# CONTENTS

## EDITORIAL

1. **BAI-WU WAN, XIAO-E RUAN, DING LIU**  
   New Results in Control of Steady-State Large-Scale Systems  
   [123]

2. **KAZIMIERZ DUZINKIEWICZ, ADAM BOROWA, KRZYSZTOF MAZUR, MICHAŁ GROCHOWSKI, MIETEK A. BRDYS, KRZYSZTOF JEZIOR**  
   Leakage Detection and Localisation in Drinking Water Distribution Networks by MultiRegional PCA  
   [135]

3. **ERIC DUVIELLA, PASCALE CHIRON, PHILIPPE CHARBONNAUD**  
   Networked Hydrographical Systems: A Reactive Control Strategy Integrating Time Transfer Delays  
   [153]

4. **FATMA TANGOUR, PIERRE BORNE**  
   Presentation of Some Metaheuristics for the Optimization of Complex Systems  
   [169]

5. **A. CHAMROO, I. SIMEONOV, C. VASSEUR, N. CHRISTOV**  
   On the Piecewise Continuous Control of Methane Fermentation Processes  
   [181]

6. **MIRCEA IVANESCU, MIHAELA CECILIA FLORESCU, NIRVANA POPESCU, DECEBAL POPESCU**  
   A Compliance Control of a Hyperredundant Robot  
   [189]

7. **MOUNIR AYADI, JOSPEH HAGGÈGE, SOUFIÈNE BOUALLÈGUE, MOHAMED BENREJEB**  
   A Digital Flatness-based Control System of a DC Motor  
   [201]

8. **FLORIN LEON, MIHAI HORIA ZAHARIA, DAN GÂLEA**  
   Emergent Dynamic Routing Using Intelligent Agents in Mobile Computing  
   [215]

9. **PASCALE ZARATÉ**  
   Decision Making Process: A Collaborative Perspective  
   [225]

10. **SVETLANA COJOCARU**  
    Computational Environment Generation for Computer Algebra Systems  
    [231]

## AUTHOR INDEX
New Results in Control of Steady-State Large-Scale Systems

Bai-Wu Wan¹, Xiao-E Ruan², Ding Liu³

¹The State Key Laboratory for Manufacturing Systems, Xi’an Jiaotong University, P. R. China.
Email: wanbw@mail.xjtu.edu.cn

²Department of Mathematics, Faculty of Science, Xi’an Jiaotong University, P. R. China
Email: wruanxe@mail.xjtu.edu.cn

³School of Automation, Xi’an University of Technology, P. R. China,
Email: liud@xaut.edu.cn

Abstract: This paper reviews what the first Author and his Group have been investigating for the past fifteen years in the on-line steady-state hierarchical intelligent control and optimization of large-scale industrial processes (LSIP), or large-scale systems (LSS), viz., the use of neural networks for identification and optimization, the use of expert system to solve some kind of hierarchical multi-objective optimization problems, the use of the fuzzy logic control, and the use of the iterative learning control. Several implementation examples and the product quality control for LSS are introduced too. Finally the paper suggests the new stage of development.

Keywords: Large-scale systems, industrial control, intelligent control, optimization, quality control

Professor Bai-Wu Wan received his B.S. degree in Electrical Engineering in 1949 and graduated from The Institute of Communication, in 1951 after two year postgraduate study both from the Jiaotong University, Shanghai (no degree system at that time). He was in charge of the Automatic Control Group (now called Department) from 1958-1978, and then he was in charge of The Large Scale Systems Group of Systems Engineering Institute, Xi’an Jiaotong University from 1978-1995. Now, as a professor of Control and Systems Engineering he works with the State Key Laboratory for Manufacturing Systems and Institute of Systems Engineering, Xi’an Jiaotong University. He is a Honorary Member of The Council of Chinese Association of Automation. He was the Member of editorial board for three top Chinese journals of automatic control, and for The Proceedings of Institution of Mechanical Engineers, Part I. Journal of Systems and Control Engineering, United Kingdom; and is the Member of Technical Committee of Large Scale Systems of International Federation of Automatic Control (IFAC). His main research interests are: large scale systems theory and application, quality control, intelligent control. More than 400 papers have been published or accepted in domestic and international journals or proceedings of conferences. He published six books including two monographs "On-line Hierarchical Steady-state Optimizing Control of Large-scale Industrial Processes" and “Optimization and Product Quality Control of Large-scale Industrial Systems”.

Associate Professor Xiao-E Ruan received B.S. and M.S. degrees in Mathematics from Shaanxi Normal University, Xi’an, P.R.China, in 1988 and 1995, respectively and PhD degree in Control science and Engineering from Institute of Systems Engineering, Xi’an Jiaotong University in 2002. Since 1995, she has been with the Department of Mathematics, Faculty of Science, Xi’an Jiaotong University, Xi’an, P.R.China. From March 2003 to August 2004, Dr. Ruan worked as a post-doctoral researcher with the Department of Electrical Engineering and Computer Science, Korea Advanced Institute of Science and Technology. Her current research field includes steady state hierarchical optimization of large-scale industrial processes, iterative learning control and fuzzy logic techniques, etc.

Professor Ding Liu was born in Shandong, China in 1957. He received the B.S and M.S degrees in Department of Automation and Information Engineering, Xi’an University of Technology, Xi’an, China in 1982 and 1987 respectively, and the Ph.D. degree in 1997 from Xi’an Jiaotong University, Xi’an, China. He was a visiting scholar from 1991 to 1993 in University of Fukui, Japan. He is currently a professor, doctoral advisor, and President of Xi’an University of Technology. He has published a book and more than 100 papers, many of them were indexed by SCI, EI, and ISTP. His research interests include complex system’s modeling and control, visual servo, intelligent robot control, digital signal processing and intelligent control theory.

1. Introduction

In the 7-th IFAC/IFORS/IMACS Symposium on Large Scale Systems Theory and Application, Beijing, China, Roberts, the first Author of this paper and Lin (1992) gave a plenary report entitled “Steady-state Hierarchical Control of Large-scale Industrial Processes: A Survey”. It considered the development of hierarchical control of LSIP in three stages: static multilevel optimization stage, steady-state hierarchical optimization stage and integrated system optimization and parameter estimation (ISOPE) stage. Fifteen years and more have passed by since then. What has been emerging in this field? And what is the fourth stage if existed?
For the past decade, the intelligent control has been a very important research direction and pushing the control science and technology forward. So do the large-scale systems theory and applications. The steady-state intelligent control of industrial processes means the application of ideas and methodologies of artificial intelligence to steady-state hierarchical control of LSIP and LSS based on human experience and knowledge in control and decision. In other words the neural networks, the expert systems (the intelligent decision unit), the fuzzy logic control, the iterative learning control, the genetic algorithms etc. and their combinations are integrated with traditional analytical approach for solving the identification, control, optimization, coordination and fault diagnosis of LSIP and LSS (Wan, 1994). Since the beginning of 1990’s the first Author and his Large-scale Systems Research Group have devoted themselves to the study of steady-state hierarchical intelligent optimizing control of LSIP and LSS for fifteen years. A brief summary of the main research results including the implementation examples in process industry and the conclusions is as follows.

2. Use of Neural Networks

2.1 Neural network modelling

The first Author’s Group has successfully applied a multi-layer BP neural network for identifying a steady-state model of a hot-cold water mixer pilot plant that includes two subprocesses, heating and levelling, and conducted hierarchical optimizing control based on the neural network steady-state model by three microcomputers in hierarchy, and hierarchical steady-state stochastic optimizing control with variance analysis even if the data are corrupted by noise (Wang, Wan and Song, 1994).

For steady-state modelling of process possessing stochastic or chaotic steady-state behaviours, Luo, Liu and Wan (1998) have proposed an adaptive fuzzy neural inferring network (AFNI network) based on Takagi-Sugeno fuzzy model. And the Group have used the neural networks for product quality model and yield model for control of LSS.

2.2 Neural network optimization

Leung, Li and Wan (1993) have used the Hopfield neural network to fit the static optimization with the interaction prediction and the interaction balance coordination methods. The Lagrange multiplier, Kuhn-Tucker multiplier and relax variable are applied to treat constraints, and an energy function $E$ is defined. Then by differentiating $E$ with respect to output $y$, set-points $c$ and the Lagrange multiplier, Kuhn-tucker multiplier and the relax variable, a set of differential equations are obtained. This set of differential equations is solved by Runge-Kutta method without iteration. It is because the differential equation represented by upper coordinative network and those represented by lower decision networks are solved step by step and simultaneously, and they interchange the integration information step by step within integration.

The first Author’s Group have proved the stability and optimality of Hopfield network (Leung, Li and Wan, 1993; Li, Wan and Leung 1994; 1995; Wan and Huang, 1998). In addition, the Hopfield network has been extended to solve the steady-state optimizing control of LSIP with global feedback or local feedback. It requires 6 on-line iterations to obtain an improved suboptimal solution or 9 on-line iterations to obtain an optimal solution for a LSIP with three subprocesses, respectively. The former result is obtained by using the output shift method, while the latter by output shift and its partial derivative compensation.

3. Use of Expert System

Based on the expert system (rule-based system) an intelligent decision (ID) unit has been suggested to solve some kinds of optimization problems of LSIP and LSS. The ID unit is composed of a knowledge base, a database, an inference engine and a learning machine (Fig.1).
By means of the ID unit an intelligent interaction prediction method has been used to flexibly choose global coordinative variables for solving the hierarchical multi-objective optimization problem (Li, Qin and Wan, 1991). In the lower level the precise traditional optimization method is used for making a decision. The ID unit ensures the convergence of the algorithm and finds a satisfactory solution for the problem.

![Diagram](image)

**Figure 1. Large-scale system using ID unit**

By the ID unit an intelligent coordination method with objective-modified has been used to steady-state optimizing control of LSIP and LSS. The idea is to use the objective or job of subprocesses or subsystem as the coordinative variable, then to combine it with traditional large-scale system theory in order to better adapt the change of the environment or the job (Li, Qin and Wan, 1992).

By means of the ID unit a hierarchical intelligent optimizing control method for steady-state multi-objective LSIP with fuzzy parameters has been suggested (Li, Qin and Wan, 1992). Generally there are several fuzzy parameters in the upper level, such as the cost parameter and the parameters in the upper global constraints, etc. A confidence level \( \alpha \) is selected by which the problem is transformed into a crisp multi-objective optimization problem depending on \( \alpha \).

### 4. Use of Fuzzy Logic Control

The use of the fuzzy logic control benefits much by treating the model as that with fuzzy parameters characterized by triangle membership functions when the model-reality difference exists. Meanwhile the controller set-point \( c \), the interaction output \( y \) and interaction input \( u \) are treated as the crisp variables. Then the programming problem formed by the steady-state optimization becomes one with fuzzy parameters. The problem is transformed into a solvable crisp programming problem after some fuzzy addition and multiplication operations. This approach requires less number of iterations and provides a better objective function to compare with those of the normal approach either in the open-loop hierarchical control case or in the closed-loop case using global feedback, and either by the interaction balance coordination, the interaction prediction coordination methods or by the mixed coordination method (Gu and Wan, 2000; 2001a; 2001b). It is necessary to point out that the double iterative algorithm for fuzzy models based on interaction balance method with global feedback requires the fewest number, about four, of on-line iteration so far.

### 5. Use of Iterative Learning Control

In hierarchical steady-state optimization closed-loop control or global feedback utilizes the real steady state output information to modify the model-based optimum. Thus the methodology needs usually several iterations to capture the applicable optimum. This means, for each iteration, every real subsystem must be stimulated simultaneously by the step-type set-point changes computed from the optimization layer without severe disturbance to LSIP. In order to improve the transient performance for the optimization programming to be applicable, an iterative learning control (ILC) strategy for LSIP in a decentralized mode has been proposed in article (Wan and Huang, 1998; Ruan, Li and Wan, 2003), i.e., there is an individual ILC unit and control action in each subsystem and further investigations have been
addressed in literatures (Ruan, Wan and Gao, 2000; Ruan et al., 2005). The control structure is shown in Figure 2 where SSI denotes the steady state information of the large-scale system. Optimization layer contains the coordinator and the local decision units. Subsystem denotes subprocess including its controller. D denotes the interaction between subprocesses. c denotes the set-point change. $y_d$ denotes the desired trajectory. $r_k$ denotes the control output of the ILC unit. $y_k$ is the subsystem output.

![Figure 2. Iterative learning control structure for a large-scale system](image)

In the studied iterative learning control strategies, the distinct magnitudes of the step-type set-point change sequence have been introduced to the proposed conventional PD-type open-loop ILC algorithm, higher-order ILC law as well as an optimal ILC rule. Instead of in the sense of $\lambda$ norm the convergence analysis in the sense of Lebesgue-p norm is derived which evaluates the output error in the Lebesgue integral over the whole operation time interval and some remarks are discussed (Ruan et al., 2005, 2008a, 2008b). The studies have concluded that the proposed ILC may efficiently improve the transient performance, such as speeding up the transient rising, decreasing the overshooting and shortening the settling time, etc, while the step-type set-point change sequence drives the system consecutively for reaching the steady-state output without any steady-state error. The studies have also discussed the influence of the inherent characteristics of the system such as the interaction, multi-dimensionality as well as the distinct magnitudes of the set-point change sequence on the convergence despite that the studies may cover the existing result for robot to track a unique desired trajectory. For one of subsystems of a linear time-invariant large-scale system, the tracking behavior is shown in Fig. 3.

![Figure 3. Output information at 8-th implementation](image)

In Fig. 3, the dashed curve denotes the predetermined desired trajectory, the dashed-dotted curve represents the output driven by the step-type decision and the solid one is that of stimulated by the ILC generated signal, respectively.

Ruan et al. (2003) suggested a local-symmetrical-double-integral type iterative learning control for
dynamics of industrial processes with time delay in the course of steady-state optimization when measurement noise is present.

This approach is a combination of study on the steady-state hierarchical optimization with that on the transient. It is evident that after seven ILC iterations the dynamic characteristics are greatly improved with two periods of set-points being provided by the coordinator in the Optimization layer of Fig.2. Furthermore, the first ILC iteration in the second optimization period is equivalent to the k+1-th iteration in the first optimization period if its number of iterations is k. Therefore the dynamic characteristics are further improved in the second period. Thus the whole set-point changes can once be fully imposed on the LSIP or LSS with little disturbance.

6. Applications in Industry

The first example is the steady-state optimizing control of a nickel flash furnace based on neural network models in a smelting plant (Wan, Wan and Yuan, 1999) which is located in Jinchang City, Gansu Province, China (Fig. 4). The quality model of the matte is based on three 5×5×1 BP neural networks.

![Figure 4. Nickel flash furnace system](image)

The inputs of the quality model are the 4 manipulating variables and 1 disturbing variable, while 3-quality indexes (properties of matte) are the outputs. The matte yield model is based on a 5×5×1 BP neural network with yield as the output. Then the objective function is minimization of the total energy consumption of the furnace subject to the inequality constraints formed by quality submodels and quality index tolerances, and product yield model with yield not less than normal. The study and the industrial site experiment for the optimization of the furnace have obtained a satisfactory result.

Liu and the first Author (Liu and Wan, 1999) have given the second example in which a multi-layer BP neural network is used to identify the steady-state model of air preheater of a big power-station boiler under different load condition (Fig. 5).

![Figure 5. Boiler preheater system](image)

The primary wind pressure and temperature, and the secondary wind pressure and temperature of the preheater are used as four inputs and the boiler-load is used as output of neural network. Both data can be measured for training neural network.

The sum of primary wind pressure and secondary wind pressure approximately represents the sum of the
two wind motor currents is selected to be the objective function. The optimization is to minimize the sum of currents under a definite load condition and is carried out by an enumerative method within feasible region latticed by small intervals. Based on this model an on-line steady-state optimizing control is successfully applied. The intelligent optimization gives considerable profit by saving electricity, is really implemented in big power-station boilers in China as well as those exported to abroad.

7. Steady-State Identification

Chen and Wan (1995; 1999) have suggested an approach to identifying a steady-state model by the dynamic data acquired from the normal set-point changes during tuning or optimization. The input is of the step-function form. They have proved, however, that under mild conditions the steady-state model obtained from the approximate dynamic model is with the strongly consistent estimates.

To a class of nonlinear slow time-varying large-scale processes, which have many subprocesses interconnected with one another, a parallel two-stage identification algorithm has been studied. The consistency of the estimates and convergence of the parallel iteration are also proved.

In addition, the Research Group has given a new steady-state identification method that provides a steady-state model of a nonlinear process only using steady-state data from several set-point changes and the estimates are strongly consistent (Huang and Wan, 1997; see Chapter 2, Wan and Huang, 1998). Besides, Huang, Wan and Han (1994) have given a method different from the above by Chen and Wan to calculate the process derivative with respect to set-point only using steady-state data acquired from several times set-point changes and its strong consistency has also been proved.

8. Robustness of Optimization Algorithm

It needs to study robustness of an optimization algorithm with respect to model parameter and noise to avoid divergence. Xu and Wan (1994) have investigated the robust stability of the algorithms for steady-state optimizing control of industrial processes, discussed the dependence of the optimal solution obtained from the algorithms on the parameters \( \lambda \) that represent the characteristic numbers of noises or process structure parameters. The Pompeiu-Hausdorff hemidistance \( H \) of two optimal solution sets is used as a measure for the robustness of the algorithm with respect to \( \lambda \). One is the optimal solution set, while the another is the optimal solution set perturbed by the parameter \( \lambda \). Actually to calculate the hemidistance is rather difficult, if not impossible. Hence \( \frac{\partial H}{\partial \lambda} \) is used as a sensitivity index to compare different optimization algorithms. The concept can be used to some simple cases (Xu, Wan and Han, 1997).

9. Generalized Steady-State of Industrial Process

Actually from the point of view of steady-state optimisation the influence of stochastic noise in process variables is often ignored due to its low level and little influence to the objective function. The problem is called stochastic optimising control of steady-state systems, or the systems are under stochastic steady-state (Lin, Han, Roberts and Wan, 1989) if the noise can not be ignored.

Luo and Wan (1999a) have extended the concept of steady-state to a generalized form, i.e., from the point of view of steady-state optimisation, a system may be under several kinds of steady-state that are constant, periodic, quasi-periodic, stochastic, chaotic steady- states in an industrial process or system. Actually some random process happens in a LSIP is often a mixture of the chaotic steady-state with a stochastic steady-state of low level. They have proposed a stringent definition about the generalized steady-state and proved that it exists when the nonlinear process satisfies some conditions. It is proved that the nominal central value of the generalized steady-state uniquely exists when the nonlinear process satisfies some conditions. The time averages of process variables uniformly converge to their respective nominal central values.

According to the definitions of generalized steady-state and the nominal central value of steady-state sets, Luo and Wan (1999a) stringently have described the problem of generalized steady-state optimizing
control of industrial processes in a finite measure space. Under certain conditions above problem is transformed into a model based equivalent deterministic problem, and an algorithm for solving the problem has been suggested. A chemical process composed of a liquid level control system (LLCS) and a continuously stirred reactor (CSTR) (Fig. 6) has been used for simulation study of generalized steady-state optimizing control.

Luo, Han and Wan (1999) have stringently given a definition for the chaotic steady-state, and proved the existence theorem under some conditions. The chaotic steady-state of a chemical process is simulated. The steady-state modelling is based on an AFNI network (Luo, Liu and Wan, 1998). The global convergence of the steady-state generalized optimizing control algorithm has been proved based on Zangwill’s Theorem of global convergence. Optimality of the optimizing control solution has been studied also (Luo and Wan, 1999b).

10. Global Convexification, Multi-Objective and Non-Separable Optimization Problems

Qian and Wan (2000) have proposed an approach through the p-th power transformation to obtain a global optimal solution by multi-objective optimization technique. The original optimization problem is embedded in a multi-objective optimization problem, then its non-inferior frontier is convexified. The original global optimal solution is picked up from the set of non-inferior solutions. Of course, this approach can be used to solve a multi-objective optimization problem for LSIP. Qian, Liu and Wan, (1999) has proposed a double iterative algorithm for non-separable multi-objective optimization problem which can satisfy the decision maker’s preference.

For non-separable steady-state systems the Group has given an approaches for solving them. It is based on transversal transmission of information among local decision units to decouple the objective function. Meanwhile the traditional longitudinal transmission of information is used to decouple the interconnection between subsystems (Qian and Wan, 1998a).

Another approach is for non-additive objective functions of general non-separable systems, Qian and Wan (1999b) have suggested a double-loop iterative algorithm. All these algorithms can be used for on-line optimizing control of LSIP.

11. Product Quality Control for Large–Scale Industrial Systems

A continuous casting and hot rolling production line in a Steel Complex in Shanghai is considered as a typical example (Fig.7). Knowledge Discovery in Database (KDD) is used to acquire data from computers of 3rd-generations for controlling the line, and from the computers in chemical analysis lab and material testing lab. To improve the product quality of the continuous casting and hot rolling, it is very necessary and important for a steel complex to find the relationship between the input variables and the product quality, i.e., to establish the steel plate static quality model. The steel plate quality model of a continuous casting furnace and hot rolling mill is a complex nonlinear function which after analysing and discussing with plant engineers Xing decides to include at least 32-input variables: 23 chemical elements variables for casting, 2 heating furnace variables, 7 rolling mill variables and 4-output variables (material testing indexes): rupture elongation rate, tensile strength, yield ratio and impact energy etc (Xing, 2000; Wan, 2002).

![Figure 6. LLCS and CSTR](image-url)
All data used for modelling are preprocessed. 15,000 useful sample data are obtained from 30,000 and more observations in the data warehouse. Among them 9,026 sample data are complete and can be used for modeling, however, they are corrupted by noise. And a high dimension input BP neural network is firstly chosen for the architecture of the steel plate quality model. To easily train and improve the accuracy the 32-input and 4-output BP neural network is decomposed as four 32-input and 1-output sub-neural network models. It is called the decomposition of the product quality modelling problem based on large-scale neural network.

The precision of modeling is expressed in the percentage of hits, i.e., the total number of hits in all 9,026 data divided by 9,026. A hit is defined as that pair of data which makes the model output within an error \( \pm 5\% \) of the real output.

Jia, Wan and Feng (2000) give a learning algorithm that each weight of BP neural network is trained separately with large inertia. The percentage of hits for suggested algorithm based on high-dimension-input BP neural network is 81.5%.

### 11.1 Modelling Based on Wavelet Neural Network

It is important to notice that the sequence of real output is in a saw tooth-like form. For this reason the wavelet neural network (WNN) is a better choice for the modelling problem.

Li and Wan (2002) choose the similar structure of the WNN as that of multi-layer perceptron (MLP), except that here the activation function of hidden nodes is replaced by a B-spline wavelet function of one dimension. Employing the MLP-like architecture, the proposed three layer WNN (1-input, 1-hidden layer and 1-output) is a powerful tool to handle high dimensional problem. The percentage of hits for this quality model is 81.5\%, while that for an ordinary BP three layer neural network with the same number of nodes provides a precision 62.7\%.

### 11.2 High-Dimension-Input Wavelet Neural Network Based on Work Procedure of the Technology and Key Inputs

The production line is a serially connected system and 32 input variables start act at different stages according to the work procedure. Therefore Li and Wan (2004a) suggest a WNN with several input layer depending on the work procedure (Fig.8). For instance, in first input layer there are 22 input variables (chemical-element variables) simulating the casting, in second input layer there are 2 variables simulating the heat furnace, and in the third layer there are 7 variables simulating the rolling. And for this special kind of steel plate there are three most important chemical elements, viz., carbon, manganese and titanium. Their corresponding input variables are connected to the input as well as the output nodes directly. A suitable learning algorithm is given also. The percentage of hits for quality model of this architecture is 93.4\%.

![Figure 7. Continuous casting and hot rolling](image-url)
Figure 8. The architecture of three-input-layer wavelet neural network based on work procedure and key inputs
- Input node; - Neuron

The further improvement of the precision can be made by clustering all the data and dividing them into 12 groups, and using the modular WNN approach (Li 2003). Then each submodel gives a precision about from 93.4–95% depending on the group of data used in modeling. And using a filled function algorithm to get a global optimum makes the above WNN result further improve 1% of precision (Li and Wan, 2004b; 2004c).

11.3 Application of Product Quality Model to New Product and New Technology Design

More occasionally it is not allowed to change all the manipulating variables in a quality model. Therefore, a new kind model, product quality control model, is suggested in which the input variables are quality index and those variables that are not allowed to change or that are assigned preliminary, and the output variables are some manipulating variables that are allowed to change. By the latter one can find the manipulating variables required from the quality index value (Xing, 2000). For instance, for a certain kind of titanium-manganese alloy steel plate the manipulating variables required or engineers hope to get from the quality control model are the amount of titanium, manganese and carbon. Sometimes the quality control model is more convenient in practice.

manipulating variables from these four submodels is a serious problem for the continuous casting and hot rolling. It is called the coordination or synthesis of the product quality modelling problem based on large-scale neural network. For simple cases the solution is the intersection set of the output manipulating variable sets from the quality control submodels. But in more complicated cases, perhaps, some kind of data fusion is necessary. How to overcome this drawback needs further study. And evidently it needs different kinds of such quality models for different design purposes.

12. Conclusions

The paper concludes that the second stage steady-state hierarchical optimization has extended to generalized steady-state hierarchical optimization and that the third stage ISOPE has extended to ISOPE and DISOPE (dynamic integrated system optimization and parameter estimation) stage, and that the fourth stage is the hierarchical intelligent control and optimization stage. Obviously, the latter is a very important one.

In the Group’s experience the neural network modelling using the data from normal set-point changes, updated by newly coming data, the optimization algorithm selected from different intelligent methods depending upon the nature of the problem, application of iterative learning control technique, and integrated with fault diagnosis.
is a good choice, it gives great potential for increasing profit. And all these functions can be integrated in intelligent agents for on-line steady-state intelligent hierarchical control of LSIP.

Acknowledgement

The work is supported by the National Natural Science Foundation of China under granted No. F030101-60274055, F030101-60574021, 9674003 and supported by the High Technology Plan (863 plan) No. 863-51-95-011 of P.R. of China. The old version of the work was presented on the 10th IFAC /IFORS /IMACS/IFIP Symposium on Large Scale Systems: Theory and Applications, July 26-28, 2004, Osaka, Japan.

REFERENCES


30. RUAN, XIAO-E, BAIWU WAN, FENGMIN CHEN, Decentralized Iterative Learning Controllers for Nonlinear Large-Scale Systems to Track Trajectories with Different Magnitudes, Acta Automatica Sinica (Accepted in Nov. 29, 2007).


39. XING, JIN-SHENG, An Investigation on Steel Plate Quality Model of Large-Scale Hot Rolling Mill Using KDD Technology. [Ph.D. Thesis]. Xi’an Jiaotong University, Xi’an, China (in Chinese).


Leakage Detection and Localisation in Drinking Water Distribution Networks by MultiRegional PCA

Kazimierz Duzinkiewicz¹, Adam Borowa¹, Krzysztof Mazur¹, Michał Grochowski¹, Mietek A. Brdys², Krzysztof Jezior¹

¹Faculty of Electrical and Control Engineering, Gdansk University of Technology, ul. Narutowicza 11/12, 80-952 Gdansk, Poland
email: [k.duzinkiewicz; aborowa; k.mazur; m.grochowski; m.brdysl; k.jezior]@ely.pg.gda.pl;
²Department of Electronic, Electrical and Computer Engineering, University of Birmingham, Birmingham B15 2TT, U.K.
email: m.brdys@bham.ac.uk

Abstract: Monitoring is one of the most important steps in advanced control of complex dynamic systems. Precise information about systems behaviour, including faults indicating, enables for efficient control. The paper describes an approach to detection and localisation of pipe leakage in Drinking Water Distribution Systems (DWDS) representing complex and distributed dynamic system of large scale. Proposed MultiRegional Principal Component Analysis (MR-PCA) skilfully takes full advantage of well known PCA method and enables not only for detecting the leakages but also supports their localisation. The main idea of MR-PCA is presented on example of small water network. Next the method is applied to DWDS in Chojnice, northern Poland. DWDS Chojnice is decomposed into suitable subnetworks what makes that the monitoring process is easier and require less sensors. The subnetworks and corresponding PCA monitoring models are selected based on the network operational knowledge and information regarding its topology.

Keywords: Monitoring, large-scale systems, network systems, fault detection algorithms, water leakage detection, statistical methods.

Kazimierz Duzinkiewicz received M.Sc. degree in Electrical Engineering and Ph.D. in Control Engineering from Faculty of Electrical and Control Engineering at Gdansk University of Technology, in 1973 and 1982, respectively. He has been employed as a university teacher starting his work in 1973 from the post of Assistant to the current position of Senior Lecturer in Department of Control Engineering. During his research work he has published over 80 scientific papers and over 50 scientific and technical reports, mainly dealing with following problems: a) production scheduling and operational control of technological systems with switchable processes (refinery type systems), b) computer control of electric power station in emergency conditions, c) safety and reliability analysis of hazardous systems, d) mathematical modelling of complex systems, e) multihorizon and multilevel optimisation, control and decision support structures and algorithms with applications to petroleum industry and environmental systems (drinking water and wastewater systems). He was NOC Chair of the 11th IFAC Symposium on Large Scale Complex Systems, Gdansk, July 23-25, 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, between 2001 and 2002 he served one's apprenticeship on Wastewater Treatment Plant at Swarzewo. From 2002 he has been a member of Intelligent Decision Support and Control System Group at Gdansk University of Technology. He was a member of organizing committee of the 11th IFAC Symposium on Large Scale Complex Systems, Gdansk, July 23-25, 2007. He is co-author of 8 publications and 4 chapters in the books. Mainly he focuses on modelling and monitoring of large scale systems, especially processes with many time scales. In 2008 he received the Ph.D. degree in Automatic Control and Robotics.

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. A Member of Intelligent Decision Support and Control System Group at Gdansk University of Technology. He was a member of organizing committee of the 11th IFAC Symposium on Large Scale Complex Systems, Gdansk, July 23-25, 2007. His research interests are in the areas of modelling, control and diagnostics of large scale systems. Co-author of 5 papers and 4 chapters in the books.

Michał Grochowski received his M.Sc. degree in Control Engineering in 2000 from Electrical and Control Engineering Department at Gdansk University of Technology. In 2001 he finished M.Sc. study at Faculty of Management and Economics of Gdansk University of Technology. From 1999 to 2003 he was a Ph.D. student in Electrical and Control Engineering Department at Gdansk University of Technology. In 2004 he received the Ph.D. degree in Automatic Control and Robotics. From 1999 is a member of Intelligent Decision Support and Control System Group at Gdansk University of Technology. From 2004 he held the posts of Assistant and Assistant Professor at Control Engineering Department at Gdansk University of Technology. He was a member of organizing committee of the 11th IFAC Symposium on Large Scale Complex Systems, Gdansk, July 23-25, 2007. He is author and co-author of about 15 refereed papers and four chapters in the books. His researches are focused on intelligent control of complex systems, supervisory control, faults detection, softly switched robustly feasible model predictive control.

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. 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 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. His research is supported by the UK and Polish Research Councils, and industry and European Commission. He is author and co-author of about 220 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 IET, a Senior Member of IEEE, a Fellow of IMA and a Chair of IFAC Technical Committee on Large Scale Complex Systems. He was IPC Chair of the 11th IFAC Symposium on Large Scale Complex Systems, Gdansk, July 23-25, 2007.

