Biblioteka Informatyki
Szkół Wyższych

Information Systems Architecture and Technology

Knowledge Based Approach to the Design, Control and Decision Support

Library of Informatics of University Level Schools

Series of editions under the auspices
of the Ministry of Science and Higher Education

The ISAT series is devoted to the publication of original research books in the areas of contemporary computer and management sciences. Its aim is to show research progress and efficiently disseminate current results in these fields in a commonly edited printed form. The topical scope of ISAT spans the wide spectrum of informatics and management systems problems from fundamental theoretical topics to the fresh and new coming issues and applications introducing future research and development challenges.

The Library is a sequel to the series of books including Multidisciplinary Digital Systems, Techniques and Methods of Distributed Data Processing, as well as Problems of Designing, Implementation and Exploitation of Data Bases from 1986 to 1990.

Wrocław University of Technology

Information Systems Architecture and Technology

Knowledge Based Approach to the Design, Control and Decision Support

Editors
Jerzy Świątek
Leszek Borzemski
Adam Grzech
Zofia Wilimowska

Wrocław 2013

Publication partly supported by

Faculty of Computer Science and Management
Wrocław University of Technology

Project editor

Arkadiusz GÓRSKI

The book has been printed in the camera ready form

OFICYNA WYDAWNICZA POLITECHNIKI WROCŁAWSKIEJ
Wybrzeże Wyspiańskiego 27, 50-370 Wrocław
http://www.oficwyd.pwr.wroc.pl;
e-mail: oficwyd@pwr.wroc.pl
zamawianie.ksiazek@pwr.wroc.pl

ISBN 978-83-7493-802-0

CONTENTS

PART 3. MODEL BASED CONTROL AND DECISION SUPPORT

INTRODUCTION

All project, decisions and control algorithms are based on the knowledge about the plant under investigation. Systems approach to the computer aided design, control and decision support requires model of the investigated process. That's why models are so important in systems research. Investigation of object of the deferent nature (technical, economical, biomedical or computational) gives us many notifications about observed processes. Based on the collected knowledge, about investigated process the model of observed reality is proposed. The mathematical model gives precise plant description. Usually the relation between values characterizing process is given.

System analysis gives us the proper tools to create further decision about investigated plant based on the collected knowledge, and consequently based on the elaborated model. Base on the model the optimization, control and management task may be formulated. Base on the knowledge about the process the prediction or diagnosis may be proposed.

The above mentioned applications of different type tasks we can recognize in selected and revived chapters which have been divided into the following groups:

- Knowledge Engineering and its Application in Decision Support Systems
- Soft computing and discrete type systems,
- Model based control and decision support.

The book provides an interesting representation of research in the area of system analysis in decision aided problems in proposed groups.

PART 1. KNOWLEDGE ENGINEERING AND ITS APLICATION IN DECISION SUPPORT SYSTEMS

Embedded systems for most of their history were seen simply as small computers with limited capabilities and functionalities. However, great progress has been made in both computer hardware and software over recent years, which enables us to have much more powerful computers in very small sizes and with many more functions. Consequently, new needs and expectations for embedded systems have increased dramatically. In the **Chapter 1** the Experience-Oriented Smart Embedded System is proposed as a new technological platform providing a common knowledge management approach that allows mass embedded systems for experiential knowledge capturing, storage, involving, and sharing. Knowledge in the Experience-Oriented Smart Embedded System is represented as set of experience knowledge structure, and organized as Decisional DNA. The platform is mainly based on conceptual principles from Embedded Systems and Knowledge Management.

During recent years, manufacturing organizations are facing market changes such as the need for short product life cycles, technological advancement, intense pressure from competitors and the continuous customers' expectation for high quality products at lower costs. In this scenario, knowledge and its associated engineering/management of every stage involved in the industrial design has become increasingly important for manufacturing companies in order to improve their performance and to take effective decisions. Knowledge based industrial design techniques have been used in the past with fair bit of success but they have their share of limitations like they may be time consuming, costly, domain specific and at times not very intelligent. In the **Chapter 2** a novel approach is proposed, where the set of experience knowledge structure and decisional DNA techniques are used for the experience based representation of engineering artefacts. The knowledge representation method is proposed; captures and reuses the product and process involved within an engineering knowledge perspective. In this chapter a new concept is introduced, called Virtual Engineering Object, which is a conceptual entity that permits a dual computerized/ real world representation of an engineering entity. It contains the embedment of the decisional model expressed within the set of experience, a geometric representation and the necessary means to relate such virtualization with the physical object being represented.

Several methods and models are currently used on the aviation market to forecast the market situation in the medium term period (6-18 months), but none of them meets all current airlines expectations. The idea of the model which should be created to meet the airlines' expectations, taking into consideration the complexity of the civil aviation market is presented in the **Chapter 3** The civil aviation market is described with its players, current conditions and its main problems. The tools which are currently used for the prognosis processes are presented. Finally, the idea of a new model (fuzzy model) for the process of prognosis is presented.

The concept of expert system called SAILBOAT, which supports advisory-decision process in the company that makes sailing boats is presented in the **Chapter 4** Basic problems in design sailboats were recognized. Foundations of systems construction and functional description of SAILBOAT system are presented. The chapter presents a computer implementation of the system and the results of testing the SAILBOAT system.

At present a great deal of research is being done in different aspects of **Content-Based Image Retrieval.** Image classification is one of the most important tasks that must be dealt with in image DB as an intermediate stage prior to further image retrieval. The **Chapter 5** shows an evolution from the simplest to more complicated classifiers. Firstly, there is the most intuitive one based on a comparison of the features of a classified object with a class pattern. The problem of finding the adequate weights, especially in the case of comparing complex values of some features is proposed. Secondly, the decision trees as another option in a great number of classifying methods. Thirdly, to assign the most ambiguous objects, fuzzy rule-based classifiers is discussed . The ranges of membership functions for linguistic values for fuzzy rule-based classifiers according to crisp attributes are proposed

PART 2. DISCRETE TYPE SYSTEMS AND SOFT COMPUTING

The theoretical prediction of the behavioural properties of mesh-like periodic structures is less difficult than in the case of nonperiodic ones. That is because all the information needed to describe the entire structure is given by the elementary substructure itself and by the manner it repeats in the whole periodic structure. In turn, cyclic schedule allows to avoid the scheduling of the whole tasks and to handle the combinatorial explosion of the problem by considering only a small pattern (cycle). That is because all the information needed to describe the entire structure is given by the elementary substructure itself and by the manner it repeats in the whole periodic system structure. So, only a small portion of the mesh-like periodic structure needs to be considered to obtain the cyclic schedule of the whole system. In that context, the **Chapter 6** provides the discussion of some system periodicity issues, and is aimed at modelling and evaluation of relationships linking features of the mesh-like structure with required system's cyclic functioning.

The studies hitherto carried out have revealed that the application of the *ShortestPathACO* strategy based on the Ant Colony Optimization (ACO) metaheuristics makes it possible to solve the shortest path problem in a way that differs from traditional approaches. Practical applications of algorithms that are based on the Ant Colony Optimization metaheuristics require accurate and deep understanding of the importance of particular parameters of the algorithm and this issue is absolutely crucial. On account of the heuristic approach to the problem to be solved, it is important then to choose appropriately the mode of operation of the algorithm to be applied as well as a determination of the parameters that will ensure proper course of operation in a given specific situation. The **Chapter 7** describes, analyse and interpret the observations provided by relevant studies and to present conclusions in relation to the conducted research work and subsequent analyses. The chapter mainly focuses on indicating the essential problems and defining the areas for further optimization of the *ShortestPathACO* approach. Limitations and constrains involved in the application of the Ant Colony Optimization metaheuristics in solving the shortest path problem are highlighted. A particular attention is given to the greediness and convergence of the algorithm, as well as to possibilities of a secondary use of available information on the pheromone level.

Many researchers have proved that Decisional DNA and Set of Experience Knowledge Structure is a technology capable of gathering information and converting it into knowledge to help decision-makers to make precise decisions in many ways. These techniques have a feature to combine with different tools, such as data mining techniques and web crawlers, helping organization collect information from different sources and using gathered knowledge to make decision or prediction. In the

Chapter 8 the focus is on another research issue, optimization problem. Although there are many algorithms that have been design to solve this problem, it still lacks efficiency and effectiveness to get optimal solutions. Therefore, the propose of these chapter is a new structure combining the Set of Experience Knowledge Structure with an evolutionary algorithm to find optimal solutions and to reuse this experience for efficient decision making support.

We discuss further contributions to The problem of creating a simulation model of a complex regional healthcare system and implementing it using real-world data is discussed in the **Chapter 9** The approach is based on modelling patients' pathways comprising diagnostic and treatment processes throughout patients' stay in the system. The general structure of the model and two main modules are outlined, with observations on problems regarding modelling logical processes within the regional system, and pitfalls when modelling large real-world systems. The approach is illustrated by a sample model of a regional system of hospital care in the Lower Silesia Region of Poland for lung cancer patients.

There is a gap between practice of accounting and simulation modeling. In the **Chapter10** a new approach to integrate concepts of discrete event simulation modeling and accounting for the generation of financial statements in production processes is discussed. The new integrative approach is the multiplication of object representing the discrete events (e.g. sale or purchase) to physical aspect object and book-entry form object. It provides the new abilities to generate and analyze financial ratios. With this approach it is possible to generate not only easy to design items like revenue income statements, cost of goods, incomes, expenses or even activity-based cost but also the balance-sheet with particular items as payables and receivables. This chapter presents a business model that uses a new approach and the relationship between supply and demand.

PART 3. MODEL BASED CONTROL AND DECISION SUPPORT

In the **Chapter 11** the problem of planning long-term training for sport performance optimization is considered. After short introduction to the problem and survey on solutions in the field of sport training support, the system to wireless sensing physiological data and their processing is presented. Designed system was presented in the details: architecture and the main elements of the proposed approach results of experimentations are presented and discussed. The work model-based algorithm to investigate plan of physical training is given. The algorithm to design plan of training is based on Banister's model describing relationship between input training (workload) and output training (physical performance).

Mathematical model of the Glucose-Insulin system allows to simulate and predict the human body response to meals and insulin intakes. It may also be useful when performing control actions for diabetic patients. In order to make the model valuable for clinical applications, its parameters need to be estimated in such a way, that the model responses resemble responses of a particular person. The problem is that the

Glucose-Insulin system is complex, whereas the number of measurement data is low. This makes estimation task ill-conditioned. The **Chapter 12** presents mathematical models describing different parts of the Glucose-Insulin system are combined into a single complex model. Moreover, parameter estimation routine based on numerical optimization methods is proposed for this complex system.

Energy and heat is necessary for civilization existence and growth. Nowadays energy needed for human communities is provided mainly as electricity in power generation processes from combustion of fossil fuels in power plants. Efforts are being made to eliminate carbon dioxide formation during combustion processes. In the **Chapter 13** the study $O_2/CO_2/H_2O$ (oxyfuel) combustion technology of coal in drop tube reactor due to its potential of allowing to mitigate CO_2 are presented. Changing of combustion atmosphere from air to $O_2/CO_2/H_2O$ results in different thermo-physical and chemical properties of heat transferring medium which finally affects the heat transfer process in combustion chamber. In oxy-combustion higher than in air combustion concentrations of CO_2 and H_2O (triatomic species playing key role in radiation) may affect especially radiation heat transfer and validated for air combustion radiation models may not be suitable. Authors by calculating absorption coefficient studied how different models of radiation (those intended for air combustion and those especially modified for oxyfuel combustion) affected the heat transfer characteristics.

Quadrocopter is a dynamic system with many variables. PID standard controller is often used to stabilisation of flight of this object. The main disadvantages of this control system is the need of tuning of other parameters arising from the length of arm, different types of propellers with motors and object's weight. The main problem for constructors is the tuning of controller parameters. In the **Chapter 14** the use of neural network quadrocopter control is discussed. The control system is divided into four subsystems. Each of them is responsible of setting the control values. The neural network is learnt by control system with standard PID controller. This approach is used for checking how neural networks cope with stabilisation of the quadrocopter. The proposed controller was tested in different structure of neural network and different states of flight: in hover, in to forward flight with constant speed, in climbing and in rotation. In all these situations the proposed controller was able to provide foreseeable behavior of the quadrocopter. Simulation results of the neural controller and PID controller working were compared to each other.

In the **Chapter 15** the problem of analysis of the pilot’s control actions is discussed. The presented approach is based on the fuzzy flow graphs idea. The concept of interactive information systems is utilized to represent the aircraft-pilot influences. The issue of selecting the appropriate attributes is considered. The final decision table with fuzzy attributes is obtained, which is analyzed with the help of fuzzy flow graphs. In consequence, fuzzy decision rules are generated which can be viewed as a part of the pilot’s control actions model.

The increasing number of wind farms results in increasing importance of wind power in the power system. It entails the necessity of preparation more and more precise forecasts of wind farms work. However, the characteristic of the source exploited to produce energy makes the task extremely difficult. The increasing amount of data is another problem, since its processing influences the pace of forecast preparation. The **Chapter 16** shows the method allowing decreasing the amount of data needed to prepare a forecast, and as a result, shorten the time of calculations. At the same time, satisfactory precision of the forecast is maintained.

The **Chapter 17** article presents the results the investigation of portfolio risk dynamics. This research, apart from the Value at Risk method uses the following method of determining the dynamics: moving average and moving standard deviation. This approach of calculating the Value at Risk will not only permit observation of changes but also allows an investigation of repeating characteristic patterns.

Wroclaw, September 2013

Jerzy Świątek

PART 1

KNOWLEDGE ENGINEERING AND ITS APLICATION IN DECISION SUPPORT SYSTEMS

Experience-Oriented, Smart,
Set of Experience Knowledge Structure, Embedded System

Haoxi ZHANG[*], Cesar SANIN[**], Edward SZCZERBICKI[***]

EXPERIENCE-ORIENTED SMART EMBEDDED SYSTEM

The Experience-Oriented Smart Embedded System (EOSES) is proposed as a new technological platform providing a common knowledge management approach that allows mass embedded systems for experiential knowledge capturing, storage, involving, and sharing. Knowledge in the EOSES is represented as SOEKS, and organized as Decisional DNA. The platform is mainly based on conceptual principles from Embedded Systems and Knowledge Management. The objective behind this research is to offer large-scale support for intelligent, autonomous, and coordinated KM on various embedded systems.

Several conceptual elements of this research have been implemented in testing prototypes, and the experimental results that were obtained show that the EOSES platform can provide active knowledge management to different embedded systems, and it can also enable various systems to learn from their daily operations in many different fields to gather valuable knowledge, assist decision making, reduce human workers' workload, and improve the system itself. As a result, the EOSES has great potential for meeting today's demands for embedded systems, and providing a universe knowledge management platform for mass autonomous mechanisms.

1. INTRODUCTION

Embedded systems for most of their history were seen simply as small computers with limited capabilities and functionalities. However, great progress has been made in both computer hardware and software over recent years, which enables us to have much more powerful computers in very small sizes and with many more functions. Consequently, new needs and expectations for embedded systems have increased dramatically

[*] The University of Newcastle, University Drive, Callaghan, 2308, NSW, Australia. Haoxi.Zhang@uon.edu.au

[**] The University of Newcastle, University Drive, Callaghan, 2308, NSW, Australia. Cesar.Sanin@Newcastle.edu.au

[***] Gdansk University of Technology, Gdansk, Poland. Edward.Szczerbicki@zie.pg.gda.pl

[1]. The prospects for applying KM technologies to embedded systems to meet these demands are very promising.

The Experience-Oriented Smart Embedded System (EOSES) is proposed as a new technological platform providing a common knowledge management approach that allows mass embedded systems for experiential knowledge capturing, storage, involving, and sharing. The platform is mainly based on conceptual principles from Embedded Systems and Knowledge Management. The objective behind this research is to offer large-scale support for intelligent, autonomous, and coordinated KM on various embedded systems.

1.1. OVERVIEW OF EOSES

Due to the diversity of embedded systems, the hardware and software of digital devices involved in these processes could be various. Moreover, different companies could use different devices for similar processes. As a consequence, the Experience-Oriented Smart Embedded System is designed to handle these issues, providing the means for organizations and individuals to have a common systematic KM tool within diverse embedded systems.

Figure 1 shows the overview of the Experience-Oriented Smart Embedded System. The platform integrates different organizations and their devices as illustrated in three major levels: Device, Process, and Knowledge. Based upon the principle of KM, the KM infrastructure is a development of the digital nervous system within an organization, which integrates the organization at a deeper level [3]. In an organization, there are many processes with the purpose of fulfilling a collection of well-defined organizational aims.

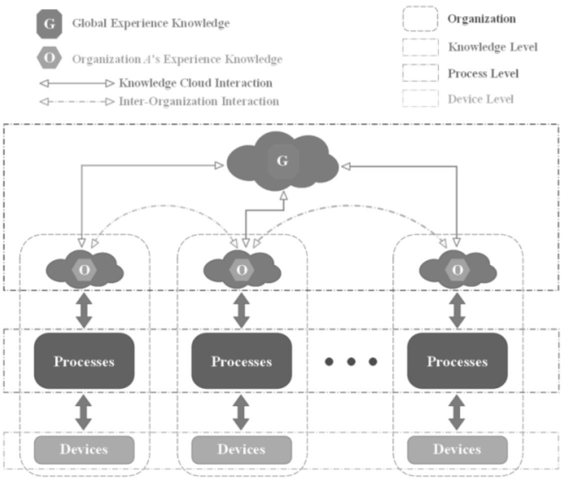


Fig. 1. Overview of the Experience-Oriented Smart Embedded System

At the bottom level of the platform, devices representing different embedded systems in knowledge-related organizational processes capture knowledge for the activities they are involved in. Each embedded system at the bottom level may be different, and may be involved in different activities. A knowledge-related process may have one or more activities. By integrating knowledge captured from a process's activities, experiential knowledge grows and evolves for that given organization process. Moreover, devices can interact with each other in order to share knowledge. At a higher level, there are processes within an organization. In a similar way, by collecting every process's knowledge, the knowledge for that organization is formed and gathered. Also, processes at the middle level can interact with each other to share knowledge. At the top, there is the knowledge level. Cutting-edge technologies, for example Cloud Computing, may be used to store and enable interaction and knowledge sharing among organizations. Organizations may create their own clouds for a specific purpose, like data security. In addition, a large-scale knowledge market could be formed, in which knowledge is treated as the main asset [16].

1.2. MAIN FEATURES OF EOSES

In order to enable different embedded systems to capture, reuse, evolve, and share knowledge in an easier and more standard way, the Experience-Oriented Smart Embedded System shall provide the following features:

• *Experience-Oriented*: experience, as one kind of knowledge learned from practice, is the ideal source for improving performance of processes, in which a lot of practical activities involve. By reusing experiential knowledge, decision makers can make decisions faster, and more efficiently base their current decisions on experiences obtained from previous similar situations. The EOSES is designed to use such experiential knowledge in supporting individuals and organizations to make decisions better and more efficiently.

• *Adaptability and Cross-platform Portability*: Since most embedded systems are special-purpose designed, the hardware and software for each of them could be distinctly different from others; thus, adaptability and cross-platform portability are critical issues for EOSES. For this reason, which causes the first problems of these three, i.e. they are too specific. In order to work with custom embedded software, most knowledge-related tools are specifically designed only for given embedded systems or even for a particular stage of a product lifecycle. Adaptability and cross-platform portability are rarely considered during the design and development of these tools. For solving this problem, EOSES is designed to work with specific designed software (applications) without re-designing or re-building. Java is used during the development of EOSES. And even though Java may be a bit slower than some other programming languages like C or C++ [11], its inherent cross-platform capability can make significant advantages on deployment and maintenance. Moreover, as processing power soars, the evolution of Java language, and optimizations of JVM [4], is not a problem anymore for running Java applications on embedded systems [18].

• *Compactness and Efficiency*: Another important element concerning the application of EOSES is that embedded systems are usually resource-limited, which means our applications have to be designed in a compact, efficient shape. As a flexible knowledge representation structure, the Set of Experience Knowledge Structure (SOEKS) and Decisional DNA can be customized compactly according to certain system goals. Moreover, it has been shown in previous case studies [22] that SOEKS is an efficient knowledge representation structure, which can indeed enhance knowledge retrieval and experience reusing applications. Thus, SOEKS is used in EOSES as an efficient compact knowledge representation.

• *Configurability*: It is common for a system to run under different situations. For example, the same system may work under different power supply conditions: connected to a power line or by using batteries. If the system performs always in the same way regardless of its power supply, it will soon run out of power and shut down when using batteries. Also, a system may work in different modes, for example it can be trained in training mode and work in an automatic self-control mode. For developers, it may have debugging mode. Thus, configurability is very desirable as it allows systems to perform properly in various scenarios according to different settings.

• *Security and trust*: it is clear that a secure environment is a key requirement for any distributed system these days, especially when the Internet is used as the primary communication channel. Also, knowledge, and knowledge sources, must be reliable to make the right decisions. Therefore, the concept of decisional trust presented by Sanin and Szczerbicki [21] is extended to include more features that reflect human-like behaviour.

2. CONCEPTUAL ARCHITECTURE OF EOSES

For carrying these five main features introduced above, the conceptual four-layer architecture is designed for EOSES. These four layers are: Application Layer, Interface Layer, Management Layer, and Knowledge Repository Layer. The platform is conceptualized on the top of embedded systems, to extend the capabilities of the latter by using DDNA and SOEKS for knowledge representation and exchange. Figure 2 illustrates the conceptual architecture proposed for the Experience-Oriented Smart Embedded System.

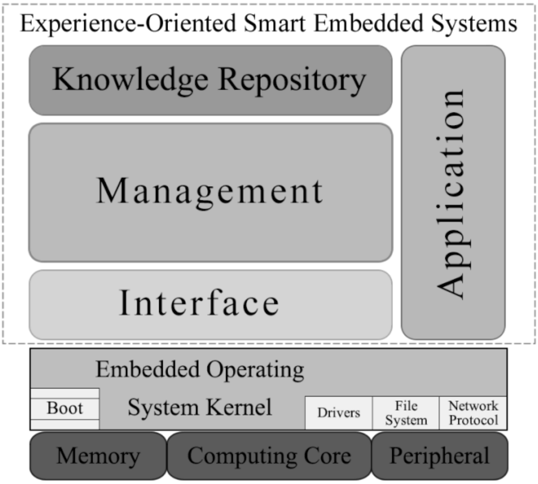


Fig. 2. Architecture of the Experience-Oriented Smart Embedded System

Each one of the layers in the conceptual architecture is engaged with a set of responsibilities and capabilities, as follows:

• *Application Layer*: This layer offers the platform's whole functionality access to the end-user. Mobile applications (APPs) may be used to help facilitate the interaction among employees or groups to solve problems, make decisions, and feed the system with information based on their daily activities. Additionally, customization and complementing technologies are fully supported in the application layer in order to fulfil different job requirements, which enable EOSES to capture experiential data from different sources.

• *Knowledge Repository Layer*: Experiential knowledge, as the most valuable source, is stored and managed at this layer. In EOSES, a single decision event is captured and represented as a SOE, and a set of SOE is organized as DDNA carrying the decisional fingerprint of the decision maker. This layer provides functionality of access, storage, and administration of knowledge.

• *Management Layer*: This layer is the control central of EOSES. In order to achieve knowledge capturing, reusing, evolving, storage, and sharing, a range of functionality and capability are attached to this layer, such as inference, analysis, knowledge extraction, knowledge retrieval, knowledge exchange, dynamic process management, inter and intra-system interactions, global policies, and cooperation amongst other mechanisms.

• *Interface Layer*: Interface is the component that connects the platform with its outer environment; which provides a range of interaction, communication, and knowledge-related services. These services provided by this layer are: system configuration, input/output, application programming interface (API) service, query service, knowledge sharing service, and updating service.

• *Embedded system hardware and operating system*: None of these two sections is a part of EOSES. However, they comprise the fundamental base for EOSES. On the top, there is the embedded operating system which manages the embedded system and runs EOSES and applications. And at the bottom, there is embedded hardware underlying, which is comprised of memory, computing elements, and peripherals.

All the layers in the conceptual architecture make extensive use of experience-oriented knowledge management to provide appropriate autonomous capabilities to embedded systems.

3. CASE STUDIES

In order to evaluate and test this conceptual platform, we applied EOSES to robotic and digital TV systems [5-9]. This section introduces the case studies of this research, and discusses the evaluation and return on experience of the EOSES.

