.2118-2127.
- [2] Boyd, S., L. El Ghaoui, E. Feron and V. Balakrishnan (1994). Linear matrix inequalities in system and control theory. *SIAM Studies in Applied mathematics*, Philadelphia.
- [3] Han, J. and R.E. Skelton (2003). An LMI optimization approach for structured linear controllers. In: *Proc. 42nd IEEE CDC, Hawaii, USA*, pp. 5143-5148.
- [4] Henrion, D., D. Arzelier and D. Peaucelle (2002). Positive polynomial matrices and improved LMI robustness conditions. In: *15th IFAC World Congress*, Barcelona, Spain.
- [5] deOliveira, M.C., J. Bernussou and J.C. Geromel (1999). A new discrete-time robust stability condition. *Systems and Control Letters*, **37**, pp. 261-265.
- [6] deOliveira, M.C., J.F. Camino and R.E. Skelton (2000). A convexifying algorithm for the design of structured linear controllers. In: *Proc. 39nd IEEE CDC, Sydney, Australia*, pp. 2781-2786.
- [7] Ge, Ming, Min-Sen Chiu, Qing-Guo Wang (2002). Robust PID controller design via LMI approach. *Journal of Process Control*, **12**, pp.3-13.
- [8] Gyurkovics, E., Takacs, T.(2000): Stabilisation of discrete-time interconnected systems under control constraints. *IEE Proceedings - Control Theory and Applications*, **147**, No. 2, 137-144
- [9] Peaucelle, D., D. Arzelier, O. Bachelier and J. Bernussou (2000). A new robust D -stability condition for real convex polytopic uncertainty. *Systems and Control Letters*, **40**, pp. 21-30.
- [10] Rosinová, D., V. Veselý and V. Kučera (2003). A necessary and sufficient condition for static output feedback stabilizability of linear discrete-time systems. *Kybernetika*, **39**, pp. 447-459.
- [11] Rosinová, D. and V. Veselý (2003). Robust output feedback design of discrete-time systems – linear matrix inequality methods. In: *2th IFAC Conf. CSD'03 (CD-ROM)*, Bratislava, Slovakia.

- [12] Skelton, R.E., T. Iwasaki and K. Grigoriadis (1998). *A Unified Algebraic Approach to Linear Control Design*, Taylor and Francis.
- [13] Veselý, V. (2003). Robust output feedback synthesis: LMI Approach. In: *2th IFAC Conference CSD'03* (CD-ROM), Bratislava, Slovakia.
- [14] Zheng Feng, Qing-Guo Wang, Tong Heng Lee (2002). On the design of multivariable PID controllers via LMI approach. *Automatica*, **38**, pp.517-526.

Danica Rosinová and Vojtech Veselý
Slovak University of Technology
Institute for Control and Industrial Informatics
Ilkovičova 3 81219 Bratislava , Slovakia
{danica.rosinova;vojtech.vesely}@stuba.sk
Received: December 24, 2006



Danica Rosinová was born in Bratislava, Slovak Republic in 1961. She received MSc. and PhD from Slovak University of Technology in 1985 and 1996. Since 1985 she has been with the Department of Automatic Control Systems, now Institute for Control and Industrial Informatics at the Faculty of Electrical Engineering and Information Technology STU Bratislava. Since 2006 she has been Associate Professor. Her research interests concentrate on robust control, large scale systems theory and optimization. students.



Vojtech Veselý was born in 1940. Since 1964 he has worked at the Department of Automatic Control Systems at the Faculty of Electrical Engineering and Information Technology, Slovak University of Technology in Bratislava. Since 1986 he has been Full Professor. His research interests include the areas of power system control, robust control, decentralized control of large-scale systems, proces control and optimization. He is author and coauthor of more then 270 scientific and technical papers, he successful supervised up today 19 PhD

Author index

Ahola T., 111

Barbu M., 132

Borne P., 174

Borowa A., 121

Caraman S., 132

Doherty B.S., 159

Hammadi S., 174

Heikkinen M., 143

Hiltunen Y., 143

Jena S.K., 149

Juuso E., 111

Kamalrulnizam Abu Bakar, 159

Krishna G.V.V., 149

Leiviskä K., 111

Liouane N., 174

Mazur K., 121

Mietek B.A., 121

Mina J., 185

Nurminen V., 143

Rosinová D., 195

Saad I., 174

Sbarciog M., 132

Verde C., 185

Veselý V., 195

Description

International Journal of Computers, Communications & Control (IJCCC) is published from 2006 and has 4 issues per year, edited by CCC Publications, powered by Agora University Editing House, Oradea, ROMANIA.

Every issue is published in online format (ISSN 1841-9844) and print format (ISSN 1841-9836). We offer free online access to the full content of the journal <http://journal.univagora.ro>. The printed version of the journal should be ordered, by subscription, and will be delivered by regular mail.

IJCCC is directed to the international communities of scientific researchers from the universities, research units and industry.

IJCCC publishes original and recent scientific contributions in the following fields:

- Computing & Computational Mathematics,
- Information Technology & Communications,
- Computer-based Control

The publishing policy of **IJCCC** encourages particularly the publishing of scientific papers that are focused on the convergence of the 3 “C” (Computing, Communications, Control).

Topics of interest include, but are not limited to the following: Applications of the Information Systems, Artificial Intelligence, Automata and Formal Languages, Collaborative Working Environments, Computational Mathematics, Cryptography and Security, E-Activities, Fuzzy Systems, Informatics in Control, Information Society - Knowledge Society, Natural Computing, Network Design & Internet Services, Multimedia & Communications, Parallel and Distributed Computing.