Krzysztof Jezior received his M.Sc. degree in Control Engineering from Electrical and Control Engineering Department at Gdansk University of Technology in 2007. Since then a Ph.D. student in this Department and a Member of Intelligent Decision Support and Control System Group at Gdansk University of Technology. Focused in his research on modelling and diagnostics of large scale systems. Co-author of 1 publication and 2 chapters in the books.

1. Introduction

Nowadays, monitoring systems besides data gathering are able to pre-process the data, to recover and estimate not directly measured variables. However, in large scale systems there is very large quantity of information that are hard to handle and sometimes almost impossible to properly process and hence to efficiently utilised it in the control process. An example of such systems is Drinking Water Distribution Systems (DWDS) the representatives of the class of network systems. The DWDS are usually, very complex (lots of pipes, connecting nodes, pumps, tanks etc.) and distributed (in space). It entails measuring of very large number of variables necessity, in order to possess information about the system state that is necessary for efficient system control. In such situations special methods enabling for analysis of large amount of data (e.g. faults detection and isolation) are required. Advanced monitoring systems should not only visualize desired data but also be able to detect devices faults and/or the unusual system behaviour. The paper proposes an approach to detecting and localisation of water leakage in pipes by using the Principal Components Analysis (PCA) method [1]. The PCA is a method that looks for multidimensional correlation between the variables and uses it to reduce the dimensionality of problems simultaneously remaining most of original information. Mostly, large amount of real data process do not provide large amount of important information. Hence, PCA explores data to find out very meaningful ones and include them into statistical models. Moreover, these models clearly indicate the abnormal state of the system thanks to specially calculated measures ($T^2$ and $SPE$). In case of DWDS such a situation might be caused by device faults (e.g. sensor or pump break down), water leakage in pipe, significant increasing of the water uptake (e.g. caused by fire brigades) etc. Detecting of fault is important however, in case of DWDS the system operator still does not know its type and localisation. Leaks detection and localisation issue is a very important and complex problem that has been widely investigated [2] – [6]. However, available active leakage control methods are basically unpractical due to costs or long leak detection and location time [4]. In the paper the novel approach the MultiRegional Principal Component Analysis (MR-PCA) method is used to detect and to locate the water leakage based on measurements from limited number of measuring devices [5]. MR-PCA tries to join operational experience of staff working in water companies and advanced mathematical analysis. Moreover, this method compromises between detection efficiency and a number of measuring devices.

The method is explained based on simple water network and followed with its application to real town case study DWDS Chojnice (northern Poland).

2. Monitoring and Diagnostics in Advanced Control Systems

Monitoring and diagnostics, which purpose is the fault detection and identification issue, are essential elements of advanced control of complex systems (Figure 1) [7] – [9].

Monitoring and diagnostics utilize a variety of methods for solving the fault detection and identification issue. Basically these methods can be divided into tree classes (Figure 2), which are quantitative model-based, qualitative model-based and process history based, also known as data driven methods [7], [10], [11].
Hybrids of monitoring and diagnostic methods can satisfy requirements imposed on a Detection and Identification Unit in a more natural way, since they utilize a set of elements, each fitted to a particular need [7]. Especially if the resulting mixture, consists of different class members, which is the case of MultiRegional Principal Component Analysis (MR-PCA) [5], [6]. MR-PCA, dedicated to Distributed Systems (DSs) following the network structure, combines Structural Decomposition (SD) and Principal Component Analysis (PCA), where the latter belongs to the Multivariable Statistics methods (Figure 2).

The main idea behind SD is to conclude about the conditions of system/process in question by means of its subsystems analysis [11]. PCA is described in the next subsection.

### 2.1 Principal Component Analysis

Principal Component Analysis is a method, which identifies linear dependencies among \( n > 1 \) variables \( x_{i=1,...,n} \), resulting in \( s \leq n \) decorrelated and linearly related variables \( t_{i=1,...,s} \) and a residuals \( \tilde{t}_{i=1,...,n-s} \) minimised in the sense of Mean Squared Error (MSE) [1]. Variables \( x_{i=1,...,n} \) are assumed to be normally distributed, with independently, identically distributed (IID) Gaussian noise contamination. Due to statistical consistency condition, PCA can model only quasi-static processes, i.e., with unnoticeable transients, because only cross-correlations between variables \( x_{i=1,...,n} \) are taken into account during the identification.

In more details, given a matrix \( X \in \mathbb{R}^{N \times n} \) consisting of data collected from the identified process (variables standardized to zero mean and unit variance) and \( N >> n \), PCA leads to the following decomposition of \( X \):

---

**Figure 1.** Monitoring and diagnostics (Fault Detection and Identification Unit) in advanced control system structure

**Figure 2.** Monitoring and diagnostic methods classification (based on [10])

Studies in Informatics and Control, Vol. 17, No. 2, June 2008 137
\[ X = TP^T + \tilde{TP}^T \]  

where \( T \in \mathbb{R}^{N \times s} \) is the scores matrix containing new data vectors \( t_{j=1,...,N} \in \mathbb{R}^s \) corresponding to original data samples \( x_{j=1,...,N} \in \mathbb{R}^n \) and residuals \( \tilde{t}_{j=1,...,N} \in \mathbb{R}^{n-s} \) collected in residual matrix \( \tilde{T} \in \mathbb{R}^{N \times n-s} \). Orthonormal block matrix \( [ P \tilde{P}] \in \mathbb{R}^{n \times n} \) plays in the decomposition (1) a key role leading to decorrelation of original cross-correlated data. Its first element, so-called loadings matrix, which column vectors \( p_{i=1,...,s} \in \mathbb{R}^n \) contain linear relations identified in the data \( X \), spans \( s \) - dimensional Principal Component Space (PCS), while column vectors \( \tilde{p}_{i=1,...,n-s} \in \mathbb{R}^n \) of the second \( \tilde{P} \in \mathbb{R}^{n \times n-s} \) span Residual Space (RS) and both spaces are orthogonal (Fig. 3). Thus \( \tilde{t}_{j=1,...,N} \) and \( t_{j=1,...,N} \) are projections of \( x_{j=1,...,N} \) on the RS and PCS respectively, where the latter (or \( P \) as its basis) is considered as the PCA model.

In order to obtain \([ P \tilde{P}]\) one can perform diagonalisation (e.g. using Eigen Decomposition (ED)) of approximated data correlation matrix \( \hat{R}_X \in \mathbb{R}^{n \times n} \):

\[
\hat{R}_X = \frac{1}{N-1}X^TX
\]

resulting in:

\[
\hat{R}_X = [ P \tilde{P}] \begin{bmatrix} \Lambda & 0 \\ 0 & \tilde{\Lambda} \end{bmatrix} [ P^T \tilde{P}^T]^T
\]

where \( \Lambda \in \mathbb{R}^{s \times s} \) and \( \tilde{\Lambda} \in \mathbb{R}^{n-s \times n-s} \) are both diagonal matrices containing eigenvalues of (2): \( \lambda_{i=1,...,s} \) and \( \tilde{\lambda}_{i=1,...,n-s} \) corresponding to appropriate block matrix of \([ P \tilde{P}]\) respectively and all eigenvalues are proportional to the variance of original data \( X \) in corresponding directions \([ P \tilde{P}]\) (Fig. 3). In the notation (3) it is assumed that \( \lambda_{i=1,...,s} \) and \( \tilde{\lambda}_{i=1,...,n-s} \) are sorted in descending order and in particular \( \lambda_1 \geq \lambda_2 \). From (2), (3) it is clear why only cross-correlation process structure can be modelled using PCA approach.

Assumption of IID Gaussian noise contamination leads to equal values of \( \tilde{\lambda}_{i=1,...,n-s} = \sigma^2 \), where \( \sigma^2 \) is the variance of noise in question in all \( n-s \) residual dimensions. This enables for clear and doubtless separation of \( \Lambda \) and \( \tilde{\Lambda} \), and in consequence \( P \) and \( \tilde{P} \). However in practice this is the rare case and one has to choose approximation \( \hat{s} \) rather than \( s \) on the basis of some of available methods [12]. The least sophisticated, though quite effective, is the Captured Percent Variance (CPV):

\[
CPV(s) = \frac{\sum_{i=1}^{s} \lambda_i}{\sum_{i=1}^{s} \lambda_i + \sum_{i=1}^{n-s} \tilde{\lambda}_i} \times 100\%
\]

where the choice of \( \hat{s} \) depend on the assumed minimal captured by the PCA model percent of data variance \( \text{CPV}_{\text{lim}} \):

\[
\hat{s} = \arg \min_s (CPV(s) \geq \text{CPV}_{\text{lim}})
\]

Because of PCA ability of modelling only the linear part of the processes, all nonlinear dependencies contained in data \( X \) would be linearly approximated minimising MSE of residuals. In such case linearization errors are included in the RS, which becomes a PCS non-fitting data container.

Data analysis after decomposition (1) into PCS and RS, can be performed by norms derived for each of subspaces separately, which are Hottelings \( T^2 \) and Squared Prediction Error (SPE) respectively [13]. The former, defined as:
\[ T^2(j = 1, \ldots, N) = (t_j)^T \Lambda^{-1} t_j = (x_j)^T P \Lambda^{-1}(P)^T x_j \]  

is the squared Euclidean distance of data sample \( x_{j=1, \ldots, N} \) projected onto PCS to the subspace origin, weighted proportionally to variance (Fig. 3). The letter norm, i.e. \( \text{SPE} \): 

\[ \text{SPE}(j = 1, \ldots, N) = (\tilde{t}_j)^T \tilde{t}_j = (x_j)^T (I - P(P)^T) x_j \]  

measures the squared Euclidean distance of the residual \( \tilde{t}_{j=1, \ldots, N} \) to the PCS (Fig. 3), where \( I \in \mathbb{R}^{n \times n} \) is the identity matrix. For both norms (6), (7) thresholds \( T_{\text{lim}}^2(\hat{s}, \beta) \) and \( \text{SPE}_{\text{lim}}(\hat{s}, \alpha) \) can be computed corresponding to data \( X \) variability in each of subspaces respectively (Fig. 3). In the first case threshold is defined as a chi-squared deviate for the level of significance \( \beta \) and \( \hat{s} \) degrees of freedom [1] 

\[ T_{\text{lim}}^2(\hat{s}, \beta) = \chi^2_{\beta}(\hat{s}) \]  

which comes from the assumption of data normal distribution. The similar assumption leads to the equation for \( \text{SPE}_{\text{lim}}(\hat{s}, \alpha) \) [13]: 

\[ \text{SPE}_{\text{lim}}(\hat{s}, \alpha) = \Theta \left( c_\alpha \sqrt{2\Theta_2 h_0^2 \frac{\Theta_3}{\Theta_1}} + 1 + \frac{\Theta_3 h_0(h_0 - 1)}{\Theta_1^2} \right) \]  

where: 

\[ \Theta_{i=1,2,3} = \sum_{j=1}^{n-i} \lambda_j, \quad h_0 = 1 - \frac{2\Theta_3}{3\Theta_2} \]  

and \( c_\alpha \) is a standard normal deviate corresponding to the upper percentile \( 1 - \alpha \). At this point should be indicated that \( \alpha \) and \( \beta \) are tuning parameters.

**Figure 3.** Example of PCA data decomposition for \( n = 3 \), \( \hat{s} = 2 \) and certain \( \alpha \) and \( \beta \).

PCA can be used in process monitoring field for abnormality detection. For this reason PCA model is to be identified from process data \( X \). Due to monitoring purposes it not only means PCS basis derivation, i.e. \( P \) (thus computing \( \hat{s} \) simultaneously), but also \( \Lambda, T_{\text{lim}}^2(\hat{s}, \beta) \) and \( \text{SPE}_{\text{lim}}(\hat{s}, \alpha) \) calculation. These last two quantities are required for bounding set of operational states \( x_{j=1, \ldots, N} \) considered as desirable/allowable, with respect to \( \alpha \) and \( \beta \) chosen values. This implies that process data should contain reasonably the largest representation of process desirable/allowable operational states (by the means of data samples \( x_{j=1, \ldots, N} \)) and \( X \), because of PCS identification source, is called training data. The quadruple \( \text{PCA}_{\text{M}} = [P, \Lambda, T_{\text{lim}}^2(\hat{s}, \beta), \text{SPE}_{\text{lim}}(\hat{s}, \alpha)] \) will be referred to as PCA Monitoring Model.
After $PCA^{MM}$ is obtained off-line from training data, the next step is the actual process monitoring performed on-line for data samples $x(k)$ constructed analogously to $x_{j=1,\ldots,N}$. Since $P_j$ corresponds to the model of the process under monitoring, current norms values $T^2(k)$:

$$T^2(k) = (x(k))^T P_j A_j^{-1} (P_j)^T x(k)$$

and $SPE(k)$:

$$SPE(k) = (x(k))^T (I - P_j (P_j)^T) x(k)$$

measure the (quadratic) distances of current operational state from their, training data based, expected values. Thus ratios $T^2(k)/T^2_{\lim}(\hat{s}, \beta)$ and $SPE(k)/SPE_{\lim}(\hat{s}, \alpha)$ can be used for abnormal operational state indication/detection in case of unity violation by either of them:

$$\text{abnormal operational state at time instant } k: \quad \frac{T^2(k)}{T^2_{\lim}(\hat{s}, \beta)} > 1 \quad \vee \quad \frac{SPE(k)}{SPE_{\lim}(\hat{s}, \alpha)} > 1$$

and the indication/detection magnitude (values of ratios $T^2(k)/T^2_{\lim}(\hat{s}, \beta)$ and $SPE(k)/SPE_{\lim}(\hat{s}, \alpha)$) depends on the abnormality magnitude measured by the (11), (12) relatively to the closest operational state concerned as a desirable/allowable, represented by the thresholds (7), (8).

As far as the monitored process fulfil PCA assumptions, there is a fundamental difference between abnormality detection by the $T^2(k)/T^2_{\lim}(\hat{s}, \beta)$ and $SPE(k)/SPE_{\lim}(\hat{s}, \alpha)$ ratios. The former value indicates abnormal operational states, which preserve cross-correlation structure of the process, hence caused mainly by the operation point changes, while the latter ratio is responsible for abnormalities detection of PCA modelled process. However in case of nonlinear process under PCA monitoring, which can be often found in practice [5], [6], [14] – [19], since PCS contains only linearised and originally linear part of dependencies among variables $x_{j=1,\ldots,N}$, current $T^2(k)/T^2_{\lim}(\hat{s}, \beta)$ values indicate an abnormality being a mixture of operation point as well as whole process changes, both not perpendicular to the $P_j$ directions. The same applies to the second ratio $SPE(k)/SPE_{\lim}(\hat{s}, \alpha)$, with an exception, that this measure detects abnormal operational states (again a mixture of operation point and process changes) not captured by the PCS.

### 3. MultiRegional Principal Component Analysis

Distributed Systems (DSs) can be decomposed into regions such that operational state of DS follow this decomposition resulting in a set of local (regional) operational states. Because DSs posses network process structure, any local operational state is mutually dependant of its neighbouring regions, where the dependencies are defined by network topology. Hence any source of a abnormal operational state of DS can be analysed through the local operational states, which indicate a local abnormality, if and only if, the abnormality in question, has significant influence on them (local operational states). Significance of analysed abnormality is given by some measure of desirability/allowability of operational state.

Moreover, it is the network topology that defines, which of local models (representing local operational states) are sensitive to abnormal operational state of DS with respect to its particularly placed source. This knowledge can be used to establish a methodology for abnormality source localization.

In case of PCA chosen as a basis monitoring method, as many regions $R$ of DS should be distinguished as it is possible, subject to $j$-th region compactness and minimal number $n_j > 1$ of monitored variables $x_{j=1,\ldots,n_j}$. For each of $R$ regions a local (regional) PCA Monitoring Model $PCA^{MM}_{j=1,\ldots,R}$ and $PCA^{MM}_j = [p_j, x_{j=1,\ldots,n_j}, T^2_{\lim,j}(\hat{s}_j, \beta_j), SPE_{\lim,j}(\hat{s}_j, \alpha_j)]$ is to be derived following the methodology stated in the previous section (2.1). Defining norms for the $j$-th region $T^2_j(k)$ and $SPE_j(k)$ analogously to (11), (12) computed at time instant $k$ with respect to local data sample $x_j(k) \in \mathbb{R}^{n_j}$, ratios: $T^2_j(k)/T^2_{\lim,j}(\hat{s}_j, \beta_j)$ and $SPE_j(k)/SPE_{\lim,j}(\hat{s}_j, \alpha_j)$ are considered as measures of desirability/allowability.
of local operational state and again (13) it assumed, that for the \( j \)-th region if either of these ratios violates unity, there is an abnormality in DS causing local abnormal operational state. Thus abnormal operational states indication by particular local models still depends on its magnitude (relatively to the closest operational state concerned as a desirable/allowable), however in this case abnormality magnitude is understood locally, i.e. corresponds to particular regional PCA Monitoring Model.

It is important to notice, that it is possible to distinguish between ‘process’ abnormality of DS and a sensor fault. While the latter is detected only by one regional PCA Monitoring Model (assigned to the specific sensor), the abnormality of DS radiates into all regions causing changes in \( T^2_j(k) \) and \( SPE_j(k) \) in more then one adjacent PCA Monitoring Model.

Network process structure of DSs can be illustrated by the means of nodes and links, which connections structure follows the network topology. From this point of view any distinguished region should consist of variables measured at a node and in all connected links. In special case of isotropic topology (Fig. 4) the highest desirable/allowable regional operational state violation, thus also the highest abnormality indication, is in the neighbourhood of abnormality source, since the greater is the distance of local model from the abnormality source localization, the less its operational state depends on abnormality ‘injected’ into the network.

Detection and localisation of abnormality source in DS can be briefly described as follows:

1. During the process operation both measures \( T^2_j(k) \) and \( SPE_j(k) \) for all \( R \) regional PCA Monitoring Models are monitored.
2. If any measure exceeds corresponding threshold \( T^2_{\text{lim},j}(\bar{s}_j,\beta_j) \) or \( SPE_{\text{lim},j}(\bar{s}_j,\alpha_j) \) respectively, then it is said that a certain abnormality (including sensor faults) occurs.
3. If it is only one regional PCA Monitoring Model that indicate abnormality, first check for sensor fault among \( n_j \) locally measured variables \( x_{i,j}=1,...,n_j \). Else, at once consider detected abnormality as affecting the process.
4. Regional PCA Monitoring Models with the largest \( T^2_j(k)/T^2_{\text{lim},j}(\bar{s}_j,\beta_j) \) or \( SPE_j(k)/SPE_{\text{lim},j}(\bar{s}_j,\alpha_j) \) values determine the localisation of the process abnormality source.

The main idea of proposed method is quite similar to Multi Block Principal Component Analysis (MB-PCA) presented in [20]. However, this name appeared earlier and was dedicated for quite different approach [21]. Therefore, it was proposed [5], [6] to name the method MultiRegional PCA (MR-PCA).

4. Drinking Water Distribution Systems

Drinking Water Distribution System (DWDS) is a good example of a DS. In this section the general description of DWDS is presented.

Nowadays, DWDS is one of the most important systems in community. Its efficient control requires advanced
method e.g. predictive control [22], [23] or adaptive control and reliable monitoring system. Proposed approach is applied to detection and localisation of failures in DWDS. Usually in DWDS, drinking water is introduced into the network by using pumps (pumping station) and transported through the network by pipes. Pipes connect in nodes where delivered water is mixed and transported farther. Flows through the pipes are enforced by nodal pressure differences. These are caused by pumps or/and by the water tanks. Tanks are used to store the water in periods when water production is greater then its consumption [23].

The network mathematical model is composed of two parts: static and dynamic. The static part is typically available in an implicit form represented by the element algebraic equalities and the interconnection equalities. This is described for water networks by Brdys and Ulanicki [25]. In general, the element algebraic equalities are described by non-linear functions. The interconnection equalities can be written based on conservation equations. Using the energy loss-gain relationships for the different elements of water distribution system, the conservation equation can be written in three forms: the node, the loop and pipe equations [26].

Unlike the node and loop equations, the pipe equations are solved for the vector of pipe flows Q and hydraulic head h simultaneously. Formulating the static part of water distribution network mathematical model we use the pipe form. The dynamic part of the network mathematical model is represented by differential equalities describing tanks. Because the measurements are available at discrete moments, the water distribution network model is formulated in discrete form.

Paper considers detection and localization of water pipe leakages. The DWDS is modelled in simulation packages Epanet [27]. The leakages are modelled as an emitter. The flow rate through the emitter varies as a function of pressure available at the node [27]:

\[ Q = C p^\gamma \]  

where: \( Q \) is the flow rate, \( p \) is the pressure and \( C \) is the discharge coefficient (emitter coefficient), finally \( \gamma \) is the pressure exponent.

5. MultiRegional Principal Component Analysis in Application to Drinking Water Distribution Systems

Most often faults in DWDS are pump breakdowns and water leakages from the pipes. The former is easy to identify while the leakage detection and localisation is harder as it is placed underground. The faults might be detected based on the system measurements.

When diagnosis of DWDS is under the consideration monitoring and diagnostics methods are directly divided into two groups [4]: ‘measurements and model based’, e.g. Inverse Transient Method [2] and ‘measurements based’ [7], [10], [11]. A group of measurement based methods utilizes statistical data analysis [28], but these methods are still in the stage of research and development [4]. Moreover available active leakage control methods are basically unpractical due to costs and time consuming or having the long leak detection and location times [4]. This is not the case of MR-PCA, which can join not only relatively uncomplicated statistical analysis dedicated to DSs, such as DWDS (due to SD) but also experience of staff working in water companies. Moreover, this method compromises between detection and localisation efficiency and a number of measuring devices [5], [6].

In the DWDS one can measure the water pipe flows, pressure in the nodes, water level in the tanks and water quality (e.g. chlorine concentration). For DWDS approach region (when regional PCA Monitoring Models are taken into consideration) means a measurement nodes together with all adjacent links, while the abnormal DWDS operational state is said to be caused by the single water leakage.

5.1 Fundamentals

Local \( j \)-th PCA Monitoring Model \( PCA_{j}^{MM} \) for each of \( R \) regions consists of the locally identified PCS basis \( P_{j} \) with corresponding modelled variance information in diagonal matrix \( \Lambda_{j} \), as well as thresholds computed for local abnormality indication. All these elements are derived with respect to chosen PCS dimensionality \( \hat{s}_{j} \) (e.g. using CPV criteria (4), (5), thus also \( CPV_{lim,j} \) and \( \alpha_{j} \), \( \beta_{j} \) parameters.

Local PCA models use nodal heads (heads is a sum of nodal pressure and its geodetic heights) and adjacent pipe flows as process variables (of course for PCA models identification as well as on-line
DWDS regions monitoring). Since each local model (PCA model or PCA Monitoring Model) in all cases includes one quantity connected to a particular node (namely nodal head), the ‘Identification number’ of the particular node will be referred to as the model name and replaces its \( j \)-th index. It means, e.g. (Fig. 5) that PCA Monitoring Model \( PCA_{88}^{MM} \) is related to the head in the node ‘88’ and flows in the pipes linking this particular node with its neighbours, which are: ‘44’, ‘55’ and ‘66’.

The minimal percent of variance captured by the \( j \)-th PCA model, \( CPV_{\text{lim},j} \), in all cases is set to 95%. However due to existing nonlinearities this ensures only capture of the linearization results by PCS. Values \( \alpha_j \) and \( \beta_j \) are selected separately to ensure abnormality detection (with respect to given local PCA model) and simultaneously false alarms reduction. This implies in the non-normal and nonlinear case of the DWDS, that one must pay special attention while tuning \( \alpha_j \) and \( \beta_j \). Authors in most of further presented case studies assigned values from the range 0.6 \( \pm \) 0.7 and 0.7 \( \pm \) 0.9 for \( \alpha_j \) and \( \beta_j \) respectively.

5.2 Simple water network case study

Paper presents MR-PCA method in application to leakage detection in Chojnice drinking water distribution network. Large numbers of potential leakages in the network as well as potential monitoring points might be investigated, there. Hence in order to explain the fundamentals of the method and to illustrate well its efficiency, the simulations were carried out on small testing network (Fig. 5), where the leakages were modelled in different but very meaningful places, as an extra node with pressure dependent demand (Emitter).

Fig. 5 presents three case studies when the leakage was modelled in different places in the testing network. First leakage was modelled in pipe between nodes ‘44’ and ‘88’ (Section 5.2.1), the second one in pipe between nodes ‘29’ and ‘30’ (Section 5.2.2) and the last one in pipe between nodes ‘25’ and ‘28’ (Section 5.2.3). All leakages took place between 46\(^{th}\) and 52\(^{th}\) hour.

![Figure 5. Simple water network. Modelled leakages marked with blue dots](image)

The following figures present the effects of monitoring process by using PCA Monitoring Models designed based on flow rates delivering water to particular node and its nodal head.

5.2.1 Water main screening effect

The dashed blue lines in Fig. 6. represents results of quality measures \( T^2 \) and \( SPE \) simulated on the training data, without any leakage. The solid red lines represents an effect of monitoring for the same simulation period but with modelled leakage (in pipe between nodes ‘44’ and ‘88’).
Figure 6. Leakage monitoring results by $PCA^{MM}$ designed at selected nodes – water main screening effect (simulation without leakage (training) - blue; simulation with modelled leakage – red)

Differences between the lines representing PCA models being identified on training data and its current (simulated) responses indicate that something unusual has took place in the network. In this particular example it means pipe water leakage. Notice how considerable increase of the $T^2$ and SPE measures is generated by PCA Monitoring Models ‘44’, ‘22’, ‘66’ and ‘88’ ($PCA^{MM}_{44}$; $PCA^{MM}_{22}$; $PCA^{MM}_{66}$ and $PCA^{MM}_{88}$). It simply means that the leakage has significantly disturbed the water flow rates and pressures being the base for particular PCA Monitoring Models ($PCA^{MM}$). There are also $PCA^{MM}$ e.g. $PCA^{MM}_{31}$ that do not indicate any abnormalities in the network in spite being in similar distance to the leakage. Such a phenomenon is rather strange at the first glance. However, these models are designed based on measurements gathered from points laying ‘behind’ the water main. Hence, the significant amount of water flowing by the water main causes that modelled leakage did not significantly affect the usual (operating states represented by training data) flow rates and nodal heads on the other side of water main. This implies that the PCA Monitoring Models are unable to detect the leakage. In the paper such an effect is called ‘screening effect’ of the water main.

5.2.2 Water tank screening effect

This section presents a case when the leakage was modelled in pipe between nodes ‘29’ and ‘30’. Similarly to the previous case study some of the PCA Monitoring Models produce quality measures significantly exceeding the assumed threshold (unity in the case of ratios $T^2(k)/T^2_{lim}(@\delta_j,\beta_j)$ and $SPE(k)/SPE_{lim}(@\delta_j,\alpha_j)$ monitoring) e.g. $PCA^{MM}_{55}$; $PCA^{MM}_{28}$; $PCA^{MM}_{29}$; $PCA^{MM}_{31}$ and $PCA^{MM}_{32}$, what indicates the leakage (Fig. 7). In opposite, there are models that generate measures similar to their training values what suggests that noting unusual happened.
Notice that these models are identified based on measurements from two sides of retention tank. This is due to the fact that significant amount of water flowing in and out of the retention tank damps the influence of the modelled leakage on flow rates and nodal heads on the other side of the tank. Hence, e.g. $PCA_{28}$ is not able to detect the modelled leakage. Since the values of components that the Monitoring Model composes of does not deviate much from the training data.

### 5.2.3 Isolated subnetwork effect

The third case (Fig. 5) illustrates the situation that the leakage is modelled (pipe between nodes ‘25’ and ‘28’) inside subnetwork that is isolated from the rest of the network (e.g. outskirts of the town). The ‘isolated subnetwork’ means that it slightly (if ever) supply any other subnetworks and hence the abnormalities taking place in its interior do not radiate outside it. The isolation effect is often enhanced by screening effects water main and/or retention tank related.

Fig. 8 shows quality measures produced by PCA Monitoring Models located inside the isolated subnetwork. Notice that all of the $PCA^MM$ indicate the abnormality. On the other hand, the example $PCA_{28}$ identified based on measurements almost entirely gathered from outside of the isolated subnetwork does not detect any symptoms of abnormal operational state.
5.2.4 Distinguishing the subnetworks

Screening effects of the water mains, water tanks and isolated subnetwork are the serious disadvantage in leakage detection and localisation process at the first glance. However, skilfully utilisation of these features allows for improving the efficiency of faults monitoring process. Possessing the knowledge about DWDS characteristics (pipe flow rates, velocities, nodal heads, diameters of pipes, tank localisations etc.) one can divide the network into several independent subnetworks. The subnetworks should be selected in such a way to enable for independent (with respect to abnormalities significance indication) leakage detection inside them. Such an approach enables for placing much less measuring devices for detecting and then localising the leakages. As the results number of simulations and confronting them with network topology, three regional subnetworks have been distinguished (Fig. 9).

![Figure 8. Leakage monitoring results by PCA\textsuperscript{MM} designed at selected nodes – isolated region effect (simulation without leakage (training) - blue; simulation with modelled leakage – red)](image)

![Figure 9. Selecting the regional subnetworks within the simple network](image)
PCA\textsuperscript{MM} in close neighbourhood to the source of leakage. Therefore, employment of at least two monitoring nodes within a single subnetwork enables for preliminary leakage localisation. Of course, the more PCA\textsuperscript{MM} are used, the more precise localisation is.

5.3 Chojnice case study

After presenting the fundamentals of described method, the MR-PCA is tested on Chojnice case study network.

5.3.1 Chojnice Drinking Water Distribution Systems

Chojnice is a city of forty thousand of citizens in northern Poland. Model of Chojnice DWDS [29] structure that sufficiently accurate for mentioned purposes is presented in Fig. 12 (Fig. 10 and Fig. 11). This model consists of 188 nodes, 284 pipes, two supply reservoirs in the system and one tank. Water is extracted from main reservoir by five pumps and provided to water treatment station. Model of Chojnice DWDS was built in Epanet simulator, while all the monitoring algorithms were implemented in Matlab. The monitoring points and leakages were selected to present the best advantages and disadvantages of proposed method. During the experiments pipe flows and nodal heads are determined by Model Predictive Control.

Based on the rules and observations described in Section 5.2 one may distinguish the subnetworks inside the network. Fig. 10 illustrates the example of such a subnetwork. Notice that PCA Monitoring Model ‘placed’ at node ‘147’ is able to detect potential leakages inside selected area, only.

The experiment has been carried out in such a way that water leakages were modelled at each of the network nodes/pipes one by one and hence selected monitoring node tried to detect these. The colours of the nodes in the figure indicate values of the ratios: $T_{147}^{\alpha}(k)/T_{147}^{lim}(\hat{\beta}_{147})$ and $SPE_{147}(k)/SPE_{147}^{lim}(\hat{\alpha}_{147})$ what determine the ability to detect the abnormality by PCA Monitoring Model ‘147’ \textsuperscript{MM}. The red nodes mean the highest detectability while the dark blue ones, the lowest.