3.1. EOSES APPLIED TO ROBOTICS

A long-standing goal of research in robotics is to build a robot that captures experience through day-to-day tasks, and reuses this experience to improve the robot's task performance ability in performing similar tasks. In recent years, more and more efforts have been made by researchers to make learning-by-doing robots. Several experience-based researches and theories in robotics can be found in literature [2], [17], [19], [20].

In our study, we used the LEGO® Mindstorms® NXT 2.0 robot, equipped with a 48 MHz ARM7 microcontroller, a 256KBytes FLASH, a 64KBytes RAM [13], a color sensor and one ultrasonic sensor. Since the Decisional DNA is defined in Java [22], we burnt the LeJOS into the NXT 2.0 as the firmware. The LeJOS is a Java programming environment for the NXT® robot [14] (see Fig. 3).

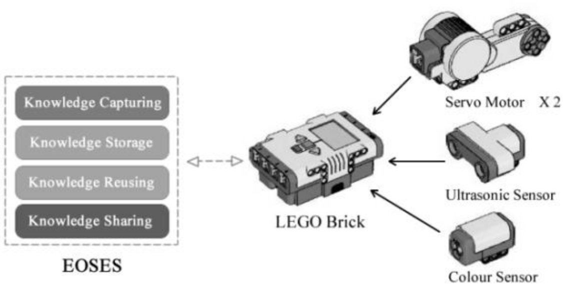


Fig. 3. System Composition of the Robotic Application

An initial map learning and path planning EOSES application, called Line Follower, is designed in this case study: the EOSES applied robot is capable of gaining experiential knowledge about maps through integrated and organized sensor data, and then reuses its knowledge for path choosing in the future. There are two modes in this application:

• *Training Mode*: The robot follows every path on a map one by one in order to learn and gain experience of these paths under the supervision of human beings. After training, the robot will have experiential knowledge about the map's paths.

• *Testing Mode*: In this mode, the robot's knowledge about the map is tested. An assumption destination will be given to the robot, and the robot should pick the right path to follow down to its destination by reusing previous knowledge learned through the Training Mode.

We tested our concepts on two 70cm×100cm maps (see Fig. 4). There are three paths drawn in different colours (green, blue, and black) on each map, and every path finishes at a terminal spot. In the experiments, the assumption is that the robot starts at position A, with spot B as the potential destination for the robot in Testing Mode. To distinguish different terminal spots, spot B is drawn in red (the destination), while the other two spots are drawn in yellow (the normal terminal spots)..

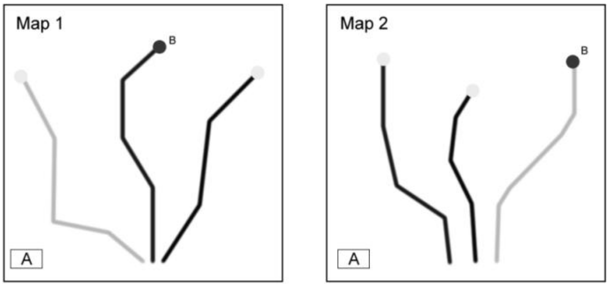


Fig. 4. Experiment Maps for Line Follower Application

A set of experiments testing the Line Follower's performance of using experience and not using experience were reported and published in our previous paper[7]. And the results showed that the time consumption is much less and more stable when using previous experience in comparison with not using previous experience.

3.2. EOSES APPLIED TO DIGITAL TV

Digital television (DTV) is the television broadcasting system that uses digital signals to transmit program contents. DTV not only delivers distortion-free audio and video signals; more importantly, it offers much higher radio spectrum efficiency than analog television does. DTV can also seamlessly integrate with other digital media, computer networks, and communication systems, enabling multimedia interactive services and data transmission [26].

In this case study, we present an approach that integrates the EOSES with DTV to capture and reuse viewers' TV watching experiences. In the implementation, a Hauppauge Nova-T USB TV stick [10] is used to receive the digital TV signal. It works with the TVHeadend [23] DTV streaming server on a laptop. The TVHeadend streams the received digital TV content to the local network. At the Android side, we use the EOSES-combined TVHGuide [24] receives and plays DTV streams on a Toshiba Shriving tablet [25] running the Android 3.2 operating system. The TVHGuide is a client for the DTV streaming server – TVHeadend (see Fig.5).

Fig. 5. Implementation of the DDNA DTV

This EOSES-based DTV captures viewer's TV watching experience by storing each single TV watching event. A single TV watching event is taken as the viewer watches a channel for more than ten minutes. And there are nine variables recorded for each single TV watching event: Channel ID, Channel Name, Programme, Watch Day, Start Time, End Time, Type, Description, and Viewer. Channel ID and Channel Name indicates which channel was watched. Programme stores the name/title of the program. Watch Day tells the day of week the channel was watched. Start Time and End Time store the start and stop time of the program. Type illustrates what kind of program it was. Description stores introduction of the program. Viewer saves the name of the user.

Once a single TV watching event is taken, those variables are gathered and sent to the Prognoser by the Client from TVHGuide; then, the Prognoser generates SOEKS statements, and sends the SOEKS statements to the DDNA Repository Manager; eventually, this TV watching event is stored as a SOEKS in XML format [15].

Furthermore, we applied fuzzy logic methods [12] to our research to model the uncertainty in generating user profiles. A user profile refers to the user's basic information, such as gender, age, and profession. It enables TV players to suggest program choices based upon the user's past viewing habits, plus program suggestions from his/her cohorts (i.e. who have similar TV viewing interests). Also, this feature provides the option of recommended programming and crafted advertising. For more details of our research, please refer to [6], [8], [9].

4. CONCLUSIONS AND FUTURE WORK

In this paper, we introduce main features, architecture, and case studies of the concept of Experience-Oriented Smart Embedded System, which is proposed as a new technological platform providing a common knowledge management approach that allows mass embedded systems for experiential knowledge capturing, storage, involving, and sharing. The platform is mainly based on conceptual principles from Embedded Systems and Knowledge Management. The objective behind this research is to offer large-scale support for intelligent, autonomous, and coordinated KM on various embedded systems. Several conceptual elements of this research have been implemented in testing proto-types, and the experimental results that were obtained show that the EOSES platform can provide active knowledge management to different embedded systems, and it can also enable various systems to learn from their daily operations in many different fields to gather valuable knowledge, assist decision making, reduce human workers' workload, and improve the system itself. As a result, the EOSES has great potential for meeting today's demands for embedded systems, and providing a universe knowledge management platform for mass autonomous mechanisms.

To continue with this idea, further research and refinement are required, some of them are:

- Continuous development of the EOSES components: Configuration Manager, Plug-In Manager, Protocol Manager, and EOSES Manager;
- Further development of the fuzzy rule base and data base;
- Refinement and further development of algorithm using in the Prognoser;
- Further development of User Profile and User Classification functionality;
- Further development of Cloud computing functionality.

REFERENCES

[1] AKHRAS G., *Smart Materials and Smart Systems for the Future*, Canadian Military Journal, Vol. 1, 2000, No. 3, pp. 25–32,.

[2] BERENSON D., ABBEEL P., GOLDBERG K., *A robot path planning framework that learns from experience*, In Robotics and Automation (ICRA), 2012 IEEE International Conference on. IEEE, 2012., pp. 3671–3678

[3] FRID Randy J., *Infrastructure for Knowledge Management*, iUniverse, 2000.

[4] GEORGES A., BUYTAERT D., EECKHOUT L., *Statistically Rigorous Java Performance Evaluation*, Proceedings of the 2007 OOPSLA Conference, 2007, Vol. 42, No. 10, Oct. 2007, pp. 57–76.

[5] ZHANG H., SANIN C., SZCZERBICKI E., *Making Digital TV Smarter: Capturing and Reusing Experience in Digital TV*, Cybernetics and Systems: An International Journal, Taylor & Francis Group, LLC. Volume 43, Issue 2, 2012, pp.127–135.

[6] ZHANG H., SANIN C., SZCZERBICKI E., *Implementing Fuzzy Logic to Generate User Profile in Decisional DNA Television: The Concept and Initial Case*, Cybernetics and Systems: An International Journal, 44:2–3, 2013, pp. 275–283.

[7] ZHANG H., SANIN C., SZCZERBICKI E., *Gaining Knowledge Through Experience: Developing Decisional DNA Applications in Robotics*, Cybernetics and Systems: An International Journal, 41:8, 2010,, pp. 628–637.
[8] ZHANG H., SANIN C., SZCZERBICKI E., *The Development of Decisional DNA DIGITAL TV*, Proceedings of Knowledge-Based Intelligent Information and Engineering Systems 16th International Conference KES 2012, in Advances in Knowledge-Based and Intelligent Information and Engineering Systems: Manuel Grana, Carlos Toro, Jorge Posada, Robert J. Howlett, Lakhmi C. Jain (Eds.), San Sebastian, Spain, September, 2012, pp. 1500–1508.
[9] ZHANG H., SANIN C., SZCZERBICKI E., *Decisional DNA applied to digital TV*, Lecture Notes in Artificial Intelligence Proceedings of Knowledge-Based Intelligent Information and Engineering Systems 15th International Conference KES 2011, Robert J. Howlett Lakhmi C. Jain (Eds.), Germany, September, 2011, Springer-Verlag Berlin Heidelberg, LNAI 6882, 2011, pp. 667–676.
[10] Hauppauge, Hauppauge Nova-T stick, Retrieved October 2012 from http://www.hauppauge.co.uk/site/products/data_novatstick.html
[11] JELOVIC D., *Why Java Will Always Be Slower than C++*, Retrieved October 2012 from http://www.jelovic.com/articles/why_java_is_slow.htm
[12] ZADEH L., *Fuzzy sets*, Information and Control, Volume 8, Issue 3, 1965, pp. 338–353.
[13] LEGO, *What is NXT*, Retrieved December 2012 fromhttp://mindstorms.lego.com/en-us/whatisnxt/default.aspx
[14] LEJOS, *Introduction*, Retrieved December 2012 from http://lejos.sourceforge.net/nxt/nxj/tutorial/Preliminaries/Intro.htm
[15] LLOYD J.W., *Logic for Learning: Learning Comprehensible Theories from Structure Data,* Springer, Berlin 2003.
[16] MANCILLA-AMAYA L., SANIN C., SZCZERBICKI E., *Smart knowledge-sharing platform for e-decisional community*, Cybernetics and Systems: An International Journal, 41(1), 2010, pp. 17-30.
[17] MEDINA J.R., LAWITZKY M., MORTL A., LEE D., HIRCHE S., *Anexperience-drivenrobotic assistant acquiring human knowledge to improve haptic cooperation*, In Intelligent Robots and Systems (IROS), 2011 IEEE/RSJ International Conference on IEEE, 2011, pp. 2416–2422.
[18] MULCHANDANI D., *Java for Embedded Systems*, IEEE Internet Computing, Vol. 2, No. 3, May/June 1998, pp. 30–39.
[19] NEMEC B., VUGA R., UDE A., *Exploiting previous experience to constrain robot sensorimotor learning*, In Humanoid Robots (Humanoids), 2011, 11th IEEE-RAS International Conference on. IEEE, 2011, pp. 727–732.
[20] PASTOR P., RIGHETTI L., KALAKRISHNAN M., SCHAAL S., *Online movement adaptation based on previous sensor experiences*, In Intelligent Robots and Systems (IROS), 2011 IEEE/RSJ International Conference on IEEE, 2011, pp. 365–371.
[21] SANIN, C., and SZCZERBICKI, E., “Implementing Decisional Trust: A First Approach For Smart Reliable Systems”, Cybernetics and Systems: An International Journal 40 (2):85 – 98, 2009.
[22] SANIN C., MANCILLA-AMAYA L., SZCZERBICKI E., CAYFORDHOWELL P., *Application of a Multi-domain Knowledge Structure: The Decisional DNA*, In Intelligent Systems for Knowledge Management, edited by N.T. Nguyen and E. Szczerbicki, Springer, Berlin–Heidelberg, Vol. 252. Original edition, 2009, pp. 65–86.
[23] Tvheadend, Overview, Retrieved October 2012 from https://www.lonelycoder.com/tvheadend/
[24] TVHGuide, About, Retrieved October 2012 from http://john-tornblom.github.com/TVHGuide/index.html
[25] PCWorld, Toshiba Thrive 16GB, http://www.pcworld.com/article/235696/toshiba_thrive_review_a_tablet_edges_closer_to_the_ideal.html
[26] WU Y., HIRAKAWA S., REIMERS U., WHITAKER J., *Overview of digital television development worldwide*, Proc. IEEE, Vol. 94, No. 1, 2006, pp. 8–21.

Decisional DNA, Set of Experience Knowledge Structure, Industrial Design, Virtual Engineering Objects

Syed Imran SHAFIQ*, Cesar SANIN*,
Edward SZCZERBICKI**, Carlos TORO***

USING DECISIONAL DNA TO ENHANCE INDUSTRIAL AND MANUFACTURING DESIGN: CONCEPTUAL APPROACH

During recent years, manufacturing organizations are facing market changes such as the need for short product life cycles, technological advancement, intense pressure from competitors and the continuous customers' expectation for high quality products at lower costs. In this scenario, knowledge and its associated engineering/management of every stage involved in the industrial design has become increasingly important for manufacturing companies in order to improve their performance and to take effective decisions. Knowledge based industrial design techniques have been used in the past with fair bit of success but they have their share of limitations like they may be time consuming, costly, domain specific and at times not very intelligent. This paper proposes a novel approach where the set of experience knowledge structure (SOEKS) and decisional DNA (DDNA) techniques are used for the experience based representation of engineering artefacts. The knowledge representation method that we propose; captures and re-uses the product and process involved within an engineering knowledge perspective. We also introduce in this paper a new concept that we call Virtual Engineering Object (VEO), which is a conceptual entity that permits a dual computerized/ real world representation of an engineering entity. It contains the embedment of the decisional model expressed within the set of experience, a geometric representation and the necessary means to relate such virtualization with the physical object being represented. VEO will act as a living representation of the object capable of adding, storing, improving and sharing knowledge through experience, in a way similar to an expert of that area.

1. BACKGROUND

The term industrial and manufacturing design used in this paper can be defined as an integration of the knowledge of product and process of an object, to demonstrate its design and manufacturing functions. Industrial design (ID) is a complex process in-

* The University of Newcastle, University Drive, Callaghan, 2308, NSW, Australia.
** Gdansk University of Technology, Gdansk, Poland.
*** Vicomtech-IK4, San Sebastian, Spain.

volving knowledge of various fields. Increasingly competitive and demanding markets are forcing companies to search for means to decrease time and costs for new product development, while satisfying customer requirements and maintaining design quality [1].

A successful industrial design is one which;

- meet a high number of product requirements,
- enable robust manufacturing with high quality,
- deliver sound profits on competitive markets,
- fulfil customer expectations,
- enable a sustainable future [2].

A well-recognized feature of design is that a large percentage of the product's life cycle time is spent on the routine tasks as it takes up-to 80% of the design time. It is noted, 'around 20% of the designer's time is spent searching for and absorbing information', and '40% of all design information requirements are currently met by personal stores, even though more suitable information may be available from other sources' [3]. *This implies that design information and knowledge is not represented in a shared and easily accessible knowledge base.*

There are a number of strong arguments for adopting computer integrated knowledge based manufacturing system to meet the above discussed features for successful ID. One of the hallmarks of the knowledge base systems is to automate repetitive, non-creative design tasks. Not only does automation permit significant time and cost savings, it also frees up time for creativity, which allows exploration of a larger part of the design envelope. Clearly, in such cases *knowledge re-use guided framework may save considerable time and effort* [1].

Due to the complex nature of modern industrial design there are few challenges of using knowledge based manufacturing system like;

- The knowledge for the desired application is not available.
- The technology in the design process is constantly changing.
- Knowledge outputs are not compatible with other systems.
- Knowledge outputs not easily understandable and readable.
- Knowledge capture is not performed in real-time, adding new knowledge to the repository is performed manually [4].

In this paper we introduce a smart knowledge based decision support tool, Set of Experience Knowledge Structure (SOEKS) and Decisional DNA (DDNA) [5]. The SOEKS has been developed to store formal decision events explicitly. It is a model based upon existing and available knowledge. *It can be described as a knowledge structure to retain explicit experiential knowledge* [6].The proposed concept of VEO will be powered by SOEKS and DDNA. That means a VEO will not only be a knowledge repository but it will have qualities like self-awareness and reflexivity embedded inside it. And all these features will enable a VEO to behave as a live object.

The structure of this paper is as follows. The section 1, describes the concepts and structure SOEKS and DDNA. In section 3, we introduce the idea of semantic in ID. In section 4 the architecture of VEO is discussed. After discussing the concepts and implementation in section 5, we discuss conclusions of this paper in section 6.

2. SET OF EXPERIENCE KNOWLEDGE STRUCTURE (SOEKS) AND DECISIONAL DNA (DDNA)

As discussed in section 1, a large amount of previous knowledge is needed to design a new component; the information may not be exactly the same but may be from the family of the related object. But it has been observed that not much effort is made to retain the knowledge. Knowledge and experience are lost indicating that there is a clear deficiency on experience collection and reuse. Some of the reasons are:

a) the non-existence of a common knowledge-experience structure able to collect multi-domain formal decision events, and
b) the non-existence of a technology able to capture, store, improve, retrieve and reuse such collected experience [7].

Sanin and Szczerbicki proposed a new smart knowledge based decision support tool, having three important elements:

a) a knowledge structure able to store and maintain experiential knowledge, that is, the SOEKS and the Decisional DNA,
b) a solution for collecting experience that can be applied to multiple applications from different domains, that is, a multi-domain knowledge structure, and
c) a way to automate decision making by using such experience, that is, retrieve collected experience by answering a query presented [5, 6].

The SOEKS is a compound of variables (V), functions (F), constraints (C) and rules (R), which is uniquely combined to represent a formal decision event. Functions define relations between a dependent variable and a set of input variables; therefore, SOEKS uses functions as a way to establish links among variables and to construct multi-objective goals (i.e., multiple functions). Similarly, constraints are functions that act as a way to limit possibilities, restrict the set of possible solutions, and control the performance of the system with respect to its goals. Finally, rules are used to represent inferences and correlate actions with the conditions under which they should be executed. Rules are relationships that operate in the universe of variables and express the connection between a condition and a consequence in the form if then else [8].

Chromosomes are groups of set of experience (SOE) that can accumulate decisional strategies for a specific area of an organization. Multiple SOE can be collected, classified, and organized according to their efficiency, grouping them into decisional chromosomes. Finally, sets of chromosomes comprise what is called the Decisional DNA of the organization as shown in Figure 1.

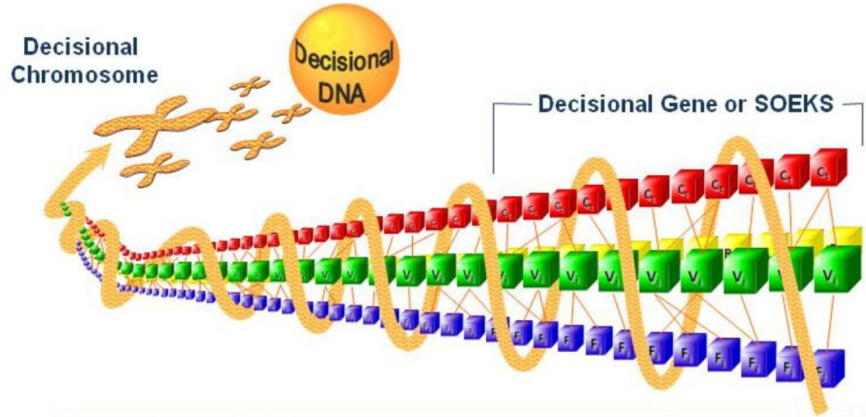


Fig. 1. SOEKS and Decisional DNA [7, 9]

3. SEMANTICS IN ID

Semantics is the discipline that studies the meaning of things. Semantic technologies compose some of the most motivating technologies resulted from the World Wide Web revolution that is frequently reviewed in different areas of knowledge engineering. Semantic web is the new-generation web that tries to represent information such that it can be used by machines not just for display purposes, but for automation, integration, and reuse [10].

Semantics in ID is based on the description of the states that can be identified in a manufacturing process (situations), the task to be performed in each situation (actions) and the rules to determine the next situation after a task is executed (decisions). The semantic representation of information enables the creation of intelligent systems, which can interpret and understand potentially automated tasks, harnessing added-value decision-making processes. Particularly, the semantic web can provide a cutting-edge formal representation and knowledge-driven set of technologies to enable automation of industrial manufacturing processes [11].

4. VIRTUAL ENGINEERING OBJECT (VEO)

According to our definition, a *Virtual Engineering Object (VEO) is a knowledge representation of an engineering artefact comprising experience models, domain and functionality along a physical attachment to the virtual object in its conceptualization.* VEO model intends to be the most complete possible model for a specified domain and can be used in multi domains.

As discussed in section 1, large portion of the design time is spent on the routine tasks, which have already been designed previously. Not much work has been done to store and reuse this information. Lot time and money is wasted in searching this information. This approach will automate repetitive and non-creative design tasks.

The objective of this study is to develop the concept of VEO using SOEKS and DDNA. A VEO should have all the possible feature of an artefact along with its past experience embedded in it.

VEO will enhance the industrial and manufacturing design, as it will be the union of knowledge and experience. It will embody information of all the aspects involved in the manufacturing of that particular object like process, resource and calculations related to it. DDNA will have the experiential knowledge of its characteristics, requirements, functionality, connections and present state of that object. VEO model is discussed in detail in the section 5.

5. CONCEPTUAL ARCHITECTURE AND FORMULATION OF VEO

As discussed in previous section that a VEO will have a knowledge base of an object then we will infuse the experience of DDNA in to it. Figure 2 show the architecture of a proposed VEO. This framework includes preliminary design and preliminary process knowledge. This model will be having manufacturing information on preliminary process planning, such as processes, sequences, parameters, cost/time etc. The experience of DDNA will provide the associated knowledge to the above parameters, which in turn make the decision making easier and intelligent [5].

Process knowledge is classified into three main types based on their forms, as shown in Fig. 2 [12]. They are:

(1) *Knowledge of process:* This kind of rule-based knowledge includes the feature process, the product process, and the typical process.

(2) A feature is the definition of a component’s basic geometric entities for manufacturing which can include cylinder, hole, plane, etc.

(3) Product process knowledge refers to process route information of a product family or similar products, which may change according to the input manufacturing data.

(4) The typical process knowledge is the mature process route information which has been validated by practice and normally used more frequently.

(5) *Knowledge of resource*: This refers to static manufacturing resource information, which includes all kinds of process resources, such as machine tools, fixtures, cutters, machining data, and materials.

(6) *Knowledge of calculation*: This refers to knowledge obtained through calculation. In process planning, the selection of working hours and material quota is a regular process.

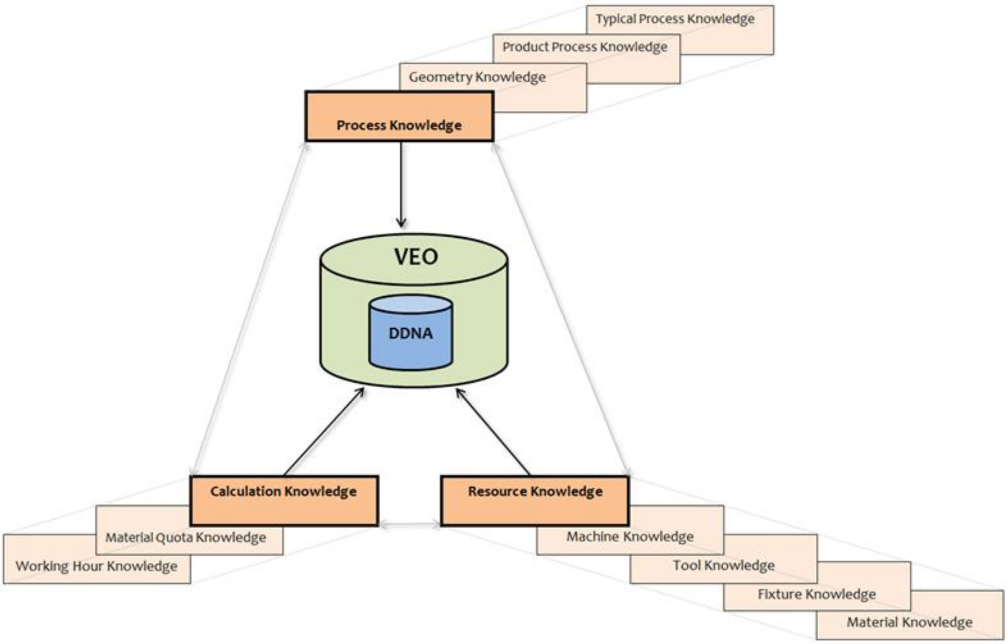


Fig. 2. Architecture of the proposed VEO

DDNA will have the following experience in it:

(1) *Characteristics* describe not only the set of expected benefits offered by the process, resource, calculation selected for a particular object, but also artefact represented by the Virtual Engineering Object. Such characteristics will depend on what kind of artefact are we considering. For example characteristics of a process can be the relative motion between a tool and a work-piece. While that of a resource, like tool is what specific actions it takes.