The articles submitted to **IJCCC** must be original and previously unpublished in other journals. The submissions will be revised independently by two reviewers.

IJCCC also publishes:

- papers dedicated to the works and life of some remarkable personalities;
- reviews of some recent important published books.

Also, **IJCCC** will publish as supplementary issues the proceedings of some international conferences or symposiums on Computers, Communications and Control, scientific events that have reviewers and program committee.

The authors are kindly asked to observe the rules for typesetting and submitting described in *Instructions for Authors*.

Instructions for authors

Papers submitted to the International Journal of Computers, Communications & Control must be prepared using a LaTeX typesetting system. A template for preparing the papers is available on the journal website <http://journal.univagora.ro>. In the template file you will find instructions that will help you prepare the source file. Please, read carefully those instructions.

Any graphics or pictures must be saved in Encapsulated PostScript (.eps) format.

Papers must be submitted electronically to the following address: ccc@univagora.ro.

The papers must be written in English. The first page of the paper must contain title of the paper, name of author(s), an abstract of about 300 words and 3-5 keywords. The name, affiliation (institution and department), regular mailing address and email of the author(s) should be filled in at the end of the paper. The last page should include a short bio-sketch and a picture of all the authors. Examples you may find in the previous issues of the journal.

Manuscripts must be accompanied by a signed copyright transfer form. The copyright transfer form is available on the journal website.

When you receive the acceptance for publication you will have to send us:

1. Completed copyright transfer form.
2. Source (input) files.
 - One LaTeX file for the text.
 - EPS files for figures - they must reside in a separate folder.
3. Final PDF file (for reference).
4. A short (maximum 200 words) bio-sketch and a picture of all authors to be included at the end of the article.

One author may submit for publication at most two articles/year. The maximum number of authors is four. The maximum number of pages of one article is 16 (including a bio-sketch). The publishing of a 10 page article is free of charge. For each supplementary page there is a fee of 50 Euro/page that must be paid after receiving the acceptance for publication.

The authors don't receive a printed copy of the journal.

The journal is freely available on <http://journal.univagora.ro>.

Order

If you are interested in having a subscription to “Journal of Computers, Communications and Control”, please fill in and send us the order form below:

ORDER FORM		
I wish to receive a subscription to “Journal of Computers, Communications and Control”		
NAME AND SURNAME:		
Company:		
Number of subscription:	Price Euro	for issues yearly (4 number/year)
ADDRESS:		
City:		
Zip code:		
Country:		
Fax:		
Telephone:		
E-mail:		
Notes for Editors (optional)		

1. Standard Subscription Rates for Romania (4 issues/2007, more than 400 pages, including domestic postal cost): 90 EURO.
2. Standard Subscription Rates for other countries (4 issues/2007, more than 400 pages, including international postal cost): 160 EURO.

For payment subscription rates please use following data:

HOLDER: Fundatia Agora, CUI: 12613360

BANK: BANK LEUMI ORADEA

BANK ADDRESS: Piata Unirii nr. 2-4, Oradea, ROMANIA

IBAN ACCOUNT for EURO: RO02DAFB1041041A4767EU01

IBAN ACCOUNT for LEI/ RON: RO45DAFB1041041A4767RO01

SWIFT CODE (eq. BIC): DAFBRO22

Mention, please, on the payment form that the fee is “for IJCCC”.

EDITORIAL ADDRESS:

CCC Publications

Piata Tineretului nr. 8

ORADEA, jud. BIHOR

ROMANIA

Zip Code 410526

Tel.: +40 259 427 398

Fax: +40 259 434 925

E-mail: ccc@univagora.ro, Website: www.journal.univagora.ro

Copyright Transfer Form

To The Publisher of the International Journal of Computers, Communications & Control

This form refers to the manuscript of the paper having the title and the authors as below:

The Title of Paper (hereinafter, "Paper"):

.....

The Author(s):

.....

.....

.....

.....

The undersigned Author(s) of the above mentioned Paper here by transfer any and all copyright-rights in and to The Paper to The Publisher. The Author(s) warrants that The Paper is based on their original work and that the undersigned has the power and authority to make and execute this assignment. It is the author's responsibility to obtain written permission to quote material that has been previously published in any form. The Publisher recognizes the retained rights noted below and grants to the above authors and employers for whom the work performed royalty-free permission to reuse their materials below. Authors may reuse all or portions of the above Paper in other works, excepting the publication of the paper in the same form. Authors may reproduce or authorize others to reproduce the above Paper for the Author's personal use or for internal company use, provided that the source and The Publisher copyright notice are mentioned, that the copies are not used in any way that implies The Publisher endorsement of a product or service of an employer, and that the copies are not offered for sale as such. Authors are permitted to grant third party requests for reprinting, republishing or other types of reuse. The Authors may make limited distribution of all or portions of the above Paper prior to publication if they inform The Publisher of the nature and extent of such limited distribution prior there to. Authors retain all proprietary rights in any process, procedure, or article of manufacture described in The Paper. This agreement becomes null and void if and only if the above paper is not accepted and published by The Publisher, or is withdrawn by the author(s) before acceptance by the Publisher.

Authorized Signature (or representative, for ALL AUTHORS):

.....

Signature of the Employer for whom work was done, if any:

Date:

Third Party(ies) Signature(s) (if necessary):