5.3.2 Selection of monitoring nodes

Based on number of simulations it has been noticed that there are several PCA Monitoring Models that are able to detect leakages in much wider area then local subnetwork, only. In these models, measurements of water flows are performed in pipes with relatively small water flows comparing to their potential possibilities (regarding its diameters) and that they are located in close neighbourhoods to main streams. The examples of such place are PCA Monitoring Models designed at nodes: ‘029’, ‘051’, ‘068’, ‘167’.

Fig. 11 illustrates the areas of potential leakages that are detectable by mentioned PCA Monitoring Models. Notice that potential leakages located in almost entire network might be detected by these nodes.

Another important observation is that any leakage at main streams are easily detectable by any PCA Monitoring Models, however it is not recommended to place the monitoring points at the main streams because of their limited ability to detect the other leakages. Nevertheless, the main streams are the crucial places of the network for the operators and hence they are most often under monitoring.
Figure 11. The leakages detectability by PCA Monitoring Models designed at nodes ‘068’ and ‘029’

In the result of such an analysis five subnetworks has been selected for Chojnice network and two monitoring places at of the regions have been indicated. These are presented at Fig. 12.

Figure 12. The Chojnice DWDS. Green ellipsoids mark the selected subnetworks while blue dots mark assumed PCA Monitoring Models.

5.3.3 Simulation results - leakage detection and localization

Fig. 13 presents the results of monitoring the Chojnice DWDS by PCA Monitoring Models designed at nodes marked in Fig. 12. The situation without leakage (training data) is marked blue lines in the figures, while simulations with modelled leakage are marked red. Simulations show that only PCA Monitoring Models build at nodes ‘152’ and ‘144’ unambiguously have indicated the failure (both $SPE$ and $T^2$ have significantly exceeded the thresholds). These models consist of nodal heads and pipe flows measurements gathered inside one selected subnetwork, namely ‘IV’ (Fig. 12). It leads to conclusions that simulated leakage is located within this area. Moreover, $SPE$ and $T^2$ produced by $PCA^{MM}_{152}$ are much greater then measures of $PCA^{MM}_{144}$, hence it might suggests that the abnormality took place closer to the node ‘152’ (and it indeed is in this case), however one cannot treat this as a straight rule without taking into account the network topology and actual water distributions and its ‘trace’. Example of tool enable for analysing the routes of water flowing inside the network is an algorithm making possible for water paths (routes) finding [30].

Besides of $PCA^{MM}_{152}$ and $PCA^{MM}_{144}$, also PCA Monitoring Model ‘029’ indicates the leakage by its $T^2$ measure but as mentioned earlier it is the model of special sensitivity.
Figure 13. Leakage monitoring results by PCA Monitoring Models designed at selected nodes (simulation without leakage (training) - blue; simulation with modelled leakage – red)

Figure 14. Fragment of the Chojnice network with broken water pipe flow sensor in pipe ‘152’
5.3.4 Simulation results - sensor fault

Very important for the monitoring process is quick and correctly distinguishing the process abnormalities (e.g. pipe water leakage, pump failure) from the sensor faults (e.g. sensor drift, outliers, missing data). The proposed method enables for distinguishing of such cases.

Following simulations present situation when flow rate sensor in pipe ‘152’ (Fig. 14) broke down. Broken sensor delivers measuring data for PCA Monitoring Model ‘123’. The modelled fault took place at about 50th hour of simulation.

Fig. 15 shows the monitoring results from all selected regional PCA Monitoring Models. The situation without failure is marked blue lines in the figures, while simulations with modelled sensor fault are marked red.

![Figure 15. Sensor in pipe ‘152’ fault - monitoring results by PCA Monitoring Models designed at selected nodes (simulation without leakage (training) - blue; simulation with modelled leakage – red)](image_url)

Based on the rules derived in the paper one might try to detect the sensor fault by analyzing the values of $SPE$ and $T^2$. If the abnormalities detected by $PCA^{MM}$ are caused by significant changing the operational state of the plant (e.g. by leakage) it should be noticeable at least by all sensors inside one of the selected subnetworks. Moreover, some of the very sensitive $PCA^{MM}$ (e.g. ‘029’ as it was in case study when leakage was simulated inside this particular region) should detect it, as well. However, only PCA Monitoring Model ‘123’ indicates the fault. It suggests that one of the measurements is abrupt. In this case $PCA^{MM}_{123}$ consists of measurements from pipes ‘152’, ‘153’ ‘154’ and nodal head ‘123’ and hence we are able to state that one of these sensors probably broke down. Having more measuring devices in this subnetwork and so PCA Monitoring Models would enable for precise indicating the broken sensor by utilizing PCA features [31] or logical elimination.

6. Conclusions and Future Work

The paper has introduced a new approach to PCA based methods utilisation in detection and localisation of pipe leakage in Drinking Water Distribution System, namely MultiRegional Principal Component Analysis. In the first place MR-PCA approach to monitoring and diagnostics of network structured Distributed Systems was stated. The key idea is to use several regional PCA models (PCA Monitoring Models) identified on the basis of spatially local, available measurements to conclude about DS operational state, instead of single ‘global’ model. In particular, this enables for abnormality detection at
least. Moreover the network topology of DS may be imbedded into the MR-PCA structure resulting in better diagnostic capabilities.

Since DWDS is a representative of network systems, its abnormal operational states detection and identification can be realised by MR-PCA. In the paper pipe leakages are assumed to be the only one process abnormalities. MR-PCA methodology illustrating simulations were presented on a simple example of water network first. These, performed for a number of demand scenarios, confirmed the ability of MR-PCA to conduct the system diagnosis with regard to leakages detection and localisation. Furthermore additional phenomena were observed, providing abnormalities localisation complexity to be reduced, due to DWDS subnetworks distinguishing criteria. These are water main and retention tank screening effects, as well as isolated subnetwork effect. In practice only small part of the network variables can be directly measured. However, this does not constrain MR-PCA capabilities of leakages detection and localisation, as long as one is able to ensure at least two local PCA Monitoring Models per distinguished DWDS subnetwork, placed in their certain regions. Choice of these regions, i.e. DWDS sensors allocation is suggested from the MR-PCA abnormalities detectability point of view. Obtained results were successfully applied to case study DWDS Chojnice (Northern Poland).

At this state of the research the accurate localisation of the leakages is supervised by a man. In the future the neural networks and/or fuzzy clustering will be used to complete the process of automatic abnormalities localisation. Moreover, MR-PCA will be a part of supervised Fault Tolerant Model Predictive Control. Another subject of research in the field of fault detection and localisation will focus on identifying the process abnormality type e.g. pipe leakage, pump or valve breakdown. Obtained results are promising and have rather generic nature, hence might be transferred into other DSs e.g. pipeline systems, telecommunication systems, power systems etc, known as a network systems.

REFERENCES


17. MAZUR, K., BOROWA, A. and BRDYS, M.A., Diagnostyka i identyfikacja procesu przy użyciu AdMS-PCA, PAK 9 (BIS), 2005. (in Polish)


Networked Hydrographical Systems: A Reactive Control Strategy Integrating Time Transfer Delays

Eric Duviella¹, Pascale Chiron², Philippe Charbonnaud²

¹ Ecole des Mines de Douai, Département Informatique et Automatique
941, rue Charles Bourseul, BP 10838, 59508 Douai Cedex, France,
E-mail: duviella@ensm-douai.fr

² Laboratoire Génie de Production
ENIT, 47 av. d’Azeréix, BP 1629
65016 Tarbes Cedex, France
E-mail: pascale.chiron@enit.fr, philippe.charbonnaud@enit.fr

Abstract: A reactive control strategy integrating time transfer delays is proposed to improve the water-asset management of networked hydrographical systems. The considered systems are characterized by large scale networks where each diffluence is equipped with a control gate and a measurement point. Modelling methods of the networked hydrographical systems with equipped diffluences are presented. The proposed strategy, based on a supervision and hybrid control accommodation approach, requires generic resource allocation and setpoint assignment rules. The simulation results show the effectiveness of the reactive control strategy.

Keywords: supervision, hybrid control accommodation, resource allocation, setpoint assignment, gridded systems, water management.

Eric Duviella was born in France in 1978. Since 2007, he is assistant professor in the Computer Sciences and Automatic Department of the Ecole des Mines de Douai, France. In 2005, he has obtained his Ph.D degree from the Institut National Polytechnique de Toulouse on ‘Reactive control of extended dynamic systems with variable time delays - Application to hydrographic networks’. His research interests include modelling, hybrid dynamical systems, supervision, reactive control strategy.

Pascale Chiron was born in France in 1961. She is assistant professor (Maitre de conférences) at ENIT (National School of Engineers in Tarbes) since 2000. She has obtained her Ph.D in 1989 at Ecole Centrale in Nantes on « matching and similarity criterion in medical imaging ». She was a postdoctoral fellow at Radiology and nuclear medicine department, Faculty of medicine, Kyoto University, Japan from 1989 to 1990. Now, she is involved in theme « Planning, Control, Supervision and Distributed Simulation» in the « Automated Production » Group of Laboratoire de Génie de Production at ENIT. Her domains of interest are modelling and simulation for system control.

Philippe Charbonnaud was born in 1962 at Angoulême, France and received his Ph.D graduation in 1991 from the University Bordeaux 1. Since 2002, he is graduated Hd.R (Habilitation à diriger les Recherches) from Institut National Polytechnique de Toulouse. Since 2003, he has been Full Professor at Ecole Nationale d'Ingénieurs de Tarbes. He is member of the IFAC TC 5.4 on Large Scale Control System. His main topic of interest concerns the real-time decision support systems, and more particularly supervision and control accommodation of distributed systems.

1. Introduction

Hydrographical system is a geographically distributed network composed of dams and interconnected rivers and channels. It is characterized by great dimensions and composed of confluenes and diffluences. In real case, each diffluence is very often equipped with a control gate and a measurement point. The flow discharges are greatly disturbed by the human activities and weather conditions. An interesting problem to address, deals with the allocation of water quantities in excess toward the catchment’s areas and of water quantities in lack amongst the users. The representations of networked hydrographical systems with equipped diffluences, as well as the determination of the discharge allocation on the network, are an essential step for the design of reactive control strategy. In [18], a hydrographical network representation considering only the diffluences is proposed. Indexed nodes represent the points of diffluence, and directed arcs, whose indexes represent the number of the node downstream, represent the hydrographical systems that connect two nodes. This model was modified and extended to the case of the confluenes in [13]. In these approaches, the control and measurement instrumentations are not taken into account. In [4, 15], object-oriented modelling techniques and a XML approach make it possible to represent the elements of the hydrographical networks and of the drinking water distribution networks. Finally, modelling approaches are proposed for the optimal water management of the irrigation networks in [17], of the drinking water distribution networks, and of sewerage networks in [3]. For these four last approaches, the representation of the control and measurement instrumentations allows integrating computation rules of the
Optimization techniques were proposed in the literature for the water-asset management. The approach proposed in [10] allows the adjustment of the criteria and the constraints of an optimization problem starting from the supervision of the network variables. However, the complexity of the hydrographical networks and the number of the instrumented points to be taken into account in the optimization problem require the use of decomposition and coordination techniques of the studied systems as proposed in [17]. These techniques are used for the optimal water management of irrigation systems. In addition, a supervision and hybrid control accommodation strategy is proposed in [9] for the water-asset management of the Neste canal located in the south western region of France. This strategy is successfully adapted for the case of dam-river networks that are characterized by non-equipped diffluences [7]. Finally, to take into account majority of the networked hydrographical systems, the supervision and hybrid control accommodation strategy has to be adapted for the case of diffluences equipped with a control gate and a measurement point.

In this paper, the water asset-management by resource allocation and setpoint assignment is considered. Hydrographical systems with confluences and equipped diffluences and their representation by a weighted digraph are presented in section 2. In section 3, the reactive control strategy is defined for the water-asset management of these systems. Finally, the effectiveness of the proposed strategy is shown by simulation within the framework of a networked hydrographical system that is part of a real network. This system is composed of two diffluences and one confluence, and supplies with water downstream dams.

2. Modelling Steps of Networked Hydrographical Systems with Equipped Diffluences

Networked hydrographical systems are composed of dams and interconnected rivers and channels. The river and channels are constituted of a finite number of Simple Hydrographical Systems (HYS), i.e. composed of one stream. A representation is proposed to be able to highlight the links between the rivers, the channels and the dams, and to locate the instrumentation, i.e. the measurement points and controlled gates (see Figure 1.a and b). Each hydrographical system is equipped with several measurement points \( M \) and controlled gates \( G \), with \( i \in [1, m] \) and \( j \in [1, n] \), where \( m \) and \( n \) are respectively the total number of measurement points and actuators. It is assumed that each diffluence is equipped with at least a control gate and a measurement point. To determine the way to distribute a water quantity measured in a place of the hydrographical network, onto the whole system downstream, the networked system is represented by a digraph of instrumented points (see Figure 1.c and d).

Step 1. A digraph of instrumented points is proposed to describe the structure of the networked system by distinguishing the confluence (see Figure 1.a and c) and the diffluence (see Figure 1.b and d). The digraph consists of a succession of two types of nodes \( M \) and \( G \), represented respectively by full circle and circle and their associated graphs, and two types of arcs \( L^d \) and \( L^c \) represented respectively by solid and dashed line, which show the links between the successive nodes and the direction of the flow (see Figure 1.c and d). The attribute of the arc \( L \) is \( D \) in the case of a diffluence, and \( S \) otherwise.

![Figure 1](image-url) 

Figure 1. (a) A confluence, (c) its associated weighted digraph, (b) a diffluence, (d) its associated weighted digraph.
Thereafter, in order to represent the possible influence of measurement points, the matrix $R$ composed of $m$ lines (measurement points) and of $n$ columns (actuators), is generated. The digraph is browsed for each measurement point $M_i$ following the algorithm given in Table I. The proposed algorithm, a classical depth-first search like algorithm [6], has not been optimized in term of numerical complexity. The algorithm is only used during the design steps of the reactive control strategy and thus numerical complexity is not a challenge. The value of $R(i,j)$ is equal to 1 if there is a direct path between the measurement point $M_i$ and the gate $G_j$, and 0 otherwise. A direct path from $M_i$ to $G_j$ is a path where no arc $L_{cd}$ can be met between $M_i$ and $G_j$.

**TABLE I.** Assignment function of R matrix.

<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Initialisation of $R$ to 0</td>
<td>For each node $h$</td>
</tr>
<tr>
<td>If $h$ is a measurement point</td>
<td>Run $(h, h, R)$</td>
</tr>
<tr>
<td>EndIf</td>
<td>EndFor</td>
</tr>
<tr>
<td>Run $(h, c, R)$,</td>
<td>For each successor $d$ of $c$</td>
</tr>
<tr>
<td>If $L_{cd}$ is $L_S$</td>
<td>Run $(h, d, R)$</td>
</tr>
<tr>
<td>EndIf</td>
<td>If $d$ is a gate</td>
</tr>
<tr>
<td>$R(h, d) \leftarrow 1$</td>
<td>EndIf</td>
</tr>
<tr>
<td>EndFor</td>
<td></td>
</tr>
</tbody>
</table>

Figure 2. Time delays between measurement points and gates.

**Step 2. The value of the transfer time delay** $T_{M_i,j}$ between the measurement point $M_i$ and the gate $G_j$ are computed only if a direct path exists between $M_i$ and $G_j$, i.e. $R(i,j) = 1$. Open-Channel Reach Section (OCRS) is a part of HYS defined between a measurement point and a gate, between a gate and a measurement point, or between two gates. The transfer delay $T_{M_i,j}$ between the measurement point $M_i$ and the gate $G_j$ (see Figure 2) can be calculated by the relation:

$$T_{M_i,j} = T_{M_i,n_i} + \delta_{R(i,j)} \sum_{g=n_i+1}^{j} t_g,$$

with $n_i + 1 \leq j \leq n$, where $n_i$ is the index of the first gate downstream the measurement point $M_i$, $n$ the total number of gates, and $\delta_{ab}$ is the Kronecker index, equal to 1 when $a = b$, and equal to 0 otherwise. The trans-
fer delay \( t_g \) associated to each OCRS is computed from the OCRS dynamics model described thereafter.

Usually, Saint Venant equations are used for the modelling of open channel dynamics. The analytic resolution of these two-coupled partial differential equations [5] is not possible. As discussed in [14, 1] discretisation methods can be used to find a solution. Otherwise, a modelling method detailed in [16] based on the simplification and linearization of Saint Venant equations can be used. This method is based on the identification for each OCRS of a transfer function plus transfer delay \( 2 \) for a reference discharge \( Q_e \), according to the OCRS geometrical characteristics.

\[
F(s) = \frac{e^{-\tau}}{1 + w_1 s + w_2 s^2}, \tag{2}
\]

where the coefficients \( w_1, w_2 \) and the pure delay \( \tau \) are computed according to the identified celerity and diffusion parameters \( C_e \) and \( D_e \), and to the adimensional coefficient \( C_L \) which is defined by:

\[
C_L = \frac{2C_e X}{9D_e}, \tag{3}
\]

where \( X \) is the OCRS length, \( C_e \) and \( D_e \) are expressed as:

\[
\begin{aligned}
C_e &= \frac{1}{L} \frac{\partial}{\partial x} \left[ \frac{\partial L}{\partial x} - \frac{\partial J}{\partial y} \right], \\
D_e &= \frac{1}{L} \frac{\partial J}{\partial Q_e},
\end{aligned} \tag{4}
\]

where \( L \) is the surface width, \( y \) the discharge depth, \( J \) the friction slope expressed with the Manning-Strickler relation as \( J = \frac{Q_e^2 P^2}{K^2 S^4} \), where \( K \) is the Strickler coefficient, \( P \) the wetted perimeter and \( S \) the wetted surface.

As displayed in Table II, the order of the transfer function depends on the \( C_L \) value. When \( C_L \leq \frac{4}{9} \), the OCRS is short and can be modelled by a first order transfer function without delay, when \( \frac{4}{9} < C_L \leq 1 \), a delay is added to a first order transfer function and when \( C_L > 1 \), the OCRS is long enough and can be modelled by a second order transfer function with delay.

<table>
<thead>
<tr>
<th>( C_L )</th>
<th>( F(s) )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( C_L \leq \frac{4}{9} )</td>
<td>( F(s) = \frac{1}{1 + w_1 s} )</td>
</tr>
<tr>
<td>( \frac{4}{9} &lt; C_L \leq 1 )</td>
<td>( F(s) = \frac{e^{-\tau}}{1 + w_1 s} )</td>
</tr>
<tr>
<td>( C_L &gt; 1 )</td>
<td>( F(s) = \frac{e^{-\tau}}{1 + w_1 s + w_2 s^2} )</td>
</tr>
</tbody>
</table>

The time delay \( t_g \) (see relation \( 1 \)) depends on the network configuration (the followed path). It is computed from the step response of the transfer function, which is identified around the reference discharge \( Q_e \), and corresponds to the time so that 50% of the step response is reached. Then, the transfer delay \( T_{M_{i,j}} \) (1) is expressed according to the sampling period \( T_s \):

\[
k d_{M_{i,j}} = \left\lfloor \frac{T_{M_{i,j}}}{T_s} \right\rfloor + 1, \tag{5}
\]

where \( \left\lfloor x \right\rfloor \) denotes the integer part of \( x \). The measured water quantity in \( M_i \) will arrive on gate \( G_j \) at the date:

\[
\mathcal{M}_{M_{i,j}} = (k + kd_{M_{i,j}}) T_s. \tag{6}
\]
Finally, the transfer time delays between the measurement point $M_i$ and each gate $G_j$ are given by the Management Objective Generation Module and expressed by the vector $T_{M_i} (n \times 1)$:

$$T_{M_i} = [\mathcal{M}_{M_i,1}, \ldots, \mathcal{M}_{M_i,j}, \ldots, \mathcal{M}_{M_i,n}]^T,$$

(7)

where $\mathcal{M}_{M_i,j}$ is null if $R(i,j) = 0$.

The complex hydrographical network representation, as well as the identification of the transfer time delays, constitutes an essential step for the design of reactive control strategies.

### 3. Reactive Control Strategy

A reactive control strategy, based on a supervision and hybrid control accommodation framework, is depicted in Figure 3. The hydrographical network is represented by a set of $m$ measurement points $M_i$ and $n$ gates $G_j$ locally controlled. For each gate $G_i$, a weekly objective discharge $q_{j, obj}$, and seasonal weights $\lambda_j$ and $\mu_j$ are given by the Management Objective Generation module according to the water contracts and to climatic events. The weekly measurement point objective discharge $Q_{M_i, obj}$ is known.

![Figure 3. Supervision and hybrid control accommodation framework.](image)

For each measurement point $M_i$, discharge supervision consists in monitoring discharge disturbances and diagnosing the resource state, simultaneously. Limnimeter measurements are conditioned by a low-pass filter on a sliding window that removes wrong data due to transmission errors for instance. Based on the discharge value $Q_{M_i}$ that is measured at each sample time $kT$, detection and diagnosis automata are used respectively to detect a discharge discrepancy and to diagnose the resource states [9]. The concurrent hybrid automaton (see Figure 4) is designed for each measurement point $M_i$. The concurrent hybrid automaton formalism is drawn from the concurrent hybrid automata proposed in [2, 11, 12].

![Figure 4. Hybrid automaton for the measurement point $M_i$.](image)
The five pertinent states retained correspond respectively to no-discrepancy state $e_0$, two states where the discharge discrepancy is either positive ($e^+$) or negative ($e^-$) and constant ($c_i$), and two states where the discharge discrepancy is either positive ($e^+$) or negative ($e^-$) and no constant ($-c_i$). Transitions between states are defined as conditions on the measured discharge values and variations:

$$\begin{align*}
\psi_i &: \left| \frac{\Delta Q_{M_i}}{Q_{M_i}} \right| < th_i, \\
\omega_i &: \left| \frac{\Delta Q_{M_i}}{Q_{M_i}} \right| > dth_i
\end{align*}$$

with $\Delta Q_{M_i} = Q_{M_i, obj} - Q_{M_i}$, where $Q_{M_i}$ is the measured discharge, $Q_{M_i, obj}$ is the management objective of the measurement point $M_i$, $\dot{Q}_{M_i}$ the estimate derivative of $Q_{M_i}$, $th_i$ and $dth_i$ respectively the detection and diagnosis thresholds.

According to the resource state and the discharge discrepancy $\Delta Q_{M_i}$, the hybrid control accommodation consists in determining the setpoints $q_j$ and in assigning them to the gates taking into account the hydrographical system dynamics. The resource allocation consists in recalculating setpoints with a goal to route resource in excess to dams and to dispatch amongst the users the resource in lack. At each sample time $kT_s$, the resource allocation leads to the determination of allocation vector $q_{M_i}$ which is composed of the new computed setpoints. The allocation vector is computed according to the resource state $e$ taking into account the seasonal weights $\lambda_j$ and $\mu_j$.

**If the resource state is** no diagnose situation (denoted $E_0$), the setpoints are the objective discharges $q_{j, obj}$. The allocation vector is such as:

$$q_{M_i} = \left[ \delta_{R(i,1)} q_{j, obj}^{1}, \ldots, \delta_{R(i,j)} q_{j, obj}^{j}, \ldots, \delta_{R(i,n)} q_{j, obj}^{n} \right]^T,$$

where $n$ is the total number of gates, and $\delta_{ab}^{ij}$ is the Kronecker index.

**If the resource state is** such as discharge is constant, in lack (denoted $e^- \land c_i$) or in excess (denoted $e^+ \land c_i$), the water resource is allocated among the gates downstream the measurement point $M_i$, according to the weights $\lambda_j$ and $\mu_j$. The allocation strategy consists in optimizing a cost function by linear programming method for each measurement point:

$$f_{M_i} = \sum_{j=1}^{n} \left( \delta_{R(i,j)} \chi_{M_i,j} (q_j - q_{j, obj}) \right)$$

with $\chi_{M_i,j} = \gamma \frac{1}{\mu_j} + (\gamma - 1) \frac{1}{\mu_j}, \gamma = \frac{1}{2} \text{sign}(\Delta Q_{M_i}) + 1$.

The optimization is carried out under constraints:

$$\begin{align*}
\sum_{j=1}^{n} \left( R(i,j)(q_j - q_{j, obj}) \right) &= \Delta Q_{M_i}, \\
q_{j, min} \leq q_j \leq q_{j, max},
\end{align*}$$

where $q_{j, min}$ and $q_{j, max}$ are respectively the minimum and maximum discharges given by gate, river or canal characteristics. In this case, the allocation vector $q_{M_i}$ is such as:

$$q_{M_i} = \left[ \delta_{R(i,1)} q_{j, obj}^{1}, \ldots, \delta_{R(i,j)} q_{j, obj}^{j}, \ldots, \delta_{R(i,n)} q_{j, obj}^{n} \right]^T,$$

**If the resource state is** such as discharge is no constant, in lack (denoted $e^- \land c_i$) or in excess (denoted $e^+ \land c_i$), in order to avoid numerous re-allocation, the water resource is allocated only on one gate,
each one in its turn, at each detection date. The selection of this gate, $G_i$, is carried out according to the weights $\lambda_j$ and $\mu_j$, and to a request criterion $S_j$, storing the gate request and associated to each gate (14). As long as the state is $-c_j$, only one gate is assigned but the selected gate changes at each detection date. Because the discrepancy is not constant, at each $kT_s$, the assigned gate $G_l$ has to absorb only the discrepancy that was not yet absorbed by the previous ones:

$$\Delta q^k_{l} = \Delta Q^k_{M_i} - \Delta Q^{k-1}_{M_i},$$

$$S_i = \min_{j \in G_i} S_j,$$

$$GG_i = \begin{cases} j \leq n \text{ and } R(i,j) = 1 \end{cases}$$

The selection of this gate, $G_i$, is carried out according to the weights $\lambda_j$ and $\mu_j$, and to a request criterion $S_j$, storing the gate request and associated to each gate (14). As long as the state is $-c_j$, only one gate is assigned but the selected gate changes at each detection date. Because the discrepancy is not constant, at each $kT_s$, the assigned gate $G_l$ has to absorb only the discrepancy that was not yet absorbed by the previous ones:

$$\Delta q^k_{l} = \Delta Q^k_{M_i} - \Delta Q^{k-1}_{M_i},$$

$$S_i = \min_{j \in G_i} S_j,$$

$$GG_i = \begin{cases} j \leq n \text{ and } R(i,j) = 1 \end{cases}$$

The allocation vector $q^k_M$ is then given by:

$$q^k_M = \begin{bmatrix} q^k_{11} & \ldots & q^k_{1n} \\ \vdots & \ddots & \vdots \\ q^k_{H_M \times n} & \ldots & q^k_{H_M \times n} \end{bmatrix},$$

$$q^k_i = q^{k-1}_i + \Delta q^k_i.$$  

At each sample time $kT_s$, the setpoint assignment matrix $A^k_{M_i} (H_{M_i} \times n)$, where $H_{M_i}$ is the allocation horizon from $M_i$ (16), is scheduled according to $T_{M_i}$ and $q_{M_i}$. $H_{M_i} = \max_j T_{M_i}(j)$.

The first row of $A^k_{M_i}$ contains the setpoints to be assigned to each gate from $M_i$ at the date $(k + 1)T_s$, the $h$-row the ones to be assigned at the date $(k + h)T_s$, following the algorithm given in Table III, the last row the ones to be assigned at the date $(k + H_{M_i})T_s$. At the initial time, the values of the setpoint assignment matrix correspond to the objective discharges, i.e. $A^0_{M_i}(h,j) = q_{jobj}$. Then, for $h$ values between the date corresponding to $T_{M_i}(j)$ and the date corresponding to the allocation horizon $H_{M_i}$, the new computed setpoints $q_{M_i}(j)$ are assigned to $A^k_{M_i}(h,j)$. For values of $h$ such as $T_{M_i}(j)$ is lower than $(k + h)T_s$, the setpoints $q_{M_i}(j)$ are not up-dated, and thus, the values of the setpoint assignment matrix $A^{k-1}_{M_i}$ at time $(k - 1)T_s$ are assigned to the new matrix $A^k_{M_i}$ at time $kT_s$, with a shift delay of one period.

**TABLE III.** Setpoint assignment function of $A^k_{M_i}$ matrix.

| Input: $H_{M_i}$ horizon, $T_{M_i}$, $q_{M_i}$ and $A^{k-1}_{M_i}$ matrices. |
|--------------------------|--------------------------|
| **Output:** $A^k_{M_i}$ matrix. |
| For each measurment point $M_i$ |
| For each gate $G_j$ |
| For each row $h$ of $A^k_{M_i}$ |
| If $T_{M_i}(j) \leq (k + h)T_s$ |
| $A^k_{M_i}(h,j) = q_{M_i}(j)$ |
| Else |
| If $h < H_{M_i}$ |
| $A^k_{M_i}(h,j) = A^{k-1}_{M_i}(h + 1, j)$ |
| Else |
| $A^k_{M_i}(h,j) = q_{jobj}$ |
| EndIf |
| EndFor |
| EndFor |

Studies in Informatics and Control, Vol. 17, No. 2, June 2008 159
TABLE IV. Assignment function of $\alpha_{M_i}$ matrix.

Input: digraph.
Output: $\alpha_{M_i}$ matrix.

Initialisation of the diagonal of $\alpha_{M_i}$ to 0,

$g \leftarrow$ first gate successor of $M_i$

Run ($M_i, g, \alpha_{M_i}$)

Run ($M_i, g, \alpha_{M_i}$),

For each successor $d$ of $c$

If $L^{\alpha,d}$ is $L_2$ and $d$ is a gate

Run ($M_i, g, \alpha_{M_i}$)

$\alpha_{M_i}(d,d) \leftarrow 1$

EndIf

EndFor

Finally, the setpoints are dispatched with the control period $T_c = \kappa T_s$, where $k$ is an integer. The control setpoint vector denoted $u$ ($1 \times n$) is updated at each date $k'T_c$, where $k' = \frac{k}{\kappa}$, thanks to the assignment matrix $A^k_{M_i}$ and the $\alpha_{M_i}$ ($n \times n$) diagonal control accommodation matrix, with $H = \frac{1}{\kappa} \max(H_{M_i})$ the control horizon. For each measurement point $M_i$, the $\alpha_{M_i}$ matrix, the role of which is to capture the actual influence of the measurement point on the gates, must be determined. In order to generate the $\alpha_{M_i}$ matrix, the weighted digraph (see Figure 1.c and d) is browsed using the algorithm given in table IV, for each measurement point $M_i$. The control setpoint vector $u'_{k'}$ ($1 \times n$) is calculated by:

$$u'_{k'}(j) = \sum_{i=1}^{m} \alpha_{M_i}(j,j)A_{M_i}^k(l,j).$$

(17)

The setpoint dispatching leads to the application of the most recently calculated setpoints. This method increases the control strategy reactivity, because discharge variations between two control dates are taken into account.

4. Management of a Networked Dam-River System

The problem addressed in this section deals with the water asset management of the Purple Dam-River System (PDRS) which is a part of a real networked hydrographical system. The PDRS is characterized by a grided network configuration with two diffluences and one confluence. The PDRS is composed of the Purple channel upstream reach which supplies the Pink river, the Blue and the Orange channels, the Yellow, Jade and Cyan rivers. The Blue channel supplies the Green, the Red and the Magenta channels and the Indigo river. The Orange channel supplies the Black and the Magenta channels and the Indigo river (see Figure 5). The Red channel and Magenta channel supply downstream dams. The channels are composed of several reach sections, i.e. a part between two measurement points, about thirty kilometres length. It is considered that all the OCRS of the reach sections have trapezoidal profiles.
Figure 5. Purple dam-river system.