(2) *Requirements* describe the set of necessities of the knowledge base and the Virtual Engineering Object for its correct working. The set of requirements depends on what kind of VEO are we considering; for a Lathe machine work holding device, tool holder, power source etc. as the essential requirements.

(3) *Connections* describe how the Virtual Engineering Object is related with other VEOs. These connections can be of different types. Some of them can be a need relationship, e.g. a gear is part of an engine or the machining parameters like feed, depth of cut and cutting speed of different machines.

(4) *The present state* of the Virtual Engineering Object indicates parameters of the VEO at the current moment. For example, questions like *how* much time has been this machine powered on? Or what is the current performance of this machine, cell, shop etc.?

As an example Figure 3 show knowledge structure of a manufacturing unit. If we want to extract VEO for a drilling, the shaded modules will be required as shown in the figure.The design information, which includes the requirements, behavior, function, form and structure of an artifact, will be provided for product design based on the manufacturing process model, and to enhance design specification.

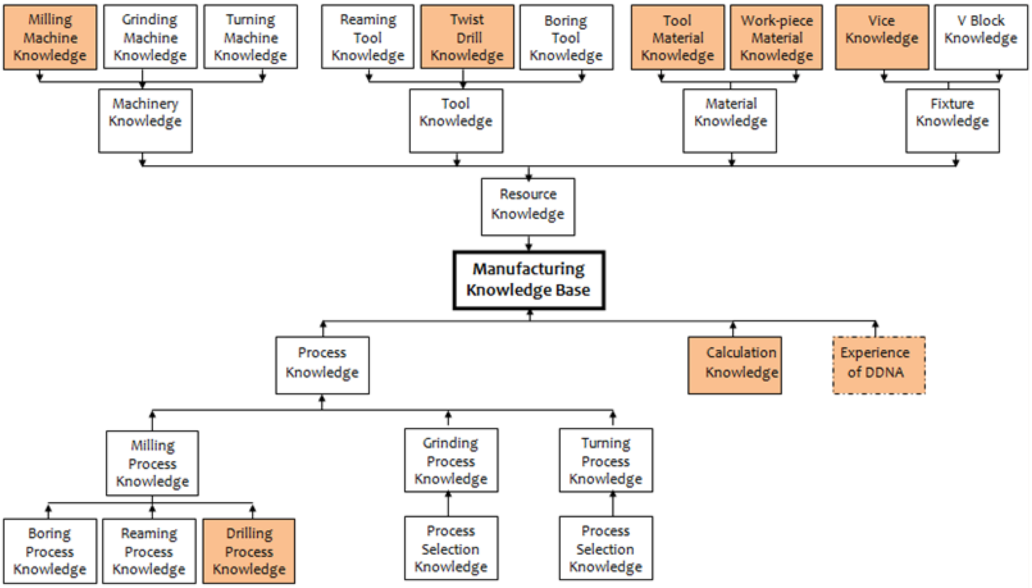


Fig. 3. Knowledge structure of a manufacturing unit

The dynamic and static knowledge and experience will be stored in the chromosome VEO of a drilling machine. Each time a manufacturing, maintenance or repair operation is done on the drilling machine the experience and knowledge is added in the VEO.

The VEO of a drilling machine will be a combination of various other chromosomes and VEO like the VEO of a tool used and the VEO of a work piece holding device. These different VEO's will be joined according to common link or functionality.

This VEO will behave as an expert of the drilling machine. When an operator will try to use this drilling machine of a particular machining operation, he can consult the VEO and it will provide all the possible knowledge, based on the previous experience on this machine.

On final integration of all the knowledge base of the design process and all the VEO, the virtual engineering object of the desired product is obtained. This VEO of the product will have experience and knowledge involved at each and each every stage of the product design process right from its inception to the final product. This VEO will keep evolving with each formal decision taken related to the particular product.

6. IMPLEMENTATION

Using, channelizing and exploiting previous knowledge in industrial design is an area which is not been researched extensively as yet. Sanin and Szczerbiwki have applied the concept Set of experience knowledge structure (SOEKS) and Decisional DNA in various other domains [13]. The challenge for this research is to apply this concept the manufacturing and design area on framework discussed in section 5.

The proposed approach to the solution of this issue is to systematically create, capture, reuse, and distribute experience in the work processes of an organization, preventing important decisional steps from being forgotten in the daily operation or research tasks, and supporting a path towards appropriate automation for recurring tasks or findings.

The main objective of this work is to apply the concept of SOEKS/DDNA on the industrial design process. On the conceptual level, set of experience will be applied on every phase of the design process i.e. mechanism to capture knowledge experience and formal decision taken in day to day operations will be developed and then it will be reused. All the important features of every design stage and its integration with SOEKS will be elaborated.

The practical implementation of SOEKS/DDNA is done on the manufacturing and production phase. Principle adopted for implementing SOEKS at every stage is the dividing a system into subsystems or modules [14]. The knowledge base of the production phase of the product design is broadly divided into Resource knowledge, Process knowledge and calculation knowledge, which are further divided into subsystems as shown in figure 2. When we integrate the experience of DDNA with this knowledge base, the concept of virtual engineering object (VEO) is developed.

In this project, we focus on the adaptation and creation of new algorithms to perform good virtual experience analysis, based on the idea of introducing semantics into the data modeling and processing in industrial design and manufacturing system. ID user knowledge gathering is still not addressed from a semantic point of view. SOEKS would able to model it. We need to accomplish four key tasks to enhance ID by utilizing SOEKS and Decisional DNA:

Task 1: Data Pre-processing and Semantic Representation. Knowledge is a combination of experiences expressed in terms of values, related information, and expert insight, which provides the framework for evaluating and incorporating new experiences and knowledge. Thus, there are four basic components that surround decision-making events: variables, functions, constraints, and rules. We need to propose new ways to pre-process and represent industrial design or manufacturing processes in the forms of Decisional DNA.

Task 2: Data Collection and Generalization. Establish techniques for collecting industrial design or manufacturing processes knowledge and transform it into Decisional DNA according to the models developed in Task 1. Afterwards, mixing of the collected knowledge and establishing a combination model for formal decision events is required.

Task 3: Evolving Knowledge Base on Industrial Design. Combined models from Task 2 open possibilities for knowledge evolving techniques based on industrial design which leads to sub-solutions. The developed approach will reduce the gap between different proposed knowledge trying to solve similar problems and will look for a holistic encounter point of their solutions.

Task 4: Establishing Techniques for Manipulating, Administrating and Sharing of Collected Engineering Design Decisional DNA. Decisional DNA as a knowledge representation for formal decision events is recognized as the fundamental component of infrastructure for advanced approaches to intelligent knowledge management and knowledge engineering automation. The experience-based knowledge structure represented by Ontologies will facilitate the achievement of this aim while allowing for an easy sharing of Decisional DNA[15].

7. CONCLUSION

Existing conceptual approach have proposed to enhance ID and manufacturing system in industrial plants by supporting decision making system and systemize organization of knowledge. This knowledge management approach introduced knowledge based structure and architecture to enhance ID by utilizing SOEKS and Decisional DNA in manufacturing system. The concept of VEO will behave a knowledge and experience repository. These VEO’s of explicit knowledge can be shared among similar organizations, industries, and partners to build up a decisional repository. This decisional repository will save significant time and money as the right information, in the right format and at the right will be available.

The main contribution of our idea is the procedure to classify types of manufacturing design knowledge and knowledge structure to support manufacturing knowledge maintenance, also an appropriate methodology to utilize manufacturing knowledge models to industrial design by utilizing SOEKS and Decisional DNA.

REFERENCES

[1] VERHAGEN Wim J.C., GARCIA P.B., v. DIJK R.E.C., CURRAN R., *A critical review of Knowledge-Based Engineering: An identification of research challenges*, *Advanced Engineering Informatics,* 2012, pp. 5–15,.
[2] STARK R., KRAUSE F.L., KIND C., ROTHENBURG U., MÜLLER P., HAYKA H.*, et al.*, *Competing in engineering design – The role of Virtual Product Creation*, CIRP Journal of Manufacturing Science and Technology ,Vol. 3, 2010, pp. 175–184.
[3] BAXTER D., GAO J., CASE K., HARDING J., YOUNG R., COCHRANE S.*, et al.*, *An engineering design knowledge reuse methodology using process modelling*, International Journal of Research in Engineering Design, Vol. 18, 2007, pp. 37–48.
[4] SUNG R.C.W., RITCHIE J.M., LIM T., KOSMADOUDI Z., *Automated generation of engineering rationale, knowledge and intent representations during the product life cycle*, Virtual Reality, 2012, Vol. 16, pp. 69–85.
[5] SANIN C., SZCZERBICKI E., *Decisional DNA and the Smart Knowledge Management System: A process of transforming information into knowledge*, A. Gunasekaran (Ed.), *Techniques and Tools for the Design and Implementation of Enterprise Information Systems,* 2008, pp. 149–175.
[6] SANIN C., SZCZERBICKI E., *Set of Experience: A Knowledge Structure for Formal Decision Events,* Foundations of Control and Management Sciences, Vol. 3, 2005, pp. 95–113.
[7] SANIN C., SZCZERBICKI E., *Towards the Construction of Decisional DNA: A Set of Experience Knowledge Structure Java Class within an Ontology System*, Cybernetics and Systems, 2007, Vol. 38.
[8] SANÍN C., *Smart Knowledge Management System*, Thesis of Doctor of Philosophy Degree From The University Of Newcastle, Department Of Mechanical Engineering, Newcastle, Australia, 2007.
[9] SANÍN C., Toro C., Haoxi Z., Sanchez E., Szczerbicki E., Carrasco E.*, et al.*, D*ecisional DNA: A multi-technology shareable knowledge structure for decisional experience*, Neurocomputing, Vol. 88, 7/1/ 2012, pp. 42–53.
[10] TORO C.A., *Semantic Enhancement of Virtual Engineering Applications*, Thesis of Doctor of Philosophy Submitted to the department of Computer Science and Artificial Intelligence At The University of the Basque Country Donostia – San Sebastian, 2009.
[11] GARCIA-CRESPO A., RUIZ-MEZCUA B., LOPEZ-CUADRADO J.L., GOMEZ-BERBIS J.M., *Conceptual model for semantic representation of industrial manufacturing processes*, Computers in Industry, Vol. 61, 9// 2010 pp. 595–612.
[12] CHEN W.L., XIE S.Q., ZENG F.F., LI B.M., *A new process knowledge representation approach using parameter flow chart*, Computers in Industry, Vol. 62, 1//2011, pp. 9–22.
[13] SANÍN C., MANCILLA-AMAYA L., SZCZERBICKI E., CAYFORDHOWELL P., *Application of a Multi-domain Knowledge Structure: The Decisional DNA*, in *Intelligent Systems for Knowledge Management,* Vol. 252, N. Nguyen and E. Szczerbicki, Eds., ed: Springer Berlin Heidelberg, 2009, pp. 65–86.
[14] STONE R.B., *Towards a Theory of Modular Design*, Doctor of Philosophy, Faculty of the Graduate School of The University of Texas at Austin, The University of Texas At Austin, 1997.
[15] TORO C., VAQUERO J., GRAÑA M., SANÍN C., SZCZERBICKI E., POSADA J., *Building Domain Ontologies From Engineering Standards*, Cybernetics and Systems, Vol. 43, 2012/02/01 2012, pp. 114–126.

civil aviation management, soft modelling, airline business

Aleksander ORŁOWSKI*, Edward SZCZERBICKI**

MODELLING PROCESSES FOR POLISH CIVIL AVIATION MARKET

The present , paper is the first step in research of Polish civil aviation market, focusing on the area of civil, legacy airlines. First, the market is described with its players, current conditions and its main problems. Next, the tools which are currently used for the prognosis processes are presented. Finally, the idea of a new model (fuzzy model) for the process of prognosis is presented.

1. INTRODUCTION

This article is the first step of works in which the authors are going to deal with a new scientific area: airline business management.

For the last five years one of the authors has concentrated on the area of financial services in the Internet, mostly the process of modeling its services using fuzzy knowledge. In time the authors found out that there are other markets that require new methods of modeling due to their fast changes. Civil aviation and airlines have been a hobby of the authors but also that market has been incredibly fast changing for 10 years. Due to that the authors decided to do research using their knowledge about the processes of modeling and the knowledge of the new market. This paper is the introduction to the new level of research.

* Gdańsk University of Technology, Faculty of Economics and Management.
** The University of Newcastle, Newcastle, Australia.

1.1. THE AIRLINE BUSINESS MARKET

The world's first airline is *Deutsche Luftschiffahrts-Aktiengesellschaft,* founded in 1909. The oldest current existing airlines are KLM and Quantas, both founded in 1920. LOT Polish Airlines was established in 1929 is one of the oldest, still operating, airlines.

In the year 2011 there were 2,8 billion passengers handled by air and more than 150 airlines which made 5,2 trillion kilometers [11].

Till the second world war, airline business was growing but still was a small part of market, a very expensive one. In the 30s of XX century it was growing so fast that for example LOT Polish Airlines was planning to open its first transatlantic (TATL) route in 1940, which did not happen due to the outbreak of World War II.

In mid 70s, British Airways started offering inexpensive transatlantic flights, which make flying popular and closed the era of transatlantic ships [10]. In the same time the deregulation of airline business has started, which led to the opening of new airlines, dropping fares and growing number of customers.

The recession in the beginning of 90s was the time in which some major airlines collapsed (PanAm, TWA). In next ten years low-cost airlines started growing in Europe changing completely the existing market. Today international airliner business is called "complex, dynamic and subject to rapid change and innovation" [1].

"A period of uncertainty" [2] is how the beginning of the twenty first century in the airline business was described by Rigas Doganis, former CEO of several airlines, currently visiting professor of the universities in the United Kingdom. Several aspects contribute to the uncertainty, first and currently most crucial are fuel prices.

1.2. FUEL PRICES

The prices of fuel have been the most significant aspect affecting overall level of costs in the airlines in last 30 years. This factor is so important that it is strongly correlated with the level of profit of the airlines. When the aviation fuel prices were falling (fig. 1) the airlines had record profits (years 1997 and 1998). Later on the prices had risen dramatically (the prices tripled between the year 2002 and 2006) and then were still rising, keeping now the high level. Basing on the observations from the past 30 years one might notice that fuel prices represent from 15% of total costs of airlines (in the years when petrol was still cheap) up to 30% of total airlines costs. The difference is significant and puts strong influence on the airlines business, which makes it one of the major aspects that should be taken into consideration in the tools which help airlines to make prognosis and plan correctly.

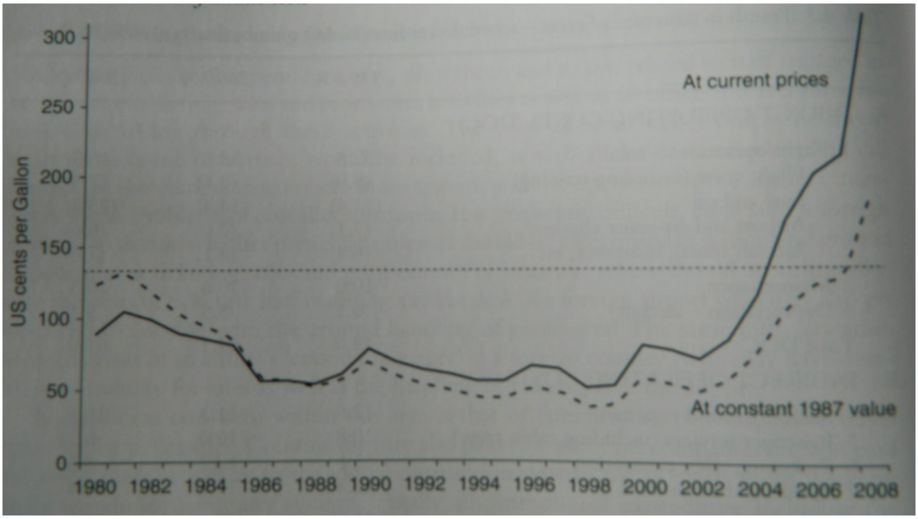


Fig. 1. Average fuel price in current and constant 1987 US cents [3]

1.3. ECONOMIC CYCLES TURNS

Economic cycle turns are very important for the world's airline business. That business is extremely related with the current economic conditions. Fig 2. illustrates the total profits of the world's airlines, starting from the good market conditions at the end of the eighties, then the crisis in the years 1990–1993, then the airlines starting to come back to profits (together with the world's economy) up to the year 2001, which marked the end of the "dot com era"- the attacks in New York on September 11th 2001.

Generally, airline business is often called the one that reflects the current economic situation. When the economy gets worst, people tend to travel less, both on business and leisure trips, which can be seen in the number of sold tickets, especially the ones in business class (called later C class) and also in falling revenues. Due to that it is crucial for the airlines to make proper prognosis about the upcoming economic conditions which will directly influence their profits.

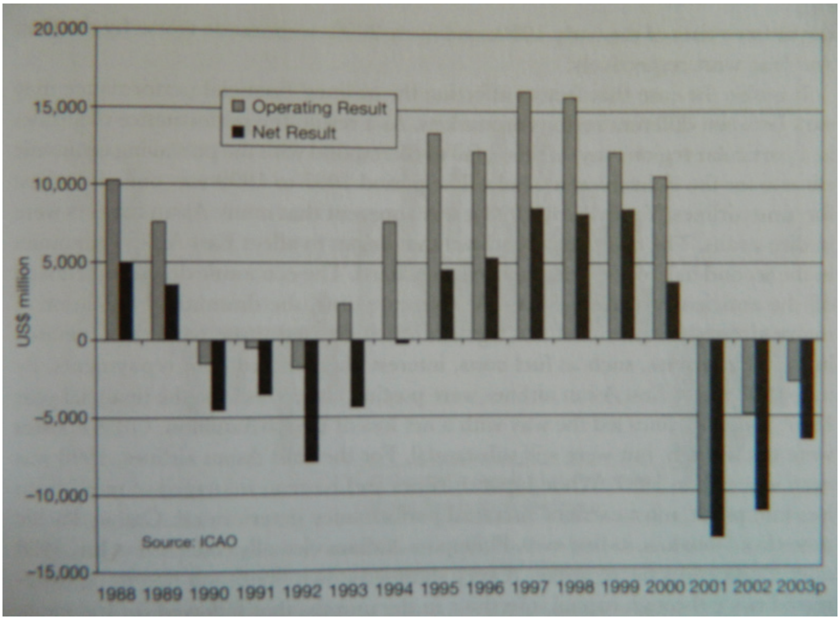


Fig. 2. Profits or loss of the world's airlines, 1998–2003.
Source: Doganis R., The airline business, 5.

1.4. EXTERNAL EVENTS

Airlines business, as mentioned in the previous part, is very sensitive to economic cycle turns but so it is to other external events. Looking at the last 13 years one might notice at least three major external events that seriously affected the aviation business. Starting from September 11th 2001, terrorist attracts influencing not only economic cycle turn but directly hitting the whole civil aviation business area. There was a huge drop in the number of passengers travelling by planes, especially in the United States where most of the airlines had to use Chapter 11, law protecting them from bankruptcy.

In the end of 2002 and 2003 the SARS (Severe Acute Respiratory Syndrome) virus appeared in Asia and resulted in a major drop in the number of passengers travelling to/from Asia which was (and still is) the fastest growing region in airline business. Later on there were smaller regional problems which resulted in a major drop of number of passengers in several countries (like the virus in Mexico).

Finally, in 2010 the volcano eruption in Island grounded for more than a week almost all airplanes in Europe, which led to major financial losses (e.g. LOT Polish Airlines lost around 30 million zloty (around 10 million US dollars).

Those three external events were all unexpected and hit mostly and strongly the airline business, which suggests the number of different aspects to be taken into consideration while managing and making prognosis for that type of business.

1.5. TECHNOLOGICAL CHANGE

Technological change which was made in the last 50 years has changed the whole industry, starting from the development of jet engines for the use in civil aviation, ending at bigger airlines with higher speeds using that jet engines (which makes higher aircraft productivity). Today those changes might be seen in a nutshell with the Boeing 787 Dreamliner consuming around 20% less fuel then similar size aircrafts existing on the market.

Another important aspect of using new technologies was the new approach to ticketing. Previously, all tickets were printed by the airline on special paper and the average cost of each paper ticket (the ticket itself together with the work involved) was around 9$ [5]. Today almost all tickets come to a customer in electronic version, which reduces the cost to about 1$ each. LCC airlines have also created their own systems of selling tickets (attempting to omit travel agents, typical offices, and large international systems for ticketing like Amadeus) which helped them to reduce ticketing costs by 30%. All changes in ticketing have been made in last 15 years, which shows how strong the market can change due to technological innovations.

1.6. STRUCTURE OF INDUSTRY

For many years one type of airlines prevailed on the market – the airline flying from one city to another, having one (or more) main airports (hub) to and from which they transported passengers. In that way the airlines created a network enabling passengers to travel from one city to another via hub (where they changed the airplane). That type of airlines is called “network”. Currently British Airways (with the hub London Heathrow), Singapore Airlines (Singapore Changi airport), Lufthansa (Frankfurt/Main and Munich) or LOT Polish Airlines (Warsaw Chopin airport) are the examples of that type of airlines.

In the year 1971 Southwest Airlines was launched in the United States. It started operating short flights in Texas- only point-to-point flights (no hubs) with high frequencies and simple product. “The aim of Southwest business model was to offer a simple, uncluttered, low-fare, point to point product using a single type aircraft fleet, while achieving very high aircraft and crew utilization by operating from secondary and uncongested airports.” [4] Because of that it can achieve operating costs 30-60 per cent lower than network airlines flying on the similar routes on the similar aircraft.

That type of airline is called Low Cost Carrier (LCC), currently there are more than 50 LCC in the world, such as Southwest Airlines or Jet Blue in the United States and Ryanair, EasyJet or Wizzair in Europe. To understand how much that type of airlines has changed the airline market it is worth to analyze the example of Poland and its airline market. LCC airlines entered the market in the year 2005 and according to the data from June 2013 the biggest airline in Poland is currently Ryanair (fig. 4), putting LOT Polish Airlines on the second place. In Poland LCC airlines control more than 52 per cent of market shares, which shows how much the new type of business has changed the market in a short time and reflects changes in current existing network airlines and their type of business. According to all prognoses LCC airlines will increase their market shares, they also plan to extend their business to new types of business (e.g. Norwegian started in June 2013 operating flights between Europe and the United States) and those changes must be taken into consideration while planning and making prognoses for each airline.

Liczba pasażerów obsłużonych w polskich portach lotniczych w krajowym i międzynarodowym ruchu regularnym w pierwszym kwartale 2012 i 2013 roku

	2013			2012		
Przewoźnik	l.p.	liczba pasażerów	udział	l.p.	liczba pasażerów	udział
Ryanair	1	1 220 212	28,11%	2	867 758	21,44%
LOT Polish Airlines + EuroLOT	2	1 215 204	28,00%	1	1 209 896	29,89%
Wizz Air	3	846 740	19,51%	3	857 981	21,20%
Lufthansa	4	311 109	7,17%	4	308 141	7,61%
EasyJet	5	100 321	2,31%	5	97 194	2,40%
Air France	6	74 950	1,73%	7	72 714	1,80%
Norwegian Air Shuttle	7	70 606	1,63%	8	64 335	1,59%
SAS	8	70 231	1,62%	9	61 681	1,52%
KLM Royal Dutch Airlines	9	47 761	1,10%	10	49 296	1,22%
EuroLOT	10	45 918	1,06%	6	95 889	2,37%
Swiss International Air Lines	11	43 521	1,00%	11	41 212	1,02%
Aeroflot-Russian Airlines	12	39 027	0,90%	15	32 332	0,80%
British Airways	13	34 577	0,80%	13	33 903	0,84%
Austrian Airlines	14	33 187	0,76%	14	33 451	0,83%
Air Berlin	15	27 640	0,64%	27	5 990	0,15%
Emirates Airline	16	22 916	0,53%	–	–	–
Finnair	17	21 227	0,49%	17	19 102	0,47%
Qatar Airways	18	18 252	0,42%	–	–	–
Turkish Airlines (THY)	19	14 856	0,34%	19	15 836	0,39%
Aer Lingus	20	13 273	0,31%	12	37 854	0,94%
Alitalia	21	12 812	0,30%	20	14 291	0,35%
Czech Airlines	22	11 082	0,26%	16	21 485	0,53%
TAP Portugal	23	10 459	0,24%	23	10 829	0,27%
EL AL Israel Airlines	24	8 601	0,20%	25	6 989	0,17%
Air Baltic	25	6 216	0,14%	24	8 470	0,21%
Pozostali przewoźnicy		19 788	0,46%		80 625	1,99%
Przewoźnicy polscy*		***1 266 864***	*29,19%*		***1 337 444***	*33,05%*
LCC**		***2 276 017***	*52,44%*		***1 907 964***	*47,14%*
SUMA		**4 340 486**			**4 047 254**	

Źródło: Opracowanie ULC na podstawie informacji uzyskanych z portów lotniczych, Warszawa, czerwiec 2013

Fig. 3. The number of passengers transported by airlines on the Polish markets, January-June 2013, source: Urząd Lotnictwa Cywilnego (National Civil aviation office).

All aspects presented above should be taken into consideration in the process of planning (and prognosing) by all airlines that intend to successfully exist on the market. To make it properly it is necessary to use dedicated models that should help managers in making decisions. Models which are currently in use by airlines will be presented in the next chapter.

2. TOOLS FOR PROGNOSIS

Forecasting seems to be the most important area of airline management. It must describe the demand with the supply which is offered for the 6-18 months time period. The decisions such as choosing the type of aircraft for a given route, opening new routes, reducing/adding frequencies, are the types of tactical decisions which are based on prognoses. All of them strongly influence the revenue generated by the airline, so they must reflect the obvious characters of every track like tourist trends, business units in the area, economic growth of the countries and number of expatriates. But the prognoses always take into consideration the category which is called “all other things”. The problem of airline business is that the influence of that category is marked by experts as stronger than in other branches of business. External factors, of which the most important were mentioned in the previous chapter, create much uncertainty in the airline business.

This chapter will present the techniques and models which are currently in use by airlines for generating their prognoses.

2.1. MODELS

There are several types of models used by airlines, out of which, the most frequently used ones will be presented in this chapter.

Regression models – most currently used ones are based on simple or multiple regression models, in which traffic is a function of different independent variables.

Therefore airlines use quite simple econometric models which help predicating the traffic on a chosen route. Historical data helps predicting the traffic on the same route in the long time run (up to ten years). Historically these methods turned out successful but currently they are found less and less effective for the reasons mentioned below.

First, it is very hard to use that method when the airline is entering a new route or a new country (no historical data). It is possible to use data gained from other airlines but it is complicated the method does not work properly when rapid changes are happening on the market (for example fast growth of the market, liberalizing the market (no restrictions)). In such cases the method does not work properly. That type of method works well when it operates on a large scale, for example it is used by companies which operate several airports (for example in the United Kingdom). In that way the company may forecast the number of passengers traveling domestic, short-houl and long-houl flights, which is the practical usage of that method. But it is not the solution for airlines trying to forecast all routes from their network.

To predict the number of passengers on new routs airlines try to use the combination of market research and executive judgment. Some airlines also try to use gravity models, a method which used to be very popular in road and railway traffic prediction. It was successfully used by the European Commission to plan air traffic between the cities that did not have direct connections, also some airlines (e.g. Lufthansa) created very complex gravity models that were expected to help forecasting potential new routes in the network. It seemed that all the necessary models were created. But later on the LCC airlines started operating hundreds of routes that were not served previously (currently used models did not show their potential) and were very successful. LCC did not take into consideration historical data, they just generated new traffic basing on their very low fares, which points to a need to create new models.

Similar methods, are used by airlines from the Far East: Qatar, Emirates and Ethiad, which started operating on large aircrafts to many medium seize airports in Europe.

For the reasons mentioned above, it seems necessary to make changes in currently used models to make it possible and take into consideration both the methods used by LCC and the airlines from the Far East which is a must for the airline to survive and develop on the current airline market.

3. AIRLINE BUSINESS MODEL

Previous chapters presented factors that make airline business market very complicated, with several different variables influencing it. Next, methods which are currently in use by airlines to forecast market reactions with its limitations in current use were presented

Basing on the experience that authors gained while creating models, especially the models for very unstable markets, it is suggested to create a special, dedicated model for airlines to aid making prognosis which will enable effective route planning. The Airline Business Model (ABM) is supposed to be a model, used by airlines to plan every single route which the airline is currently operating (to check the demand and due to that the size of aircraft that will fit that route) or is interested in operating (with current competitors or no airline operating).

Basing on the experience gained in previous attempts to modeling [6, 7] it seems necessary to choose a method of modeling. Looking at the currently used techniques (presented in the previous chapter) and the number of variables which are hard to define in numbers (e.g. current economic situation) there is a need to take those variables into consideration. Authors suggest using the fuzzy modeling [8], with traditional process of building fuzzy model, consisting of three steps: fuzzification, fuzzy inference, and defuzzification [9].

Creating a list of variables requires an in-depth analysis and was not a part of this paper. Basing on the current level of works it should take into consideration variables like:

- Traffic – historical traffic recorded on the route, if it existed before
- GDP growth in the countries weightened in proportion of shares of the nationalities in total traffic
- Touristic potential of the cities
- Business potential of the cities
- Expatriates from one region recorded in the second destination

The initial variables will be chosen in cooperation with the managers of airlines, later on the variables will be tested in the model and the changes will be introduced.

4. CONCLUSIONS AND FUTURE WORK

This paper is an initial step in which the authors are moving into the new area of research: civil aviation management. The goal of the paper was to research the market, find its problems and describe future scientific aims.

The first part of the paper presents the description of the market with its history. Next, it is explained why the market tends to be called very complicated, with the presentation of different aspects influencing the market in the last 20 years. The second chapter consists of the presentation of methods and models which are currently used on the market to forecast the market situation in the medium term period (6-18 months). All presented methods are currently used by the airlines but none of them meets all current airlines expectations. Because of that the authors decided, to present in the third chapter, the idea of the model which should be created to meet the airlines' expectations, taking into consideration the complexity of the market which was presented in chapter 1.

Future research will concentrate on choosing the variables for the model, which is a crucial aspect for every fuzzy model. This step will involve direct contact with the airlines. Next, there is a need to build the model and to sign an agreement with the airline to gain the data necessary to fill the model and then to verify the model when it is made.

REFERENCES

[1] DOGANIS R., *Flying Off Course, Airline Economics and Managent,*Fourth Edition, London and New York, Routledge 2010, 1.
[2] DOGANIS R., *The Airline Business,* Second Edition, London and New York, Routledge, 2005, 258.
[3] DOGANIS R., *Flying Off Course, Airline Economics and Managent,*Fourth Edition, London and New York, Routledge 2010, 76.
[4] DOGANIS R., *Flying Off Course, Airline Economics and Managent,*Fourth Edition, London and New York, Routledge 2010, 134.
[5] DOGANIS R., *Flying Off Course, Airline Economics and Managent,*Fourth Edition, London and New York, Routledge 2010, 145.
[6] ORLOWSKI A., SZCZERBICKI E., *The role and concept of sub-models in the smart fuzzy model of the internet mortgage market*, Springer Verlag, Kaiserslautern 2011.
[7] ORLOWSKI A., SZCZERBICKI E., *Toward Fuzzy Model of Polish Internet Mortgage Market*, Cybernetics and Systems: An International Jurnal, 44:2–3, Taylor & Francis, 2013, 264–274
[8] PIEGAT A., *Modelowanie i sterowanie rozmyte,* Warszawa 1999.
[9] YAGER R., FILEW D*., Podstawy modelowania i sterowania rozmytego*, WNT, Warszawa 1995.
[10] http://traveltips.usatoday.com/history-airline-industry-100074.html , 15.07.2013.
[11] IATA Annual Report 2012, http://www.iata.org/about/Documents/annual-review-2012.pdf, 15.-7.2013.

expert systems, knowledge base, sailing boats,
computer advisory-decision process assist

Zbigniew BUCHALSKI*

SELECTION OF SAILING BOATS CONFIGURATION FOR VARIOUS USER GROUPS

This paper presents certain concept of expert system called SAILBOAT, which supports advisory-decision process in the company that makes sailing boats. At the beginning basic problems in design sailboats were recognized. These researches resulted in essential foundation of analyzed system establishment. Fundamental establishments of systems construction and functional description of SAILBOAT system is presented. The paper presents a computer implementation of this system and the results of testing the SAILBOAT system. The tests SAILBOAT system showed that this system successfully meets the goals and objectives put in its construction. It is easy to use, allows the efficient configuration of sailing boat, and above all the configuration of sailing boat takes much less time than the traditional way.

1. INTRODUCTION

In the era of information and knowledge, when there is an unlimited access to electronic information, contrary to common belief the process of making decisions has become more complex. The use of cheap and easily accessible computer techniques to solve complex decision problems has turned out to be extremely important.

Making decisions is an act of choosing one option (direction) of action out of a given set. This choice can be made based on a specific course of action which leads to finding the most advantageous (optimal) alternative. Intelligent informatics systems, such as expert systems play an important role in the process of supporting the decision-making process [1, 2, 5, 6, 7, 11, 12, 14, 15, 16, 17, 19, 20].

---

* Institute of Computer Engineering, Control and Robotics, Wrocław Univeristy of Technology, Wrocław, Poland, e-mail: zbigniew.buchalski@pwr.wroc.pl

Expert systems play a big role in informatics and they are used widely in many fields. They are successfully used as diagnostic, consulting, forecasting, classification and monitoring systems [3, 4, 8, 9, 10, 13, 18, 21, 22].

The purpose of this paper is to present some concepts of the expert system called SAILBOAT, supporting the work of a company producing sailboats. This system is designed as a tool to assist in decision-making process related to the creation of the appropriate configuration of sailboats tailored to customer requirements.

2. BASIC PRINCIPLES OF THE CONSTRUCTION OF THE SAILBOAT SYSTEM

Sailboats are a group of individuals with a long tradition, much diversity of design solutions and different purposes. Grading the sailboats different basis of division can be used. Among applied, the attention should be put on the destination of a boat (sports registration criterion, the criterion of use of the boat, the criterion of the class) and construction criteria (hull construction criterion, sailing equipment type).

Taking into account operation assumptions, boats can be divided:

- racing boats,
- tourist-racing boats,
- tourist boats,
- boats intended for training.

Racing boats – are designed mainly for racing in the Olympic triangle or another route for short distances or for long-distance race: inland, sea and ocean. The main common feature of all racing boats is to achieve maximum speed. Therefore, the racing boats are: large sail areas relative to displacement, relatively light hulls and equipment, and devices allowing the crew ballasted centerboard boat. Characterized by specialized and expensive facilities for the handling and navigation in racing conditions.

Tourist-racing boats – are designed for racing and for tourism. The design meets the requirements of the provisions of rule measurement formulas that take into consideration the long-distance and offshore races and also need to ensure adequate living conditions for the crew.

Tourist boats – boats of this group include mainly those, to design of which was not taken into account the advantages of racing, but only the safety and convenience of shipping. They can swim in the waters offshore, coastal or inland.

Boats for training – serve as a small training ship. They should be as cheap to build and operate as possible, while ensuring the greatest possible safety of navigation. They provide a large number of seats for passengers, sailing equipment should be easy to use, provided with an auxiliary mechanical drive, the aesthetic body shape, with good sea features and modest equipment.

Yacht hull shapes vary greatly in size and different parts of the construction and general proportions. The most characteristic parameters are: the height of the freeboard, the shape of the upper edge of the sides, the width of the hull and the shape of the bow and stern. Hull construction determines the appearance and qualities of a sailing boat, and thus its type.

Sailboat have to have many features, most important of which are:

- the ability to afloat safely, which ensures properly resolved hull design, especially the internal bonds and deck plating,
- tendency to resist heeling moment, resulting from the wind pressure on the sail,
- resisting the force causing the drift.

There are four basic hull design solutions applied that distribute sailing boats on the following types:

- ballast boats, also called keel boats, which provides sufficient stability ballast for lower center of gravity relative to the center of buoyancy,
- centerboard boats with wide hulls, which provides a relatively wide stability waterline, high freeboard and the ability to ballast by the crew,
- ballast-centerboard boats as a combination of the previous generation,
- multiple hull boats (catamarans and trimarans) exhibiting high stability thanks to the transverse distance between the light and narrow hulls connected by a platform representing the deck.

Sailing equipment types used over the centuries have been subject of multiple transformations before it took the present form. To this day, the following types of equipment are still in use: ket, sloop, cutter, yawl, ketch, schooner, brig, brigantine, bark, bark, frigate. Each of these has its own advantages and disadvantages. While the equipment listed in the first six positions are commonly used as weapons sailboats, the final five types of rigging are used almost exclusively on large sailing vessels plying the seas and oceans as a school ship.

The sailing boats equipment is relatively simple and has unsophisticated name. On the other hand the names of the school board equipment are complex. Although it is characterized by facility of building names thanks to transmitting them from the masts, yardarms and belly stays.

The basic idea of the construction of the SAILBOAT expert system is to obtain effective tool to supporting the process of setting up sailing boats, depending on the version, destination and cost of building a boat.

Next steps in building the SAILBOAT system are as follows:

- development of concept of the system, so the creation of the conceptual model of the system and specification requirements,
- design of the system, namely the creation of the logical and physical model of the system,
- implementation of the system, the representation of the system architecture and the scope of implementation,

- functional testing of the system to ensure the proper operation and correct any errors in its implementation.

The SAILBOAT system is to serve as a tool to assist design and improve the production of omega-class sailboats. It is aimed primarily for manufacturers, dealers and potential buyers of sailboats.

The SAILBOAT system is adapted to the needs of the user who wants to complete the omega-class sailing boat in one of three configurations: tourism, racing and tourism-racing. In addition, it informs the person that uses the application of suitable mass boats adapted to the weight given by the crew. A major advantage of the system is simple and intuitive user interface which allows to use the software without need for training in the implementation of the application.

3. COMPUTER IMPLEMENTATION OF THE SAILBOAT SYSTEM

3.1. USER INTERFACE

The user interface is based on bookmarks. Each of them provides a functional group. Home page “Configuration of the reference amount” provides the ability to generate a set of elements of a sailing boat based on the specified parameters. In the upper part of the panel, there are two groups of selectors, so that you can identify with what options you want. The first one specifies the type of configuration. The user can select one of three configurations:

- by desired amount,
- maximum,
- minimum.

For the first configuration, according to the desired amount it is required to fill in the fields below the "Amount". For a given value SAILBOAT system selects the appropriate components of a sailing boat. Maximum configuration option allows the user to choose the most expensive items in each group, while the last one (minimum) will choose the cheapest components.

Another user interface elements are used to determine the weight limit of a boat depending on the weight of the crew. For each building, there are separate rules defining the weight of the boat.

For racing boats, the following rules apply:

- when the weight of the crew is not more than 200 kg, the weight of the boat should be 270 kg,
- when the weight of the crew is in the range between 200 kg and 220 kg, the weight of the boat should be 265 kg,
- when the weight of the crew is in the range between 220 kg and 240 kg, the weight of the boat should be 260 kg,

- when the weight of the crew is in the range between 240 kg and 260 kg, the weight of the boat should be 255 kg.

For tourist-racing boats, the following rules apply:

- when the weight of the crew is not more than 200 kg, the weight of the boat should be 320 kg,
- when the weight of the crew is in the range between 200 kg and 250 kg, the weight of the boat should be 300 kg,
- when the weight of the crew is greater than 250 kg, the weight of the boat should be 280 kg.

The building type of a tourist boat has been based on a single principle:

- regardless of the weight of the crew, the weight of the boat should be 320 kg.

The second group of parameter selection is responsible for determining building, for which the system will select the components of a sailing boat. The person using the system can select one of three possible types:

- racing,
- tourist-racing,
- tourist.

Each item in the knowledge base is intended for one type of building. The default is the first-racing building.

The second tab "Settings rates" is associated with the functionality of the home page. It allows you to specify what percentage of the given amount is to be entered for the specific component of the boat.

When you start, the default settings are selected. They are as follows:

- mast – 23%,
- derrick – 5%,
- standing rigging – 6%,
- centerboard – 5%,
- control – 5%,
- running rigging – 1%,
- cleats – 4%,
- blocks – 11%,
- sails – 36%,
- spreaders – 3%.

The user can freely manipulate percent, but must be careful not to exceed 100%. Otherwise, a message appears that the selected values are incorrect.

Another tab, the third one, named "Manual Setup" provides the ability to manually select all elements of interest. The top list contains all the elements, while the bottom - items selected by the user. Moving components between lists is possible with the aid of two buttons located in the center. "Add" button allows you to select the item (transfer of selected records from the list of items to selected), while the second one – "Delete” button allows you to move the selected item from the list of selected items to available. After each transfer of records between lists the cost of selected items is calculated and it is displayed on the label located below the table of records.

The fourth and last tab called “Available Products” provides functions that allow you to modify the application of the knowledge base. “Delete” button allows you to delete the selected record in a list that contains all the database products. “Add” button adds a new element to the knowledge base of the expert system. It is required that you provide all the defining features of the added component. First, it must define its type by selecting one of three possible values:

- racing,
- tourist-racing,
- tourist.

In addition, it is required that you have selected the type of component of the sailing boat. This is made by specifying an item from the list that contains the following items: mast, spreaders, derrick, trapezoids, topmast rigging, shrouds, backstay, forestay, mainsail, jib, spinnaker, centerboard, bow, running rigging, cleats, pulley. Additionally, the person using the app introduces a price and description of the item in text boxes.

These four tabs mentioned above are the user interface, through which all options are available to the SAILBOAT system.

3.2. MANUAL PARAMETRIZATION OF THE SAILBOAT SYSTEM

This application allows you to manually configure a sailing boat with the help of interface elements contained in the third tab. By selecting a row on the top of the list and clicking the “Add” button, the user selects the elements that he wants his boat to be built from. Each time after adding an item to the list the total cost of the configuration is shown. It is displayed at the bottom of the window.

While compiling a set, a user can withdraw an item by selecting it and clicking the “Remove” button. Selected elements and components are available on the two lists. Together they form the entire knowledge base. After each change, the contents of the tables is refreshed. A user always have confidence that the information is updated.

3.3. AUTOMATIC PARAMETRIZATION OF THE SAILBOAT SYSTEM

One of the major tasks of the application is the automatic configuration of a sailing boat based on the user-selected parameters. The task of the person using the system is to determine the configuration, the type of building and the amount to be allocated for the construction of the boat. If you select a minimum, the application finds the elements with the lowest price for this type of development. Accordingly, selecting a maximum means that the application will find items with the highest price for the selected type of development.

The last configuration, according to the desired amount allows the selection of elements by percentage distribution of the money for the necessary components. This distribution is determined in the second tab.

The whole setup is as follows:

- the total amount deducted by the amount for the execution of the hull, respectively, for each type of sailing boat:
  - – racing type – 10 000 PLN,
  - – tourist-racing type – 7 000 PLN,
  - – tourist type – 5 000 PLN,
- the necessary elements are selected in order from the least important to the most important,
- for each component the purchase amount is calculated,
- the application searches the list of available items and choose an item that is designed for the type of development and its price is most similar to that calculated in the previous step,
- the difference between the amount allocated to the purchase of the component, and its real price is passed on to the pool of money for the purchase of another item.

There is a possibility that the application does not find the desired item. Then, the user is informed of such situation by an appropriate message and auto configuration process is interrupted. Each type of building has a set of components necessary to complete the boat. Based on the principles described above automatic configuration of the boat is created.

The above-described steps allow the completion of professional sailing boat components according to specified parameters. A key limitation appears to be the amount that you are willing to spend to buy a boat.

4. TESTING OF THE SAILBOAT SYSTEM

The design of the construction and equipment of omega-class sailing boat is restricted by numerous technical regulations by Polish Yachting Association and the Polish Omega Class Association, which is mandatory to follow. To test the SAILBOAT system, omega-class athletes have been selected due to the best knowledge of the design of its yachts. The best opportunity for doing so were the Polish Championships in that class.

Eight players were divided into two groups. The first group of four was asked to create a configuration of omega-class race boat with a standard pencil and paper. For the second group the SAILBOAT system configuration was presented to create a sailing boat. After a brief presentation of the functionality of the system they were asked to create a configuration of omega-class racing boat.

Players with the “Available Products” tab could see the available system components with a brief description. Then, using the “Manual Setup” tab they were choosing the item in the “Available Items” and press the “Add” button can freely configure the boat. The selected configuration could be tracked on the ongoing basis in the “Selected Items”. During configuration, the system automatically calculated the cost of selected items. This functionality has been positively assessed by the panellists of the SAILBOAT system because it allows for controlling the cost of boat building by the specified configuration.

The results proved that the group using the SAILBOAT system created more complete configuration during incomparably less time. The system made it possible not only to make the correct decisions in a short time, but also allowed full control of the cost of building a boat.

The SAILBOAT system was also tested by a group of people with less knowledge of the configuration options of a sailboat. For these people, the opportunity to generate summaries was presented prepared by the “configuration by desired amount”. After administration of basic information such as the type of configuration, type of building, the weight of the crew and the amount of a sample configuration and pressing the “Generate” button, a sample configuration of the boat and its price were presented.

After checking the statement offered by the system, some panelists found that the offered configuration is appropriate. Among the testing groups, however, some people were not satisfied with the generated configuration and tried their abilities to create new list using the “Manual Setup”, where they found the other choices of a sailing boat components. The whole group confirmed that the SAILBOAT system was very useful for sailboats configurations.

5. SUMMARY

Tests of the SAILBOAT system shown that the system successfully meets the goals and objectives of its construction. It is easy to use, allows a user to configure a smooth sailing boat, especially a sailing boat configuration takes much less time than in the traditional way.

Information technology used in the construction of the SAILBOAT system, delivered a tool that can be operated without need for specialized and unavailable or hard to find software. This for example allows work of sales representatives at the customer.

In the application the previously established functionality has been implemented. Testing of the system shown the correctness of its operation, and above all, a good estimate of the effectiveness in achieving the pursued objective which is undoubtedly the increase in sales of sailboats.

The SAILBOAT system has many useful features, however, there are many opportunities that can be implemented. One of them may be a function of internet communication between the client and the designer, which would facilitate contact between users and the client would save time.

REFERENCES

[1] BROWN C. E., *Introduction to Artificial Intelligence and Expert Systems*, University of Southern California, 1995.

[2] BUCHALSKI Z., *An model of knowledge acquisition and inference on the frames knowledge base elements of expert system*, In: Knowledge engineering and expert systems, t. 2, Z. Bubnicki and A. Grzech (eds.), Oficyna Wydawnicza Politechniki Wrocławskiej, Wrocław 2003, 78–86.

[3] BUCHALSKI Z., *Komputerowe wspomaganie podejmowania decyzji z wykorzystaniem regułowego systemu ekspertowego,* In: Komputerowo zintegrowane zarządzanie, t. 1, R. Knosala (ed), WNT, Warszawa 2004, 156–164.

[4] BUCHALSKI Z., *Zastosowanie systemu ekspertowego do komputerowego wspomagania sieci komputerowych*, In: Komputerowe Wspomaganie Badań Naukowych, t. XII, Wrocławskie Towarzystwo Naukowe, Wrocław 2005, 185–190.

[5] BUCHALSKI Z., *Zarządzanie wiedzą w podejmowaniu decyzji przy wykorzystaniu system ekspertowego,* In: Bazy danych. Struktury, algorytmy, metody, WKiŁ, Warszawa 2006, 471–478.

[6] BUCHALSKI Z., *Knowledge Management of Expert System Based on the Symbolic Representation of Natural Language Sentences*. In: Information Systems Architecture and Technology, L. Borzemski, A. Grzech, J. Świątek, Z. Wilimowska (eds.), Oficyna Wydawnicza Politechniki Wrocławskiej, Wrocław 2006, 75–85.

[7] BUCHALSKI Z., *The Role of Symbolic Representation of Natural Language Sentences in Knowledge Acquisition for Expert System.* Polish Journal of Environmental Studies, Vol. 16, No. 4A, 2007, 40–43.