The PDRS is equipped with telecontrol system to satisfy at best all users and to preserve the resource. The diffluences are equipped at least by a controlled gate and a measurement point. Thus, the discharge flows, downstream the diffluence, are controlled. The PDRS’s instrumentation consists of four measurement points $M_i$ to $M_4$, and of nine controlled gates $G_i$ to $G_9$. The gate characteristics, i.e. objective discharge $q_{j,ob}$, maximum and minimum discharges $q_{j,max}$ and $q_{j,min}$, and their associated weights $\lambda_j$ and $\mu_j$, and the objective discharges of the Cyan river and the Indigo river, respectively denoted $q_{10}$ and $q_{11}$, are given in Table V. The objective discharges of $M_1$, $M_2$, $M_3$, and $M_4$ correspond, respectively, to 15 m$^3$/s, 2.5 m$^3$/s, 5 m$^3$/s and 2.5 m$^3$/s.

**TABLE V. Gate characteristics.**

<table>
<thead>
<tr>
<th>Gate</th>
<th>$q_{j,ob}$ [m$^3$/s]</th>
<th>$q_{j,min}$ [m$^3$/s]</th>
<th>$q_{j,max}$ [m$^3$/s]</th>
<th>$\lambda_j$</th>
<th>$\mu_j$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$G_1$</td>
<td>2</td>
<td>0.5</td>
<td>4</td>
<td>4</td>
<td>10</td>
</tr>
<tr>
<td>$G_2$</td>
<td>5</td>
<td>2.5</td>
<td>12</td>
<td>10</td>
<td>4</td>
</tr>
<tr>
<td>$G_3$</td>
<td>2.5</td>
<td>2</td>
<td>8</td>
<td>10</td>
<td>4</td>
</tr>
<tr>
<td>$G_4$</td>
<td>3</td>
<td>1</td>
<td>5</td>
<td>4</td>
<td>10</td>
</tr>
<tr>
<td>$G_5$</td>
<td>1</td>
<td>0.5</td>
<td>3</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>$G_6$</td>
<td>0.5</td>
<td>0.2</td>
<td>2</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>$G_7$</td>
<td>2</td>
<td>1</td>
<td>5</td>
<td>10</td>
<td>4</td>
</tr>
<tr>
<td>$G_8$</td>
<td>1</td>
<td>0.5</td>
<td>3</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>$G_9$</td>
<td>1.5</td>
<td>0.5</td>
<td>7</td>
<td>10</td>
<td>4</td>
</tr>
<tr>
<td>$q_{10}$</td>
<td>1.5</td>
<td>0.5</td>
<td>4</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>$q_{11}$</td>
<td>2.5</td>
<td>1</td>
<td>5</td>
<td>–</td>
<td>–</td>
</tr>
</tbody>
</table>
Figure 6. (a) PDRS representation, (b) its associated digraph representation.

The PDRS is subjected to disturbances upstream the Purple Channel. The considered management scenario consists in:

- Allocating the water quantities in lack, due to withdrawals, amongst the Pink and the Yellow rivers,
- Allocating the water quantities in excess, due to water restitution, toward the Red channel and Magenta channel which supply dams.

Thus, the weights $\lambda$ that are associated to the gates G2, G3, G7 and G9 are the maximum ones (see Table V). In addition, the maximum weights $\mu$ are associated to the gates $G_1$ and $G_4$ that correspond to minor priority uses.

The first step of the proposed method is based on the digraph representation of the PDRS, shown in Figure 6, which leads to the determination of the $R$ and $\alpha_M$ matrices according to the algorithms given in Tables I and IV, respectively. The matrix $R$ is given by relation (18). The values of $R(1,j)$, i.e. from the measurement point $M_1$ to the gates, is equal to 1 for the gates $G_1$ to $G_5$ because there is a direct path between this measurement point and these gates, and 0 otherwise.

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

(18)

The diagonal matrices $\alpha_{M_1}$ to $\alpha_{M_4}$ are given by (19). The fifth value of $\alpha_{M_1}$ is equal to 0 because measurement point $M_1$ is located upstream gates $G_i$.

$$ \alpha_{M_1} = diag\{1, 1, 1, 0, 0, 0, 0\}, $$
$$ \alpha_{M_2} = diag\{0, 0, 0, 0, 1, 0, 0, 0\}, $$
$$ \alpha_{M_3} = diag\{0, 0, 0, 0, 1, 1, 0, 1\}, $$
$$ \alpha_{M_4} = diag\{0, 0, 0, 0, 0, 1, 1, 1\}. $$

(19)
The second step aims at determining the transfer time delays $T_{M,i,j}$. The HYS have been modelled according to the specific length and profile section of each OCRS. The OCRS are numbered from 1 to 11 (see Figure 6.a). The OCRS with trapezoidal profile is characterized by the bottom width $B$, the average fruit of the banks $f$, the profile length $X$, the discharge depth $y$ and the slope $I$ (see Figure 7). The geometrical characteristics of the OCRS are given in Table VI.

For trapezoidal profiles, the celerity and diffusion parameters $C_e$ and $D_e$ are expressed as:

$$
\begin{align*}
C_e &= \frac{Q_e}{L^2} \left[ -f + \frac{L}{3} \left( \frac{2B}{P_y} + \frac{5L}{S} - \frac{2}{y} \right) \right] \\
D_e &= \frac{Q_e}{2LJ}.
\end{align*}
$$

(20)

with $L = B + 2fy$, $S = yB + fy^2$, $P = B + 2y\sqrt{1 + f^2}$, and the slope $J$ is equivalent to the reach slope $I$ for a non critical discharge.

In the studied case, the transfer function is estimated for one operating point for each OCRS. Parameters of the transfer functions identified for reference discharges $Q_e$ are given in Table VII. The response times $t$ are computed from the step response of every identified model, as the time such that 50% of the step response is reached. Then, the transfer delays $t_g$ are computed according to response times $t$ and to the PDRS configuration (see Figure 8). The measurement points $M_2$, $M_3$ and $M_4$ are located close to their respective upstream gates $G_4$, $G_2$ and $G_3$.

**TABLE VI.** Geometrical characteristics of the OCRS.

<table>
<thead>
<tr>
<th>OCRS</th>
<th>$B$ [m]</th>
<th>$f$</th>
<th>$X$ [m]</th>
<th>$I$</th>
<th>$K$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>5.4</td>
<td>0.8</td>
<td>1200</td>
<td>5.10^{-4}</td>
<td>70</td>
</tr>
<tr>
<td>2</td>
<td>5.4</td>
<td>0.95</td>
<td>3700</td>
<td>5.10^{-4}</td>
<td>70</td>
</tr>
<tr>
<td>3</td>
<td>5.2</td>
<td>0.95</td>
<td>1500</td>
<td>5.10^{-4}</td>
<td>70</td>
</tr>
<tr>
<td>4</td>
<td>5</td>
<td>0.9</td>
<td>2400</td>
<td>5.10^{-4}</td>
<td>70</td>
</tr>
<tr>
<td>5</td>
<td>4.8</td>
<td>0.9</td>
<td>1200</td>
<td>5.10^{-4}</td>
<td>70</td>
</tr>
<tr>
<td>6</td>
<td>4</td>
<td>0.95</td>
<td>11000</td>
<td>2.10^{-4}</td>
<td>70</td>
</tr>
<tr>
<td>7</td>
<td>4</td>
<td>0.95</td>
<td>15000</td>
<td>2.10^{-4}</td>
<td>70</td>
</tr>
<tr>
<td>8</td>
<td>4</td>
<td>0.95</td>
<td>2625</td>
<td>4.10^{-4}</td>
<td>70</td>
</tr>
<tr>
<td>9</td>
<td>3</td>
<td>0.95</td>
<td>10000</td>
<td>4.10^{-4}</td>
<td>70</td>
</tr>
<tr>
<td>10</td>
<td>3</td>
<td>0.95</td>
<td>5000</td>
<td>3.10^{-4}</td>
<td>70</td>
</tr>
<tr>
<td>11</td>
<td>3</td>
<td>0.95</td>
<td>2000</td>
<td>2.10^{-4}</td>
<td>70</td>
</tr>
</tbody>
</table>
Finally, the transfer time delays between the measurement points $M_i$ and each gate $G_j$ are expressed according to the sample time $T_s$ (equal to 120 s) using equation (5). The vectors $\mathbf{kd}_{M_i}$ are given by equation (21). Then, at each sample time $kT_s$, the vectors $\mathbf{T}_{M_i}$ are computed using (6).

\[ \mathbf{kd}_{M_1} = [4, 16, 21, 31, 37, 0, 0, 0, 0]^T, \]
\[ \mathbf{kd}_{M_2} = [0, 0, 0, 0, 6, 0, 0, 0, 0]^T, \]
\[ \mathbf{kd}_{M_3} = [0, 0, 72, 175, 0, 201]^T, \]
\[ \mathbf{kd}_{M_4} = [0, 0, 0, 0, 65, 115]^T. \]

![Figure 8](image_url)

Figure 8. Time delays between measurement points and gates for (a) Purple channel stream, (b) Blue channel stream, and (c) Orange channel stream.

**TABLE VII.** Parameters of the transfer functions.

<table>
<thead>
<tr>
<th>OCRS</th>
<th>$Q_{fr}$</th>
<th>$w_1$</th>
<th>$w_2$</th>
<th>$r_{fr}$</th>
<th>$r_f$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>15</td>
<td>570</td>
<td>0</td>
<td>0</td>
<td>395</td>
</tr>
<tr>
<td>2</td>
<td>13</td>
<td>1246</td>
<td>0</td>
<td>624</td>
<td>1480</td>
</tr>
<tr>
<td>3</td>
<td>8</td>
<td>778</td>
<td>0</td>
<td>69</td>
<td>607</td>
</tr>
<tr>
<td>4</td>
<td>5.5</td>
<td>1057</td>
<td>85800</td>
<td>419</td>
<td>1180</td>
</tr>
<tr>
<td>5</td>
<td>2.5</td>
<td>673</td>
<td>0</td>
<td>240</td>
<td>707</td>
</tr>
<tr>
<td>6</td>
<td>5</td>
<td>7425</td>
<td>109800</td>
<td>2590</td>
<td>8590</td>
</tr>
<tr>
<td>7</td>
<td>4.5</td>
<td>9397</td>
<td>218200</td>
<td>4540</td>
<td>12400</td>
</tr>
<tr>
<td>8</td>
<td>2.5</td>
<td>1619</td>
<td>443200</td>
<td>540</td>
<td>1800</td>
</tr>
<tr>
<td>9</td>
<td>2.5</td>
<td>4314</td>
<td>531200</td>
<td>3970</td>
<td>7710</td>
</tr>
<tr>
<td>10</td>
<td>1.5</td>
<td>3610</td>
<td>314900</td>
<td>1650</td>
<td>4660</td>
</tr>
<tr>
<td>11</td>
<td>4</td>
<td>1967</td>
<td>0</td>
<td>0</td>
<td>1360</td>
</tr>
</tbody>
</table>
The hydrographical system is subjected to disturbances upstream the measurement points $M$, (see Figure 9.a). Figure 9 shows discharges measured on $M$, $M'$, and $M''$, and the discharges $q_0$ and $q_t$ in the case where no reactive strategy is used (dotted line) and where the reactive strategy is used (continuous line). The results on $M$, and the setpoints dispatched on gates $G$, are not detailed herein.

For each canal reach, Figures 10, 11 and 12 show measured discharges in (a), the corresponding resource states diagnosis in (b), and the new setpoints that have been dispatched at the gates. As an example, the discharge measured on $M$, is depicted in Figure 10.a, the corresponding diagnosed resource state in Figure 10.b, the setpoints dispatched on gates $G_1$, $G_3$, $G$ and $G$, respectively, in Figures 10.c, 10.d, 10.e, and 10.f, and the resource at the end of the channel $G$, in Figure 10.g. The setpoint dispatched on gate $G$, is resulting from the addition of the new setpoints calculated from $M$ and $M$ (see Figures 11.e and 12.d). Thus, in Figure 11.e, and in Figure 12.d, the setpoint allocated starting from $M$, and, respectively, from $M$, are represented in dashed line, and the resulting setpoint dispatched on gate $G$, in continuous line.

![Figure 9](image-url) Discharges (a) $Q_{M_1}$, (b) $Q_{M_2}$, (c) $Q_{M_3}$, (d) $q_{10}$, (e) $q_{11}$, with control accommodation (continuous line) and without (dotted line).

![Figure 10](image-url) (a) Discharge $Q_{M_1}$, (b) diagnosed states from $M$. Setpoints assigned to (c) gate $G_1$, (d) gate $G_2$, (e) gate $G_3$, (f) gate $G_4$, and (g) gate $G_{10}$.

When no strategy is applied, the water quantities in lack or in excess are not allocated on the gates $G_1$ to $G_4$ (see dotted line in Figure 10.c, d, e and f), as shown by the values of the measured discharges on $M$, and $M$, (see dotted line in Figure 9.b and c). The volumes of water due to the disturbances are propagated upstream to downstream on the Purple channel, and the discharge at the end of the channel $q_0$, is far from its discharge objective $q_{10, obj}$, from the 6th hour to the 20th hour (see dotted line in Figure 9.d).

When the reactive strategy is applied, as defined by the management scenario, the water quantities in lack are allocated amongst the gates $G_1$ and $G_3$ (see Figure 10.c and f), and the water quantities in excess are allocated amongst the gates $G_1$ and $G_3$ (see Figure 10.d and e) and finally amongst the gates $G$ and $G_4$ (see...
Figures 11.d, e and 12.d). The discharges at the end of the hydrographical system are closed to the objective values, respectively, of 1.5 m³/s for $q_{10}$ and 2.5 m³/s for $q_{11}$ (see Figure 10.g and Figure 11.f). The discharge discrepancies at the end of the channels are lower than 0.1 m³/s. The positive water discrepancy upstream $M_1$ correspond to a volume of 15000 m³ during 7 hours. The proposed strategy leads to the allocation of a great part of the water volume measured upstream to $M_1$ in the downstream dams, by the control of the gates $G_6$, $G_7$, $G_9$ and $G_{11}$. A volume of 12000 m³ is directed from $M_1$ to the catchment areas, i.e. 80 % of the water volume in excess.

5. Conclusion

The water-asset management of networked hydrographical systems which are characterized by great dimensions and composed of confluences and equipped diffluences is improved by using the supervision and hybrid control accommodation strategy proposed here. In order to implement the strategy, a weighted digraph representation of dam-river systems was proposed and the resource allocation and setpoint assignment rules were defined. The reactive control strategy aims at detecting discrepancies, diagnosing the resource state and accommodating the discharge setpoints sent to the gates. The strategy was evaluated in the case of a dam-river system composed of two diffluences and one confluence, which supplies with water downstream dams. The simulation results show that the reactive strategy allows valorising the water by resource allocation and setpoint assignment. The strategy proposed in this paper is a generic tool for water resource valorisation whatever the configuration of the dam-river networks is. An interesting extension of the strategy would be the integration of fault detection and isolation methods for sensors and actuators in the supervision scheme.

Figure 11. (a) Discharge $Q_{M_1}$, (b) diagnosed states from $M_1$. Setpoints assigned to (c) gate $G_6$, (d) gate $G_7$, (e) gate $G_9$, and (f) gate $q_{11}$.

Figure 12. (a) Discharge $Q_{M_4}$, (b) diagnosed states from $M_4$. Setpoints assigned to (c) gate $G_6$, (d) gate $G_9$, and (e) gate $q_{11}$. 
REFERENCES


7. DUVIELLA, E., P. CHIRON, and P. CHARBONNAUD, Hybrid Control Accommodation for Water-asset Management of Hydraulic Systems Subjected to Large Operating Conditions, ALSIS06, 1st IFAC Workshop on Applications of Large Scale Industrial Systems, Helsinki, Finland, August 30-31, 2006.


Presentation of Some Metaheuristics for the Optimization of Complex Systems

Fatma Tangour1,2 and Pierre Borne1

1LARA, École Nationale d'Ingénieurs de Tunis, BP 37, Le Belvédère 1002 Tunis, Tunisie
2LAGIS, École Centrale de Lille, Cité scientifique, BP 48, 59651 Villeneuve d'Ascq Cedex, France
fatma.tangour@laposte.net, PIERRE.BORNE@EC-LILLE.FR

Abstract: Some metaheuristics for the optimization of complex systems are proposed in this paper. The metaheuristic approaches can be separate in two classes: the local search techniques and the global ones. An important difficulty which appears in complex optimization problems is the existence of constraints which can be strict and inviolable or soft. To resolve these problems, some hybrid approaches are considered.

Keywords: Optimization, complex systems, metaheuristics, tabu search, simulated annealing, genetic algorithms, ant colony optimization, particle swarm optimization, tunneling Algorithms.

Fatma Tangour graduated from Ecole Normale Superieure de l'Enseignement Technique de Tunis, received the Master degree of automatic and signal treatment in 2004 and Doctoral degree from the Ecole Nationale d'Ingénieur de Tunis and the Ecole Centrale de Lille in 2007. She is currently researcher in LARA Automatique at the Ecole Nationale d'Ingénieur de Tunis and working with the higher institute of sciences applied and technology of Kairouan, Tunisia. Her research interests are in the area of optimization methods for discrete events systems, computer science and operational research.

Pierre Borne 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, of the University of Waterloo (Canada) in 2006 and of the Polytechnic University of Bucarest (Romania). He is author or co-author of about 200 Journal articles and book chapters, and of 34 plenary lectures and of more than 250 communications in international conferences. He has been the supervisor of 68 PhD thesis and is author of 20 books. He is Fellow of IEEE and 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.

1. Introduction

Many optimization problems such as combinatorial optimization [3, 31] ones are usually N-P hard problems which prevent the implementation of exact used solving methodologies. It is the reason why engineers prefer to use metaheuristics which are able to produce good solutions in a reasonable computation time. The metaheuristic approaches can be separate in two classes: the local search techniques and the global ones. Among the local search techniques the Tabu search [14] is the more known. The other methods usually involve a part of stochastic approach, like the Simulated Annealing [1, 6, 27], the Genetic or Evolutionary Algorithms [28, 29, 36], the Ant Colony Optimization [11] or the Particle Swarm Optimisation [4, 8]. An important difficulty which appears in complex optimization problems is the existence of constraints which can be strict and inviolable or soft but with penalization which increase strongly with the degree of violation.

A possible acceleration of the convergence can be obtained by using tunnelling algorithms.

The multiobjective optimization is considered at the end of this paper with presentation of the OWA approach, of the Choquet integral and the Pareto optimality.

2. Tabu Search

Tabu search [12, 15, 16] is a local optimization method which enhances the performance of a local search method by using memory of the previous obtained solutions in order to permit to escape to a local optimum. It is an iterative local search procedure which enables to move from a solution to another solution in its neighbourhood until stopping criterion is satisfied. In practice the main Tabu search approach consists to determine, starting from a solution, the best solution in its immediate neighbourhood with interdiction to go to one of the N previous obtained solutions. Let us denote N(x) the list of the N solutions that have been visited in the recent past, at each step of iteration we eliminate the oldest solution of the list and we add the new one. With this method we avoid to have a cyclic evolution. It can appear that during some time we can have a degradation of the solution but it enables us to get out of a local optimum and to enlarge the search space.
Another type of Tabu search corresponds to another definition of the Tabu list: which can prohibit solutions that have certain attributes or which can prevent certain solutions that contain prohibited attributes.

An example of Tabu search corresponds to the research of the smallest value in the following Notice Board. The initial solution corresponds to a fitness equal to 12 and the Tabu list comports 5 elements, figure 1.

![Figure 1. The research of smallest value](image)

3. Simulated Annealing

3.1 Principle

Simulated annealing [23, 32] is a generic probabilistic algorithm developed to solve global optimization problems for a function defined in a large search space. The simulated annealing has obtained excellent results in various complex problems known for their important combinatorial properties. It is inspired of the physical thermic annealing. At each step of calculation of this algorithm the current solution is replaced by a nearly one chosen with a probability that depends of the variation of a fitness function (called the energy function by analogy with the physical process) via a parameter T (called the temperature) which gradually and regularly decrease during the process. In this approach the solution changes almost randomly for the large value of T and tends globally to obtain the minimum of the energy function as T tends to zero. The random evolution enables motions in which the energy can sometime increase which avoid falling and being trapped in a local minimum which can appear with usual downhill methods as the gradient method.

This algorithm can be presented as follows: let us denote $s$, $T$ and $e$ respectively the current state, temperature and energy, and $s_n$ and $e_n$ respectively the new state and energy.

The process is initialize with $s = s_0$ and $e = e_0$ which correspond to the initial state $s_0$ of energy $e_0$ at time $k = 0$.

While the stopping condition is not satisfied (time $k < k_m$ and energy $e > e_m$), pick some state in the neighbourhood and compute its energy.

3.2 Example of simulated annealing algorithm

Initialization:

$s = s_0$, $e = E(s)$, $k = 0$, $T = T_0$

while $k < k_m$ and $e > e_m$

$s_{k+1} = \text{neighbour}(s_k)$

$e_{k+1} = E(s_{k+1})$, $\Delta e_k = e_{k+1} - e_k$

if random $[0, 1] < \exp(-\Delta e_k / T_k)$

then $s_k := s_{k+1}, e_k := e_{k+1}, T_{k+1} := T_k, k = k + 1$

return while
After stabilization, decrease T and return to s.

We always save the best solution that will be the final solution given by the simulated annealing algorithm.

4. Genetic Algorithms

4.1 Principle

Genetic algorithms [9, 13, 17, 18, 26] are iterative algorithms whose aim is to optimize a fitness function. These exploration algorithms are a particular class of evolutionary algorithms [2, 19, 30, 37] based on natural biological evolution of a population who evolve by selection, crossover between individuals and mutation.

4.2 Crossover

The crossover corresponds to an exchange of genes usually between two individuals of the population. For example if we have two parents, the two points crossover corresponds to the exchange of genes represented in Figure 2.

The two parents are selected according to their fitness via a probability defined by the roulette wheel and the crossover points can be decided with a stochastic approach or using special rules.

![Two point's crossover](Image)

4.3 Example of mutation

Several positions are randomly chosen in the chromosome and the corresponding genes are randomly modified. The mutation enables to keep a sufficient diversity in the population and enables to acquire a chromosome gene value which was not already present in the population.

4.4 General algorithm

With each generation a new set of individuals (population) is created by using best parts of the precedent generation as well as innovating parts. The genetic algorithms are not purely random. They effectively exploit information obtained previously to speculate in the choice of new solutions to explore, with the hope to improve the performance. The main difficulty in the implementation of the genetic algorithms to solve an optimization problem is to determine a good coding of the problem called chromosome. Each chromosome represents an individual (a solution). An individual can have a binary representation using number 0 and 1, but any other alphanumerical encoding can be used. The algorithm is initialized by a population that can be determined by another approach or whose individuals are randomly generated. Starting from this initial population new generations are created from which the fitness of every individual is evaluated.

An example of implementation of genetic algorithm can be summarized as follows, figure 3:

1. Create an initial population
2. Evaluate the fitness of each individual of this population
3. While the terminating condition is not satisfied, repeat:
   - Select best ranking individuals to reproduce.
- Breed new generation through crossover and mutation to create offspring.
- Evaluate the fitness of the new individuals.
- Replace worst ranked individuals of the population by the new ones.

In practice, the algorithm needs to be adapted to the specificities of the studied problem and in particular crossover and mutation are to be defined in order to create viable individuals satisfying all the conditions needed for the specific problem.

The more important is to choose the good chromosome for the encoding of the solution. The mutations which correspond to the random change of some characteristics of individual insure to maintain a sufficient diversity in the population and avoid converging prematurely towards local optimum rather than to the global optimum of the problem.

### 4.5 Example of coding for planning and scheduling optimization

\[ O_{ij} : \text{operation } i \text{ of job } j \]
\[ k_{ij} : \text{machine use to achieve operation } O_{ij} \]
\[ t_{ij} : \text{starting time of the operation } O_{ij} \]

For example for three jobs, two with three operations and one with two operations:

<p>| | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>( O_{11}, k_{11}, t_{11} )</td>
<td>( O_{21}, k_{21}, t_{21} )</td>
<td>( O_{31}, k_{31}, t_{31} )</td>
</tr>
<tr>
<td>( O_{12}, k_{12}, t_{12} )</td>
<td>( O_{22}, k_{22}, t_{22} )</td>
<td>( O_{32}, k_{32}, t_{32} )</td>
</tr>
<tr>
<td>( O_{13}, k_{13}, t_{13} )</td>
<td>( O_{23}, k_{23}, t_{23} )</td>
<td></td>
</tr>
</tbody>
</table>

The various operations have to satisfy various constraints:
- precedence,
- resource disponibility,
- preemption possible or not,
- earliest starting time,
- due date,
- perisability,…

For optimizing various criteria
- makespan,
- maximum workload,
- maximum delay penalty,…

5. Ant Colony Optimization

5.1. Principle

Initially, the ant colony optimization [10, 25] was inspired by the ability and the organisation of real ant colony using external chemical pheromone trails acting as means of communication, figure 4.

![Figure 4. Pheromone trail for ants’ communication](image_url)

The main idea is that of a parallel search over several constructive computational solutions based on characteristics problem data and on a dynamic memory structure containing information on the quality of the previous obtained solutions. Generally the behaviour of an ant system optimization mechanism depends on many unsure parameters, incomplete knowledge of the real ant system attitude and the imprecise information for the identification of the relationship, the strategy of choice of the parameters and the global behaviour of the ant system metaheuristics. In practice, ants looking for food at the beginning wander randomly and having found food return to their colony while laying down pheromone trails. So other ants finding such a path are likely not to keep travelling at random but prefer generally to follow the trail, returning and reinforcing it. In fact pheromone trail slowly evaporate, reducing its attractive strength, so the more time it takes for an ant to achieve its trip to the food, the more time the pheromones have to evaporate. So the shortest path keep the highest density of pheromone trail as pheromone is laid on the path faster as it can evaporate. The evaporation is necessary in order to avoid the premature convergence to a locally optimal solution. Once an ant has found a short path from the colony to a food source, other ants prefer to follow that path, which involves a positive feedback such as finally all the ants will follow this single path. The ant colony optimization algorithms mimic this behaviour with simulated ants evolving on the graph representing the problem to solve.

5.2. Example of application of ACO to planning and scheduling

\[ P_{ijk}^f \] : Probability for the ant \( f \) to assign the operation \( i \) of job \( j \) \( (O_{ij}) \) to machine \( k \) \( (O_{ijk}) \)

\[ p_{ijk} \] : Processing time of \( O_{ij} \) with the machine \( k \)

\[ \tau_{ijk} \] : Pheromone trail related to \( O_{ijk} \)

\( D \) : Set of none performed operations

\( \alpha, \beta, \rho \) : Parameters of the algorithm (positive)

\[ L_{ijk}^f \] : Minimum value of the criterion obtained by the ant \( f \) performing \( O_{ijk} \)

\[ L_{\text{min}} \] : Global obtained minimum
• Example of ACO Algorithm

\[ \tau_{jk}(t+1) = \rho \tau_{jk}(t) + \sum_j \Delta \tau_{jk}(t+1) \]  
\[ \Delta \tau_{jk}(t) = \frac{L_{\text{min}}}{L_{jyk}(t)} \]  
\[ P_{jk}^f = \frac{(\tau_{jk})^\alpha \cdot (P_{jyk})^\beta}{\sum_{j \in D} (\tau_{jk})^\alpha \cdot (P_{jyk})^\beta} \quad \text{if} \quad j \in D \]  
\[ P_{jk}^f = 0 \quad \text{otherwise} \]

\( \alpha, \beta \) positive parameters of the algorithm.

5.3 Application

Let us consider a flexible job shop scheduling problem composed by three jobs \( J_j \) (j=1,2,3) and six machines \( M_k \), \( k = 1;::;6 \).

The objective is to optimize the completion time of scheduling, the makespan.

Table 1 depicts, for each job, the operation ordering and the processing time required by each machine.

<table>
<thead>
<tr>
<th></th>
<th>( M_1 )</th>
<th>( M_2 )</th>
<th>( M_3 )</th>
<th>( M_4 )</th>
<th>( M_5 )</th>
<th>( M_6 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( J_1 )</td>
<td>( O_{1,1} )</td>
<td>10</td>
<td>7</td>
<td>6</td>
<td>13</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td>( O_{2,1} )</td>
<td>4</td>
<td>5</td>
<td>8</td>
<td>12</td>
<td>7</td>
</tr>
<tr>
<td></td>
<td>( O_{3,1} )</td>
<td>9</td>
<td>5</td>
<td>6</td>
<td>12</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td>( O_{4,1} )</td>
<td>7</td>
<td>8</td>
<td>4</td>
<td>10</td>
<td>15</td>
</tr>
<tr>
<td>( J_2 )</td>
<td>( O_{1,2} )</td>
<td>15</td>
<td>12</td>
<td>8</td>
<td>6</td>
<td>10</td>
</tr>
<tr>
<td></td>
<td>( O_{2,2} )</td>
<td>9</td>
<td>5</td>
<td>7</td>
<td>13</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>( O_{3,2} )</td>
<td>14</td>
<td>13</td>
<td>14</td>
<td>20</td>
<td>8</td>
</tr>
<tr>
<td>( J_3 )</td>
<td>( O_{1,3} )</td>
<td>7</td>
<td>16</td>
<td>5</td>
<td>11</td>
<td>17</td>
</tr>
<tr>
<td></td>
<td>( O_{2,3} )</td>
<td>9</td>
<td>16</td>
<td>8</td>
<td>11</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td>( O_{3,3} )</td>
<td>6</td>
<td>14</td>
<td>8</td>
<td>18</td>
<td>21</td>
</tr>
</tbody>
</table>

Applying the ant colony optimization meta-heuristic, the results simulation propose different scheduling with \( C_{\text{max}} = 19 \) ut (unit of time), table 2 and table 3.

<table>
<thead>
<tr>
<th></th>
<th>( S_1 )</th>
<th>( O_1 )</th>
<th>( O_2 )</th>
<th>( O_3 )</th>
<th>( O_4 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( J_1 )</td>
<td>( M_1; [0,1] )</td>
<td>( M_2; [1,5] )</td>
<td>( M_3; [5,11] )</td>
<td>( M_4; [10,13] )</td>
<td></td>
</tr>
<tr>
<td>( J_2 )</td>
<td>( M_1; [0,6] )</td>
<td>( M_2; [6,11] )</td>
<td>( M_3; [11,19] )</td>
<td>***</td>
<td></td>
</tr>
<tr>
<td>( J_3 )</td>
<td>( M_1; [0,5] )</td>
<td>( M_2; [5,8] )</td>
<td>( M_3; [8,14] )</td>
<td>***</td>
<td></td>
</tr>
</tbody>
</table>

The solution given in the table 2 has a makespan equal to 19 ut. The machine \( M_3 \) is the cause of this value of makespan. To solve this problem, the tabu search optimisation work is applied for this solution. This method finds the operation \( O_{1,1} \) for job \( J_1 \) on \( M_2 \) that can be swapped with other machines which will reduce makespan to 18 ut. And this method finds that the operation \( O_{1,1} \) 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 \( J_2 \). The obtained solution by the tabu search method presented in table 3, [25].
### Table 3. Tabu search optimisation solution.