[8] BUCHALSKI Z., *Estimation of Expert System's Efficiency for the Computer Networks Design Assist*, In: Grzech A. and others (eds), Information Systems Architecture and Technology. Information Systems and Computer Communication Networks, Wrocław, Oficyna Wydawnicza Politechniki Wrocławskiej, 2007, 37–44.

[9] BUCHALSKI Z., *Computer Advisory-Decision System for the Logistics Services Support*. Polish Journal of Environmental Studies, Vol. 18, No. 3B, 2009, 53–57.

[10] BUCHALSKI Z., *Analysis of expert system application possibilities for the needs of doctors decision in various disease diagnostics*, In: Information Systems Architecture and Technology, J. Świątek, L. Borzemski, A. Grzech, Z. Wilimowska (eds.), Oficyna Wydawnicza Politechniki Wrocławskiej, Wrocław 2009, 49–58.

[11] CHROMIEC J., STRZEMIECZNA E., *Sztuczna inteligencja. Metody konstrukcji i analizy systemów* eksperckich, Akademicka Oficyna Wydawnicza PLJ, Warszawa 1994.

[12] JAGIELSKI J., *Inżynieria wiedzy w systemach ekspertowych,* Lubuskie Towarzystwo Naukowe, Zielona Góra 2001.

[13] KLONOWSKI Z.J., *Systemy informatyczne zarządzania przedsiębiorstwem. Modele rozwoju i właściwości funkcjonalne*, Oficyna Wydawnicza Politechniki Wrocławskiej, Wrocław 2004.

[14] KRISHNAMOORTHY C.S., RAJEEV S., *Artificial Intelligence and Expert Systems for Engineers*, CRC Press, London 1994.
[15] LIEBOWITZ J., *The Handbook of Applied Expert Systems*. CRC Press, London 1996.
[16] NIEDERLIŃSKI A., *Regułowo-modelowe systemy ekspertowe,* Pracownia Komputerowa Jacka Skalmierskiego, Gliwice 2006.
[17] OWOC M., *Elementy systemów ekspertowych,* cz. 1, *Sztuczna inteligencja i systemy ekspertowe,* Wydawnictwo Akademii Ekonomicznej im. Oskara Langego we Wrocławiu, Wrocław 2006.
[18] RADZIKOWSKI W., *Komputerowe systemy wspomagania decyzji,* PWE, Warszawa 1990.
[19] RUTKOWSKI L., *Metody i techniki sztucznej inteligencji,* PWN, Warszawa 2012.
[20] STEFANOWICZ B., *Systemy eksperckie. Przewodnik,* PWN, Warszawa 2003.
[21] TWARDOWSKI Z., *Inteligentne systemy wspomagania decyzji w strategicznym zarządzaniu organizacją gospodarczą,* Wydawnictwo Akademii Ekonomicznej w Katowicach, Katowice 2007.
[22] ZIELIŃSKI J., *Inteligentne systemy w zarządzaniu. Teoria i praktyka,* PWN, Warszawa 2000.

CBIR, classification, decision trees,
fuzzy rule- based classifiers

Tatiana JAWORSKA*

TOWARDS FUZZY CLASSIFICATION IN CBIR

At present a great deal of research is being done in different aspects of Content-Based Image Retrieval (CBIR). Image classification is one of the most important tasks that must be dealt with in image DB as an intermediate stage prior to further image retrieval. The issue we address is an evolution from the simplest to more complicated classifiers. Firstly, there is the most intuitive one based on a comparison of the features of a classified object with a class pattern. We propose a solution to the problem of finding the adequate weights, especially in the case of comparing complex values of some features. Secondly, the paper presents decision trees as another option in a great number of classifying methods. Thirdly, to assign the most ambiguous objects we have built fuzzy rule-based classifiers. We propose how to find the ranges of membership functions for linguistic values for fuzzy rule-based classifiers according to crisp attributes. In this paper, we present the promising results of the three above-mentioned classifications. Experiments demonstrate the precision of each classifier for the crisp image data in our CBIR. Furthermore, these results are used to construct a search engine, taking into account data mining. If the classification precision appears insufficient for the search engine requirements, in the next step fuzzy decision trees will be introduced.

1. INTRODUCTION

In recent years, the availability of image resources and large image datasets has increased tremendously. This has created a demand for effective and flexible techniques for automatic image classification and retrieval. Although attempts to construct the Content-Based Image Retrieval (CBIR) in an efficient way have been made before, a major problem in this area, which is the extraction of semantically rich metadata from computationally accessible low-level features, still poses a tremendous scientific challenge. Images and graphical data are complex in terms of visual and semantic contents. Depending on the application, images are modelled using their

- visual properties (or a set of relevant visual features) [3],

---

* Systems Research Institute, Polish Academy of Sciences, 01-447, 6 Newelska Street, Warsaw, Poland, e-mail: Tatiana.Jaworska@ibspan.waw.pl

- semantic properties [2], [14],
- spatial or temporal relationships of graphical objects [4].

The classification problem is crucial for multimedia information retrieval in general, and for image retrieval in particular. There are a number of standard classification methods in use such as: k-NN [5], SVM [7], naïve Bayes classifier [18], neural network [20], and others [1]. Having surveyed these methods, we started our classification from the simplest algorithm, namely, the similarity to the pattern which compares the features of a classified object with the set of pattern features which define classes.

Object classification is so important in the context of CBIR because it is used for several purposes, for example [10]:

1. to compare whole images. Specifically, an algorithm which describes a spatial object location needs classified objects.
2. to help the user form a query in the GUI. The user forms a query choosing graphical objects semantically collected in groups.
3. to compare image objects coming from the same class as a stage in the image retrieval process. Details are presented in sec. 5.

Generally, the classification problem can be defined as follows. Let Ω be a complete set of objects which we want to automatically recognize, hence we want to define a division into k separate classes $c_1,\ldots,c_k$. It means that there must be a division function Θ, such as:

$$\Theta\colon \Omega \rightarrow L = \{1, \ldots, k\} \tag{1}$$

which assigns to each object of the set Ω a particular class. We do not know the assignment rules, but only know the Ω subset that we call the learning or training subset.

1.1. CBIR CONCEPT OVERVIEW

In general, our system consists of five main blocks (see Fig. 1):

1. the image preprocessing block (responsible for image segmentation), implemented in Matlab, (cf. [12]);
2. the database, which is implemented in the Oracle Database (DB), stores information about whole images, their segments (here referred to as graphical objects), segment attributes, object location, pattern types and object identification, (cf. [11]);
3. the classification module, which is used by the search engine and the GUI, is implemented in Matlab. The algorithms applied in this module will be described in the following sections.

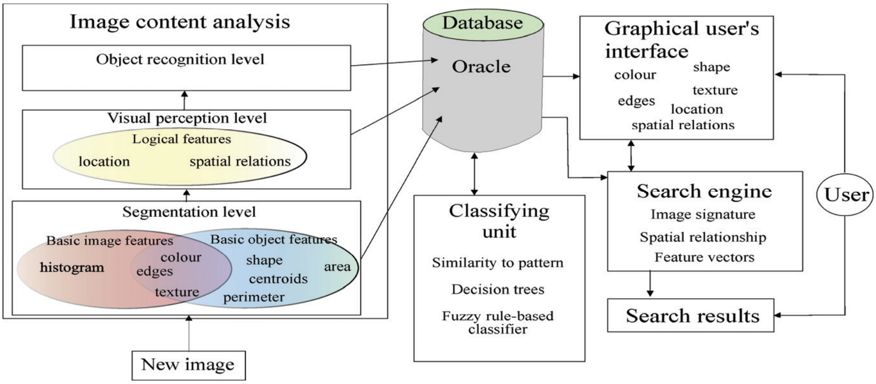


Fig. 1. Block diagram of our content-based image retrieval system

4. the search engine responsible for the searching procedure and retrieval process based on feature vectors of objects and spatial relationship of these objects in an image, implemented in Matlab;
5. the graphical user's interface (GUI), which allows users to compose their own image, consisting of separate graphical objects as a query. Classification helps in the transition from rough graphical objects to human semantic elements. We have had to create a user-friendly semantic system, also implemented in Matlab.

1.2. REPRESENTATION OF GRAPHICAL DATA

In our system, a new image is segmented, yielding as a result a collection of objects. Both the image and the extracted objects are stored in the database. Each object, selected according to the algorithm presented in detail in [12], is described by some low-level features. The features describing each object include: average colour k_{av}, texture parameters T_p, area A, convex area A_c, filled area A_f, centroid $\{x_c, y_c\}$, eccentricity e, orientation α, moments of inertia m_{11}, m_{12}, m_{21}, m_{22}, bounding box $\{bb_1(x,y), ..., bb_4(x,y)\}$, major axis length m_{long}, minor axis length m_{short}, solidity s, Euler number E, Zernike moments $Z_{00}, \ldots, Z_{33}$ [19], and some others.

Let F_O be a set of features where:

$$F_O = \{k_{av}, T_p, A, A_c, ..., E\} \tag{2}$$

Hence, for an object, we construct a feature vector: $\mathbf{x} = [x_1, x_2, \ldots, x_r]$, where n is the number of the above-mentioned features, in our system $r = 45$.

2. SIMILARITY TO PATTERN

The simplest approach to the classification is the comparison of an object feature vector **x** to the previously prepared patterns P_k for each class. Patterns can be created in different ways. The simplest method is the calculation of the average value of each vector component. The subsets of objects used to define particular patterns are also used as learning subsets. In order to compare the object vector with a pattern we apply the Euclidean metrics:

$$d(\mathbf{x}, P_k) = \sqrt{\sum_{i=1}^{r} \xi_{P_k}(x_i) \left| \mathrm{x}(x_i) - P_k(x_i) \right|^2} \qquad (3)$$

where: k – pattern number, $1 \le i \le r$. All pattern vectors are normalized. A new object is classified to a class for which d is the minimum [3], [10].

We also assume weights $\xi_k(i)$ for all pattern features where: i is the number of feature, $1 \le i \le r$. Weights for real features are the coefficients of variation

$$\xi_{P_k} = \frac{\sigma(i)}{\overline{x}(i)} \in [0,1] \qquad (4)$$

in order to reflect the dispersion of each feature in the subset selected as a pattern (where σ – standard deviation and $\bar{x}$ – mean value for each feature). However, Zernike's moments are complex features, hence to obtain the real weight we apply the formula [6]:

$$\xi(i) = \sqrt{\frac{\sigma_{\mathrm{Re}}^2 + \sigma_{\mathrm{Im}}^2}{\overline{x}_{\mathrm{Re}}^2 + \overline{x}_{\mathrm{Im}}^2}} \qquad (5)$$

where standard deviations and means are calculated separately for real and imaginary parts of complex moments.

For all these classes we have created the pattern library (also stored in the DB) which contains information about pattern types, weights and objects belonging to learning subsets [11].

We decided to classify separately objects with and without texture to reduce the misclassification between these two groups. This division diminishes the number of classification errors resulting from the fact that the pattern for non-textured objects gave a smaller d in spite of introducing weights. All results are presented in sec. 5.

3. DECISION TREES

A decision tree represents a function that takes as an input a vector of attribute values and returns a single output value as a "decision". We consider a list of attributes of our objects $\{x_1, x_2, \ldots, x_r\}$ and classes $C = \{c_1, \ldots, c_k\}$. A learning subset contains examples associated with both values of the attributes and a class [15].

Each attribute x_j can be either symbolic, numerical, or fuzzy. In our case, attributes are numerical: real and complex. The aim of the inductive process is to find a general rule to point out the relation between values of attributes and classes in C. The inductive method is based here on a decision tree from the learning subset.

In the construction of decision trees, a measure of discrimination is used in order to rank attributes and select the best one. Each vertex of a binary tree is associated with an attribute [16]. We construct our trees using the Matlab function `ClassificationTree.fit(training_set,classes)`.

In order to avoid high error rates resulting from as many as 24 classes we use the hierarchical method. The more general division is created by dividing the whole data set into four clusters applying *k*-means clustering. The most numerous classes of each cluster constituting a meta-class are assigned to four decision trees, which results in 6 classes for each one.

The second stage of the method, after constructing the trees, is the classification of a new object on the basis of its values of the feature vector. This stage is also realized by the Matlab function `predict(tree,X_new)`.

4. FUZZY CLASSIFICATION

The results presented in sec. 5 indicate that there are objects difficult for classification. Some difficulties arise from the fact that there are imbalanced classes and mistakes in object segmentation. All this motivated us to use the fuzzy rule-based classifiers.

4.1. FUZZY RULE-BASED CLASSIFIERS

Let us consider an M-class classification problem in an n-dimensional normalized hyper-cube $[0, 1]^n$. For this problem, we use fuzzy rules of the following type [8]:

$$\text{Rule } R_q : \text{If } x_1 \text{ is } A_{q1} \text{ and ... and } x_n \text{ is } A_{qn} \text{ then Class } C_q \text{ with } CF_q, \tag{6}$$

where R_q is the label of the q^{th} fuzzy rule, $\mathbf{x} = (x_1, \ldots, x_n)$ is an n-dimensional feature vector (2), A_{qi} is an antecedent fuzzy set ($i = 1,\ldots,n$), C_q is a class label, CF_q is a real number in the unit interval [0,1] which represents a rule weight. The rule weight can be specified in a heuristic manner or it can be adjusted, e.g. by a learning algorithm introduced by Ishibuchi et al. [17], [9].

We use the n-dimensional vector $\boldsymbol{A}_q = (A_{q1}, ..., A_{qn})$ to represent the antecedent part of the fuzzy rule R_q in (6) in a concise manner.

A set of fuzzy rules S of the type shown in (6) forms a fuzzy rule-based classifier. When an n-dimensional vector $\mathbf{x}_p = (x_{p1},..., x_{pn})$ is presented to S, first the *compatibility grade* of $\mathbf{x}_p$ with the antecedent part $\boldsymbol{A}_q$ of each fuzzy rule R_q in S is calculated as the product operator

$$\mu_{A_q}(\mathbf{x}_p) = \mu_{A_{q1}}(\mathbf{x}_{p1}) \times ... \times \mu_{A_{qn}}(\mathbf{x}_{pn}) \quad \text{for} \quad R_q \in S, \tag{7}$$

where $\mu_{A_{qi}}(.)$ is the membership function of A_{qi}. Then a single winner rule $R_{w(\mathbf{x}_p)}$ is identified for $\mathbf{x}_p$ as follows:

$$w(\mathbf{x}_p) = \arg\max_q \{CF_q \times \mu_{A_q}(\mathbf{x}_p) \mid R_q \in S\}, \tag{8}$$

where $w(\mathbf{x}_p)$ denotes the rule index of the winner rule for $\mathbf{x}_p$.

The vector $\mathbf{x}_p$ is classified by the single winner rule $R_{w(\mathbf{x}_p)}$ belonging to the respective class. If there is no fuzzy rule with a positive *compatibility grade* of $\mathbf{x}_p$ (i.e., if $\mathbf{x}_p$ is not covered by any fuzzy rules in S), the classification of $\mathbf{x}_p$ is rejected. The classification of $\mathbf{x}_p$ is also rejected if multiple fuzzy rules with different consequent classes have the same maximum value on the right-hand side of (8). In this case, $\mathbf{x}_p$ is on the classification boundary between different classes. We use the single winner-based fuzzy reasoning method in (8) for pattern classification.

An ideal theoretical example of a simple three-class, two-dimensional pattern classification problem with 20 patterns from each class is considered by Ishibuchi and Nojima [8] (Fig. 2.a)). There three linguistic values (*small*, *medium* and *large*) were used as antecedent fuzzy sets for each of the two attributes, and 3×3 fuzzy rules were generated. S_1 was the fuzzy rule-based classifier with nine fuzzy rules shown below:

S_1: fuzzy rule-based classifier with nine fuzzy rules

R_1: If x_1 is *small* and x_2 is *small* then Class2 with 1.0,
R_2: If x_1 is *small* and x_2 is *medium* then Class2 with 1.0,
R_3: If x_1 is *small* and x_2 is *large* then Class1 with 1.0,
R_4: If x_1 is *medium* and x_2 is *small* then Class2 with 1.0,
R_5: If x_1 is *medium* and x_2 is *medium* then Class2 with 1.0,
R_6: If x_1 is *medium* and x_2 is *large* then Class1 with 1.0,
R_7: If x_1 is *large* and x_2 is *small* then Class3 with 1.0,
R_8: If x_1 is *large* and x_2 is *medium* then Class3 with 1.0,
R_9: If x_1 is *large* and x_2 is *large* then Class3 with 1.0.

For simplicity, the rule weight is 1.0 in S_1. The location of each rule is shown in (Fig. 2b)).

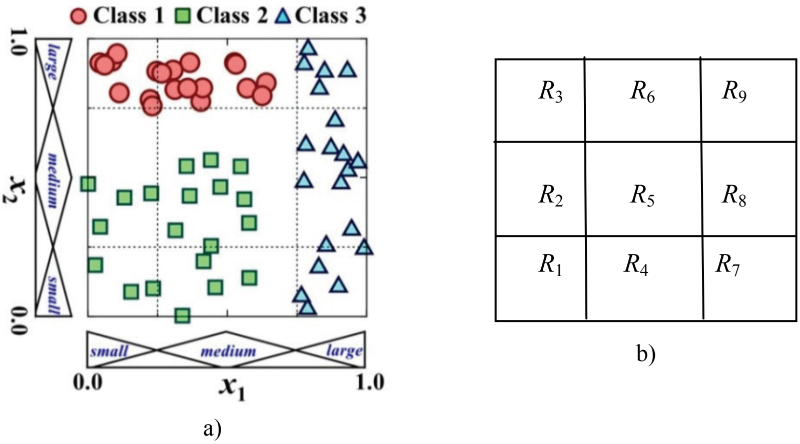


Fig. 2. a) An ideal example of a fuzzy rule-based classifier S_1 developed by Ishibuchi and Nojima [8]; b) Classification boundaries for a fuzzy rule-based classifier.

4.2. CONSTRUCTION OF MEMBERSHIP FUNCTIONS

The theoretical method presented by Ishibuchi does not answer the question how to construct membership functions, especially those corresponding to the linguistic values. We solved this problem calculating the mean value $\bar{x}$ and standard deviation σ for the elements of each of the three classes. The membership function of each class is constructed as a trapezoidal function (see Fig. 3), where points *b* and *c* are in the $\pm\sigma/2$ distance from the mean value $\bar{x}$, and the basis points *a* and *d* are $\pm\sigma$ distant from the mean value.

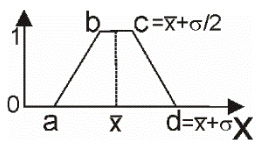


Fig. 3. Exemplification of a membership function calculated on the basis of statistical class parameters

Then, we divide the ranges of features x_1 and x_2 into three equal intervals. Next, we assign the mean value of a particular class to correspondent intervals. The effect is visible in Fig. 4 for the horizontal and vertical axes.

In each case, the fuzzy rule-based classifier is constructed automatically by matching the membership function related to the proper linguistic value, resulting in the right class for each rule. The classifier S_2 corresponds to the example seen in Fig. 4:

***S*₂: fuzzy rule-based classifier with nine fuzzy rules**

R_1: If x_1 is *small* and x_2 is *small* then non-defined with 1.0,

R_2: If x_1 is *small* and x_2 is *medium* then balkon with 1.0,

R_3: If x_1 is *small* and x_2 is *large* then arc with 1.0,

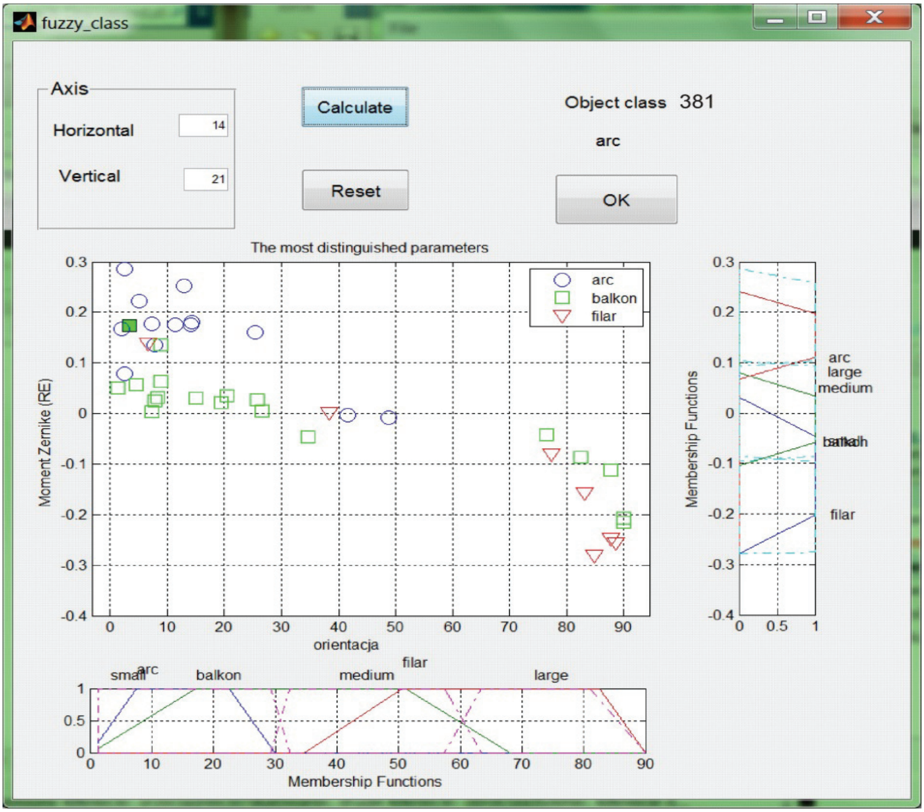


Fig. 4. Classification example. The new element marked by the full green square is recognized as an arc. Membership functions are represented by solid colour lines and linguistic intervals are drawn in dashed lines

R_4: If x_1 is *medium* and x_2 is *small* then non-defined with 1.0,
R_5: If x_1 is *medium* and x_2 is *medium* then balkon with 1.0,
R_6: If x_1 is *medium* and x_2 is *large* then non-defined with 1.0,
R_7: If x_1 is *large* and x_2 is *small* then filar with 1.0,
R_8: If x_1 is *large* and x_2 is *medium* then non-defined with 1.0,
R_9: If x_1 is *large* and x_2 is *large* then non-defined with 1.0.

The winner is the rule for which the product operator is maximum (cf. (7)), as follows:

$$\mu_{R_3}(\mathbf{x}_p) = \mu_{small}(\mathbf{x}_1) \times \mu_{large}(\mathbf{x}_2) = \mu_{small}(8.6383) \times \mu_{large}(0.1506) = 1 \times 1 = 1$$

The fuzzy rule-based classifier is stable, irrespective of attribute selection. We treat it as a “decisive voice” in the case of differences between Euclidean and decision tree classifications.

5. RESULTS

Our learning set consists of 472 objects, which gives about 20 objects per each of the 24 classes. Based on it we classified 532 new objects of all classes and we obtained the total precision of 21.5% for the similarity to pattern algorithm, 68.6% for decision trees and 88% for the fuzzy rule-based classifier FRBC (see Tab. 1).

Table 1. Classification precision

Precision	Similarity to pattern	Decision trees	FRBC
Total (for 24 classes)	**21.5%**	**68.6%**	**88%**
Window-pane	16.1%	72%	89.7%
Window	46.7%	61%	57.6%
Brick wall	9%	45.5%	90.9%
Arc	63.6%	68.2%	58%
Roof edge	8.4%	86.7%	93.9%

The high rate of false classification in the similarity to pattern algorithm results from extensive aggregation of information. Although the weights are used, all the features are involved in eventual class assignment, whereas, in the case of trees, only the most informative features are selected. Our classification process is divided into four trees due to the number of meta-classes.

The FRBC is in 'the best situation' because it is used to distinguish only among tree classes.

An additional problem, which we avoided in the learning set construction, arises from imbalanced classes. In the proper classification, however, it is inevitable.

6. CONCLUSIONS

The results presented here seem to be encouraging to move forward to the next stages of the CBIR system preparation, namely, to the description of spatial object location, the GUI and the search engine. The methods already implemented will be also evaluated in terms of the addition of new classes to the system. GUI development will also enforce introducing subclasses to some of the most numerous classes.

If classification precision turns out insufficient, we will have to apply fuzzy decision trees [13] or other more sophisticated methods.

REFERENCES

[1] ALI J.M. *Content-Based Image Classification and Retrieval: A Rule-Based System Using Rough Sets Framework*. Chap. IV, Artificial Intelligence for Maximizing Content Based Image Retrieval, Zongmin Ma (ed.), New York, 2009, 68-82.
[2] BERZAL F., CUBERO J.C., KACPRZYK J., MARIN N., VILA M.A. and ZADROŻNY S. *A General Framework for Computing with Words in Object-Oriented Programming*. In: Bouchon-Meunier B. (ed.), International Journal of Uncertainty. Fuzziness and Knowledge-Based Systems, Vol. 15, (Suppl.), Feb., World Scientific Publishing Company, Singapore, 2007, 111-131.
[3] CANDAN K.S. and LI W-S. *On Similarity Measures for Multimedia Database Applications*. Knowledge and Information Systems, Vol. 3, 2001, 30-51.
[4] CUBERO J.C., MARIN N., MEDINA J.M., PONS O. and VILA, M.A. *Fuzzy Object Management in an Object-Relational Framework*, Proceedings of the 10th International Conference IPMU, Perugia, Italy, 2004, 1775-1782.
[5] DASARATHY B.V. (ed.) *Nearest Neighbor (NN) Norms: NN Pattern Classification Techniques*, IEEE Computer Society Press, 1991.
[6] ALBERT A., ZHANG L. *A novel definition of the multivariate coefficient of variation*, Biomedical J., Vol. 52 (5), Wiley-VCH, 2010, 667-675.
[7] FEI, B. and LIU, J. *Binary Tree of SVM: A New Fast Multiclass Training and Classification Algorithm*. IEEE Transaction on Neural Networks, Vol. 17(3), 2006, 696–704
[8] ISHIBUCHI H. and NOJIMA Y. *Toward Quantitative Definition of Explanation Ability of fuzzy rule-based classifiers*, IEEE International Conference on Fuzzy Systems, June 27–39, 2011,Taipei, Taiwan, 2011, 549–556.
[9] ISHIBUCHI H. and YAMAMOTO T. *Rule weight specification in fuzzy rule-based classification systems*, IEEE Transactions on Fuzzy Systems, Vol. 13(4), 2005, 428-435.
[10] JAWORSKA T. *A Search Engine Concept Based on Multi-feature Vectors and Spatial Relationship,* LNAI Vol. 7022, Springer, 2011, 137-148.
[11] Jaworska T. *Database as a Crucial Element for CBIR Systems*. Proceedings of the 2nd International Symposium on Test Automation and Instrumentation, Vol. 4, World Publishing Corporation, Beijing, China, 2008, 1983-1986.
[12] JAWORSKA T. *Object extraction as a basic process for content-based image retrieval (CBIR) system*. Opto-Electronics Review, Association of Polish Electrical Engineers (SEP), Vol. 15(4), Warsaw, 2007, 184-195.
[13] BOUSHON-MEUNIER M. and MARSALA Ch. *Fuzzy decision tree and databases*. Chap. 14, 277-288, *Flexible Query Answering Systems*, T. Andreasen et al. (eds.), Kluwer Academic Publisher, 1997.
[14] LIU Y., ZHANG D., LU G. and MA W-Y. *A survey of content-based image retrieval with high-level semantics*. Pattern Recognition, Vol. 40, Elsevier, 2007, 262-282.
[15] FAYYAD U.M. and IRANI K.P. *The attribute selection problem in decision tree generation*, Proceedings of the 10th National Conference on artificial Intelligence, AAAI, 7, 1992, 104-110.
[16] BREIMAN L., FRIEDMAN J.H., OLSHEN R.A., STONE C.J. *Classification and Regression Trees*, Chapman and Hall, New York, 1984.
[17] NOZAKI K., ISHIBUCHI H. and TANAKA H. *Adaptive fuzzy rule-based classification systems*, IEEE Transactions on Fuzzy Systems, Vol. 13(4), 1996, 238-250.
[18] RISH I. *An empirical study of the naive Bayes classifier*, Proceedings of IJCAI-2001 workshop on Empirical Methods in AI, 2001, 41-46.
[19] TEAGUE M.R.: *Image analysis via the general theory of moments*, JOSA, Vol. 70, ed. 8, 1980, 920-930.
[20] ZHANG G.P. *Neural Networks for Classification: A Survey*, IEEE Transactions on Systems, Man and Cybernetics, part C: Applications and reviews, Vol. 30(4), 2000, 451-462.

PART 2

DISCRETE TYPE SYSTEMS AND SOFT COMPUTING

Structural periodicity, behavioral periodicity, Diophantine problem, cyclic scheduling, multimodal processes

Grzegorz BOCEWICZ*, Robert WÓJCIK**, Zbigniew BANASZAK***

MULTIMODAL PROCESSES SCHEDULING IN MESH-LIKE NETWORKS COMPOSED OF PERIODIC SYSTEMS

The theoretical prediction of the behavioral properties of mesh-like periodic structures is less difficult than in the case of nonperiodic ones. That is because all the information needed to describe the entire structure is given by the elementary substructure itself and by the manner it repeats in the whole periodic structure. In turn, cyclic schedule allows to avoid the scheduling of the whole tasks and to handle the combinatorial explosion of the problem by considering only a small pattern (cycle). That is because all the information needed to describe the entire structure is given by the elementary substructure itself and by the manner it repeats in the whole periodic system structure. So, only a small portion of the mesh-like periodic structure needs to be considered to obtain the cyclic schedule of the whole system. In that context, the contribution provides the discussion of some system periodicity issues, and is aimed at modeling and evaluation of relationships linking features of the mesh-like structure with required system's cyclic functioning.

1. INTRODUCTION

An examples of mesh-like periodic structure systems composed of a set of elementary substructures providing "guideways" for local cyclic processes follows from networks composed of Personal Rapid Transit or Automated Guideway Transit [6] systems as well as flow shops with rotary and/or carousel transportation. In that context, we are interested in how the geometry of a periodic sub-structure, providing local

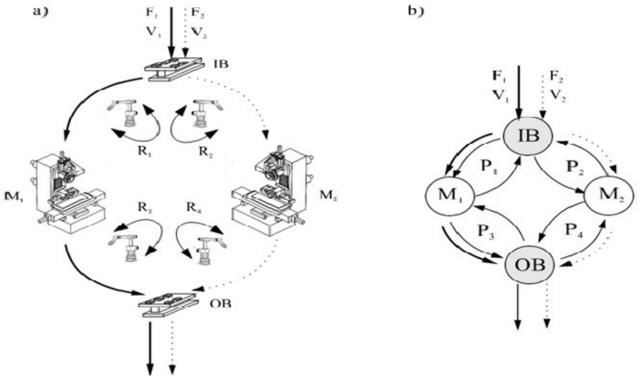
* Dept. of Electronics and Computer Science, Koszalin University of Technology, Koszalin, Poland, bocewicz@ie.tu.koszalin.pl

** Institute of Computer Engineering, Control and Robotics, Wrocław University of Technology, Wrocław, Poland, robert.wojcik@pwr.wroc.pl

*** Dept. of Business Informatics, Warsaw University of Technology, Warsaw, Poland, Z.Banaszak@wz.pw.edu.pl

cyclic processes guideways, determines the resultant behavioral properties of the whole system.

Consider the flexible manufacturing [14] cell composed of four industrial robots, two machine tools, and the input and output buffers shown in Fig. 1. Robots transporting work pieces from input buffer or machine tool to another machine tool or output buffer serve for work pieces handling – two kinds of work pieces processed along two different production routes. Production routes can be seen as cyclic repetitively executed multimodal processes supporting the work pieces flow. In general case the cells considered can be organized in much more complex mesh-like structures, e.g. see the flexible manufacturing systems shown in Fig. 2. In such structures the typical cyclic scheduling problem concerns of work pieces routing, in which a single robot has to make multiple tours with different frequencies. The objective is to find a minimal makespan schedule in which the robots repeat their handlings with assumed frequencies [8].



Legend:

R_i – the i-th robot; M_j – the j-th machine tool; IB, OB – the input and output buffers,

P_i – the i-th cyclic work piece transportation/handling process,

F_i – the i-th production route,

V_i – the batch processed along the i-th - production route,

→ - the cyclic process flow direction,

┈► - the i-th production route.

Fig. 1. Robotic cell: a) the cell layout, b) the model of cyclic processes

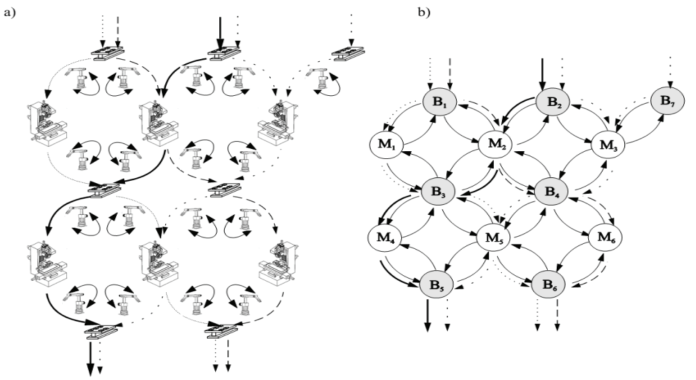


Fig. 2. Structure of the Flexible Manufacturing System; a) the mesh-like layout structure, b) the graph model of the mesh-like layout structure

The question stated above is usually considered under the assumption that a repetitive manufacturing system consists of a set of processes sharing common resources synchronized by a distributed mutual exclusion protocol. The issue is whether the production order can be accepted for execution when the availability of some resources is constrained in time.

We belief that our mesh-like driven approach to cyclic scheduling can be useful for cycle time minimization in the repetitive manufacturing flow-line system with no store constraints, which provides the fixed mix of various products on the system output. Note, that minimal cycle time is not a regular criteria and the considered cyclic job shop scheduling problem, is strongly NP-hard [9]. Similarly to the above mentioned AGVs fleet match-up scheduling problem providing conditions allowing one to adjust the AGVs fleet schedule due to the timetable of operations executed in an assumed multi-product manufacturing environment can be considered as well [3, 4].

In other words we believe our approach provides an promising alternative for many models and methods have been considered so far [9]. Among them, the mathematical programming approach [1, 15], max-plus algebra [10], constraint logic programming [3, 4, 5, 11], evolutionary algorithms [7, 12], Petri nets [13] frameworks belong to the more frequently used. Most of these are oriented at finding of a minimal cycle or maximal throughput while assuming deadlock-free processes flow.

It seems to be obvious, that not all the behaviors (including cyclic ones) are reachable under constraints imposed by the system's structure. The similar observation concerns the system's behavior that can be achieved in systems possessing specific structural constraints. Since system constraints determine its behavior, both the system structure and the desired cyclic schedule have to be considered simultaneously. In that context, our contribution provides a discussion of some solubility issues concerning cyclic processes dispatching problems, especially the conditions guaranteeing solvability of the cyclic processes scheduling. Their examination may replace exhaustive searches for solution satisfying required system capabilities.

The rest of the paper is organized as follows: Section 2 provides illustrative examples of structure depended system behaviors. Section 3 formulates a problem statement. Then the mesh-like approach to concurrent process cyclic scheduling is considered in the Section 4. Concluding remarks are presented in the Section 5.

2. ILLUSTRATIVE EXAMPLES

Let's consider the tram routes from Fig. 3 and the following problem related to them. The region S, representing the layout of a city, divided into square sectors, respectively, $s_{1,1}$, $s_{1,2}$, …,$s_{1,8}$, $s_{2,1}$,…, $s_{4,8}$, where: $s_{i,j}$ is the sector located in the i-th row and j-th column. Four transportation routes (e.g. tram routes) denoted as I, II, III, IV, together with stops. Assuming that each transportation route constitutes a closed loop, and that transportation in this city can take place only along transportation routes and along the direction denoted in the picture as well as that changes can take place only in sectors common for the given routes and the travel time $t(s_{a,b}, s_{c,d})$ between two sectors $s_{a,b}, s_{c,d}$ is calculated as a number of sectors that must be traversed from sector $s_{a,b}, s_{c,d}$ using an unlimited number of routes the response to the following question is sought: What is the shortest travel time between sectors $s_{4,4}$ and $s_{1,4}$ (denoted by), as well as between sectors $s_{3,1}$ and $s_{3,8}$ (denoted by)?

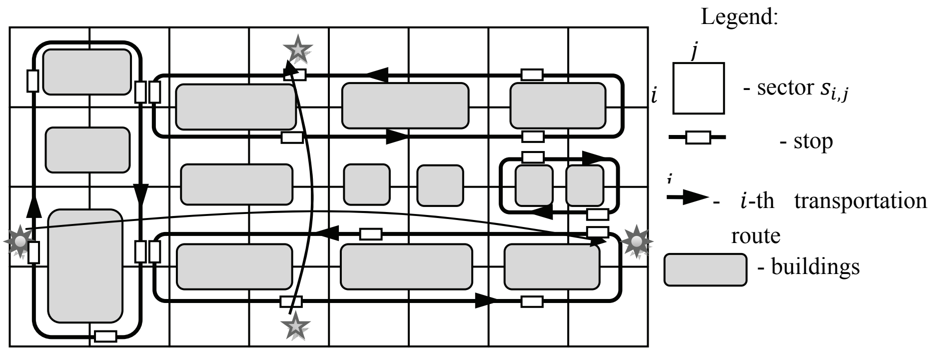


Fig. 3. Tram routes in a built-up area

Assuming the duration of a change is assumed to be equal to zero, we can easily find the answer by means of a quick analysis of available connections (there aren't many). The shortest travel time between sectors $s_{4,4}$ and $s_{1,4}$ is 13: $t(s_{4,4}, s_{1,4}) = 13$, (the route denoted by a dot-dashed line, Fig. 4 a)), whereas between sectors $s_{3,1}$ and $s_{3,8}$ – 11: $t(s_{3,1}, s_{3,8}) = 11$ (the route denoted by a dashed line, Fig. 4 b).

Since we already have some idea about the performance of the available communication system – is there any way to improve it? Let's assume that for some egoistic reasons, we wish to improve the connection between sectors $s_{4,4}$ and $s_{1,4}$. These sectors are separated by only 4 sectors, but it takes 13 arbitrary units of time to travel between them.

For the area S as in Fig. 3 – is it possible to reorganise traffic so that expected travel time is $t(s_{4,4}, s_{1,4}) < 13$, with the number of tram routes kept? constant?

This means that we allow changing the tram routes, but not their number. We can find the right solution almost immediately. Moving route IV as in Fig. 4 b) we obtain the travel time $t(s_{4,4}, s_{1,4}) = 7$ (the route denoted by a dot-dashed line, Fig. 4 b)). Unfortunately, this comes at a price of making the route between $s_{3,1}$ and $s_{3,8}$ longer, from 11 to 13 and also, what is even worse, increasing the travel time between sectors $s_{2,7}$ and $s_{3,8}$ to 14 (instead of 2).

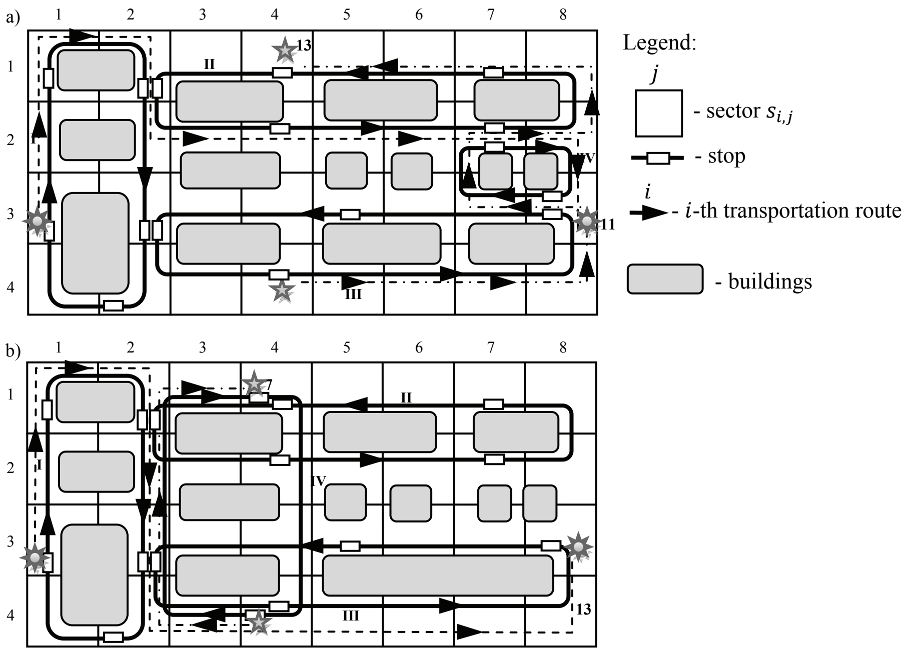


Fig. 4. Example of traffic organisation

Ideally, it would be possible to move from each sector (in which there is a stop) to another sector in a time not longer than dictated by the distance between the sectors. Therefore, is it possible, by varying only the tram routes and not their number, to organise traffic in such a way that the travel time between any two stops would be equal to the number of sectors that separate them?

In other words, the traffic must be organised so that: $t(s_{4,4}, s_{1,4}) = 4$, $t(s_{3,1}, s_{3,8}) = 8$, $t(s_{2,7}, s_{3,8}) = 2$, etc. How can this be achieved with only 4 tram routes? Is this not impossible? In the presented example we dealt with a situation in which the improvement of one connections resulted in making another worse. Moreover, the obtained solutions are correct only for the structure for which they were determined. Therefore, we can hardly expect ideal solutions – unfortunately, in this case, only compromises could constitute universal solutions.

Fig. 5 shows an example of a transportation system for Flexible Manufacturing System (FMS). Three transportation trucks P_1, P_2 and P_3 move between sectors according to assigned transportation routes in a warehouse. At a given time, only one truck can be loaded or unloaded in a given sector. Access to shared sectors R_1, R_3, R_5 is determined by priority dispatching rules.

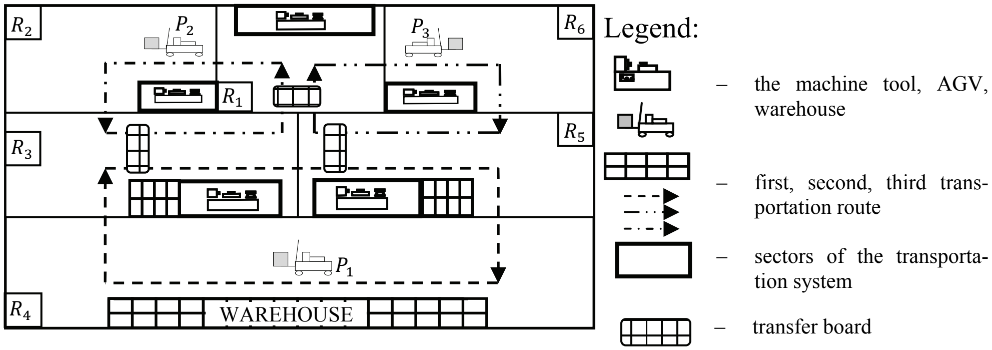


Fig. 5. Transportation subsystem of flexible production systems

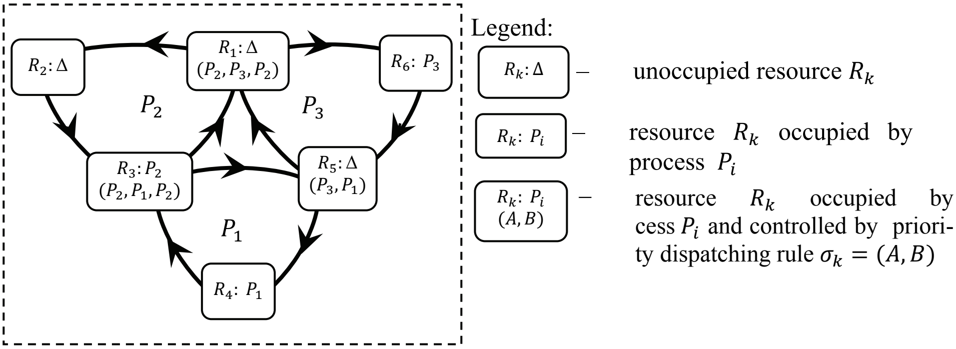


Fig. 6. System of concurrent cyclic processes

Systems of this type can be represented as Systems of Concurrent Cyclic Processes (SCCP) [4], see Fig. 6, where processes (P_1, P_2, P_3) are carried out according to assigned transportation routes, and the order in which processes acquire shared resources is determined by the priority dispatching rules σ_k. For instance, the order in which resource R_3 from Fig. 6 will be acquired is P_2, P_1, P_2 (and then, again, P_2, P_1, P_2). For such systems, we usually try to answer the following – is it possible to cyclically carry out these concurrent processes? If so, what are their time periods?

3. PROBLEM STATEMENT

Consider SCCP system shown in Fig. 7 composed of the set of resources $\{ {}^{(i)}R_1, \ldots, {}^{(i)}R_{19} \}$ and four cyclic digraphs passing on by the following sequences of resources:

$$ {}^{(i)}p_1 = \left({}^{(i)}R_1, {}^{(i)}R_2, {}^{(i)}R_3, {}^{(i)}R_4, {}^{(i)}R_5, {}^{(i)}R_6 \right), $$

$$^{(i)}p_2 = \left(^{(i)}R_3, ^{(i)}R_7, ^{(i)}R_8, ^{(i)}R_9, ^{(i)}R_{10}, ^{(i)}R_{11} \right),$$
$$^{(i)}p_3 = \left(^{(i)}R_4, ^{(i)}R_{11}, ^{(i)}R_{12}, ^{(i)}R_{13}, ^{(i)}R_{14}, ^{(i)}R_{15} \right),$$
$$^{(i)}p_4 = \left(^{(i)}R_{10}, ^{(i)}R_{16}, ^{(i)}R_{17}, ^{(i)}R_{18}, ^{(i)}R_{19}, ^{(i)}R_{12} \right).$$

Four cyclic processes $^{(i)}P_1$, $^{(i)}P_2$, $^{(i)}P_3$, $^{(i)}P_4$, are executed in this network along the guideways determined by sequences $^{(i)}p_1$, $^{(i)}p_2$, $^{(i)}p_3$, and $^{(i)}p_4$, respectively. The priority dispatching rules assigned to shared resources $^{(i)}R_3$, $^{(i)}R_4$, $^{(i)}R_{10}$, $^{(i)}R_{11}$, and $^{(i)}R_{12}$ are as follows: $^{(i)}\sigma_3 = \left(^{(i)}P_2, ^{(i)}P_1 \right)$, $^{(i)}\sigma_4 = \left(^{(i)}P_3, ^{(i)}P_1 \right)$, $^{(i)}\sigma_{10} = \left(^{(i)}P_4, ^{(i)}P_2 \right)$, $^{(i)}\sigma_{11} = \left(^{(i)}P_2, ^{(i)}P_3 \right)$, $^{(i)}\sigma_{12} = \left(^{(i)}P_4, ^{(i)}P_3 \right)$.

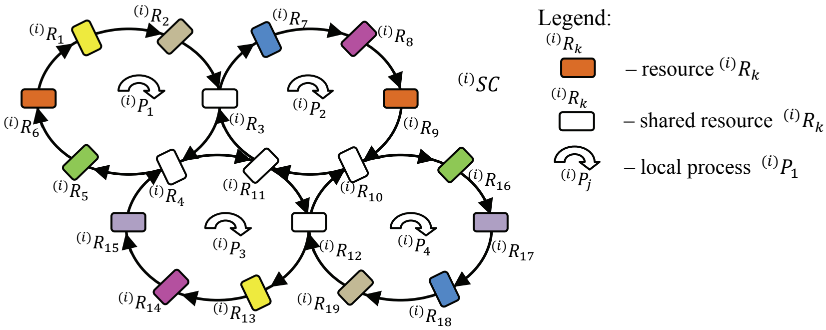


Fig. 7. The SCCP structure of a periodic structure object

The Gantt's chart of the processes cyclic execution is shown in Fig. 8. Since the all operation times are the same and equal to one unit of time (t.u. for short), the length of the cycle time is equal to 6.

Consider the periodic structure shown in Fig. 9 made of the basic object (structure) from Fig. 7. The response to the following question is sought: What is the cycle time of the periodic structure SCCP, while assuming the operation times are the same and equal to the one unit of time?

Then the next one, assuming a given set of round trip itineraries, concerns the cycle time of the itineraries network.

4. IN MESH-LIKE BASED MULTIMODAL PROCESSES SCHEDULING

Response to the above stated questions follow from the observation regarding the way the periodic structure shown in Fig. 9 is composed of a set of elementary substructures shown in Fig. 7.