<table>
<thead>
<tr>
<th>$S_j$</th>
<th>$O_1$</th>
<th>$O_2$</th>
<th>$O_3$</th>
<th>$O_4$</th>
</tr>
</thead>
</table>

### 6. Particle Swarm Optimization

#### 6.1 Principle

Particle swarm optimization [21, 22] is a population based stochastic optimization technique. It is founded on the notion of cooperation between agents (the particles) that can be seen as animals with limited intellectual capacities: small memory and small intelligence. The exchange of information between them permits nevertheless that globally they succeed to solve difficult problems as it appears with bees, fishes or birds. It appears that social sharing of information among individuals in competition offers an evolutionary advantage. In the particle swarm optimization algorithm, particles move in multidimensional space and are characterized by a position and a velocity. They have two essential reasoning capabilities: the memory of their own best position and the knowledge of their neighbourhood best position.

The standard version of the algorithm can be summarized as follows:

At the beginning the particles of the swarm have a random repartition in the search space and a random velocity [5].

Then at each time step:

- Each particle evaluates the quality of its position and memorizes the best position it has reached at this time and its quality.
- Each particle exchanges information with other particles in its neighbourhood in order to know the best performance of each of them.
- At each time step each particle choose the best performance it knows and modify its velocity according to the whole data it has, to define its moving as a compromise between three tendencies: to keep on its present velocity, to go back towards its own previous best position and to move towards the overall best position it knows, figure 5.

#### 6.2 Particle swarm optimization algorithm

\[
x_i(t+1) = f(x_i(t), v_i(t), x_{im}(t), x_M(t))
\]

Very often this relation is linear

![Figure 5. Particle swarm optimization principle](image-url)
7. Tunneling Algorithms

7.1. Principle

Tunneling algorithms [24] enable to escape to local optima. The idea is the following: each time a local optimum is reached the algorithm bore a tunnel towards a new valley of the objective function $f$, figure 6.

At the origin tunneling approaches have been defined for problems with continuous variables and have been adapted to combinatorial problems later.

Two main strategies have been proposed: the stochastic tunneling and tunneling with penalty functions:

![Figure 6. Tunneling algorithm](image)

7.2 Stochastic tunneling

The stochastic tunneling [34] was initially defined to escape to local minima when implementing the simulated annealing algorithms at low temperature. The idea was to circumvent the slow dynamics of ill-shaped energy function by applying a non-linear transformation to the objective function.

7.3 Tunneling with penalty function

The tunneling with penalty functions modifies the value of local optima by adding penalty values in order to facilitate the algorithm to escape to these local optima.

7.4 Example of penalty function for Tunneling

If $f(x)$ is the fitness function we can choose the new fitness function $f_{\text{new}}(x)$:

$$f_{\text{new}}(x) = 1 - \exp\left(-\gamma \left(f(x) - f(x^*)\right)\right)$$

where $x^*$ corresponds to the best known solution and $\gamma > 0$.

8. Multiopjective Optimization

When we have several criteria in competition to optimize, various approaches can be considered:

8.1 Ordered Weighted Operator (OWA)

The operator of aggregation $C_{\text{OWA}}$ has been introduced by Yager [35]. After normalisation, the various criteria $C_i(x)$ are aggregate in a single one with weight coefficients
\[ C_{\text{owa}}(x) = \sum_{j=1} w_j C_j(x) \]  
(6)

With \( w_j \in [0,1], \sum_{j=1} w_j = 1 \)  
(7)

### 8.2 Choquet Integral [7]

It is an OWA type approach in which the weights \( w_j \) are calculated according to the interaction between various criteria.

In order to be self-contained as far as possible, necessary definitions are given in this section, adapted for multi-criteria decision making.

Let consider a finite interval set \( N_c = \{1, \ldots, n_c\} \), which can be thought as an index set of the given criteria.

- **Definition 1**

A fuzzy measure over \( N_c = \{1, \ldots, n_c\} \) is a set function \( \mu : \mathcal{P}(N_c) \to [0,1] \), such that:

1. \( \mu(\emptyset) = 0, \quad \mu(N_c) = 1 \)
2. \( \mu(A) \leq \mu(B) \) whenever \( A \subset B \subset N_c \)

The meaning attributed to \( \mu(A) \) is usually the importance or the power of the coalition \( A \) (e.g., for decision making)

- **Definition 2**

Let \( \mu \) be a fuzzy measure over \( N_c \) and \( a = (a_1, \ldots, a_{n_c}) \) the vector of criteria. The discrete Choquet integral \( C_{\mu} \) with respect to \( \mu \) is defined by:

\[ C_{\mu}(a_1, \ldots, a_{n_c}) = \sum_{i=1}^{n_c} (a_i - a_{i-1}) \mu(\{i, \ldots, n_c\}) \]  
(8)

with \( a_0 = 0 \) and \( a_i \leq \ldots \leq a_{n_c} \).  
(9)

- **Definition 3**

Let \( \mu \) be a fuzzy measure over \( N_c \). The shapely index \( I_k \), for every \( k \in N_c \), is defined by:

\[ I_k = \sum_{k \in N_c \setminus \{i\}} \frac{\mu_k - |k| - 1)!|k| !}{n_c !} \left( \mu(k \cup \{i\}) - \mu(k) \right) \]  
(10)

where \( |k| \) indicates the cardinal of \( k \) and \( 0! = 1 \).

- **Definition 4**

The average interaction index \( I_{ij} \) between two criteria \( i \) and \( j \), with respect to a fuzzy measure \( \mu \), is defined by:

\[ I_{ij} = \sum_{k \in N_c \setminus \{i,j\}} \frac{(n_c - |k| - 2)!|k| !}{(n_c - 1)!} \left( \mu(k \cup \{i,j\}) - \mu(k \cup \{i\}) - \mu(k \cup \{j\}) + \mu(k) \right) \]  
(11)

The interaction index, ranged in \([-1,1]\), is negative in the case of redundancy, and positive in the case of synergy.

- **Definition 5**

The Choquet integral formulation in terms of interaction representation is reduced to an easily interpretable form
in the case of (at most) 2-additive measures, which is for any \( a = (a_1, ..., a_n) \), as following:

\[
C_\mu(a) = \sum_{i,j=0} |a_i \land a_j| I_y + \sum_{i,j=0} |a_i \lor a_j| I_y + \sum_{i=1}^n a_i \left( I_i - \frac{1}{2} \sum_{i,j=0} I_y \right)
\]

with \( \land \) and \( \lor \) denote min and max respectively.

### 8.3. Pareto optimality approach

Pareto optimality [20, 33] is a measure of efficiency in multicriteria problems. In this approach, a non dominated solution is such that there is no other solution that performs at best as well on every criterion and which is strictly better on at least one of criteria. For a Pareto optimal solution a criterion cannot be improved without damaging at least one of the other criteria. The set of Pareto-optimal solutions corresponds to the Pareto optimal curve also called front of Pareto, figure 7.

If we have the possibility to determine lower bounds of the various criteria the Pareto optimality approach can be associate with the OWA approach. After normalisation of the criteria, we realize an aggregation of the various criteria with adaptive weights which enable a dynamic search in the direction of the lower bounds point.

For example if the optimisation is realised with a genetic algorithm, figure 8 represents the evolution of the population.
9. Conclusions

The various metaheuristics which have been presented here have been implemented in various applications such as the optimization manufacturing problems but in each case the formulation of the problem have to be adapting to the chosen algorithm.

Very often hybrid approaches are implemented using simultaneously several metaheuristics and usual local search like the hill climbing methods.

REFERENCES

On the Piecewise Continuous Control of Methane Fermentation Processes

A. Chamroo1, I. Simeonov2, C. Vasseur3, N. Christov3

1 Laboratoire d’Automatique et d’Informatique Industrielle, Université de Poitiers, 40, avenue du recteur Pineau, 86022 Poitiers, France
2 Institute of Microbiology, Bulgarian Academy of Sciences, Acad. G. Bonchev Str., 1113 Sofia, Bulgaria
3 Laboratoire d’Automatique, Génie Informatique et Signal, Université des Sciences et Technologies de Lille, 59655 Villeneuve d’Ascq, France

Abstract: The paper deals with a new approach for computer-based control of a methane fermentation process in stirred tank bioreactors. The process is described by a non-linear mathematical model based on one-stage reaction scheme. The process control is realized using a piecewise continuous regulator which enables tracking of a variable set point trajectory and overperforms the existing anaerobic digestion control schemes. Computer simulation examples illustrate the performance of the proposed approach.

Keywords: Anaerobic digestion, Non-linear systems, Piecewise continuous control

Afzal Chamroo received the PhD degree in Automatic Control from the Lille University of Science and Technology, Lille, France, in 2006. Since 2007, he has been an Associate Professor with the ESIP (Ecole Supérieure d’Ingénieurs de Poitiers) and member of the LAII (Laboratoire d’Automatique et d’Informatique Industrielle), Poitiers, France. His research interests include real time identification and control using switched systems.

Ivan Simeonov received the MSc and PhD degrees in Control Engineering from the Technical University of Sofia, Bulgaria, in 1971 and 1981, respectively. Since 1986 he has been a Senior Research Associate with the Bulgarian Academy of Sciences, where he is presently the Head of the research group “Mathematical Modeling and Computer Science” at the Institute of Microbiology. He has been an Associate Professor at the Technical University of Sofia (1988-1999), Technical University of Gabrovo (2000-2005) and New Bulgarian University (since 2005). He is the author/co-author of over 160 scientific papers published in international journals and conference proceedings.

Christian Vasseur (MSc'70, PhD'72) is a Professor of Computer Engineering, Automatics, and Signal Processing at the Lille University of Science & Technology (USTL), Lille, France. He was Director of the ‘Ecole Nationale Supérieure des Arts et Industries Textiles’ in Roubaix, France (1988-1997), and Head of the I3D (Interaction, Image et Ingénierie de la Décision) Research Laboratory at USTL (1998-2004). He is co-author of nearly 200 scientific publications in the field of automatic classification applied to real-time signal and image analysis.

Nicolai Christov (MSc'71, PhD'78) is a Professor of Computer Engineering, Automatics, and Signal Processing at the Lille University of Science & Technology (USTL), Lille, France. From 1978 to 2004, he held different research and faculty positions in the Technical University of Sofia (TUS), Bulgaria. He was Vice-Head of the Joint Laboratory of Systems Optimization, TUS – Moscow Power Institute (1988-1990), Head of the TUS Control Systems Research Laboratory (1990-1997) and Co-Chairholder of the UNESCO Chair of Engineering for Development (1997-2004). His research interests are in linear systems theory, optimal and robust control and estimation, and sensitivity analysis and numerical methods for linear systems, in which areas he has published a large number of journal and conference papers and several books.

1. Introduction

During the last two decades the anaerobic digestion (methane fermentation) has been widely used in the biological wastewater treatment and depolution of highly concentrated wastes from animal farms and agroindustries [1]. Anaerobic processes result in the production of a biogas, an important energy source that can replace fossil fuel sources and thus contribute to the greenhouse gas reduction.

Anaerobic digestion processes are very complex and their efficient control is still an open problem. Due to the very limited on-line information, the process control is usually reduced to the regulation of the biogas production rate (energy supply) or of the concentration of polluting organic matter (depolution control) at a desired value in presence of disturbances [2]. In this case, however, the classical linear regulators do not have good performances because of the strongly non-linear input-output process characteristics [3]. More sophisticated linearizing controllers have been proposed in [2,4], but due to some implementation difficulties they haven’t found practical application.

In this paper a new control scheme for anaerobic digestion processes is proposed, based on the piecewise continuous regulator developed in [6-8]. The proposed control scheme can be efficiently implemented and overperforms the existing anaerobic digestion controllers both in disturbance rejection and set point...
tracking. The proposed controller is particularly adapted in this context due to its original property which allows itself to be independent of the plant’s model. Thus, it shows robustness against the non linear nature of processes. The paper gives the theory of piecewise control leading to the robust control method. Moreover, simulation examples are proposed so as to show the performance of the approach.

2. Mathematical Model of the Process and Formulation of the Control Problem

Consider the model of anaerobic digestion in a stirred tank bioreactor, based on a one-stage reaction scheme [2, 5]:

\[
\frac{dX}{dt} = \mu X - DX
\]

\[
\frac{dS}{dt} = -k_1 \mu X + D(S_{0i} - S)
\]

\[Q = k_2 \mu X\]

with

\[
\mu = \frac{\mu_{\max} S}{k_S + S},
\]

where \(X\) is the biomass concentration (g/l), \(\mu\) - the specific growth-rate of the methane producing bacteria (day\(^{-1}\)), \(D\) - the dilution rate (day\(^{-1}\)), \(S\) - the concentration of the soluble organic (g/l), \(S_{0i}\) - influent organic pollutant concentration (g/l), \(Q\) - the biogas flow rate (litre biogas for litre of the medium per day), and \(k_1, k_2, k_S\) and \(\mu_{\max}\) are coefficients.

To avoid the washout of microorganisms, the variations of \(D\) and \(S_{0i}\) are limited in some admissible ranges:

\[0 \leq D \leq D_{\max}, \quad S_{0i}^{\min} \leq S_{0i} \leq S_{0i}^{\max}.
\]

For a laboratory-scale bioreactor the following estimates of the model parameters \(k_1, k_2, k_S\) and \(\mu_{\max}\) have been obtained [5]:

\[\hat{k}_1 = 6.7, \quad \hat{k}_2 = 16.8, \quad \hat{k}_S = 2.3, \quad \hat{\mu}_{\max} = 0.35.\]  

The non linear mathematical model (1) together with the parameter values (2) are used in the computer simulation of the anaerobic digestion control systems.

The problem of optimal control of anaerobic digestion may be decomposed in three sub problems [4]:

a) static optimisation;

b) optimal start-up;

c) dynamic optimisation.

In turn, the dynamic optimisation problem is reduced to regulation of:

1) the biogas production rate \(Q\) (energy supply),

or

2) the organics concentration \(S\) (depolution control),

at a prescribed value \((Q^*\) and \(S^*\) respectively) by acting upon the dilution rate \(D\).

In this paper the regulation of both biogas production rate and organics concentration is considered for the anaerobic digestion process (1). New, efficient solutions for these regulation problems are obtained using the piecewise continuous control scheme presented in the next section.
2. Piecewise Continuous Control

In [6,7] a class of hybrid systems called Piecewise Continuous Systems (PCS) has been introduced. These systems, characterized by autonomous switchings and controlled impulses can be used as regulators: Piecewise Continuous Controllers (PCC). PCC are easily implemented on digital calculators and allow set point tracking by the plant’s state. Though the standard PCC requires a linear model of the plant to be controlled, it is shown in [8] that an adaptation of the PCC gives rise to a particular regulator that allows control without knowledge of the plant’s model, thus suitable for some time-varying or non linear plants.

2.1 Piecewise continuous controller

The behaviour of a PCC can be summarized as follows:

- The state \( \lambda(t) \in \Sigma^\hat{n} \) of the PCC is switched to forced values at regular intervals of period \( t_e \). The corresponding switching set is represented by \( S = \{ k t_e, k = 0, 1, 2, \ldots \} \).

- The equations of the controller are

\[
\begin{align*}
\dot{\lambda}(t) &= \alpha \lambda(t), \forall t \in \left] k t_e, (k + 1)t_e \right], \\
\lambda(k t_e^+) &= \delta \psi(k t_e), \forall k = 0, 1, 2, \ldots, \\
w(t) &= \gamma \lambda(t), \forall t.
\end{align*}
\]

Equation (3a) describes the continuous evolution of the controller’s state \( \lambda(t) \in \Sigma^\hat{n} \) upon \( \left] k t_e, (k + 1)t_e \right] \), \( \alpha \in \mathbb{R}^{\hat{n} \times \hat{n}} \) being the state matrix of the controller. The only parameter that defines the behaviour of the controller’s state in this interval of time is \( \alpha \) which can take an arbitrary value. Usually, it is fixed such that the PCC is stable between switching instants.

Equation (3b) defines the controller’s state at switching instants, by means of a bounded discrete input \( \psi(k t_e) \in V^* \), and according to the linear relationship characterized by the matrix \( \delta \in \mathbb{R}^{\hat{n} \times \hat{n}} \).

Equation (3c) is the output equation of the controller, characterized by the full rank matrix \( \gamma \in \mathbb{R}^{\hat{n} \times \hat{n}} \). The output \( w(t) \in Y^\hat{m} \) constitutes the input command to be fed to the plant.

Figure 1a gives the realization diagram of a PCC and figure 1b shows its state’s evolution.

Note that from now on, the discrete values of every function will be considered as being sampled at \( t_e \) period and to simplify the notations, any time function \( f(t) \) at a given \( k t_e \) instant will be written as \( f(k t_e) = f_k \ \forall k = 0, 1, 2, \ldots \). Moreover, if any signal \( f(t) \) is discontinuous, we shall consider the right value at the discontinuity since the switchings at each \( k t_e \) imply consequences occurring at every \( k t_e^+ \). However, for simplification sake, the notation \( f_k^+ \) will be used, instead of the strict one: \( f_k^+ = f(k t_e^+) \).

2.2 Control strategy

In order to illustrate the functioning of the PCC in its standard form, we consider that the plant to be controlled is described by the usual linear state representation

\[
\begin{align*}
\dot{x}(t) &= Ax(t) + Bu(t), \\
y(t) &= Cx(t),
\end{align*}
\]
with \( A \in \mathbb{R}^{n \times n} \), \( B \in \mathbb{R}^{m \times n} \) and \( C \in \mathbb{R}^{m \times n} \) being real constant matrices, and \( x(t) \in \Sigma^m \), \( u(t) \in U^r \) and \( y(t) \in Y^m \) representing respectively the state, the input and the output of the plant.

The aim is to define the PCC parameters \( \psi(t) \) and \( \delta \) so as to achieve discrete tracking of a \( c(t) \) state trajectory by the plant’s state \( x(t) \) at each switching instant and with one sampling period delay: \( x((k+1)t_c) = c(kt_c) \), \( \forall k = 0, 1, 2, \ldots \).

To do so, we consider the closed loop system whose equations are

\[
\begin{align*}
\dot{x}(t) &= Ax(t) + Bu(t), \quad \forall t, \quad (5a) \\
\dot{\lambda}(t) &= \alpha \lambda(t), \quad \forall t \in \left] k t_c, (k+1)t_c \right[, \quad (5b) \\
u(t) &= \gamma \lambda(t), \quad \forall t, \quad (5c) \\
\lambda_k &= \delta \psi_k, \quad \forall k = 0, 1, 2, \ldots. \quad (5d)
\end{align*}
\]

By integration, the first three equations allow us to write in a sampled format, the next step value \( x_{k+1} \) of the state as a function of its previous one \( x_k \):

\[
x_{k+1} = f x_k + M \lambda_k, \quad (6)
\]

with \( f = e^{At_c} \) and \( M = \int_0^{t_c} e^{-At} B \gamma e^{At} d\tau \).

In order to realize the discrete tracking which is defined above, we only have to fix down the tracking condition which is \( x_{k+1} = c_k \), where \( c(t) \) is the desired state trajectory. Thus, from (6) we have

\[
\lambda_k = M^{-1} \left\{ c_k - f x_k \right\} \quad (7)
\]
Equation (7) gives the switching value of the controller’s state, under the condition that $M^{-1}$ exists [6]. Hence, in this case, we are able to define the PCC with

$$\delta = M^{-1} \text{ and } \Psi(t) = c(t) - f(t).$$

### 2.3 Adaptation of the PCC

According to [8], it is possible to enhance the performance of a PCC by enabling switching at high frequencies, i.e. $t_e \rightarrow 0^+$. The author shows that in this case, a PCC allows control of time varying plants or even non-linear systems, using state or output reference trajectories for state and output feedback respectively.

To understand the effect of fast switching, let’s rewrite the equation of the closed loop structure in the case of an output feedback:

$$y_{k+1} = C_f x_k + C_f \left\{ \int_0^{t_e} e^{-\alpha t} B \Psi(t) d\tau \right\} \lambda_k$$

with $\dim(\lambda(t)) = m$, $\dim(\alpha) = m \times m$, $\dim(\gamma) = r \times m$.

By realizing $t_e \rightarrow 0^+$, it is possible to simplify (8) by

$$y_{k+1} = \hat{Y} x_k + \hat{Y} \left\{ \int_0^{t_e} (I_n - A\tau) B \Psi(I_n + \alpha t_e) d\tau \right\} \lambda_k$$

with $\hat{Y} = C(I_n + At_e)$ and $I_n$ being the $n$-th order identity matrix.

We can thus write (9) as

$$y_{k+1} = y_k + CAt_e x_k + (CB\Psi t_e + \varepsilon(t_e^2))\lambda_k,$$

$\varepsilon(t_e^2)$ being negligible when $t_e \rightarrow 0^+$.

In order to evaluate the initial condition of the PCC state at each switching instant, we fix the tracking condition as $y_{k+1} = (c_o)_k$, where $c_o(t)$ is the output’s desired trajectory, such that the closed loop structure becomes

$$(CB\Psi t_e + \varepsilon(t_e^2))\lambda_k = (c_o)_k - y_k - CAt_e x_k$$

In order to solve (11) numerically, we rewrite the latter as

$$\lambda_k - \lambda_k + (CB\Psi t_e + \varepsilon(t_e^2))\lambda_k = (c_o)_k - y_k - CAt_e x_k$$

With fast switching ($t_e \rightarrow 0^+$), equation (12) becomes

$$\lambda_k = I_m^{-1} \lambda_k + (c_o)_k - y_k$$

(13)

Note that $I_m^{-1}$ is the $m$-th order diagonal matrix whose non zero terms are strictly less than 1 and tend to 1.

Equation (13) can be interpreted algorithmically by an iterative evaluation of $\lambda_k$ at each calculation step:

$$\lambda_k \leftarrow I_m^{-1} \lambda_k + (c_o)_k - y_k$$

(14)

The calculation of the initial condition of the PCC state at each switching instant is thus highly simplified.

Moreover, the structure of the PCC is simplified by the fact that $t_e \rightarrow 0^+$. In fact, in this condition, the
evolution of the controller’s state is negligible, such that the integrator setup of figure 1a acts as a Zero Order Holder (ZOH). Furthermore, if we consider that switching occur at each calculation step of a digital calculator, the ZOH can be replaced by a short circuit.

Figure 2 illustrates the structure of such a regulator.

![Diagram of a regulator](image)

Fig.2. Adaptation of a PCC (fast switching)

The major advantage of this adaptation of the usual PCC is the fact that the model of the plant is unnecessary. It thus allows control of non linear systems.

3. Simulation Experiments

By means of its mathematical model, the anaerobic digestion process has been simulated in Matlab/Simulink® and controlled with the presented regulator.

**Regulation of the output $Q$ (energy supply)**

The performances of the piecewise continuous control have been evaluated for step changes of:

- the set point $Q^*$ in the interval from 0.1 to 0.5 (Fig. 3a),
- the disturbance $S_{0i}$ in the interval from 3 to 4 (Fig. 3b).

**Regulation of the output $S$ (depolution control)**

In the same way, the controller’s performance has been tested with step changes of:

- the set point $S^*$ in the interval from 1 to 0.5 (Fig. 4a),
- the disturbance $S_{0i}$ in the interval from 3 to 4 (Fig. 4b).

According to [5], the input $D$ is limited in the interval $0 \leq D \leq 0.35$ in all cases.

The simulation results show that the proposed piecewise continuous controller presents very good performances of set point tracking (Figs. 3a, 4a) of a discontinuous reference trajectory. It also ensures regulation (Figs. 3b, 4b) of the plant’s output, rejecting disturbances in a large range.
4. Conclusion and Future Works

Control of the anaerobic digestion process presents difficulties due to its non linear nature. However, the proposed method uses a controller, which is independent of the plant's model, thus overcoming the non linear effects in the present working conditions. Moreover, the regulator can be very easily implemented and requires few step calculations.
As a perspective of our study, works are presently being carried out to realize control of the anaerobic digestion process with a wider range of the set points and disturbances.

ACKNOWLEDGEMENT

This work is supported by the European Union under Grants 15010/02Y0064/03-04 CAR/Presage N 4605 Obj. 2-2004:2 - 4.1 – N 160/4605.

REFERENCES

A Compliance Control of a Hyperredundant Robot

Mircea Ivanescu¹, Mihaela Cecilia Florescu¹, Nirvana Popescu², Decebal Popescu²

¹ University of Craiova, Romania
e-mail: mirceaivanescu@yahoo.com

² University Politehnica Bucharest, Romania

Abstract: The grasping control problem for a hyperredundant arm is studied. First, the dynamic model of the arm is analyzed. The control problems are divided into the subproblems: the position control in a desired reaching area, the control of the arm around the object-load and the force control of grasping. The difficulties determined by the complexity of the non-linear integral-differential equations are avoided by using a very basic energy relationship of this system. First, the dynamic control of the arm for a desired reaching area is inferred. Then, the position control and the force control for grasping are discussed. Numerical simulations are presented.

Keywords: distributed parameter systems, force control, grasping, hyperredundant robots.

Mircea Ivanescu received his B.Sc. degree in Automatic Control in 1965 from the University Politehnica Bucharest. He received his Ph.D. degree in Control Systems in 1975. Now, he is a Professor at the Department of Mechatronics, University of Craiova. His research interests are focused on: Distributed Parameter Systems, Discrete Optimization Problems, Fuzzy Systems, Nonlinear Systems, and Robotics.

Mihaela Cecilia Florescu received her B. Sc. and M. Sc. degrees in Electromechanical Engineering in 1995 and 1996, respectively, from University of Craiova. Now she is a Lecturer at the Mechatronics Department, University of Craiova. Her research interests are focused on: Distributed Parameter Systems, Fuzzy Systems, and Robotics.

Nirvana Popescu received her B.Sc. and M. Sc. degrees in Computer Science in 1997 and 1998, respectively, from University Politehnica, Bucharest. She received the Ph.D. degree in Computer Science in 2003 from University Politehnica, Bucharest. Now she is a Lecturer at the Computer Department, University Politehnica, Bucharest. Her research interests are focused on: Fuzzy Systems, Digital Circuits, Intelligent Adaptive Systems, and Vision Control.

Decebal Popescu received his B. Sc. and M. Sc. degrees in Computer Science in 1997 and 1998, respectively, from University Politehnica, Bucharest. He received the Ph.D. degree in Computer Science in 2003 from University Politehnica, Bucharest. Now he is a Lecturer at the Computer Department, University Politehnica, Bucharest. His research interest is focused on: Fuzzy Systems, Digital Circuits, VLSI circuits, and Robotics.

1. Introduction

The control of a hyperredundant manipulator is very complex and a great number of researchers have tried to offer solutions for this difficult problem. In [7] it was analyzed the control by cables or tendons meant to transmit forces to the elements of the arm in order to closely approximate the arm as a truly continuous backbone. In [6], Gravagne analyzed the kinematical model of “hyper-redundant” robots, known as “continuum” robots. Important results were obtained by Chirikjian and Burdick [3, 4], which laid the foundations for the kinematical theory of hyper-redundant robots. Mochiyama has also investigated the problem of controlling the shape of an HDOF rigid-link robot with two-degree-of-freedom joints using spatial curves [11]. In [11, 14] it is presented the “state of art” of continuum robots, outline their areas of application and introduce some control issues.

The difficulty of the dynamic control is determined by integral-partial-differential models with high nonlinearities that characterize the dynamic of these systems. In [8] the dynamic model for 3D space is inferred and a control law based on the energy of the system is analyzed.

In this paper, the problem of a class of hyperredundant arms with continuum elements that performs the grasping function by coiling is discussed. The difficulties determined by the complexity of the non-linear integral-differential equations, that represent the dynamic model of the system, are avoided by using a very basic energy relationship of this system. Energy-based control laws are introduced for the position control problem. A force control method is also proposed.

The paper is organized as follows: section 2 presents the basic principles of a hyperredundant structure with continuum elements; section 3 studies the dynamic model; section 4 discusses the both problem of grasping by coiling, the position control and force control; section 5 verifies by computer simulation the control laws.
2. Background

A. Technological Model

The paper studies a class of hyperredundant arms that can achieve any position and orientation in 3D space, and that can perform a coil function for the grasping. The arm is a high degree of freedom structure or a continuum structure.

Figure 1. The force sensors distribution

The general form of the arm is shown in Figure 1. It consists of a number (N) of elements, cylinders made of fibre-reinforced rubber. There are four internal chambers in the cylinder, each of them containing the ER fluid with an individual control circuit. The deformation in each cylinder is controlled by an independent electrohydraulic pressure control system combined with the distributed control of the ER fluid (Figure 2).

The last \( m < N \) elements represent the grasping terminal. These elements contain a number of force sensors distributed on the surface of the cylinders. These sensors measure the contact with the load and ensure the distributed force control during the grasping.

Figure 2. The cylinder structure

B. Theoretical model

The essence of the hyperredundant model is a 3-dimensional backbone curve \( C \) that is parametrically described by a vector \( r(s) \in \mathbb{R}^3 \) and an associated frame \( \phi(s) \in \mathbb{R}^{3x3} \) whose columns create the frame bases (see Figure 3).
Figure 3. (a) The backbone structure; (b) The backbone parameters

The independent parameter $s$ is related to the arc-length from the origin of the curve $C$, $s \in [0, L]$, where $L = \sum_{i=1}^{N} l_i$, where $l_i$ represent the length of the elements $i$ of the arm in the initial position.

The position vector on curve $C$ is given by

$$r(s, t) = [x(s, t) \ y(s, t) \ z(s, t)]^T$$

where

$$x(s, t) = \int_{0}^{s} \sin \theta(s', t) \cos q(s', t) \, ds'$$

$$y(s, t) = \int_{0}^{s} \cos \theta(s', t) \cos q(s', t) \, ds'$$

$$z(s, t) = \int_{0}^{s} \sin q(s', t) \, ds'. \quad s' \in [0, s]$$

For an element $dm$, kinetic and gravitational potential energy will be:

$$dT = \frac{1}{2} dm \left(v_x^2 + v_y^2 + v_z^2\right), \quad dV = dm \cdot g \cdot z$$

where $dm = \rho \cdot ds$, and $\rho$ is the mass density.

The elastic potential energy will be approximated by the bending of the element:

$$V_e = k \frac{d^2}{4} \sum_{i=1}^{N} \left(q_i^2 + \theta_i^2\right)$$

We will consider $F_q(s, t), F_\theta(s, t)$ the distributed forces on the length of the arm that determine motion and orientation in the $\theta$-plane, $q$-plane. The mechanical work is:
The energy-work relationship will be
\[
\left( T(t) + V^*(t) \right) - \left( T(0) + V^*(0) \right) = \int_0^t \left( F_\theta(s, \tau) \partial(s, \tau) + F_q(s, \tau) \dot{q}(s, \tau) \right) ds \, d\tau
\]
where \( T(t) \) and \( T(0) \), \( V^*(t) \) and \( V^*(0) \) are the total kinetic energy and total potential energy of the system at time \( t \) and 0, respectively.

3. Dynamic Model

In this paper, the manipulator model is considered as a distributed parameter system defined on a variable spatial domain \( \Omega = [0, \, L] \) and the spatial coordinate \( s \).