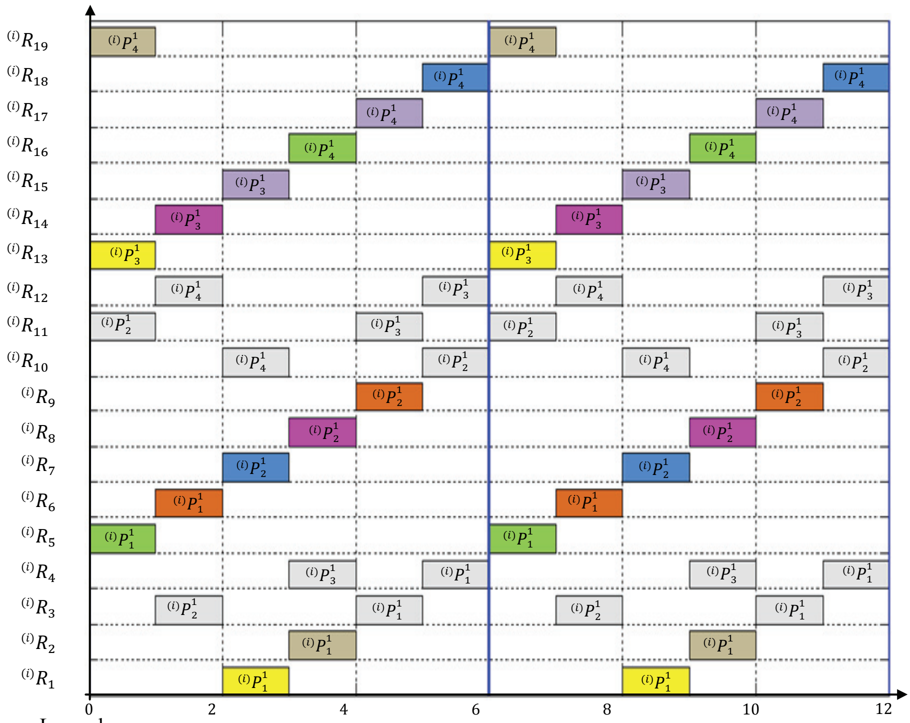


Fig. 8. The Gantt's chart of the SCCP from Fig. 7

In order to respond to the question: What is the cycle time of the periodic structure SCCP from Fig. 9? let us consider the Gantt's chart shown in Fig. 10 encompassing the behavior of SCCP composed of seven subsystems $^{(i)}SC$, $^{(i+1)}SC$,…, $^{(i+6)}SC$. The considered diagram can be seen as composition of the Gantt charts shape shown in Fig. 11 and following $^{(i)}SC$, $^{(i+1)}SC$,…, $^{(i+6)}SC$. That is easy to observe, that cycle time is equal to 6 t.u., i.e. to the cycle time of the elementary periodic structure from Fig. 7. Using the same argumentation, in general case it can be shown that applied modeling framework can be helpful in course of cycle time evaluation in any periodic structure SCCP.

Regarding the set of round trip itineraries routes distinguished by bold: dashed-dotted, dashed, doted, and solid lines in Fig. 11, the corresponding travel cycles of round trips are calculated due to the formulae (1) and are equal to: $Tt_R = 54, Tt_G = 66, Tt_B = 66, Tt_V = 156$, respectively.

$$Tt = \sum_{t_i \in t(\rho)} t_i + \sum_{tp_i \in tp(\rho)} tp_j \quad (1)$$

where: ρ – the guideway of the multimodal process linking distinguished resources,

Λ – the set of all admissible guideways ρ;

t_i – the execution time of i-th operation from path ρ,

$t(\rho) = \{t_1, \dots, t_i, \dots, t_g\}$ – the set of all execution times of operations belonging to guideway ρ;

tp_i – the awaiting time (covering suspension and transfer times) required to process change (e.g. metro lines etc.);

$tp(\rho) = \{tp_1, \dots, tp_i, \dots, tp_q\}$ – the set of all awaiting times from the guideway ρ .

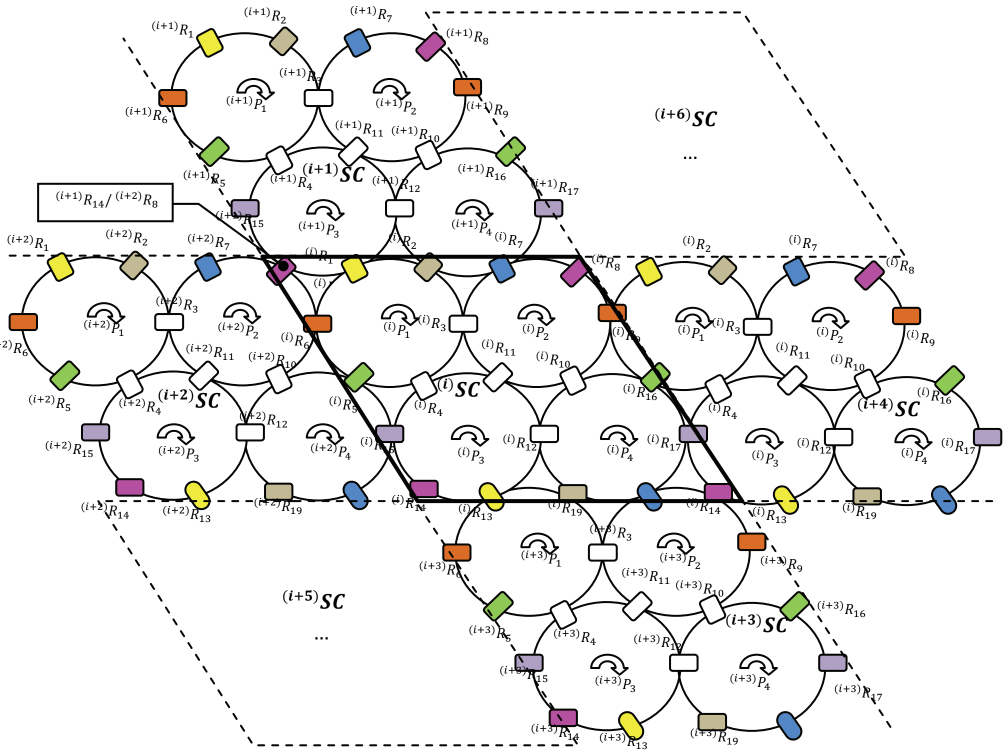


Fig. 9. The way the periodic structure is composed of elementary object from Fig. 7

The formulae follows from observation that that the itinerary considered can be seen as result of a multimodal process realization. The passenger's itinerary including different metro lines encompass a plan of multimodal process execution within a considered metro network can be seen as its example.

Consequently, the resultant cycle time of the round trip itineraries, calculated due to formulae (2) is equal to 15444 t.u.:

$$T = LCM(Tt_R, Tt_G, Tt_B, Tt_V). \quad (2)$$

In general case, the operation times can be assumed as belonging to $\mathbb{N}^+$ as well as the shape and size of an elementary substructure can be different than shown in Fig. 7.

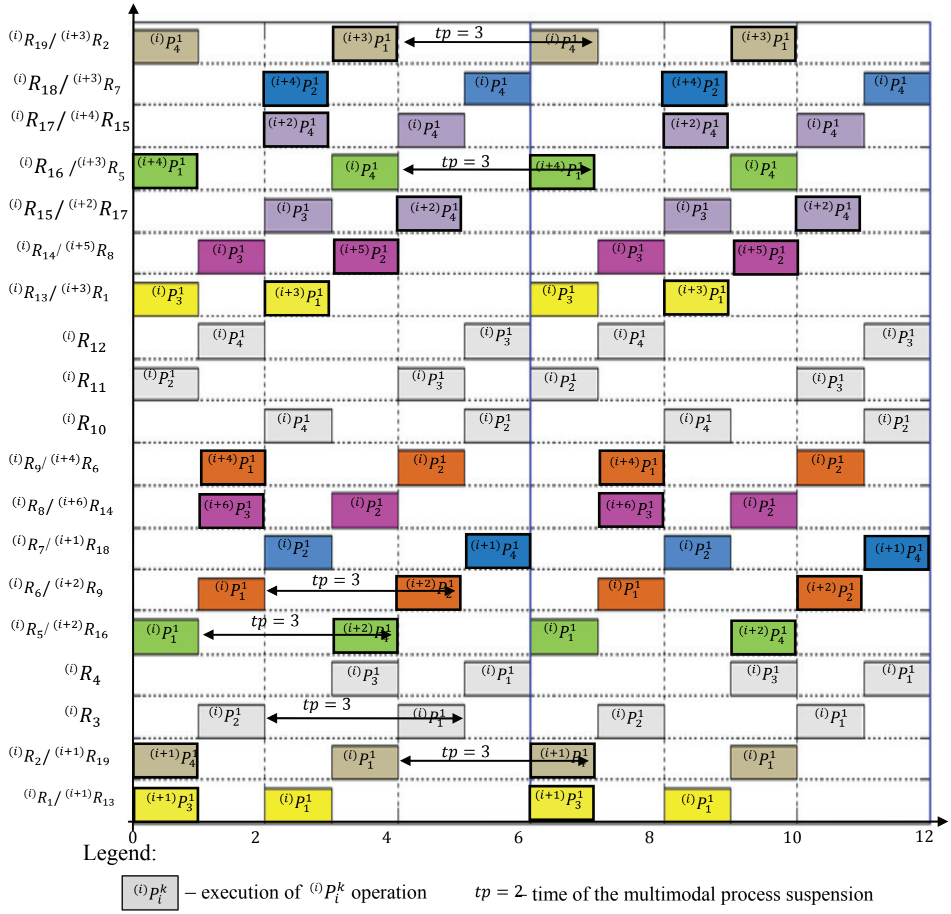


Fig. 10. The Gantt's chart of the SCCP composed of ${}^{(i)}SC$, …, ${}^{(i+6)}SC$ as shown in Fig. 9

Therefore, the considered periodic structure SCCP can be seen as homeomorphic model of many real life systems such like passengers' city transportation, and FMS's material handling networks. In this context, the approach presented leads to solutions allowing the traveler to reschedule his itinerary in case of unforeseen transportation network malfunction. Moreover, allows one to redesign, for instance an urban public transport system in such a way as to obtain its assumed robustness specified in terms of required (e.g. awaited by passengers) quantitative and qualitative behavioral (functional) features. So, the tools implementing this approach can support ex ante evaluation of spatial urban passengers/freight planning and spatial plans of urban public transport system.

Another interpretation of considered periodic structure SCCP assumes a given network of local cyclic acting AGV services. In such a regular network, i.e. composed of elementary and structurally isomorphic sub-networks, the work pieces pass their origin-destination routes among workstations using local AGVs. Since an AGVs fleet scheduling problem can be seen as a blocking job-shop NP-hard problem where jobs might block either workstations or AGVs, hence AGVs fleet scheduling in mesh-like environments also belongs to NP-hard problems. The solution sought assumes that schedules of locally acting AGVs match-up the given, i.e. already planned, schedules of work pieces machining.

Note, that since cyclic substructures generate cyclic behaviors, and since each local cyclic schedule match-up the cyclic schedule generated by assigned elementary structures, hence the behavior of the whole mesh-like structure is also cyclic. Moreover, since the cyclic behavior of the mesh-like structure can be easily evaluated from the behaviors of its component substructures, hence the size of searched states space [3, 4, 5] can be substantially reduced that results in shortening of AGVs fleet scheduling problem solution.

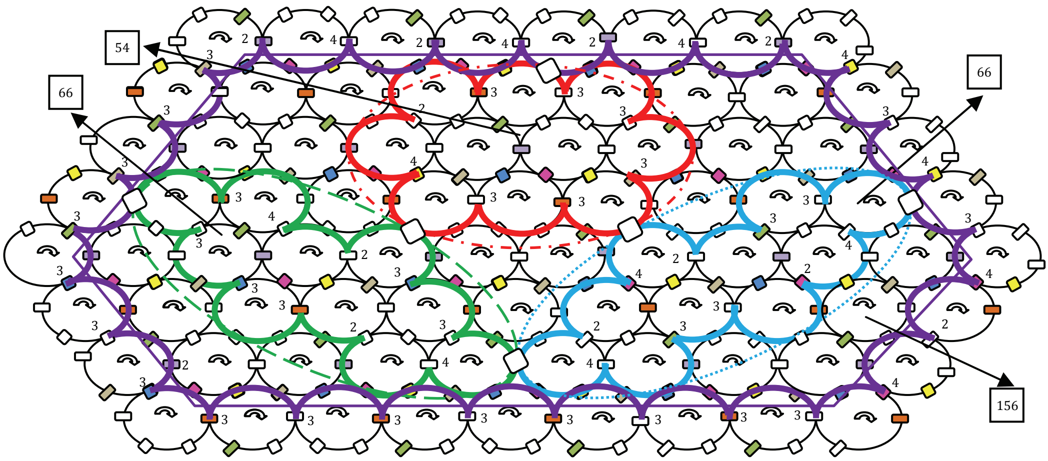


Fig. 11. Four cyclic multimodal processes executed within environment of periodic structure SCCP

5. CONCLUDING REMARKS

The paper introduces to the structural periodicity concept assuming a system structure is set-up of several isomorphic substructures. The structure considered can be seen as a digraph composed of a set of isomorphic sub-digraphs. Assuming vertices of each sub-digraph model the workstations and the arcs linking vertices model the material handling operations, the considered structure can be treated as graphical model of a mesh-like multimodal (i.e., composed of different means such as: automated guided vehicles, roller conveyors, tow lines, shuttle cars etc.) transportation network. In that context, work piece flows are treated as multimodal processes passing through a mesh-like layout of FMS. So, in opposite to traditional approach we assume the given network of local cyclic acting AGV services, i.e. corresponding to distinguished isomorphic sub-networks of FMS layout, where work pieces pass their origin-destination routes among workstations using local transportation means. The main objective is to provide conditions guaranteeing solvability of the cyclic processes scheduling, i.e. guaranteeing the right match-up of local cyclic acting AGV schedules to a given work pieces machining schedules. In that context the goal of further work is to provide a declarative model enabling to state a constraint satisfaction problem aimed at multi-modal transportation processes scheduling encompassing production flows.

REFERENCES

[1] ABARA J., *Applying integer linear programming to the fleet assignment problem*. In: Interfaces, 19, 1989, 4–20.

[2] BANASZAK Z., BOCEWICZ G., *Decision Support Driven Models and Algorithms of Artificial Intelligence*, Management Sciences Series, Warsaw University of Technology, 2011, 237 pp.

[3] BOCEWICZ G., BANASZAK Z., *Declarative approach to cyclic scheduling of multimodal processes*. In: *EcoProduction and Logistics*, Golińska P., (Eds.), Vol. 1, Springer-Verlag, Berlin-Heidelberg, 2013, Vol. 1, Springer Heidelberg, 2013, 203–238

[4] BOCEWICZ G., BANASZAK Z., *Declarative approach to cyclic steady states space refinement: periodic processes scheduling*. In: *International Journal of Advanced Manufacturing Technology, SI: Advanced Dispatching Rules for Large-scale Manufacturing Systems*, DOI 10.1007/s00170-013-4760-0.

[5] BOCEWICZ G., WÓJCIK R., BANASZAK Z., *Cyclic Steady State Refinement. In: International Symposium on Distributed Computing and Artificial Intelligence*, series: *Advances in Intelligent and Soft Computing*, Abraham A., Corchado J.M., Rodríguez González S., de Paz Santana J.F. (Eds.), Vol. 91, Springer-Verlag, Berlin Heidelberg, 2011, 191–198.

[6] BOYLE D.K. *Fixed-Route Transit Ridership Forecasting and Service Planning Methods.Transportation Research Board National Academies,* WASHINGTON, D.C., 2006.

[7] DERIS S., OMATU S., OHTA H., KUTAR L.C.S., SAMAT P.A., *Ship maintenance scheduling by genetic algorithm and constraint-based reasoning,* European Journal of Operational Research, 112, 1999, 489–502.

[8] JOLAI F., FOUMANI M., TAVAKOLI-MOGHADAM R., FATTAHI P., *Cyclic scheduling of a robotic flexible cell with load lock and swap*, Journal of Intelligent Manufacturing , 2012, 23:1885–1891.

[9] LEVNER E., KATS V., ALCAIDE D., PABLO L, CHENG T.C.E, *Complexity of cyclic scheduling problems: A state-of-the-art survey.* Computers & Industrial Engineering, Vol. 59, Issue 2, 2010, 352–361.

[10] POLAK M., MAJDZIK P., BANASZAK Z., WÓJCIK R., *The performance evaluation tool for automated prototyping of concurrent cyclic processes*, Fundamenta Informaticae. Vol. 60, No.1–4, 2004, 269–289.

[11] RELICH M., *An evaluation of project completion with application of fuzzy set theory*, Management, 2012, vol. 16, no. 1, pp. 216–229.

[12] SLOWIK A., *Application of an Adaptive Differential Evolution Algorithm with Multiple Trial Vectors to Artificial Neural Networks Training*, IEEE Transactions on Industrial Electronics, Vol. 58, Issue 8, 2011, 3160–3167.

[13] SONG J.-S., LEE T.E., *Petri net modeling and scheduling for cyclic job shops with blocking*, Computers & Industrial Engineering,Vol. 34, No.2, 1998, 281–295.

[14] ŚWIĆ A., GOLA A. *Economic analysis of the production of parts of casings in a flexible manufacturing system*, Actual Problems of Economics/Актуальні Проблеми Економіки 2013, 3(141), 526–533.

[15] Von KAMPMEYER T., *Cyclic scheduling problems, Ph.D. Dissertation*, Fachbereich Mathematik/Informatik, Universität Osnabrück, 2006.

shortest path problem, ant colony optimization, ACO, metaheuristics

Mariusz GŁĄBOWSKI, Bartosz MUSZNICKI,
Przemysław NOWAK, Piotr ZWIERZYKOWSKI*

AN IN-DEPTH DISCUSSION OF CHALLENGES RELATED TO SOLVING SHORTEST PATH PROBLEMS USING ShortestPathACO BASED ALGORITHMS

The studies hitherto carried out have revealed that the application of the *ShortestPathACO* strategy based on the Ant Colony Optimization (ACO) metaheuristics makes it possible to solve the shortest path problem in a way that differs from traditional approaches. Practical applications of algorithms that are based on the ACO metaheuristics require accurate and deep understanding of the importance of particular parameters of the algorithm and this issue is absolutely crucial. This involves deep understanding of both their influence and the degree to which they are expected to offer a possibility of the modification in their implementation. On account of the heuristic approach to the problem to be solved, it is important then to choose appropriately the mode of operation of the algorithm to be applied as well as a determination of the parameters that will ensure proper course of operation in a given specific situation.

This chapter aims then to describe, analyse and interpret the observations provided by relevant studies and to present conclusions in relation to the conducted research work and subsequent analyses. The chapter mainly focuses on indicating the essential problems and defining the areas for further optimization of the *ShortestPathACO* approach. Limitations and constrains involved in the application of the ACO metaheuristics in solving the shortest path problem are highlighted. A particular attention is given to the greediness and convergence of the algorithm, as well as to possibilities of a secondary use of available information on the pheromone level. In addition, movement patterns of ants around the graph and the choice of their number depending on the class of graphs are analysed and are followed by a proposition of alternative approaches.

1. INTRODUCTION

As a result of extensive studies on the application of the Ant Colony Optimization (ACO) metaheuristics [1] in solving variants of the shortest path problem, the present

---

* Poznan University of Technology, Faculty of Electronics and Telecommunications, Chair of Communication and Computer Networks, Polanka 3, 60-965 Poznan, Poland.
e-mail: bartosz@musznicki.com, przemyslaw.nowak@inbox.com

authors have developed a general concept of the *ShortestPathACO* algorithm [2]. Using this concept as a starting point in successive research studies, the authors have developed further relevant solutions such as, for example, the strategy to find the shortest path between two nodes, described in detail in [3]. Further research studies have confirmed the authors' presumption that, with an appropriate selection of the parameters for the operation of the heuristic algorithm, it is possible to significantly shorten the running time for necessary calculations and at the same time retaining almost a 100 per cent probability of obtaining the optimum results. Successive studies, based on the concept presented in [2], deal with the construction of Shortest Path Tree (SPT) [4]. The *ShortestPathTreeACO* algorithm makes it possible to solve the shortest path problem with one initial vertex using the *ShortestPathACO* algorithm. Furthermore, at the initial stage in the assessment of the potential in the applications of the ACO algorithm, the authors decided to start an in-depth analysis of those algorithms that represented a more traditional approach to the problem, such as, Dijkstra's, Dial's or Bellman-Ford algorithm. The efficiency evaluation of 12 algorithms published in [5] enabled the authors to not only perform a comparison of some selected algorithms, but also to establish a reference level that can be used in heuristic evaluation methods based on *ShortestPathACO*.

The observations presented in this chapter are a continuation of a series of earlier studies presented in [2]-[5]. The paper discusses the observations and conclusions, as well as indicates key problems and areas adequate for further optimization of both the *ShortestPathACO* algorithm and the ACO metaheuristics. Limitations and constraints following the application of the ACO metaheuristics in solving the shortest path problem are also identified in the chapter. The above imposes the particular structure of the chapter. Sections 1.1 and 1.2 briefly describe the shortest path problem and the *ShortestPathACO* algorithm. Then, in Section 2, the influence of some selected parameters and features of the algorithm upon its greediness, convergence and possible capacity of using information on the pheromone level between iterations in the construction of SPT are discussed. Section 3 provides considerations concerning the movements of ants on the graph, with particular attention given to the method for depositing of the pheromone trail and the directions of ants' movements. Final thoughts and conclusions are presented in the summary.

1.1. SHORTEST PATH PROBLEM

For the directed graph $G = (N, E)$, where N is the set of nodes (vertices) and E is the set of edges (arcs), let a_{ij} be the length (cost) of the edge $(i, j) \in E$. Then, the length a_P of the path $P = (n_1, n_2, \ldots, n_k)$ can be expressed by Formula (1).

$$a_P = \sum_{i=1}^{k-1} a_{n_i n_{i+1}} \tag{1}$$

A path is called the shortest path if it has the shortest length from among all paths that begin in *s* node and terminate in *t* node. Therefore, the basic aim of solving the shortest path problem is related to finding the shortest path between a pair of nodes. Another variant is directed at finding the Shortest Path Tree (SPT) in a graph, which solves as well the Single-Source Shortest Paths problem (SSSP) related to finding the shortest paths from a single initial *s* node to all other nodes in a weighted graph [6]. The solutions for the problems find their applications in various fields such as satellite navigation, pipeline transport or routing in computer networks [7], [8].

1.2. SHORTESTPATHACO ALGORITHM

The *ShortestPathACO* algorithm is a method and a set of recommendations to ensure that the Ant Colony Optimization metaheuristics for solving the shortest path problem is properly applied. A detailed introduction to the methodology, discussion on the context and methods for finding paths, as well as a discussion on the classes of parameters and the methods for updating pheromones, are introduced in [2]. The subsequent papers [3], [4] present thorough analyses of the *ShortestPathACO*-based approaches to finding the shortest path between two nodes and to constructing the Shortest Path Tree. Before executing the algorithm, one should have a full awareness that the ACO metaheuristics has been constructed to seek solutions of *NP*-hard problems [1]. As such, it does not always guarantee finding the most optimum solution. Therefore, the obtained results may be both accurate (optimal) and approximations that depend on the degree of fitness of the algorithm itself for each individual problem to be solved. Moreover, in particular situations a solution may not even be found. Because of this particular feature, it is extremely important to first analyse a given task and to properly select the operations running parameters to be executed and then to perform their optimization.

2. PARAMETERS OF THE ALGORITHM

As mentioned earlier, in any application of the algorithms based on the ACO metaheuristics a deep understanding of the importance of individual parameters of the algorithms is of key importance. This refers to both their influence and a viability of applicable modifications in the methods for their application. The following parameters of the algorithms will be discussed:

- s – initial node,
- t – end node (required when the shortest path between a pair of nodes is to be calculated),
- m – the number of ants,
- α – the parameter that defines the influence of pheromones on the choice of the next vertex,
- β – the parameter that determines the influence of remaining data on the choice of the next vertex.

2.1. GREEDINESS OF THE ALGORITHM

The inclusion of the length of the path in the calculation of the edge coefficient (in the procedure of the choice of the next vertex) can eventually lead to a situation where the algorithm becomes greedy, always selecting the shortest edges. This, in turn, is very disadvantageous because this strategy, though seemingly advantageous in generating short paths, is in fact conducive to making a totally adverse effect. Quite frequently the set of the shortest paths includes an edge whose length is decidedly deviated from the rest of edges that originate at the vertex under consideration. When this is the case, finding the optimum (the shortest) path is virtually impossible. On the other hand, however, it can be assumed that the adopted degree of greediness makes it possible to orientate the algorithm towards shorter paths and thus to find shorter paths quicker without omitting a given subset of longer edges.