From (5), (6), (7), the distributed parameter model becomes,
\[
\rho \int_0^L \left( \dot{q}'(\sin q' \sin q'' \cos(q' - q'') + \cos q' \cos q'') - \dot{\theta}' \cos q' \sin q'' \sin(\theta' - \theta') \right) + \\
\frac{1}{2} (\dot{q}')^2 (\cos q' \sin q'' \cos(\theta' - \theta'') - \sin q' \cos q'') + (\dot{\theta}')^2 \cos q' \sin q'' \cos(\theta' - \theta'') - \\
- \dot{q}' \dot{q}'' \sin(q'' - q') ds ds'' + \rho_0 \int_0^L \cos q' ds' + k^* q = F_q
\]
\[
\rho \int_0^L \left( \dot{q}' \sin q' \cos q'' \sin(\theta' - \theta') + \dot{\theta}' \cos q' \cos q'' \cos(\theta' - \theta') - (\dot{q}')^2 \cos q' \sin q'' \sin(\theta' - \theta') \right) + \\
+ (\dot{\theta}')^2 \cos q' \cos q'' \sin(\theta' - \theta') - \dot{\theta}' \dot{\theta} \sin q' \cos q'' \cos(\theta' - \theta') ds ds'' + k^* \theta = F_\theta
\]

The control forces have the distributed components along the arm, \( F_\theta(s, t), \, F_q(s, t), \, s \in [0, \, L] \) that are determined by the lumped torques,
\[
\begin{align*}
F_\theta(s, t) &= \sum_{i=1}^{N} \delta(s - s_i) \tau_{\theta_i}(t) \\
F_q(s, t) &= \sum_{i=1}^{N} \delta(s - s_i) \tau_{q_i}(t)
\end{align*}
\]
where $\delta$ is Kronecker delta, $l_1 = l_2 = \ldots = l_N = l$, and

$$\tau_{q_i}(t) = \left(p_{q_i}^1 - p_{q_i}^2\right)S \cdot d/8 \quad (12)$$

$$\tau_{q_i}(t) = \left(p_{q_i}^1 - p_{q_i}^2\right)S \cdot d/8, \ i = 1,2,\ldots, N \quad (13)$$

In (12), (13), $p_{q_i}^1$, $p_{q_i}^2$, $p_{q_i}^1$, $p_{q_i}^2$ represent the fluid pressure in the two chamber pairs, $\theta$, $q$ and $S$, $d$ are section area and diameter of the cylinder, respectively (Figure 4). The pressure control of the chambers is described by the equations:

$$a_{ki}(\theta) \frac{dp_{ki}^k}{dt} = u_{qki}$$

$$b_{ki}(q) \frac{dp_{ki}^q}{dt} = u_{qki}, \ k = 1,2; \ i = 1,2,\ldots, N$$

where $a_{ki}(\theta)$, $b_{ki}(q)$ are determined by the fluid parameters and the geometry of the chambers and $a_{ki}(0) > 0$, $b_{ki}(0) > 0$ \quad (16)

4. Control Problem

The control problem of a grasping function by coiling is constituted from two subproblems: the position control of the arm around the object-load and the force control of grasping.

We consider that the initial state of the system is given by

$$\omega_0 = \omega(0, s) = [\theta_0, \ q_0]^T$$

corresponding to the initial position of the arm defined by the curve $C_0$

$$C_0 : (\theta_0(s), \ q_0(s)), \ s \in [0, \ L]$$

The desired point is represented by a desired position of the arm, the curve $C_d$ that coils the load,

$$\omega_d = [\theta_d, \ q_d]^T$$

$$C_d : (\theta_d(s), \ q_d(s)), \ s \in [0, \ L]$$

Figure 5. (a) The grasping position; (b) The grasping parameters
In a grasping function by coiling, only the last $m$ elements ($m < N$) are used. Let $l_g$ be the active grasping length, where $l_g = \sum_{i=m}^{n} l_i$.

We define by $e_p(t)$ the position error

$$e_p(t) = \int_{L-l_g}^{L} \left( (\theta(s,t) - \theta_b(s)) + (q(s,t) - q_b(s)) \right) ds$$

It is difficult to measure practically the angles $\theta$, $q$ for all $s \in [0, L]$. These angles can be evaluated or measured at the terminal point of each element. In this case, the relation (21) becomes

$$e_p(t) = \sum_{i=m}^{N} \left( (\theta_i(t) - \theta_{bi}) + (q_i(t) - q_{bi}) \right)$$

The error can also be expressed with respect to the global desired position $C_d$

$$e_p(t) = \sum_{i=1}^{N} \left( (\theta_i(t) - \theta_{di}) + (q_i(t) - q_{di}) \right)$$

The position control of the arm means the motion control from the initial position $C_0$ to the desired position $C_b$ in order to minimize the error.

A. Desired Area Reaching Control

An area reaching control problem is discussed. The desired area is specified by the inequality function:

$$f(\delta r) \leq 0$$

where $f$ is a scalar function with continuous first partial derivates, $\delta r = r_F - r_0$, $r_0 \in R^3$ is a reference point of the desired area and $r_F$ is the position vector (3) of the terminal point.

The potential energy function for the area reaching control has the form [1]:

$$V_p(r) = \begin{cases} 
0, f(\delta r) \leq 0 \\
\frac{1}{2} k_p f^2(\delta r), f(\delta r) > 0 
\end{cases}$$

Theorem 1. The closed-loop control system for the desired reaching area problem is stable if the control forces are

$$\tau_{\theta_i}(t) = -k_{\theta_i} e_{\theta_i}(t) - k_{\theta_i} e_{\theta_i}^2(t) - \max \left\{ 0, \frac{\partial V_p^T}{\partial \theta} \cdot k_{\theta_i} a^*(\theta_i, q_i) \right\}$$

Proof. See Appendix.
B. Fluid Pressure Control

**Theorem 2.** The closed-loop control system of the position (9), (10), (14), (15) is stable if the fluid pressures control law in the chambers of the elements given by:

\[ u_{\theta i}(t) = -a_{ji}(\theta)k_{\theta i}^{j1}\dot{\theta}_i(t) + k_{\theta i}^{j1}\dot{\theta}_i(t) \]  
\[ u_{qi}(t) = -b_{ji}(\theta)k_{qi}^{j1}\dot{\theta}_i(t) + k_{qi}^{j1}\dot{\theta}_i(t), \]

where \( j = 1,2 \); \( i = 1,2,\ldots,N \), with initial conditions:

\[ p_{\theta i}(0) - p_{\theta i}^2(0) = (k_{\theta i}^{11} - k_{\theta i}^{21})\dot{\theta}_i(0) \]  
\[ p_{qi}(0) - p_{qi}^2(0) = (k_{qi}^{11} - k_{qi}^{21})\dot{\theta}_i(0) \]  
\[ \dot{\theta}_i(0) = 0 \]  
\[ \dot{\theta}_i(0) = 0, \ i = 1,2,\ldots,N \]

and the coefficients \( k_{\theta i}^{11}, k_{\theta i}^{21}, k_{\theta i}^{12}, k_{\theta i}^{22} \) are positive and verify the conditions

\[ k_{\theta i}^{11} > k_{\theta i}^{21}, \ k_{\theta i}^{12} > k_{\theta i}^{22} \]
\[ k_{qi}^{11} > k_{qi}^{21}, \ k_{qi}^{12} > k_{qi}^{22}, \ i = 1,2,\ldots,N \]

**Proof.** See [9].

C. Force Control

The grasping by coiling of the continuum terminal elements offers a very good solution in the face of uncertainty on the geometry of the contact surface. The contact between an element and the load is presented in Figure 6. It is assumed that the grasping is determined by the chambers in \( \Theta \)-plane.

The relation between the fluid pressure and the grasping forces can be inferred for a steady state from [6],

\[ \int_0^l k \dddot{\theta} ds + \int_0^l f(s)\dddot{\dot{\theta}}(s) + \int_0^s \dddot{\theta}(s)ds = (p_1 - p_2)S \frac{d}{8} \]
where \( f(s) \) is the orthogonal force on the curve \( C, \ f(s) \) is \( F_\theta(s) \) in \( \theta \)-plane and \( F_q(s) \) in \( q \)-plane, respectively.

For small variation \( \Delta \theta_l \) around the desired position \( \theta_{id} \), in \( \theta \)-plane, the dynamic model (22) can be approximated by the following discrete model [9],

\[
\begin{align*}
\dot{m}_l \Delta \dot{\theta}_l + c_i \Delta \dot{\theta}_l + H_i (\theta_{id} + \Delta \theta_l, \ \theta_{id}, \ q_d) - H(\theta_{id}, \ q_d) &= d_i (f_i - F_{el}) , \\
\end{align*}
\]

where \( m_i = \rho S \Delta, \ i = 1,2,\ldots, I_1 \). \( H(\theta_{id}, \ q_d) \) is a nonlinear function defined on the desired position \( (\theta_{id}, \ q_d) \), \( c_i = c_i (\nu, \ \theta_l, \ q_d) \), \( c_i > 0, \ \theta, q \in \Gamma(\Omega) \), where \( \nu \) is the viscosity of the fluid in the chambers.

The equation (37) becomes:

\[
\begin{align*}
m_i \Delta \dot{\theta}_l + c_i (\nu, \ \theta_l, \ q_d) \Delta \dot{\theta}_l + h_i (\theta_{id}, \ q_d) \cdot \Delta \theta_l &= d_i (f_i - F_{el}) \\
\end{align*}
\]

The aim of explicit force control is to exert a desired force \( F_{id} \). If the contact with load is modelled as a linear spring with constant stiffness \( k_L \), the environment force can be modelled as \( F_{el} = k_L \Delta \theta_l \).

![Figure 7. The block scheme of the control system](image)

The error of the force control may be introduced as

\[
e_{fi} = F_{ie} - F_{id}
\]

It may be easily shown that the equation (38) becomes

\[
\frac{m_i}{k_L} \ddot{e}_{fi} + \frac{c_i}{k_L} \dot{e}_{fi} + \left( \frac{h_i}{k} + d_i \right) e_{fi} = d_i f_i - \left( \frac{h_i}{k} + d_i \right) F_{id}
\]

**Theorem 3.** The closed force control system is asymptotic stable if the control law is

\[
f_i = \frac{1}{k_L d_i} \left( h_i + k_L d_i + m_i \sigma^2 \right) \dot{e}_{fi} - \left( h_i - k_L d_i \right) F_{id}
\]

\[
c_i > m_i \sigma
\]

**Proof.** See [9].

**5. Simulation**

A hyperredundant manipulator with eight elements is considered. The mechanical parameters are: linear density \( \rho = 2.2 \text{ kg/m} \) and the length of one element is \( l = 0.05m \). The control problem in the \( \theta \)-plane
will be analyzed. The initial position is defined by \( C_0 \) where \( \theta_0(s) = \frac{\pi}{2} \). First, a reaching desired area control is introduced where the area is defined by the circle \( \delta(\theta) = (r - r_0)^2 - R^2, \quad r_0 = (3, 2), \quad R = 1 \) and the control law (20) is applied. The position error \( \Delta r \) is computed and the phase error is presented in Figure 8. Then, the grasping function is performed for a circular load defined by \( C_b : (x^* - x_0^*)^2 + (y^* - y_0^*)^2 = (r^*)^2 \), where \( (x^*, y^*) \) represent the coordinates in \( \theta \)-plane. A discretisation for each element with an increment \( \Delta = l/3 \) is introduced. A control law (28) is used. The result is presented in Figure 9.

**Figure 8.** The position error phase portrait

**Figure 9.** The position control
A force control for the grasping terminals is simulated. The phase portrait of the force error is presented in Figure 10. First, the control (28), (29) is used and then, when the trajectory penetrates the switching line the viscosity is increased for a damping coefficient $\xi = 1.15$.

![Force control phase portrait](image)

**Figure 10.** The force control phase portrait

### 6. Conclusion

The paper treats the control problem of a hyperredundant robot with continuum elements that performs the coil function for grasping. The structure of the arm is given by flexible composite materials in conjunction with active-controllable electro-rheological fluids. The dynamic model of the system is inferred by using Lagrange equations developed for infinite dimensional systems.

The grasping problem is divided in two subproblems: the position control and force control. The difficulties determined by the complexity of the non-linear integral-differential equations are avoided by using a very basic energy relationship of this system and energy-based control laws are introduced for the position control problem. Numerical simulations are presented.

### Appendix

We consider the following Lyapunov function:

$$W(t) = T(t) + V(t) + \frac{1}{2} \sum_{i=1}^{N} \left( k_{q_i} e_{q_i}^2 (t) + k_{\dot{q}_i} e_{\dot{q}_i}^2 (t) \right) + V_p (r_T(t))$$ (A.1)
where $T$, $V$ represent the kinetic and potential energies of the system. $W(t)$ is positive definite because the terms that represent the energy $T$ and $V$ are always $T(t) \geq 0$, $V(t) \geq 0$.

By using (19), the derivative of this function will be:

$$
W(t) = \sum_{i=1}^{N} \left( \tau_{\theta_i}(t) \dot{\theta}_i + \tau_{q_i}(t) \dot{q}_i \right) + k_{\theta} \epsilon_{\theta_i}(t) \dot{\theta}_i + k_{q_i} \epsilon_{q_i}(t) \dot{q}_i + k_{p_{q_i}} a^*(\theta_i, q_i) + k_{p_{\theta_i}} a^*(\theta_i, q_i) + k_{p_{q_i}} a^*(\theta_i, q_i) + k_{p_{\theta_i}} a^*(\theta_i, q_i) + k_{q_i} b^*(\theta_i, q_i) + k_{q_i} b^*(\theta_i, q_i) + k_{q_i} b^*(\theta_i, q_i) + k_{q_i} b^*(\theta_i, q_i) \right)
$$

(A.2)

For a constant desired position $\theta_{d_i}$, $q_{d_i}$, from (20)-(21), the relation (A.2) can be rewritten as:

$$
\dot{W}(t) = -\sum_{i=1}^{N} \left( k_{\theta_i} \dot{\theta}_i^2 + k_{q_i} \dot{q}_i^2 \right) \leq 0
$$

(A.3)

and, from [1], the closed-loop system defined by (20) converges to the desired position $\theta_i \rightarrow \theta_{d_i}$, $q_i \rightarrow q_{d_i}$ and the terminal point $r_F$ converges to the desired area $f(\delta r) \leq 0$ as $t \rightarrow \infty$.

REFERENCES


A Digital Flatness-based Control System of a DC Motor

Mounir Ayadi, Joseph Haggège, Soufiène Bouallègue and Mohamed Benrejeb

Research Unit LARA Automatique, ENIT
BP 37, Le Belvédère, 1002 Tunis, Tunisia
{mounir.ayadi, joseph.haggege, mohamed.benrejeb}@enit.rnu.tn
boualllegesoufiene@yahoo.fr

Abstract: In this paper, an approach to design and implement a real-time system based on a RISC microcontroller dedicated to a DC motor speed control, is proposed. A polynomial RST controller based on the flatness property of linear systems is implemented in C/C++ embedded programming language. The flatness property is used in order to design a robust controller with high performance in terms of tracking. The proposed controller is then applied to control a DC motor. The simulation and experimental results underline the efficiency of the flatness-based polynomial controller in a real-time control framework.

Keywords: Flatness control, RST polynomial controller, robustness, trajectory planning and tracking, microcontroller, AC-DC power converter.

Mounir Ayadi graduated from Ecole Nationale d’Ingénieurs de Tunis in 1998 and received his PhD degree in Automatic Control from the Institut National Polytechnique de Toulouse in 2002. He was a post-doctoral fellow at the Ecole Supérieure d’Ingénieurs en Génie Electrique de Rouen in 2003. He is currently Maître-Assistant at the Ecole Nationale d’Ingénieurs de Tunis. His research interests are in the area of control system theory, predictive and adaptive control, and flat systems.

Joseph Haggège graduated from Ecole Nationale d’Ingénieurs de Tunis and received the PhD degree in Electrical Engineering in 1998 and 2003, respectively. He is currently Maître-Assistant at the Ecole Nationale d’Ingénieurs de Tunis. His research interests are in the area of fuzzy, neural and neuro-fuzzy control, and embedded systems.

Soufiène Bouallègue received the Engineer Diploma degree in Electrical Engineering and the Master in Automatic Control from Ecole Nationale d’Ingénieurs de Tunis in 2006 and 2007, respectively.

Mohamed Benrejeb has obtained the Diploma of “Ingénieur IDN” (French “Grande Ecole”) in 1973, the Master degree of Automatic Control in 1974, the PhD in Automatic Control of the University of Lille in 1976 and the DSc of the same University in 1980. He is currently a full Professor at the Ecole Nationale d’Ingénieurs de Tunis and an invited Professor at the Ecole Centrale de Lille. His research interests are in the area of analysis and synthesis of complex systems based on classical and non conventional approaches.

1. Introduction

The possibility of varying or reversing the rotation speed of motors constitutes a necessity in industrial domain and in process automation. In this framework, the use of a DC motor can be a good solution. However, the speed control presents some delicate problems, particularly in terms of robustness, of perturbation rejection and especially when tracking trajectory. In this case, it is interesting to consider adequate control laws for the drive. Several techniques of control theory systems using conventional and non conventional methods have been applied on such type of motor [16][22][24][28].

The flatness property, introduced in 1992, presents a new point of view in the control theory domain [1-10]. This property, developed initially in the nonlinear continuous-time case, defines a class of systems well known as flat systems. The existence of a variable called flat or linearizing output allows to define all other system variables. In the linear case [3][4], it is sufficient to consider the Brunovsky's outputs of the canonical controllability form like the flat outputs [30][31]. Thus, the dynamic of such a process can be deduced without solving differential equations. Therefore, it is possible to express the state, as well as the input system, as differential functions of the flat output. The main contribution of the flatness that will be exploited in this paper is the effective trajectory planning and tracking solutions with high performances specification.

The flatness property of DC motor model will be used to obtain of a closed-loop system with high performances by designing a robust polynomial RST controller with an argued choice for its design. The tracking of a reference trajectory, function of the flat output system, as well as the rejection of disturbances and noises of measure, will be the goal of this controller.

On the other hand, the evolutions in the micro-electronics field, especially in the domain of digital signals processors are at the origin of many progresses in power electronics. Many works in the control theory domain have shown the performances of DSP (Digital Signal Processor) integration, microcontrollers (PIC, AVR...) or as digital components like FPGA or ASIC. The development of built-in digital controller on programmable target PIC 16F876 presents an advantage of fast real-time implementation of the...
control algorithms as well as the reduction in terms of components number and bulk [12-14].

In this paper, an approach to design such a control system as well as its hardware architecture is presented. A software integration of the developed RST algorithm, relative to the computer tools bound to this type of built-in circuits in embedded programming C/C++ language is implemented in a real-time environment.

2. Central Algorithm

2.1. Flatness-based control

The dynamic linear discrete system can be represented by equation (1) where the signals \( u_k \) and \( y_k \) represent the discrete systems input and output, respectively:

\[
A(q)y_k = B(q)u_k
\]  

(1)

In this equation, the polynomials \( A(q) \) and \( B(q) \) are the denominator and the numerator of the transfer function respectively:

\[
A(q) = q^n + a_{n-1}q^{n-1} + \ldots + a_1q + a_0
\]

\[
B(q) = b_{n-1}q^{n-1} + \ldots + b_1q + b_0
\]  

(2)

where \( q \) is the forward operator.

Applied in a real-time framework on industrial processes [5-10], the flatness property consists in designing a control signal \( u_k \) that allows to the flat output system \( z_k \) to track asymptotically a desired trajectory \( z_k^d \). The discrete flat output \( z_k \) is related to the control signal \( u_k \) and to the output system \( y_k \) by the following relations:

\[
u_k = A(q)z_k
\]

\[
y_k = B(q)z_k
\]  

(3)

The flatness-based control law can be then given by the following expression:

\[
u_k = z_{k+n}^d + \sum_{i=0}^{n-1} k_i(z_{k+i}^d - z_{k+i}) + \sum_{i=0}^{n-1} a_i z_{k+i}
\]  

(4)

which can be rewritten in the following polynomial expression:

\[
u_k = K(q)z_k^d + (A(q) - K(q))z_k
\]  

(5)

The tracking error dynamics is defined by the choice of the \( k_i \) coefficients of the \( K(q) \) polynomial given by:

\[
K(q) = q^n + k_{n-1}q^{n-1} + \ldots + k_1q + k_0
\]  

(6)

which must be Hurwitz [7].

2.2. Flatness-based RST controller

To design a polynomial controller in a linear case, we are going to follow the method given in [5] [6]. The proposed method is based on the design of a direct observer of the flat output and its forward values.

By considering the state vector \( Z_k = (z_k \ z_{k+1} \ \ldots \ z_{k+n-1})^T \) of the controllable Luenberger realization of system (3), the system of the equations (3), in which the variable \( z_k \) is the partial state [30],
is considered in a canonical state space representation of the controllable LUENBERGER realization:

\[ Z_{k+1} = AZ_k + Bu_k \]
\[ y_k = CZ_k \]  

(7)

where the matrix \( A \), \( B \) and \( C \) are given by:

\[
A = \begin{bmatrix}
0 & 1 & 0 & \ldots & 0 \\
0 & 0 & 1 & \ddots & \vdots \\
\vdots & \ddots & \ddots & \ddots & 0 \\
0 & \ldots & 0 & 0 & 1 \\
-a_0 & -a_1 & \ldots & -a_{n-2} & -a_{n-1}
\end{bmatrix}, \quad B = \begin{bmatrix} 0 & \ldots & 0 & 1 \end{bmatrix}^T, \quad C = \begin{bmatrix} b_0 & b_1 & \ldots & b_{n-1} \end{bmatrix}
\]

By using the flatness property, the control law \( u_k \), given by the equation (5) can be expressed as:

\[ u_k = KL(q)z_k^d + (a - k)Z_k \]  

(8)

where \( a \) and \( k \) are two constant vectors constituted by the \( a_i \) and \( k_i \) coefficients of the \( A(q) \) and \( K(q) \) polynomials given by:

\[
a = (a_0, a_1, \ldots, a_{n-1}), \quad k = (k_0, k_1, \ldots, k_{n-1})
\]

The realisable structure of the RST controller in terms of the \( q^{-1} \) operator can be obtained by:

\[ S(q^{-1})u_k = KL(q)z_k^d - R(q^{-1})y_k \]  

(9)

with

\[
S(q^{-1}) = 1 + (a - k)\left( A^{n-1}Q^{-1}M - (A^{n-2}B \ldots AB) \right)Q^*
\]

(10)

\[
R(q^{-1}) = -(a - k)A^{n-1}Q
\]

(11)

and

\[
Q^* = \begin{bmatrix} q^{-(n-1)} & q^{-(n-2)} & \ldots & q^{-1} \end{bmatrix}^T
\]

(12)

\[
Q = \begin{bmatrix} q^{-(n-1)} & q^{-(n-2)} & \ldots & 1 \end{bmatrix}^T
\]

where \( M \) and \( O \) are the controllability and the observability matrices respectively. Thus, the closed loop dynamics is defined by the tracking polynomial \( K(q) \) of the desired flat trajectory \( z_k^d \).

*Figure 1. Flatness-based RST Polynomial Controller for DC Motor Model*
In order to obtain a robust controller, we must introduce in the model equation (3), the pre-specified parts $H_S$ and $H_R$, given by equation (13). In [32], it is shown that the polynomial $H_R$ is used to eliminate the high frequency noises on the input signal and the $H_S$ polynomial to allow the rejection of static disturbance present on the output signal.

$$H_s(q^{-1}) = 1 - q^{-1}$$  \hspace{1cm} (13)$$

$$H_r(q^{-1}) = 1 + q^{-1}$$

While taking into account the pre-specified parts $H_S$ and $H_R$, the polynomials of the re-calculated RST controller can be seen as in equation (14).

$$\hat{S}(q^{-1}) = H_s(q^{-1})S(q^{-1})$$

$$\hat{R}(q^{-1}) = H_r(q^{-1})R(q^{-1})$$  \hspace{1cm} (14)

The additive static disturbances lead generally to the increase of the control signal magnitude applied to the system exceeding admissible values. The actuator saturation can have an adverse effect upon the behaviour of the control signal and, in particular, when the controller contains an integrator [32]. Thus, the use of an anti-windup technique is mandatory.

Let's note by $u_{sat}^{inf}$ and $u_{sat}^{sup}$ the lower and higher limits of the control saturation, respectively. The effective control law applied to the plant becomes:

$$\overline{u}_k = \begin{cases} u_k & \text{if} \quad u_{sat}^{inf} \leq u_k \leq u_{sat}^{sup} \\ u_{sat}^{inf} & \text{if} \quad u_k > u_{sat}^{sup} \\ u_{sat}^{sup} & \text{if} \quad u_k < u_{sat}^{inf} \end{cases}$$  \hspace{1cm} (15)

It is also possible to impose a specified dynamics when the system leaves the saturation. The desired dynamics is defined by the polynomial $P_S(q^{-1})$ given by equation:

$$P_s(q^{-1}) = 1 - \exp\left(-\frac{T}{\tau_{sat}}\right) q^{-1}$$  \hspace{1cm} (16)

where $T_{sat}$ and $\tau_{sat}$ represent respectively the sampling period and the time constant of a first order system both chosen equal to 10 ms in the DC motor's case.

The new control law of the flatness-based polynomial controller in presence of saturation will be given by the following complete algorithm:

$$P_S(q^{-1})u_k = K(q)\varepsilon_k - \hat{R}(q^{-1})y_k - \left(\hat{S}(q^{-1}) + \exp\left(-\frac{T}{\tau_{sat}}\right)\right)\overline{u}_{k-1}$$  \hspace{1cm} (17)

The polynomial $\hat{S}(q^{-1})$ which introduces delays into the control signal, is given by:

$$\hat{S}(q^{-1}) = 1 - q^{-1}\hat{S}(q^{-1})$$  \hspace{1cm} (18)

Thus, this digital control algorithm takes into account the presence of a disturbance due to the load as well as to effects of saturation.

As in [32], the sensitivity functions which used to study the robustness of the designed RST controller are given by:

$$S_{yp}(q^{-1}) = \frac{A(q^{-1})\hat{S}(q^{-1})}{K(q^{-1})}$$  \hspace{1cm} (19)
\[ S_{ac}(q^{-1}) = -\frac{A(q^{-1}) \hat{R}(q^{-1})}{K(q^{-1})} \quad (20) \]

The designed control law is then applied to drive the DC motor shown on figure 2, which is a separated excitation motor with the following nominal features \(U_n = 180\text{ V} \), \(I_n = 2.1\text{ A}\) and \(\Omega_n = 3000\text{ rpm}\).

2.3. Process identification

The studied plant, constituted by the DC motor, the AC-DC power converter and the tachometer can be considered as a second order system given by the following continuous-time transfer function:

\[ H(s) = \frac{G}{(1 + \tau_m s)(1 + \tau_e s)} \quad (21) \]

where \(G\) is the static gain of the global plant, \(\tau_m\) and \(\tau_e\) are the mechanical and electrical constant times respectively, and \(s\) is the LAPLACE operator.

![The DC Motor Benchmark](image)

In order to identify these characteristic parameters, experimental measures have been achieved. The obtained numerical values are given in Table 1.

**Table 1. Model’s Parameters**

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Measured Numeric Values</th>
</tr>
</thead>
<tbody>
<tr>
<td>(G)</td>
<td>0.05 (V/V)</td>
</tr>
<tr>
<td>(\tau_m)</td>
<td>300 ms</td>
</tr>
<tr>
<td>(\tau_e)</td>
<td>14 ms</td>
</tr>
</tbody>
</table>

The discrete-time transfer function \(H(q)\), given by the equations (22), is obtained by sampling the continuous-time model (21). The sample period \(T_s = \frac{1}{f_s}\) is chosen equal to 10 ms as mentioned previously.

\[ H(q) = \frac{y_k}{u_k} = \frac{B(q)}{A(q)} \quad (22) \]

where \(A(q)\) and \(B(q)\) are co-prime polynomials.
3. Simulation Results

As in [32], the robustness of the designed RST controller can be guaranteed while observing figures 3 and 4. Indeed, the module of the output sensitivity functions $S_{yp}$ remains inside the predefined template and the one of the input sensitivity function $S_{up}$ presents attenuation in high frequencies.

![Figure 3. Output sensitivity function: case of static disturbance rejection](image)

![Figure 4. Input sensitivity function: case of high frequency noises elimination](image)

The implementation of the flatness-based RST controller is illustrated by simulation results shown in figure 5. In this application, the speed's DC motor varies from 0 rpm to 1000 rpm. The desired continuous-time flat trajectory $z^d(t)$ can be computed following the method described in [6]. One obtains:

$$z^d(t) = \begin{cases} 0 & \text{if } 0 \leq t \leq t_0 \\ P_1(t) & \text{if } t_0 \leq t \leq t_1 \\ \frac{1}{B(1)} & \text{if } t_1 \leq t \leq t_2 \\ P_2(t) & \text{if } t_2 \leq t \leq t_3 \\ 0 & \text{if } t \geq t_3 \end{cases}$$

(23)

where $t_0 = 10\, s$, $t_1 = 20\, s$, $t_2 = 50\, s$ and $t_3 = 60\, s$ are the instants of transitions, $B(1)$ the static gain between the flat output $z_k$ and the output signal $y_k$, and $P_1(t)$ and $P_2(t)$ are polynomials given by equation (24).

$$P_1(t) = 5.6844(t-t_0)^3 - 0.8526(t-t_0)^2 + 0.0341(t-t_0)^5$$

$$P_2(t) = -5.6844(t-t_2)^3 + 0.8526(t-t_2)^4 - 0.0341(t-t_2)^5 + 568.413$$

(24)

From these equations, the desired discrete-time trajectory $z_k^d$ can be obtained easily by sampling $z^d(t)$. The obtained polynomials of RST controller are numerically given by:

$$S(q) = 1 - 0.7102q^{-1} - 0.2025q^{-2} - 0.0873q^{-3}$$

$$R(q) = 242.3 - 83.72q^{-1} - 223.5q^{-2} + 102.6q^{-3}$$

$$P_3(q) = 1 - 0.1353q^{-1}$$

(25)

where the $k_i$ coefficients of the tracking $K(q)$ polynomial are chosen as:
The simulation results in figure 5, show that the effective speed of DC motor tracks a desired trajectory with high performance. The tracking error is very small in the transient regime and equal to zero in steady-state.

\[ K(q) = q^3 - 2.02q^2 + 1.313q - 0.259 \]  

(26)
After the simulation phase, the proposed digital control algorithm will be applied in the real-time framework.

4. Design Aspects

4.1. Digital control system hardware

Starting from the idea consisting in varying the supply voltage of the motor, several arguments led us to choose a structure of hardware control system based on an AC-DC one phase asymmetric rectifier \cite{17,22,23,24,26,27,29}.

Presently, the choice of such a converter rather than another solution, as DC-DC converter for example, is based essentially on the possibility to vary the supply motor voltage from 0V to 250V while directly using the one phase distribution network as alternative voltage source. Such an interval of variation is very sufficient to satisfy our needs since the machine to drive requires 180V as nominal supply voltage.

The present AC-DC converter, based on thyristors, represents only the power circuit of the developed hardware control system. For the thyristor's gate control circuit, a digital control based on a RISC microcontroller PIC16F876 produced by MICROCHIP, is used \cite{12}.

![Figure 7. Hardware Architecture of the Digital Control System](image)

This microcontroller will be the framework for the implantation of the designed control algorithm in order to generate gate impulses control signal for the thyristors. The choice of the PIC microcontroller is justified by the simplicity of implementation, using a flexible embedded programming language such as C/C++, in addition to a material architecture of peripherals allowing to satisfy integration constrains of control for an electrical system.

In order to complete the structure of the hardware control system of the separated excitation DC drive, two other hardware modules are designed: a circuit of acquisition and conditioning of the speed measurement, and a circuit providing the excitation of the DC motor. Thus, the complete architecture of the proposed hardware control system is given in Fig. 7.

4.2. AC-DC power converter control

This electronic control interface is designed and developed around an impulse transformer \cite{16,18,20,21,23,25,26,27} which allows to control by gate impulses, the BT152 thyristors used in
the bridge rectifier, while ensuring a galvanic insulation between the control circuit, based on the PIC microcontroller, and the power circuit. The use of the impulse transformer in this application, instead of optocouplers for example, is due to the possibility of working in the high frequencies and possibility to provide an important current able to fire the considered thyristors, as well as the good behaviour with high voltage and the simplicity of the use. The principle of the control circuit is given in Fig. 8.

Figure 8. AC-DC Power Converter Control

4.3. The microcontroller

In order to fire the BT152 thyristors, the used microcontroller PIC16F876 must provide the control signal as gate impulses [12-14].

The PIC microcontroller includes several variants. Indeed, the PIC16F876 is chosen because it has the necessary peripherals to implement the control algorithm. In addition, this type of microcontroller can operate with clock frequencies up to 20 MHz, improving performances of the digital controller in terms of speed signal processing.

The various peripherals considered in our application are given in the following section.

4.3.1. Timer1

Through registers T1CON, INTCON, PIE1 and TMR1 [12], the timer1 is configured to operate as a 16-bit counter which can increment from 0x0000 to 0xFFFF. With 4 MHz quartz and 1:8 prescale value, a \(8\,\mu\text{s}\) clock signal is obtained.

The content of the timer1, which is the value of registers TMR1H and TMR1L, represents the value of the thyristors firing angle. As the timer 1 will be launched at each occurrence of the external interrupt that finished when the AC supply voltage crosses the zero, every 10 ms, this time can count until 1250, corresponding to angle variation from 0° to 180°. Then, one must load the content of registers TMR1H and TMR1L by the value \((64285+\text{Alpha})\) where \(\text{Alpha}\) takes a numerical value between 0 and 1250. Thus, the timer1 will count from the value \((64285+\text{Alpha})\) to 65536 to represent the angle \(\alpha\) of thyristor firing.

An interruption, flag TMR1IF, will be set when timer1 overflows, in order to start the timer0.

4.3.2. Timer0

The role of the timer0 is to generate the gate impulses used for firing SCR after the delay given by the timer1. It is an 8-bit peripheral configured in a mode through the associated registers [12]. With the same 4 MHz quartz and with 1:2 prescale value, a 500 KHz count frequency is obtained.

After loading the content of the register TMR0 by the value 246 obtained by a simple calculation to
have one square signal of $40\,\mu s$ period as typical value, the timer0 generates the impulses after the firing delay, until the occurrence of the next external interrupt.

As RB4 pin is chosen to generate this control signal, initially forced to zero, the operation consists simply in an exclusive-OR with logic "1" of this PIC output signal every overflow of the timer0, in order to generate the gate-impulses.

4.3.3. Analog-to-digital converter

The Analog-to-Digital Converter (ADC) of PIC16F876 microcontroller has 5 inputs (AN0 ... AN4). The conversion of an analog input signal results in a corresponding 10-bit digital number. This peripheral is used to acquire the speed measurement of DC motor through tachometer. As all other peripherals, the initialization of the ADC is made by configuring the associated registers [12].

The developed electronic circuit is shown in Fig. 9.

![Figure 9. Developed Electronic Circuit](image)

5. Software Implementation

After the configuration of the PIC peripherals (timers, ADC converter, I/O interfaces) as well as the general validation of interrupts (GIE=1), the setting of a bit flag-Te relative to the sampling time, triggers the execution the following subroutines, in this order:

1) Subroutine of the set point acquirement
2) Subroutine of acquirement and treatment of the measure
3) Subroutine of the control computation
4) Subroutine of adaptation of the obtained control
5) Subroutine of application of the control signal

Indeed, the setting of this flag is made at each occurrence of the external interruption INT/RB0 (flag INTF=1), chosen as time basis for the whole application.
A graphic and normalized representation of such a built-in system of control is given by the flow chart in Fig. 10.

**Figure 10.** Flow Chart of the Control algorithm

6. Experimental Results

The real-time implementation with C/C++ embedded programming language of the flatness-based RST polynomial controller in PIC 16F876 target, led to the experimental results shown in figures 11, 12, 13 and 14. Here, it can be notice that experimental results given by figure 11 are comparable with those obtain by simulations, in figure 6, which valid the designed control system.

**Figure 11.** Experimental Response-DC Motor Speed  
**Figure 12.** Synchronization and Control Signals
As we will essentially be interested by the tracking trajectory in the transient regime, it is sufficient to consider solely the part of the response illustrated by figure 15.

While observing the figure 15, the superposition of the experimental response with the one simulated show the high performance in terms of tracking. Indeed, the tracking error is small in the transient regime and it is nearly equal to zero in the permanent regime.

7. Conclusions

An application of concepts, methods and specific tools to the design and the implementation of real-time systems to control processes has been considered in this paper. The feasibility of the real-time controller with a high performance in terms of tracking, is the principal contribution.

While basing on analytic procedures, the flatness control of the DC drive has led to a robust polynomial controller. However, the exploitation of performances in tracking a reference trajectory and in rejection of additive load disturbances as well as the simplicity of implementation represent arguments that justify the choice of such technique of advanced automatic control. The design of the RST polynomial controller using the flatness property of the plant, is guided by the choice of the tracking dynamics. The robustness of the developed controller with regards to modelling uncertainties as well as the intervention of disturbance and noises in the control, is guaranteed.

The implemented control algorithm led to satisfactory experimental results close to those obtained by simulation. It remains to illustrate effects of additive disturbance using validation experiment in real-time framework.


Emergent Dynamic Routing Using Intelligent Agents in Mobile Computing

Florin Leon, Mihai Horia Zaharia, Dan Gâlea

“Gh. Asachi” Technical University of Iaşi
Faculty of Computer Science and Engineering
e-mail: florinleon@gmail.com

Abstract: Routing begins to have an important place in the context of high performance distributed systems, with an increasingly notable influence on the overall performance of the system under analysis. While many global algorithms have been proposed for the routing problem, in this paper we demonstrate how a relatively simple agent-based approach, based on ideas inspired from complex systems and reinforcement learning, can generate a highly complex set of local interactions between individual agents, whose emergent behaviour results in the desired routing effect.

Keywords: emergence, routing, intelligent agents, mobile computing

Florin Leon holds a PhD in Computer Science and is presently a Lecturer in the Department of Computer Science and Engineering of the “Gh. Asachi” Technical University of Iaşi. His main research interests are: artificial intelligence, simulations using intelligent agents, and data mining. He is the author or co-author of 7 books and over 50 articles.

Mihai Horia Zaharia is an Associate Professor in the Department of Computer Science and Engineering of the “Gh. Asachi” Technical University of Iaşi and has a PhD in distributed systems. His main research areas are: distributed computing, distributed artificial intelligence, knowledge-based geographical systems, computer systems security, and computer architectures. He is the author or co-author of 9 books and over 40 articles.

Dan Gâlea is a Professor in the Department of Computer Science and Engineering of the “Gh. Asachi” Technical University of Iaşi, and a PhD Supervisor. His main research areas are: artificial intelligence, knowledge-based geographical systems, image processing, and natural language processing. He is the author or co-author of 10 books and over 100 articles.

1. Introduction

Nowadays, one of the important problems of information and communications technology is related to re-designing interactions between isolated computing elements in order to create large, powerful, and fault tolerant distributed systems. There are already many approaches that try to solve this problem, from e-government to virtual universities and organizations, peer to peer communities, GRID computing, and e-science computing. If the approaches offered by the traditional engineering cannot keep up with this convergence, those conglomerates of computing power can create many problems due to an inherent instability and mostly to their non-predictive behaviour if some limits are reached. The analysis of the natural complex systems leads us to alternative approaches. The examples under scrutiny refer to animal brain, immunity systems, ants colonies, and so on. All these examples have powerful, complex, aggregated behaviours obtained from the combination and interactions of simple elements or individuals. As a result, this kind of problem solving techniques begins to be used for various specific problems.

Distributed systems are usually analyzed from the points of view of load balancing, geographical distribution of the control, and databases. In most models, the communications are presumed to be very good and fault tolerant. Most of the time, these attributes are valuable but in some cases there are serious quality of service degradation at the communication line hot spots where bottlenecks appear due to the used routing algorithms. Therefore the routing problem begins to have equal importance with the previously mentioned basic characteristics of distributed systems.

In this paper a distributed framework based on intelligent agents was used to propose and test a new approach in routing for mobile devices. According to Wooldridge (2000), an agent is a computer system that is situated in its environment and is capable of autonomous action in order to meet its design objectives. Agent Oriented Programming, AOP, is a fairly new programming paradigm that supports a societal view of computation. In AOP, objects known as agents interact to achieve individual goals. Agents can exist in a structure as complex as a global internet or one as simple as a module of a common program. Agents can be autonomous entities, deciding their next step without the interference of a user, or they can be controllable, serving as an intermediary between the user and another agent.

Intelligent agents retain the properties of autonomous agents, and in addition show a so-called “flexible” behaviour (Wooldridge & Jennings, 1995): reactivity (the ability to perceive their environment, and respond in a timely fashion to changes that occur in it), pro-activeness (the ability to exhibit goal-directed behaviour by taking the initiative), and social ability (to interact with other agents and possibly humans).
2. Routing in Mobile Computing

Due to the inherent mobility of the nodes, the routing protocols at both the ad hoc and internet network levels are different from the classical approach. As a result, proactive schemes are used to continuously improve the routing tables for mobile nodes. Unfortunately, they use large amounts of available broadband and sometimes computing power. On the other hand, the on-demand routing protocols spend a lot of time for route discovery and they are sometime not suitable for real-time communication or processing.

There is a third approach which implies mobile agents. They can be used to establish an efficient route and also topology discovery, and increase the node connectivity (Marwaha, Tham & Srinivasan, 2005). The agent-based approach has some disadvantages related to the total node dependencies on knowledge provided by the agent. It is possible that routing can fail when the network topology is dynamic and the time to live for a route is short. In agent-based routing, the mobile nodes must wait until the routing agent provides the required routes to begin a communication. In some situations, it is possible that the nodes that carry the agents to be unexpectedly disconnected from the network. The causes may vary from a too long radio signal to entering the sleep mode or shut down. With any closed carrying node, the number of routing agents will decrease and so will the routing efficiency.

There are numerous ad-hoc routing algorithms that accept various conditions to connect in the network. They can be classified as proactive or reactive. As we previously mentioned, the proactive ones are more efficient but sometimes they can generate unjustified overload of computing and communication resources, and it may be an idea to create route only on demand like the other class of algorithms. The reactive approach can produce undesired delays with route establishment, especially when they cannot reach the destination and usually a proactive approach is preferred (Carrillo, Marzo, Vilà & Mantilla, 2004). Even if the proactive protocols offer a better image of the network when the topology is highly dynamic, the routing effort also creates overloads in communication and computation resources levels. Some proactive protocols are DSDV (Destination-Sequenced-Distance-Vector Routing Protocol, Perkins & Watson, 1994) and WRP (Wireless Routing Protocol, Murthy & Garcia-Luna-Aceves, 1995). They differ by the number of the routing tables, table content, and methods used to send the table updates between the nodes. Some reactive routing protocols are AODV (Ad Hoc-On Demand-Distance-Vector Routing, Das, Perkins & Royer, 2002) and DSR (Dynamic Source Routing, Johnson & Maltz, 2002). This protocols are source routing-based, where each packet knows the entire route to be followed or hop by hop-based, where each packet knows only the next jump in the net. As expected, some hybrid approaches were proposed to improve the performance and decrease disadvantages. The result seems to have a better scalability that their ancestors because the near nodes collaborate so the overload of route discovering is reduced. Typical examples are ZRP (Zone Routing Protocol, Hass & Pearlman, 1999) and DDR (Distributed Dynamic Routing, Nikaein, Laboid & Bonnet, 2000). Each type of protocols can use a plain or hierarchic routing approach.

3. The Model of the Distributed Framework

The mobile agents framework provides the end-user with the possibility to easily build a virtual cluster that fits its needs and also to find and use existing services. This virtual cluster guarantees to contain only the nodes that are capable of overall performance increase. The framework was entirely developed using the Java programming language. It provides support for developing any kind of agent, such as load balancing agents, security agents, communication agents, mobile or stationary agents, intelligent agents or state agents and so on. As for the communication protocol, the framework uses Remote Method Invocation, RMI.

In order to create a feasible infrastructure of agents, we must first consider some basic models for the entities involved in the system (Zaharia & Leon, 2006; Zaharia, Leon, Calistru & Gâlea, 2008).

3.1. The Task Model

In the proposed model a task is defined as a quadruple \( (d_c, d_r, t_m, c_m) \), where:

- \( d_c \): request size (characteristics of the code to be solved);
- \( d_r \): refers to the estimated size of the result. This is needed to compute the returning costs. Of course this is an estimation because there are situations when the user cannot accurately predict this factor (e.g. PSpice simulations);
• $t_m$: is the maximum time allowed for the task to be executed. This is required because some code error can drive to infinite loops, or the task has an undesired high complexity design. The latter situation can lead to unjustified higher loads at runtime especially for the NP classes of problems. As the result $t_m$ was conceived as an estimation of maximum execution time for the task (if the code is correct) over a computing node from the lowest speed class. As a result, uniformity in task handling is ensured no matter what the computing nodes power is. If the task crosses over this time limit, the task will be terminated, and the owner will be notified;

• $c_m$: represents the maximum cost that a user can afford so as to execute the task. The agents will use this cost to place the tasks on the servers. Those with a higher cost available will be placed on powerful nodes that have higher costs, and the others will be placed only on lower computing power but cheaper nodes.

3.2. The Server Model

In this approach we take into account the IP number of the server and also we associate three different queues for task receiving and manipulation for each server as follows:

• *An entry queue:* receives the task to be solved that comes from users. The user injects a new task in the system on a server of his/her choice. The task is entered into the entry queue of the server;

• *An execution queue:* stores the tasks selected, using cost analyses, to be executed on the server;

• *An output queue:* stores the results that must be returned to the users.

Because in heterogeneous distributed systems the computing node characteristics may vary throughout a large domain, we considered five performance classes. The MIPS factor was used as the parameter to scale servers into domains. As a result, we have the following categories:

- Server with a speed around 1000 MIPS - $V_S^1$, that is equivalent to a node with a Pentium II processor with 600 MHz working frequency;
- Server with a speed around 2000 MIPS - $V_S^2$, that is equivalent to a node with an AMD Athlon processor with 900 MHz working frequency;
- Server with a speed around 3000 MIPS - $V_S^3$, that is equivalent to a node with a Pentium III processor with 1.13 GHz working frequency;
- Server with a speed around 4000 MIPS - $V_S^4$, that is equivalent to a node with a Pentium IV processor with 2 GHz working frequency;
- Server with a speed around 5000 MIPS - $V_S^5$, that is equivalent to a node with a Pentium IV processor with 2.53 GHz working frequency.

Of course, these domains can be modified in accordance with the newest computer architecture available on the market but this is not relevant to the way that agents will interact; therefore no changes are needed for the analysis.

In fact only the price of the resource and the number of MIPS must be changed, because this computational price is related to the computer power.

As a result, we consider the following price classes:

- Price class $S_1$, related to speed category $V_S^1$, around 2.7¢/s;
- Price class $S_2$, related to speed category $V_S^2$, around 9¢/s (≈ 4.5¢/s/computational unit);
- Price class $S_3$, related to speed category $V_S^3$, around 24¢/s (≈ 8¢/s/computational unit);
- Price class $S_4$, related to speed category $V_S^4$, around 48¢/s (≈ 12¢/s/computational unit);
- Price class $S_5$, related to speed category $V_S^5$, around 85¢/s (≈ 17¢/s/computational unit).

The prices were chosen using a nonlinear manner because that conforms to an economic approach. An older machine has already reached its amortisation rates and there will also be fewer requests to be used because most of the users need more and more additional computer power in the shortest possible time.

It is important to note that these prices are only some general recommendations, and processors within the
same speed class can have different prices. It is the agents’ responsibility in this case to find the cheapest
server, whenever possible.

Most of the available machines on the Internet have operating systems capable of measuring their own
loads of all the resources. As a result, the agent framework will directly ask this information in order to
fulfil agent requirements. Due to the fact that we cannot predict which servers will be used in the dynamic
cluster created by the agents to solve one request, it will not be optimal for the servers to share
information about their states because this will mean all-to-all continuous communications.

3.3 The Connection Model

Another aspect of the heterogeneous distributed systems is related to various available broadband at
different communication levels. In this approach we used only three types of network connections:

- Connection class $VC_1$, with a 10 Mb/s broadband;
- Connection class $VC_2$, with a 100 Mb/s broadband;
- Connection class $VC_3$, with a 3200 Mb/s broadband.

Due to the discrete and geographically distributed nature of heterogeneous distributed systems we must
also analyze the inter-node connection cost. In developed societies these costs begin to be insignificant
over small and medium clusters. However, these costs must be used in the formula of task execution costs
especially when the task is time consuming and the amount of returned data is higher. For our model we
consider the following cost domains:

- Price category $PC_1$, corresponding to speed class $VC_1$, around 1 ¢/MB;
- Price category $PC_2$, corresponding to speed class $VC_2$, around 1.5 ¢/MB;
- Price category $PC_3$, corresponding to speed class $VC_3$, around 5 ¢/MB.

As a result, only the information about local load, speed price class, and communication price class is the
most interesting about any node in the network.

4. The Emergent Routing Algorithm

Most routing solutions are dedicated to obtaining a better performance. Unfortunately, if one takes into
account other criteria, such as the communicational price, additional constraints need to be considered,
and therefore several opposite criteria need to be optimized simultaneously. This situation can be
encountered especially in the case of mobile computing, when a system can be simultaneously connected
to more than one network. Therefore, the routing decision cannot be made only by a mono-criterial
analysis. Specific algorithms such as Mobile IP (Perkins, 2002; Johnson, Perkins & Arkko, 2004) are
based on the idea of finding the “way home”, i.e. wherever he is, the user can connect first to the initial
provider, and solve the problems from there. If we analyse this method through the associated costs, one
may want to avoid this method, by using similar entities in the neighbourhood.

In our present approach, routing was obtained in an implicit, not explicit, fashion by the use of a
dedicated algorithm, as the result of the aggregation of the local behaviours of individual agents. In order
to do that each server will maintain a list of agent probabilities to be moved from a server to its
neighbours. This list has the means of strengthening or weakening its values based on reinforcement
learning, and are therefore dynamically modified on each transport. Those values exist only for active
neighbours. If a server is down its transport probability automatically become zero for all neighbours. If a
new server enters the network, its transport probabilities are automatically initialized by its neighbours
with random little values which will gradually increase with the time and the use of its services.

The main idea in transport probabilities updates is depicted in figure 1. When an agent follows a path $A-
B-C$ and passes from server $C$ on server $D$, at each transport it memorizes the required time to be
transported on the next server. On $D$ server it will update the reinforcement values that are equivalent to
transport probabilities corresponding to previously visited servers. This procedure is based on the fact that
a connection is considered symmetric: the time needed to transport from server $A$ to server $B$ is equal to
the time needed to transport the agent from server $B$ to server $A$.  

To compute the reward, \( r \in [0,1] \), the following equation is used:

\[
r = \frac{Minimum\ Transport\ Time}{dt},
\]

(1)

where \( Minimum\ Transport\ Time \) is a system constant and \( dt \) is the time required by an agent to move between two servers. In the implementation of the proposed system a \( Minimum\ Transport\ Time \) of \( 2.8 \cdot 10^{-6} \) s was considered.

The equation for updating the probabilities is derived from the reinforcement learning paradigm, as follows:

\[
P = P + r \cdot (1 - P).
\]

(2)

In figure 1, the updated rewards on server \( D \) are presented in the following formulas:

\[
r_{D-C} = \frac{2.8 \cdot 10^{-6}}{0.36} = 7.78 \cdot 10^{-6}
\]

(3)

\[
r_{D-C-B} = \frac{2.8 \cdot 10^{-6}}{0.42} = 6.67 \cdot 10^{-6}
\]

(4)

\[
r_{D-C-A} = \frac{2.8 \cdot 10^{-6}}{0.75} = 3.73 \cdot 10^{-6}
\]

(5)

For example, the probabilities from the reinforcement list of server \( C \) will be updated by the agents that come back from \( D \) or by the agents sent from it.

To analyze the routing process using the probabilities of the reinforcement list, let us consider an alternative connection between servers \( A \) and \( D \), as presented in figure 2.

If an agent comes using this connection, it will receive the following reward:

\[
r_{D-A} = \frac{2.8 \cdot 10^{-6}}{0.4} = 7 \cdot 10^{-6}
\]

(6)
Let us consider the simplest case when only two agents reach server D, one of them using the direct route and the other using the alternative route. We assume that the initial probabilities on D are zero. In order to make a decision to move from D to A, if we normalize the probabilities, we will have the situation presented in figure 3. The decision is stochastically made with determined probabilities, and thus the agent will use the direct route with a probability of 65.24% or will follow the alternative route, through C, with a probability of 34.76%.

As more agents will continue to come and follow the direct route, the associated probability will increase, therefore the agents from server D will choose the direct and quickest route increasingly often.

![Figure 3. Route selection probabilities](image)

5. Case Study

We will consider a scenario that consists of a system with 3 different servers, 5 initial agents, and 10 tasks to be deployed and executed in the system. Its structure is presented in figure 4. Servers are represented using ellipsoids that contain information about naming, IP, speed, and cost. The connections are arcs, and the associated speed and costs for the communication channels are displayed within the rectangles.

![Figure 4. The structure of the distributed system](image)

Below we show the way agents execute in this particular configuration.

**Configuration**

**Servers**
- Server 1: IP=192.168.0.10 Speed=2979 MIPS Cost=22.733 c/s
- Server 2: IP=192.168.0.11 Speed=1005 MIPS Cost=2.431 c/s
- Server 3: IP=192.168.0.12 Speed=5019 MIPS Cost=84.264 c/s

**Agents**
- Agents 1 starts on 192.168.0.10
- Agents 2 starts on 192.168.0.10
- Agents 3 starts on 192.168.0.11
- Agents 4 starts on 192.168.0.11
- Agents 5 starts on 192.168.0.12

**Connections**
- Connection 192.168.0.10 - 192.168.0.11: Speed=10 Mb/s Cost=1.099 c/MB
- Connection 192.168.0.11 - 192.168.0.12: Speed=101 Mb/s Cost=1.506 c/MB
Starting framework

Running agent 2 on server 192.168.0.11
Agent 2 takes task 9 (1854KB, 220KB, 204s, 34.88$)

Running agent 3 on server 192.168.0.11

Running agent 4 on server 192.168.0.12
Agent 4 takes task 0 (354KB, 227KB, 7s, 0.21$)

Running agent 1 on server 192.168.0.10
Agent 1 takes task 1 (396KB, 247KB, 78s, 2.41$)

Running agent 0 on server 192.168.0.10
Agent 0 takes task 3 (1082KB, 274KB, 159s, 19.23$)

Running agent 1 on server 192.168.0.10
Agent 1 waiting for peer answers on server 192.168.0.10

Running agent 0 on server 192.168.0.10
Agent 0 waiting for peer answers on server 192.168.0.10

Running agent 4 on server 192.168.0.12
Agent 4 waiting for peer answers on server 192.168.0.12

Running agent 3 on server 192.168.0.11
Agent 3 moving randomly to server 192.168.0.12

Running agent 2 on server 192.168.0.11
Agent 2 waiting for peer answers on server 192.168.0.11

Running agent 4 on server 192.168.0.12
Agent 4 tries to solve task 0 (354KB, 227KB, 7s, 0.21$) on 192.168.0.11
Task 0 (354KB, 227KB, 7s, 0.21$) to be sent from 2 to 1 by 1

Running agent 3 on server 192.168.0.12
Agent 3 waiting for scouts on server 192.168.0.12

Running agent 1 on server 192.168.0.10
Agent 1 tries to solve task 1 (396KB, 247KB, 78s, 2.41$) on 192.168.0.11
Task 1 (396KB, 247KB, 78s, 2.41$) to be sent from 0 to 1 by 1

Running agent 2 on server 192.168.0.11
Agent 2 tries to solve task 9 (1854KB, 220KB, 204s, 34.88$) on 192.168.0.12
Task 9 (1854KB, 220KB, 204s, 34.88$) to be sent from 1 to 2 by 2
Server 192.168.0.10 begins solving task 3 (1082KB, 274KB, 159s, 19.23$)
Link 1-2 sending task 0 (354KB, 227KB, 7s, 0.21$) to server 192.168.0.11. Cost 0.23c Time 62.954ms
Task 0 (354KB, 227KB, 7s, 0.21$) to be solved on server 192.168.0.11
Link 1-2 sending task 9 (1854KB, 220KB, 204s, 34.88$) to server 192.168.0.12. Cost 1.20c Time 329.709ms
Task 9 (1854KB, 220KB, 204s, 34.88$) to be solved on server 192.168.0.12
Link 0-1 sending task 1 (396KB, 247KB, 78s, 2.41$) to server 192.168.0.11. Cost 0.35c Time 359.610ms
Task 1 (396KB, 247KB, 78s, 2.41$) to be solved on server 192.168.0.11

Running agent 2 on server 192.168.0.10
Agent 2 tries to solve task 8 (133KB, 288KB, 35s, 4.23$)
Agent 2 tries to solve task 6 (133KB, 288KB, 35s, 4.23$) on 192.168.0.12
Task 8 (133KB, 288KB, 35s, 4.23$) to be sent from 0 to 2 by 1
Server 192.168.0.11 continues solving task 7 (906KB, 1370KB, 113s, 3.50$)
Server 192.168.0.12 continues solving task 4 (1878KB, 358KB, 252s, 43.09$)
Server 192.168.0.12: task 4 (1878KB, 358KB, 252s, 43.09$) was successfully solved. Cost 29.74$ / 43.09$

Link 0-1 sending task 8 (133KB, 288KB, 35s, 4.23$) to server 192.168.0.11. Cost 0.26c Time 261.535ms
Task 8 (133KB, 288KB, 35s, 4.23$) to be sent from 1 to 2 by 2
Link 1-2 sending task 8 (133KB, 288KB, 35s, 4.23$) to server 192.168.0.12. Cost 0.19c Time 25.608ms
Task 8 (133KB, 288KB, 35s, 4.23$) returned home on server 192.168.0.12

Running agent 4 on server 192.168.0.10
Agent 4 waiting for scouts on server 192.168.0.10

Running agent 3 on server 192.168.0.10
Agent 3 moving randomly to server 192.168.0.11

Running agent 0 on server 192.168.0.10
Agent 0 moving randomly to server 192.168.0.11
Link 0-1 sending task 6 (390KB, 253KB, 174s, 5.39$) to server 192.168.0.10. Cost 0.22c Time 229.751ms
Task 6 (390KB, 253KB, 174s, 5.39$) returned home on server 192.168.0.10

Running agent 3 on server 192.168.0.12

Running agent 1 on server 192.168.0.11
Agent 1 moving randomly to server 192.168.0.10

Running agent 4 on server 192.168.0.10
Agent 4 moving randomly to server 192.168.0.11

Running agent 2 on server 192.168.0.11

Running agent 0 on server 192.168.0.11
Report: Server 192.168.0.10: task 6 (390KB, 253KB, 174s, 5.39$) was successfully solved. Cost 3.26$ / 5.39$

Finished

The statistics referring to task execution are presented in table 1.

Table 1. Task execution

<table>
<thead>
<tr>
<th>Task ID</th>
<th>Origin server</th>
<th>Executing server</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>9</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>3</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>1</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>8</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>4</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>7</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>5</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>6</td>
<td>0</td>
<td>1</td>
</tr>
</tbody>
</table>

In table 2 the number of tasks that originate from a server and are executed on other server in the system is presented.
Table 2. Statistics regarding task execution

<table>
<thead>
<tr>
<th>Origin</th>
<th>Execution</th>
<th>0 (192.168.0.10)</th>
<th>1 (192.168.0.11)</th>
<th>2 (192.168.0.12)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0 (192.168.0.10)</td>
<td></td>
<td>1</td>
<td>3</td>
<td>1</td>
</tr>
<tr>
<td>1 (192.168.0.11)</td>
<td></td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>2 (192.168.0.12)</td>
<td></td>
<td>1</td>
<td>3</td>
<td>0</td>
</tr>
</tbody>
</table>

6. Conclusions

In this paper we described a system of autonomous agents, in the sense that each agent decides its actions based on its internal state and the state of the environment, without an explicit external command, other than the task to be executed, which is the actual goal of the agent. However, the autonomy is not incompatible to the global order. The system manifests a form of self-organization, resulting in the routing behaviour of the multiagent system as a whole. The proposed emergent routing algorithm has a dual effect. An agent discovers the routes but also uses the previous information about the routes stored by the previous agents that have already travelled through the local neighbourhood of a server. Since the agent-based approaches are scalable and robust by nature, the research can be continued in the future by considering a larger number of servers and agents, which will allow the study of the real life problems on massive infrastructures, such as the GSM internet servers.

REFERENCES


Decision Making Process: A Collaborative Perspective

Pascale Zaraté
Université de Toulouse, INPT-IRIT
118 route de Narbonne, 31062 Toulouse Cedex 9, France
Pascale.Zarate@irit.fr

Abstract: The introduction of Information and Communication Technologies (ICT) in organisations implies an evolution of the decision making processes. This evolution is analysed on a cognitive and organisational point of view. These processes are analysed and described. The necessity to cooperate is then shown and we finally underline the fact that new systems must be designed in order to support the decision making process.

Keywords: Decision Making, Decision Support Systems, Cooperative Decision Support Systems.

1. Introduction

[1] has shown that the Information and Communication Technologies (ICT) introduction in organisations obligatory leads to a time fragmentation. He also shows that the information overload is both a source and a consequence of this overload.

Nevertheless, with the great development of the ICT, the Decision Support Systems become partially usable: only when the decision makers meet each other. The collaborative work among several actors involved in a decisional process is reinforced. It becomes then essential to model the cooperation for the collaborative work.

Due to the economic evolution, the decisional processes in organisations present a great evolution. The reactivity obligation for companies and the technological evolution face the consequences that the cognitive and organisational decisional processes have been considerably modified.

2. Organisational Decision Making Process

The organisational decisional processes evolve and more actors participate to the decision: the responsibility and the initiative are more distributed.

On the other side, it is more and more necessary to have a report of activities. A great part of the managers’ activity aims to imply, motivate a great number of actors concerned by the problem to solve.

Even when the decision is organisationally made by one person, it is rather always prepared by a collaborative work. The cooperative processes increase and contribute more and more to decisions to make. Two fundamental conditions of the cooperation are facilitated by this evolution:

- The interactions are intensified, then the possibilities to establish cooperative behaviours increase and these cooperative behaviours effectively increase (see [2]; [3]).
- This exchange increase leads to better know the other actors’ objectives, which also are a preponderant issue to establish a cooperative behaviour (see [4]).

The organisational context is very present for decision making. It is fully part of the individual decisions through the actor involvement but also of the collective decisions. The organisational decisions by definition build up the organisation and at the same time determine its action. In other words, the decision making is often diffuse in the organisational and collaborative activity. The decision is
made step by step, circulating in several circles of exchanges among actors. When a decision is made, by a kind of way, an organisational change of state is operated.

For the collective activity, the way of decision making also is the place of negotiations among involved actors. In this exchange among actors called by [5] problem setting / problem solving, for which the problem is set until its easy solving; there is negotiation and confrontation among actors on the way how to set the problem.

This is indeed a strategic action from all points of view that gives several problem settings, especially the contextual elements having strong relationships with the problem definition and for the consequences of several choices.

All these issues allow us not only to reconsider our model of decision making, but also to design supports in another way underlining three points: information overabundance, innovation and design processes must be reinforced, and finally the choice step of the decision making cognitive process must be made relative to the introduction of negotiation processes or other collective processes.

### 3. Cognitive Decision Making Process

The actor in the organisation is in front of a new situation characterised by a large access to information even information overabundance; a great participation to several processes, meetings of transversal groups. He then must face the prospect of shortened deadlines of decision making and a set of situations that often conduct to the cognitive overflow syndrome (see [6]).

Among the cognitive processes, that become very important for managing the new situation, we find the efficiency selection and the vigilance that it implies. The need to see the efficiency implies a great involvement of a person in her task but also in the organisation. It is indeed obvious that for an actor, rapidly feeling what is efficient for the organisation, it implies a shrewd and pregnant consciousness of the organisation interests, needs and perspectives. The more pertinent is a fact, the lower the cognitive cost. The vigilance is a willing attitude of the actor that allows him to improve his chances to perceive pertinent elements in their environment.

In order to perceive the information pertinence and to take the opportunities that they could represent for the organisation, the actor must fulfil a vigilance function.

Based on these initial issues, we propose to revisit the decision making cognitive process proposed by Simon, having for hypothesis that the reflexion process of an actor is less than 10^th seconds [5]. An abstract of the previous description would be that the environment evolving very rapidly, each actor is led to sharpen its own intelligence and vigilance processes. Among the several steps distinguished by [5] (figure 1), the intelligence step becomes more active and complex, because the environment that must be taken into account in this step is more complex. The way to apprehend the environment also is modified; the actor has a pre-eminent role of efficiency research. Instead of research information, in order to not forget important pieces of information, he must now operate a very rapid outranking in a plethora of information. The design step also becomes more frequent, because each environment apprehension implies to measure the efficiency.

The choice step is fundamentally not modified because there really is not a generation of similar alternatives, a systematic comparison among them and a selection of one of them by rational evaluation process reasoning, in the very rapid outranking process described behind.

![Figure 1. The revisited decision-making process of [5]](image-url)
The cognitive iterative process of decision making is then modified (figure 1): the two first steps are more often visited that the third one, the return loop from the choice to the intelligence, but also from the design to the intelligence are reinforced. Several iterations are necessary before the choice step (for more details on this process see [7]).

The evolution of the organisational and cognitive processes are strongly connected and operate in a simultaneously, causal and imbricate way.

4. Cooperation Necessity

In organisations, the great majority of decisions are made after an intensive consultation of several actors but not by individual decision makers working in larger organisations (see [8]). [8] have shown that the more complex organisations are, the less decisions are made by individual actors. For [9] the decision making processes in organisations generally imply several actors interacting each others. This interaction implies the information communication and a shared understanding by decision makers involved in these processes. They analyse this interaction along three issues: the used knowledge meaning must be common to every one; an artificial or not agent must have the necessarily authority for regulating the workloads; then the users must have enough confidence in the used technologies, that eventually could be seen through several visualisations of shared knowledge.

The participants to a decision making process must join their efforts in order to have a common goal where they must integrate multiple points of view not necessarily in harmony each others. They must work together not necessarily at the same place and at the same time. They are engaged in a coordination effort in order to solve the problem for which they must divide the decision making in several sub-tasks which will be assigned to individual participants.

5. Groups’ Evolution

Several authors analysed the collective decisional process along several issues without really introduce the context notion to take into account in the design of tools able to support the decision makers. We noticed that these analyses are established in several communities and more particularly those of the social psychology. In this last field, [10] has shown that the group behaviour is difficult to establish a generic analyse. The work context also must be taken into account. He also shows that it is necessarily to understand the nature of the task assigned to the group and its characteristics.

[11] analyses the collective decisional process under five dimensions: the group structure, the group roles, the group processes, the group style and the group norms. This analysis is for us interesting because it proposes a definition of the context of the group work.

The group structure is defined by several structures linked to the number of persons and the type of the group: individual, team with a hierarchy notion, and committee with a necessity of consensus among the group members, group less structured.

The group roles are assigned to one or several persons. Each member of the group can play one or several roles, which are easy to implement in a Decision Support Systems (user, analyst…).

The group processes used for the decision making could considerably influence it. If the information or decision making flows circulating are generally well-known, and could be analysed as workflows for example; the way how the process must emerge is generally not explicit et it could influence the decision in one or in another way. Must a consensus be raised? If yes, is there a deadline?

The group style is composed by the decision makers’ styles that could influence the process, its behaviour under several conditions and the decision making outputs quality.

The group norms are certainly the most important dimension. The social psychology of the decision making is in that sense very important.

It is necessary to be careful to use a common sense among participants; towards the collective or individual social pressures; towards the gender (balance male, female) and to the prescriptions about the group behaviour, the personal believes, the potential sanctions that constitute the decision making environment.

Following [12] the shared knowledge building among the group members dynamically arises and is strongly connected to a context proceduralisation, which is an activation of a general context part of
the task implemented by interactions among individuals. Otherwise [13] defined a different context granularity from the decision making or task focus. He shows that the decision making context is dynamic and changes in accordance with the point of view adopted for the group observation: the group in itself or several individuals or the project context.

[13] published a case study on the publication of daily and weekly periodicals in Ireland. They have shown that the Information and Communication Technology (ICT) introduction in this enterprise has considerably modified the production process improving the enterprise effectiveness.

We analysed this case study through the five dimensions defined by [11].

The group structure is more diffuse, the group is generally constituted in an ad hoc way without a real formalisation of its structure. The group doesn’t know if it is structured as a team, a committee or in a hierarchical way.

The group style includes several styles coming from several groups constituting the decision making group; the group established only for a precise need, the members have not enough time for establishing a new style to the group. The group style is generally built through a long period.

The norms and the processes of the group become from several different groups and are then mixed one to each others.

The group roles include several roles coming from different groups, a group member will then create his own role that could be a mixed of several roles played in the past in other groups.

In other words, we could assume that the context dynamic of decision making influenced by the introduction of ICT is more reactive and volatile. The cognitive load of decision makers is more important because they share more information and they must remember more information (for more details see [14]).


In order to support in a proper way this decisional process, new tools have to be designed. These systems are generally defined by several authors as frameworks integrating several tools. The main point of these frameworks is that the system must be able to support dynamically decision-makers by proposing an “Intelligent” assignment of tasks among the involved actors: decision-makers and software seen as agents.

[15] has shown that Cooperative Decision Support Systems must be developed. She defined these tools as frameworks composed by several tools. Therefore, we propose a Cooperative Decision Support framework. This framework is composed by several packages:

1. an interpersonal communication management system,
2. a task management system,
3. a knowledge management tool,
4. a dynamical man/machine interactions management tool (for more details see [15]).

7. Conclusion

According to [16] electronic communication has some effects on the group behaviour and she concludes by:

- An electronic meeting can be used as an ante-meeting to gather information and solicit opinions before a decision to be made face-to-face;

- Face-to-face decision making probably is the best when a decision requires complex thinking and subtle multiparty negotiations, and when problems are ill-defined (p. 191).

[17] concludes his study by assuming that: the study of the effectiveness of decision making could be studied, in order to highlight that too much technologies could generate negative effects, in particular the decline of consensus and personal participants implication.
All these effects led us to think that a methodology of Collective DSS use has also to be defined. [17] concludes his study by assuming that face to face meeting must be planned during all along the group work, with a sufficient number, by using classical synchronous electronic meeting systems.

We are convinced by the fact that meeting points have to be introduced all along the group decision making process. This process has to be managed as a classical project and methodologies coming from project management domain could also be used. The meeting points have the same utility than the milestones in project management and offer the advantages of the face to face situation.

Our perspective of this work is to develop a methodology of Management for Collective Decision Making Processes supported by Cooperative Decision Support Systems.

REFERENCES


Computational Environment Generation for Computer Algebra Systems

Svetlana Cojocaru

Institute of Mathematics and Computer Science, Academy of Sciences of Moldova
Str. Academiei, 5, Chişinău, MD 2028, Republic of Moldova

Abstract: This paper deals with the problem of user interface development for Computer Algebra Systems. An algorithm of computational environment generation based on inferences from users’ actions is proposed.

Keywords: intelligent interface, computer algebra systems, adaptation

Svetlana Cojocaru – Dr. h., is the deputy director of the Institute of Mathematics and Computer Science, Academy of Science of Moldova. S. Cojocaru has published 3 books, 5 chapters in books, more than 100 papers. Her research interests include natural language processing, computer algebra, biological computing.

1. Introduction

The software complexity is permanently increasing and it requires from the user more vast and deep knowledge about systems functioning. Interfaces, which realize the human-computer interaction, take upon themselves to itself also a part of user’s tasks. But in the process of interface developing one can observe the same tendency as in general software evolution – they also become more and more complicated and insufficiently flexible. A solution of this problem might be intelligent interfaces, which are able to adapt themselves on user’s necessities, can learn, can take over the initiative in their interaction with user, are able to guide him in order to facilitate his objectives achieving in a fast and comfortable way.

Interface elaboration for Computer Algebra Systems (CAS) was and remains a topic to be investigated. There are a number of overviews treating this subject, among them we can mention (Kajler 1998, Kajler, Soiffer, 1998, Grabmeier, Kaltofen, Weispfenning, 2003). There is also a large list of concrete CAS descriptions, the corresponding interface (in a rudimentary or developed form) being one of the presented components.

Analysing the CAS interfaces one can conclude that they respect the requests specific for intelligent interfaces (Ross, 2000, Filip, 2007, Waern, 1997) in a reduced measure only:

• The user is guided in the process of his problem solution,
• It is not necessary to study preliminarily how the system works, often it is possible to learn the functioning mode directly in the process of interaction with the system,
• The system makes some errors prevention (partially) by checking the problem’s formulation correctness and verification of initial data.

Less significant is the attention paid to user modelling. Because CAS are from the very beginning oriented on a pre-established users class, their model (in relation to the class of problems, data type, non-ambiguous interpreting) is beforehand included into the system.

Therefore in the case of CAS it would be more relevant to speak about user’s personalisation. We should mention that the problem of personalisation, being realised, is not sufficiently examined. User can apply some possibilities of individualisation, but he should adapt the system to his needs or preferences by himself.

We will note also that the possibility to interact in natural language in the frame of CAS is rather limited.

Thus, in order to increase the intelligence level of CAS interfaces one can fix the necessity to implement some additional features, including user’s personalisation, adaptation to his individual requests and preferences. One should note, that the user not always is able to formulate his preferences.
In the next sections the process of adaptation in CAS interfaces will be examined. We will respect the following principles:

- The goal of adaptive systems is to facilitate user’s objectives achieving in the fastest and easiest way, assuring, as much as possible, a high degree of satisfaction (Ross, 2000).
- It is important that system shall not irritate the user, not reduce by its interventions the speed of his work, it shall not initiate unbidden operations in some inappropriate moments, creating in such a way obstacles to achieve the goal immediately.

2. Adaptation to the User

The more widely is used a software product the more difficult is to elaborate a comfortable environment, which will be able to assure a high performance for every user. A series of studies performed since 1980 had been finished in conviction that there is no possibility to specify all varieties of users, programs and equipment in order to work out interfaces ideally convenient for each case. The only resource comparable with the number of users are the users themselves, therefore namely they should be involved in the process of the personalised interfaces developing (Weld, Anderson, Domingos, 2003).

Analysis of many CAS interfaces convinced us that even in the case of the most developed ones users have some wishes, which demand an adaptation to their interests and individual habits (Cojocaru, Malahova, Colesnicov, 2006). As it was mentioned above, adaptation to user is one of the distinctive, indispensable features of intelligent interfaces. We will describe the process of its implementation in general case and the specific for CAS in particular.

Let us recall that the adaptation process might be realised in the following three modes:

- Adaptation is performed in advance, by a design oriented to a specific type of users.
- The system permits a dynamic adaptation, which is performed by user in the process of his interaction with the system.
- The adaptation is made by system in the process of its interaction with the user.

In the last case the adaptation is made through a dialogue initiated by the system or through “spying” user’s actions and making adjustments to his preferences basing on inferences from the collected knowledge.

In some sense we can consider that all CAS are from the very beginning oriented on a fixed class of problems and a determined group of users, thus the last two methods of adaptation will be the subject of our interest. The case, when the system permits a dynamic adaptation realised according to user’s request (and with his contribution) is frequently implemented in modern software and is called “customization”. In (Weld, Anderson, Domingos, 2003) the following ways of implementing of that adaptation method are underlined:

- Visible menu selection,
- Adding buttons in menu bar,
- Definition of macros,
- Using specialised programming languages to write own scripts.

Of course, only a limited number of users poses the necessary abilities and have sufficient time (and patience) to apply this methods (especially, the last two), therefore, keeping in mind all these possibilities, we will deal with dynamic adaptation initiated by the system. Our approach will be explained basing on the example of Computer Algebra system Bergman (www.servus.math.su.se/bergman).

3. Interface of Computer Algebra System Bergman

Bergman is a computer algebra system calculating non-commutative and commutative Gröbner basis, Hilbert series, Anick's resolution, and Betti numbers for ideals over rings and for modules. Bergman allows a multitude of choices for various kinds of input, operation, output, etc. One may let predefined
“top of the top” procedures make the choices for him, or specify details by means of a large number of lower level mode setting procedures.

To perform some computations in Bergman you should set up the polynomial ring and its surrounding. This action includes selection of the algebra type (commutative or non-commutative), the variables, the ordering, the characteristic, type of coefficients, and variable weights. One can also select the strategy of computation, and input-output mode. There are some minor mode selections too. For a complete description see (Backelin, Cojocaru, Ufnarovski, 2005).

Note, that different kinds of calculations demand different parameters settings, they should be compatible with each other and compatible with the corresponding calculation procedure. Because Bergman has a high degree of flexibility in modes selection, sometimes it might be difficult for user to make a correct and efficient selection. These problems are avoided by developing of graphical user interface, which permits to operate with Bergman in a manner close to the usual mathematical surroundings.

The main principles respected in the process of interface design correspond both to common requests and to those specific for intelligent interfaces, namely:

- the interface reflects specific mathematical object and notions,
- it is possible to use the system without preliminary studying of its description,
- some tasks are taken over from the user and executed by the system,
- the interface simplifies access to different Bergman’s functions,
- the interface operates the data control before calculations,
- input and output can be performed in bi-dimensional format,
- it is possible to save the environment in which the computations were performed and to repeat the calculations in the same conditions, with the same or modified data.

Moreover, the interface can serve also as a prospectus for Bergman giving the possibility to see on the graphical panel the most important facilities without reading a manual or applying the corresponding help function.

The shell consists of the main menu at its top, then the toolbar, and then three panels (top-left, bottom-left, and the right one). The main menu consists of items that are used to organize calculations, and the top-left panel contains drop-down menus and other elements that specify mathematical parameters of the problem to be solved. The bottom-left and right panels display information, or are used to enter some data (e.g., variables and relations are input from the right panel). Main menu items are those ordinary (File, Tools, Options, Windows, Run etc.), but there are two specific ones named Sessions and Environment. These items are less obvious and we shall explain them.

Our observations on CAS using permit to conclude that most of the mathematicians deal for a more or less long period with research of a certain class of problems, which need calculations in a specific mathematical environment, determined by a series of parameters. These parameters remain mostly unchanged during several computation sessions, only a few of them may be modified (for example, in the case of Bergman many calculations are performed only by variations related to input polynomials or modules). We introduced the notion of Environment just to establish this repeatable surrounding. It fixes the preferable setup immediately after starting Bergman. It is possible to have a number of environments for each user, depending on the class of problems one is interested to solve. Moreover, it is possible to make some changes in the environment, permanent or valid during the current session. The same environment might be used by different users if they solve the problems belonging to the same class.

A possibility with a higher degree of determinism is that realized by the mechanism of sessions. A session is a set of parameters that fully defines the problem to be solved. Session is implemented as a directory where all data are saved, Bergman input files are generated, and Bergman output files are produced. Sessions serve to return to previous Bergman calculations, modify them, and experiment with them. Informally, sessions give to a mathematician the possibility to use the previous experience of Bergman's users (own or others) and to save the current setup for the future calculations.

When the user wants to create a new session, he will be asked which environment should be used as the
base to create this session. An environment is a partial set of data common for several sessions. It corresponds to the group of mathematical problems the user investigates during different sessions. E.g., after Bergman’s installation the environment directory contains a profile called "commutative" that fixes a single parameter, the commutativity. There is also a pre-defined environment containing the default settings. For example, by default, if no other settings are made, Bergman will calculate Gröbner basis in the commutative ring with characteristics equal to zero.

4. Environment Generation

The menu items Environment and Session serve to implement the mechanism of adaptation. There are three methods implementing environment creation in Bergman.

The first one is based on a dialogue operated by the system. It puts a number of questions, collecting and analyzing the answers so that it becomes possible to fix necessary data permitting to generate the computational environment. Each question is followed by a list of possible answers. A fragment of such a dialogue is presented below:

> (dialogue)
What kinds of objects do you plan to work with?
- One of the following: grc, ngrc, grnc, ngrnc
> grc
The suggested answers mean correspondingly "graded commutative", "non-graded commutative" etc. Obviously, this method can be used only by persons having some experience in work with the system.

The second method is based on session editing; it might be the current session or any of the saved ones. The system displays the list of parameters with corresponding values and proposes to select the wanted ones. The selected values can be edited if necessary.

The third method is based on inferences from user’s actions. The series of previously saved sessions are analyzed and on their base a new environment is created. In adaptation based on inferences, as a rule, one of two possible criteria is used: that of recent actions or that of frequent actions (Gajos et al., 2006). We propose an heuristic algorithm which will take in account both criteria. It is necessary to mention that solving of this problem for CAS has its peculiarities: the parameters’ values, selected from different sessions according to the frequency criterion might be contradictory.

Let $A_1, \ldots, A_n$ be a sequence of attributes associated with Computer Algebra System $S$. Each attribute $A_i$ has a corresponding set of values $V_i = \{v_{i1}, \ldots, v_{in}\}$ ($i = 1, \ldots, n; k = 1, \ldots, r$). These values form a multidimensional array

$$F_S = [v_{11}, v_{12}, \ldots, v_{1n}] X [v_{21}, v_{22}, \ldots, v_{2n}] X \ldots [v_{m1}, v_{m2}, \ldots, v_{mn}] .$$

The element $f_{i_1, i_2, \ldots, i_n}$ has the value 1 if there exists a sequence of values $v_{i_1}, v_{i_2}, \ldots, v_{i_n}$ the attributes $A_{i_1}, A_{i_2}, \ldots, A_{i_n}$ can have in system $S$, otherwise $f_{i_1, i_2, \ldots, i_n} = 0$. This array will be called a compatibility function.

Basing on these attributes and values we construct the cortege

$$C_l = \{(A_{j1}v_{j_{m_j}}), \ldots, (A_{j_{l-1}}v_{j_{m_{j_{l-1}}}})\}, l = 1, \ldots, N.$$

Each cortege will have the corresponding weight $\ell_l$.

Our aim is to construct a cortege $C$ following both criteria of the most frequent and most recent actions.
adding one by one attributes $A_i$ with compatible values. Initially we will take into account the frequency
criterion. Let $\varphi$ be the initial threshold of frequency.

1. An attribute $A$ is included into cortege $C$ if it is present at least in $\varphi$ of corteges $C_i$.

2. The value $v$ of this attribute is established equal to the most frequent value of the attribute $A$
ocorr{occurred} in corteges $C_i$. If there are two or more values with equal frequencies go to step 3, if not
– go to step 4.

3. Use the recent action criterion: among the values with equal (maximal) frequencies select as a
value of attribute $A$ that, which belongs to the cortege with the greatest weight.

4. Verify if values from the obtained cortege satisfy the compatibility function. If yes, then $C$ is the
result, else go to step 5.

5. From the sequence of values select that component (or those components) which cause the
incompatibility. Replace them with other values having lower frequencies. Repeat step 4 if such
values exist, otherwise go to step 6.

6. It was not possible to construct a cortege with those parameters. Replace them, namely the sequence of
initial corteges $C_i$, the threshold of frequency $\varphi$. Repeat the process starting with step 1.

The proposed algorithm does not produce always a resulting cortege. In an unsuccessful case one can vary
the initial parameters (the sequence of corteges $C_i$, the threshold of frequency $\varphi$ etc.) and repeat the
process until a result satisfying the compatibility function will be obtained.

5. Conclusions

N. Kajler (Kajler, 1998) considers computer-human interaction in symbolic computation “a wide and
interdisciplinary research area”, including different aspects, which “current user interfaces either do not
handle well or do not address at all”. We tried to cover only a small part of them, namely the problem of
adaptation to users using the mechanism of computational environments and sessions. Our experience
shows that in many cases about 5 saved sessions are enough to be able to understand user’s preferences
and to generate for him a convenient environment reflecting his domain of interests.

REFERENCES

1. BACKELIN, J., S. COJOCARU, V. UFNAROVSKI, Mathematical Computations Using Bergman,
Lund University, Centre for Mathematical Science, 2005, 206 p.

2. COJOCARU, S., L. MALAHova, A. COLESNICOV, Providing Modern Software Environments
pp. 120-140.

3. COJOCARU, S., A. COLESNICOV, L. MALAHova, Interfaces to Symbolic Computation
Systems: Reconsidering Experience of Bergman, Computer Science Journal of Moldova, ISSN


5. GAJOS, K., M. CZERWINSKI, D. TAN, D. WELD, Exploring the Design Space for Adaptive
Graphical User Interfaces, In: Proceedings of the working conference on Advanced visual interfaces

6. GRABMEIER, J., E. KALTOFEN, V. WEISPfENING (eds.) Computer Algebra Handbook,


10. WAERN, A., What is an Intelligent Interface? Available at www.sics.se/~annika/papers/intint.html

Large-scale complex systems (LSS) have been traditionally represented in control literature by models which are characterized by a large number of state and control variables, strong nonlinearities and uncertainties. Their decomposition into smaller, more manageable subsystems, possibly organized in a hierarchy, has been associated with intense and time-critical information exchange and with the need for efficient decentralization and co-ordination mechanisms.

The last decade of the past Millennium and the early years of the 21st century have revealed new characteristic features of industrial and non-industrial large-scale and complex systems. The enterprise of the present time is to operate in a highly networked environment. Also there is an ever more increased concern for new aspects such as: a) integration of various technologies, b) economic, environmental, and social aspects, and c) security issues. Consequently, the design of control systems must take into account more aspects and needs additional skills and tools. At the same time, the recent advances in computer and communication technologies have shown that they can provide effective tools and adequate technical infrastructures to support the design and enable the implementation of advanced control systems for the large-scale and complex applications of the present time.

The TC (Technical Committee) 5.4 on "Large-Scale Complex Systems" of IFAC (International Federation of Automatic Control) has already a more than 30-year tradition. The main events of this TC are the symposia entitled "Large- Scale Systems: Theory and Applications". The inaugural edition was held in Udine, Italy in 1976. Since then, at a three-year succession, the Symposium took place in Toulouse, France (1980), Warsaw, Poland (1983), Zurich, Switzerland (1986), Berlin, GDR (1989), Beijing, China (1992), London, UK (1995), Patras, Greece (1998), Bucharest, Romania (2001), Osaka, Japan (2004), and Gdansk, Poland (2007). The next edition of Symposium will be held in Lille, France, in 2010.

In 2002, at the IFAC World Congress and General Assembly held in Barcelona, the IFAC Technical Committee 5.4 was placed under the Coordinating Committee "Manufacturing and Logistic Systems", chaired by Professor Shimon Nof. Since then, the scope and objectives of the TC have been adapted accordingly to reflect the current trends in control systems and the new place of this TC within the IFAC organization. The TC focuses on manufacturing and related systems which are characterized by complexity and/or a networked structure of interconnected subsystems. It aims at developing new decentralized and hierarchical control methods, decision-making, and risk analysis techniques together with practical solutions based on new advances in computer and communication tools. A practical consequence of the new place of the TC is organizing the MCPL (Management and Control of Production and Logistics) conference starting with 2004. MCPL conferences were held in Santiago de Chile (2004) and Sibiu, Romania (2007). The next MCPL conference is to be organized in Coimbra, Portugal in 2010.

It can be appreciated now that several subfields that are traditional for LSS domain remain of increasing interest to the scientific community, such as decentralized and hierarchical control, model reduction, optimization, and complex system analysis. Traditional applications of LSS methods, such as power, gas, transportation, manufacturing, environmental systems are more than ever of great interest. The new issues such as collaborative control, intelligent network systems, risk management and decision support in LSS are well received by the LSS community. Among main ideas and trends one can include: a) large systems become ever "larger" (more complicated and networked) and b) there is still a need for practical solutions, possibly including new technologies with a view to attaining a correct balance between control techniques and computer and communication tools.

At the LSS symposium in Gdansk a formal meeting of the TC5.4 was held and Prof. Brdys of Birmingham University, UK, was elected for nomination as a chair of the Committee. According to Prof. Brdys, complexity is anticipated to be a central problem in modern system theory and practice. Complexity issues in control of large scale and complex systems such as manufacturing systems and related systems require a deeper and systematic continuation in the development of a relevant theory of complex systems including its extension to recent communication issues. Variety of theoretic results motivated by practice and successfully developed within various long term activities of this TC could offer a deep background for such continuation. Multidisciplinary approaches appear to be adequate in modelling and controlling large-scale and complex systems.

At the end of my second term as a chair of the TC 5.4, I intended to publish a special issue of this journal addressing several relevant aspects and results in LSS domain. Consequently, the present issue of the
Wan, Ruan, and Liu of China review their investigating efforts and results over the past fifteen years in the on-line steady-state hierarchical intelligent control and optimization of large-scale industrial processes, a research direction initially proposed by Roberts three decades ago. Borowa and colleagues from UK and Poland describe an approach for detecting and localisation of pipe leakage in Drinking Water Distribution Systems (DWDS) which are large-scale and complex distributed dynamic systems. A practical DWDS is decomposed into suitable subnetworks which make the monitoring process easier while fewer sensors are required. The subnetworks and corresponding PCA (Principal Component Analysis) monitoring models are selected based on the network operational knowledge and information regarding its topology. Duviella, Chiron and Charbonnaud of France propose a weighted digraph representation of dam-river systems to solve the water asset-management by resource allocation and setpoint assignment problem. Tangour and Borne of EC Lille (France) give a review of several various metaheuristics optimisation methods which have been implemented in various applications such as the optimization manufacturing problems and recommend hybrid approaches. Chamroo and colleagues from France and Bulgaria propose a new control scheme for a highly non-linear anaerobic digestion process. Ivanescu and Florescu of Romania present a solution for the control problem of a hyperredundant robot with continuum elements that performs the colli function for grasping. Ayadi and his colleagues from Tunisia propose a practical approach to design and implement a real-time system based on a RISC microcontroller dedicated to a DC motor speed control.

Leon, Zaharia, and Galea of Romania propose a system of autonomous agents for dynamic routing in high performance distributed systems.

Pascale Zarate of France presents several methodological issues in an effort to develop Cooperative Decision Support Systems (CDSS) for collective decision making processes. In order to make the information systems more usable, Svetlana Cojocaru from the Moldavian Academy of Sciences propose intelligent interfaces, which are able to adapt themselves to user’s learning needs, take over the initiative in their interaction with user, are able to guide him so that the objectives may be achieved in a fast and comfortable way.

F. G. Filip

Chair, IFAC TC 5.4 (2002-2008)
AUTHOR INDEX

Ayadi, M., 201
Benrejeb, M., 201
Borne, P., 169
Borowa, A., 135
Bouallègue, S., 201
Brdys, M. A., 135
Chamroo, A., 181
Charbonnaud, P., 153
Chiron, P., 153
Christov, N., 181
Cojocaru, S., 231
Duviella, E., 153
Duzinkiewicz, K., 135
Florescu, M. C., 189
Gâlea, D., 215
Grochowski, M., 135
Haggège, J., 201
Ivanescu, M., 189
Jezior, K., 135
Leon, F., 215
Liu, D., 123
Mazur, K., 135
Popescu, D., 189
Popescu, N., 189
Ruan, X.-E., 123
Simeonov, I., 181
Tangour, F., 169
Vasseur, C., 181
Wan, B.-W., 123
Zaharia, M. H., 215
Zaraté, P., 225
INSTRUCTIONS TO AUTHORS

SIC welcomes original contributions on all aspects of research and practice in IT. Authors should not submit to SIC material that is under consideration elsewhere. Papers should preferably not exceed 15 standard typewritten pages. However, while concision improves chances of publication, serious consideration will be given to articles in which the nature and quality of the content justify greater length. Articles should be sent if possible in disk form, accompanied by two printouts and a cover letter signed by the (principal) author to the following address: National Institute for R&D in Informatics, Publishing Department, 8-10 Averescu Blvd., Bucharest 011455 Romania

The manuscript should be arranged in the following order: title; author's name and forename; name and address of the organization authors belong to; abstract; keywords and phrases; short biographical notes, main body of the paper; references in alphabetical order by first author's surname; tables and figures. Headings should be numbered by Arabic numbers. Tables and figures should be titled, numbered consecutively, put on separate white sheets. Legends are to be typewritten on a separate page. References in the text must have in the order of their occurrence a consecutive number included in brackets. For references, quotations etc., please follow the conventions used in recent issues of SIC.

We always acknowledge receipt of contributions submitted to us. After a preliminary triage to eliminate submissions that are unsuitable for the SIC journal, the remaining science-based papers, will enter a review process which typically takes 4-6 weeks from the time we acknowledge the receipt of the paper. Following the confirmation to the corresponding author of acceptance for publication of the submitted article (presumably sent as a hard copy), a quick editing and printing work would be possible if agreed to provide a soft copy, on diskette or via e-mail, under Windows (Microsoft Word 2003).

It is a condition of publication in SIC that authors grant an exclusive licence to the Journal, which is operated on the Journal's behalf by the Publisher, by completing a copyright licence form. This ensures that requests from third parties to reproduce articles are handled efficiently and consistently and will also allow the article to be as widely disseminated as possible. When an article is published, a PDF of the final version is emailed to the corresponding author. It is the author's responsibility to obtain permission to reproduce verbal or visual material from copyright sources.