In the *ShortestPathACO* algorithm, a preliminary (initial) stage of the edge selection has been introduced to the *ComputeCoefficient((i,j),k)* function [3] that uses different variations of the formula for the calculation of edge coefficients q_{ij} or q'_{ij} for the k-th ant for the edge between nodes i and j. The length of a given edge a_{ij} is a basic variable on which the formula is based, whereas the parameters α and β have modification function. In order to compensate adverse effects of the greediness of the algorithm, the form of the selected variants of the formula has been made dependent on the values of variables *vi_nodes* and *vi_edges* that store information on already visited nodes and edges. One of the methods for the calculation of the q'_{ij} coefficient is presented in Formula (2), while more variants are included in, for example, [2].

$$q'_{ij} = a_{ij}^{\alpha}\left(1+\beta\right)^2 \quad \text{for} \quad \begin{cases} vi\_nodes_j = false \\ vi\_edges_{ij} = false \end{cases} \tag{2}$$

Thanks to the adopted strategy it is possible to make ants select as great number of edges and vertices that have not been visited earlier as possible, even if the length of these edges are large. This solution, however, has some limitations. This is clearly visible in the following example. Let us consider a situation where 10 ants are in one vertex from which 5 edges originate. The first 5 ants will select consecutively an edge that has not been hitherto selected by any other ant, but the remaining 5 ants will travel along the shortest edge. As a result of this behaviour of the ants, the pheromone trail on this shortest path will be enhanced (by the deposited pheromone trail) and, in consequence, this will influence decisions made by other ants that, in successive iterations, will be exploring paths from this vertex. Therefore, the algorithm includes mechanisms that prevent it from manifesting greediness of his type (e.g. an increase in the amount of pheromones left by an ant if the path found is the shortest one from the beginning of the operation of the algorithm), though in some situations the mechanisms may turn out to be insufficient and then an application of auxiliary methods should be taken into consideration.

2.2. THE CHOICE OF THE VALUE OF THE PARAMETER β

The considerations on the method for a calculation of edge coefficients presented in Section 2.1 clearly indicate the significance of the parameter β. This parameter defines the degree of the increase in values of the applied edge coefficients by available data. An appropriate choice of the values for the β parameter is crucial for the proper operation of the mechanism preventing all ants from selecting the same edge. By examining the formulas for edge coefficients presented in [2] one can establish what value should the β parameter take to guarantee all ants to explore and check all edges originating from a single vertex. These conclusions are addressed in (5), in which the B parameter denotes the shortest length of an edge in the graph (Formula (3)), while C is equal to the longest length of edge in a given graph (Formula (4)). For example, in the graph for the edge length $a_{ij} \in \langle 1;9 \rangle$ and $\alpha = 1$, the recommended value is $\beta = 2{,}1$.

$$B = \min_{(i,j)\in\varepsilon} a_{ij} \tag{3}$$

$$C = \max_{(i,j)\in\varepsilon} a_{ij} \tag{4}$$

$$\beta > \left(\frac{C}{B}\right)^{\frac{\alpha}{2}} - 1 \tag{5}$$

A decrease in the value β below the threshold determined by (5) may lead to a situation in which the algorithm would use only paths with their lengths that do not exceed a given constant, which would then translate into the algorithm's partial greediness. In some graphs, this allows the algorithm to find the shortest path very quickly because long edges are not taken into account at all by ants in their selection of the next vertex. In other cases, however, this results in reaching the convergence to non-optimum paths by the algorithm. As it has been proved in the studies on the *ShortestPathTreeACO* algorithm, an appropriate selection of the β parameter in itself does not guarantee that the algorithm will construct the SPT (i.e. will solve the Single-Source Shortest Paths problem). This means that there are still situations where the algorithm is too greedy. The chances for finding the optimum solution are also influenced by the values of the remaining parameters of the algorithm, which directly results from the heuristic character of *ShortestPathACO*.

2.3. WASHING OFF PHEROMONES (PHEROMONE TRAIL) WITH A CHANGE OF THE DESTINATION NODE

The *ShortestPathTreeACO* algorithm in its successive iterations changes the end node, while the data obtained earlier are not used any more. In addition, the fixed number of ants with searches of different paths may turn out to be questionable. Here, not taking advantage of information on the amounts of pheromones deposited on the edges of the graph in earlier iterations can be also an ineffective solution. It is highly possible that the data obtained while searching a path to a given end vertex could turn out to be useful with another search of a path to another vertex. If in this particular application the *ShortestPathACO* algorithm had information on the level of pheromones obtained from the previous iteration, it is conceivable that the first stage of its operation could be omitted and it would be possible to directly proceed immediately to an attempt at achieving the convergence. In this way, the time required for finding a path to a given vertex would be shortened greatly, as well as the total time necessary to find all paths included in the Shortest Path Tree. There is, however, a certain risk involved in the application of such a strategy. Assuming that the shortest path to one of the vertices is composed of a given set of edges, there is no certainty that a path to yet another vertex includes any part of this set. There is a pitfall then, resulting from the fact that if that algorithm had some initial pheromone level on individual edges, ants would then move along those edges that have the highest amount of pheromones. Inasmuch as in the case of some of graphs this would turn up to be beneficial, in the case of the majority of structures ants would rather attempt to find the shortest path using edges that would not necessarily be included in this path. The strategy under consideration requires then a particularly accurate verification and a testing phase to be carried out to establish unequivocally whether it indeed brings the expected advantages.

2.4. THE NUMBER OF ANTS USED IN SEARCHING PATHS

By analysing the results of the operation of the *ShortestPathTreeACO* algorithm it is possible to observe that there is a possibility of the occurrence of a certain computational redundancy related to the application of the fixed number of ants in the process of finding successive paths. This problem may occur only with graphs characteristic for a multi-stage problem of the shortest path or other graphs that do not have back-ward edges that would allow ants to withdraw to vertices that are located closer to the initial vertex, that is those ones whose edges are kind of directed towards one destination vertex. An exemplary multi-stage graph is presented in Fig. 1. The said redundancy results from the fact that the vertices in the close vicinity of the destination vertex require a greater number of ants to find the shortest path, whereas those that are closer to the initial vertex, a lower number. If, however, one of the methods for searching the graph to establish the distance (expressed in the number of vertices) between a given vertex and the initial or destination (end) vertex is used, there is a scope for the application of the strategy presented in this section. When it is known additionally that the graph is divided into stages or that the distance between a vertex and the initial vertex mentioned earlier is known, then it is possible to make the number of ants dependent on, appropriately, either the stage involved or the distance. The advantage of the proposed mechanism is that the running time of the algorithm can be significantly reduced. This is, however, at the expense of the necessity of possessing or calculating additional data on the graph. Depending on the stage, the time needed for a path to be generated can be shortened even several times.

In order to illustrate the proposed concept, the following generalized and simplified analysis will be carried out. Consider a graph with the initial node *s* and the end (destination) node *t*. The graph is composed of *z*-stages, whereas in each of the stages *u* vertices are involved. Let us assume that the time needed to find one path, with the application of *m* ants, is proportional to the number of ants and is t_m. With a different approach, where for each stage the number of ants $w=i\ m/z$ will be used, where $i \in <1;z>$ is proportional to the number of a given stage, i.e. respectively m/z, $2\ m/z$, ..., $(z-1)m/z$, m ants for the following stages, the aggregate operation time of the algorithm t_{total} will be expressed by Equation (6) in which t_{pr} is the averaged time of finding a path to a single vertex in a given stage.

$$t_{total} = \sum_{i=1}^{z} t_{pr} u = \sum_{i=1}^{z} \frac{t_m w}{m} u = t_m \frac{u}{z} \sum_{i=1}^{z} i = t_m \frac{(1+z)}{2} u \qquad (6)$$

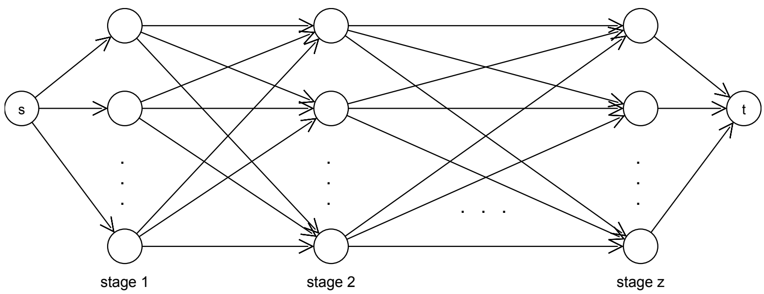


Fig. 1. A multi-stage graph

For example, using the standard approach for the *ShortestPathTreeACO* algorithm, for a multi-stage graph with 5 stages composed of 10 vertices, the running time of the algorithm is $50t_{m,}$ whereas while using the number of ants proportional to each of the stages, according to Formula (6), the running time can be shortened to $30t_m$.

It should not be forgotten though that the above estimates are only generalized estimates that simply mean to justify the purposefulness of the discussion on the choice of a strategy. Therefore, details concerning, for example, the method of choice for successive vertices to which paths are to be found as well as the dependence of the path on the structure of the graph (which may significantly influence the running time of the algorithm) have been omitted here. Any practical application should be then preceded with a detailed analysis and a thorough preliminary study.

3. MOVEMENT OF ANTS

The characteristic features as well as the operation of the *ShortestPathACO* algorithm depend largely on the applied rules and principles for ants to use the deposited pheromone trail. The biggest obstacle in the operation of the ACO metaheuristics in its application to solving the shortest path problem is, however, the speed at which the pheromone level on the edges originating from the vertex from which the routes of the ants are commenced is established. As early as after the first routes are covered by ants it turns out that one edge has the concentration of pheromones high enough so that, without introducing additional modification strategies, the remaining edges soon become unused. If a non-optimum edge is not part of the shortest path, it is rather difficult to internally change the algorithm from a state in which the pheromone level in each of the routes is increased (including non-optimum paths), which with every iteration reduces the chance of finding the shortest path. It is worthwhile then to have a closer look at the selected assumptions.

3.1. THE FIRST ITERATIONS OF THE ALGORITHM

While analysing various aspects of movements of ants one should pay particular attention to the place where ants begin their routes, i.e. the initial vertex (the above, however, does not apply to solving All-Pairs Shortest Path Problem). Because of the fact that an ant, after the path is generated, every time starts its next route from this vertex, the pheromone trail on the edges that originate from it, as well as the pheromone trail on some part of the edge that is in close vicinity of the vertex, is enhanced decidedly more often than on the edges that are farther and farther from it, which is intuitively understandable. During the first iteration of the algorithm, ants select edges along which they are to reach the next vertex. Because the number of m ants is most frequently far greater than the number of edges originating from the initial vertex, each of the edges will be selected by a number of ants. When ants reach the end vertex and start to deposit pheromones on the edges, it is just the edges that are closer to the beginning of their routes that will have the largest amount of the deposited pheromone trail. During the second, and then all the successive routes, the ant colony will usually select just one of all available edges originating from the initial vertex – the one that includes the largest amount of pheromones. If ants were to start each of their routes from a different vertex, then this situation would never ever happen. However, in the case of the shortest path problem, the starting point for ants has to be the same and invariable, which implies that the above mentioned problems cannot be avoided. However, it is worthwhile to attempt to counteract this phenomenon. One of the mechanisms that can be applied in this situation involves storing information on vertices and edges in the graph that have not yet been visited and increasing for them (or in the case of vertices – for the edges that terminate in them) the values of the edge coefficient. Such a mechanism allows us, at least to a certain degree, to increase the number of edges along which ants move, though only during initial iterations. Another possible strategy involves an introduction of a transitory period during which explorative features of ants are enhanced. Within the duration of this period, defined as the first three routes of each of the ants, a totally different edge coefficient can be applied, or even a chance variation (random character) in the selection of successive edges can be introduced. The application of a transitory period is, however, followed by lengthening of the running time of the algorithm. Without this period, in some of the graphs the algorithm is capable of reaching the convergence as early as during the second route performed by the ants, but its introduction doubles the necessary time (necessary for the ants to cover four routes from the source node to the destination (target) node). It is worthwhile as well to mention that without the additional modifications the *ShortestPathACO* algorithm can initiate reaching the convergence as early as after the first route of each of the ants [3].

The observations done so far allow us to put forward the following theses:

- Procedures based on *ShortestPathACO* do not always check all edges in the graph,
- If there are more edges in the graph that originate from the initial vertex than ants, then the edges that are not included in the first path that has been found by each of the ants will never be used,
- If all ants select the same edge that originates from the initial vertex during their first route, then the remaining edges will never be used.

The above statements are true then and only then when no mechanism influencing the explorative features of ants and those that increase the probability of checking as great number of edges as possible are applied.

3.2. EXCESSIVE DEPOSITION OF PHEROMONES ON NON-OPTIMUM PATHS

The problem of pheromone deposition (accumulation of pheromones) on non-optimum paths is connected with the issue considered in Section 3.1. While analysing the proposed statements it can be noticed that it is just the first route taken by ants that causes pheromones to be excessively cumulated (deposited) on edges that are not included in the shortest path. The reason for this operation of the algorithm is the lack of information on the quality of particular paths. What is more, except the lengths of edges that originate in the vertex in which ants are currently at, the ants do not basically have in their first routes any additional information. The pheromone trail on the edges is created as late as during the return of ants to the initial vertex, which makes it visible only for other ants that begin their successive routes and for those ants that have not yet finished their first route. This phenomenon cannot be counteracted though because during the initial operation of the algorithm ants have to explore as great number of edges as possible to be ready to select the shortest path in the next stage. It is worthwhile, however, to draw our attention to the following problems:

- in the case of the selection of a longer edge originating in the initial vertex (or an edge that causes a longer path to be selected) by a greater number of ants, they can deposit a greater amount of pheromones than just one ant that has selected the shortest path,
- a frequent selection of a sub-optimal path may eventually lead to a situation where finding the shortest path is not followed by an introduction of changes in the amount of pheromones on its edges that are large enough to have any influence on the selection of next paths.

The first problem concerns mainly the first iterations of the algorithm. By taking into account the influence of the amount of pheromones on the edges originating in the initial vertex, it is easy to observe that without additional enhancement of the pheromone level on the edge included in the shortest path found so far, ants do not necessarily have to select this particular path. Hence, in the *ShortestPathACO* algorithm, additional enhancement of currently optimum path has been introduced [3]. It turns out that the simplest way to guarantee the largest amount of pheromones on such a path is to establish the pheromone level on a sufficient level that would be larger even if all other ants would select another path. For this purpose, the product of the inverse of the length of the shortest path $\Delta\tau$ found by the number of ants m so far is used. Apparently, this strategy performs quite well [3].

The situation looks somewhat different in the case of the other problem that involves a situation where ants are already busy finding successive routes. At this moment, the amount of pheromones on edges is already established and hence it is difficult to determine to what degree their level should be enhanced. A determination of the value of the pheromone level that would force ants to move along the path would significantly prolong the operation time of the algorithm. As a rule, the *ShortestPathACO* algorithm terminates the searching process for new paths and initiates the process of convergence as early as after the first routes of the ants, and therefore the problem has not yet been addressed nor has a solution to it has been expected [3].

3.3. A CHANGE IN THE DIRECTION OF THE MOVEMENT OF ANTS

Yet another interesting issue is the reversal of the operation of the algorithm. This would mean a commencement of a path search from the destination vertex and an exploration towards the initial vertex. When this is the case, an appropriate selection of the structure of the graph would be necessary (a change in the direction or the weights of edges to give converse directions or weights). It is preferable to select a reversal of the edges of the graph. If, additionally, the vertex of the commencement of the route of an ant were selected randomly, then it might be possible to counteract a too fast determination of the pheromone level on edges. At the same time, it might become feasible to avoid a situation where ants start moving along only a given subset of edges. This would make it possible to avoid problems related to excessive accumulation of pheromones on non-optimum edges. After termination of the operation of the algorithm, edges with the biggest concentration of pheromones would automatically create the Shortest Path Tree. The above assumptions, however, require a new algorithm that would behave completely different from *ShortestPathACO* to be worked out from the scratch. If, additionally, we take into account the application of a total independence of ants, then it could turn out that a completely different method for finding paths has to be applied, as well as different ways of selection of the next vertex and the method for updating the pheromone trail. This, however, does not change the significance of the presented approach and in the authors' opinion should be also addressed in further studies.

4. CONCLUSIONS

This chapter discusses some selected specific features of approaches to finding solution to the shortest path problem that are based on the *ShortestPathACO* algorithm. The factors that may lead to ineffective operation of the algorithm or to the convergence to non-optimum results are indicated and discussed. The phenomenon of an excessive greediness of the algorithm is discussed and methods for its counteracting are proposed. The presented considerations allow us to state that there is still a room for further optimization of the operation of the *ShortestPathACO* algorithm. Further improvements can be obtained by improving the formulas already in use that determine, for example, edge coefficients and the parameter β. It can be adopted that a modified strategy will make it possible to find optimum paths also in such graphs for which the algorithm is currently not performing well. A change in the way pheromone information is utilized may also prove advantageous. The same applies to the use of information on the pheromone level obtained in earlier iterations. Various aspects related to movement patterns manifested by ants on the graph are discussed. A relevant modification is proposed and a proposition of a choice of the number of ants for multi-stage graphs is presented. It should be stressed that the implementation of some of the proposed solutions can lead to a significant change in the operation of the algorithm as we know it and even lead to a development of a completely new algorithm.

REFERENCES

[1] DORIGO M., STÜTZLE T., *Ant Colony Optimization*, Cambridge, Massachusetts: The MIT Press, 2004.

[2] GŁĄBOWSKI M., MUSZNICKI B., NOWAK P., ZWIERZYKOWSKI P., *Shortest Path Problem Solving Based on Ant Colony Optimization Metaheuristic*, International Journal of Image Processing & Communications, Vol. 17, No. 1–2, 2012, 7–17.

[3] GŁĄBOWSKI M., MUSZNICKI B., NOWAK P., ZWIERZYKOWSKI P., *ShortestPathACO based strategy to find the Shortest Path between two nodes*, In: Proceedings of 2013 IEICE Information and Communication Technology Forum (ICTF), Sarajevo, Bosnia and Herzegovina, 29–31 May 2013.

[4] GŁĄBOWSKI M., MUSZNICKI B., NOWAK P., ZWIERZYKOWSKI P., *An Algorithm for Finding Shortest Path Tree Using Ant Colony Optimization Metaheuristic*, (to be published) In: Image Processing and Communications Challenges 5, Advances in Intelligent Systems and Computing, Vol. 233, Choraś R.S. (Ed.), Switzerland, Springer International Publishing, 2014, 317–326.

[5] GŁĄBOWSKI M., MUSZNICKI B., NOWAK P., ZWIERZYKOWSKI P., *Efficiency Evaluation of Shortest Path Algorithms*, In: Proceedings of AICT 2013, The Ninth Advanced International Conference on Telecommunications, Rome, Italy, 23–28 June 2013, 154–160.

[6] CORMEN T.H., LEISERSON C.E., RIVEST R.L., STEIN C., *Introduction To Algorithms*, Cambridge, Massachusetts: The MIT Press, 2009

[7] STACHOWIAK K., WEISSENBERG J., ZWIERZYKOWSKI P., *Lagrangian relaxation in the multicriterial routing*, In: Proceedings of IEEE AFRICON 2011, Livingstone, Zambia, September 2011, 1–6.

[8] PRAGYANSMITA P., RAGHAVAN S.V., *Survey of Multicast Routing Algorithms and Protocols*, In: Proceedings of ICCC 2002, the Fifteenth International Conference on Computer Communication, Washington, DC, USA, 2002, 902–926.

Decisional DNA (DDNA), Set of Experience Knowledge Structure (SOEKS), Evolutionary Algorithm, Genetic Algorithm, Heuristics, Optimization Problem

Peng WANG*, Cesar SANIN*, Edward SZCZERBICKI**

DECISIONAL DNA AND OPTIMIZATION PROBLEM

Many researchers have proved that Decisional DNA (DDNA) and Set of Experience Knowledge Structure (SOEKS or SOE) is a technology capable of gathering information and converting it into knowledge to help decision-makers to make precise decisions in many ways. These techniques have a feature to combine with different tools, such as data mining techniques and web crawlers, helping organization collect information from different sources and using gathered knowledge to make decision or prediction. In this paper, our focus is on another research issue, optimization problem. Although there are many algorithms that have been design to solve this problem, it still lacks efficiency and effectiveness to get optimal solutions. Therefore, we propose a new structure combining the SOEKS with an evolutionary algorithm to find optimal solutions and to reuse this experience for efficient decision making support.

1. INTRODUCTION

With the rapid development of computer networks and multimedia technologies, Internet has changed human lifestyle on an everyday basis and became the major supplier of vast amounts of information. However, it is difficult to quickly and easily find desired information from websites due to unstructured and dynamical changes in information presentation and content [1]. Furthermore, Knowledge and users' experience play an important rule on today's organization market competition. The ability to learn from their experiences and adapt to rapidly changing environment, determines which organizations will win the battle in today's global economy. Many different applications offer solutions for complex problems by the means of integrating knowledge into computer system. It uses many computer science domains such as

* School of Engineering, Faculty of Engineering and Built Environment The University of Newcastle, Australia.

** Gdansk University of Technology, Gdansk, Poland.

knowledge representation, artificial intelligence, data mining and decision support systems to make powerful, more efficient and effective system for learning, reasoning and forecasting from current knowledge and experience.

Knowledge engineering is defined as a discipline that offers solutions for optimization problems, decision making problems and market predictions by integrating knowledge into computer systems [2]. Set of Experience Knowledge Structure (SOEKS or SOE) as an experience-based knowledge structure is able to store uncertain and incomplete data and to make qualitative and quantitative extractions of knowledge from available data that can often be unstructured, semi structured, fuzzy, and vague. Additionally, SOEKS can be shaped in an extensive understandable and transportable language such as Extensible Markup Language (XML) or Ontology Web Language (OWL) allowing knowledge to be exchanged quickly and securely between applications and systems [3]. Our research presented by this article introduces a novel adaptable knowledge structure to work with Heuristic and to promise simplicity and improvement of the decision-making processes. This structure utilizes advanced advantages of the both techniques to find optimal solutions from optimization problems and to store the gathered results as the formal decision events to be reused in the future.

2. BACKGROUND

2.1. HEURISTIC

Heuristic which derives from the Greek word "find" or "discover" [4] applies to experience-based techniques for problem solving, learning and discovering. It is a rule of thumb, an educated guess, an intuitive judgment or simply common sense. A heuristic method is utilized to rapidly approach to a solution that is expected to be close to the optimal solution. It is a general way of solving a problem. In real world problems, heuristic copes with optimal problems solving in human beings and machines [4].

Combinatorial optimization

Combinatorial optimization is a subset of optimization usually employed to find an optimal solution from a finite set of objects [4]. Its domain is to cope with optimization problem where the set of feasible solution is discrete or can be reduced to a discrete one. Its aim is to find the best solution [4, 5]. It has important applications in several fields, including artificial intelligence, machine learning, mathematics, auction theory, and software engineering. Combinatorial optimization problems can be treated as searching for the best element of some set of discrete items.

The Travelling Salesman Problem (TSP) is a typical example of combinatorial optimization problem [6]. The goal of this problem is to find a shortest possible path that is enable tour to visit each city exactly once. In order to find the best path in the worst case, the choices will be a factorial of number of cities. Therefore, there is no efficient algorithm for solving TSPs [7]. Even though it is impractical to find the best solution for a problem likes TSPs, heuristics algorithms are employed to discover close to optimum solutions such as genetic algorithms[8]. These algorithms can find good enough solutions to the problem.

Multi-objective Genetic Algorithm

Multi-objective Genetic Algorithm is used to minimize multiple objective functions that are tasks to a set of constraints and to identify the set of evenly distributed non-dominated optimal solutions [7, 9, 10]. The advantage of the multi-objective genetic algorithms is that it generates the holistic pareto-optimal solutions in a single simulation. A number of multi-objective genetic algorithms are developed in last few years. Some of them are, Multi-Objective Genetic Algorithm (MOGAs) [11], Niched Pareto genetic algorithm (NPGA) NPGA [12], Pareto Archived evolution strategy (PAES), Non-dominated Sorting Genetic Algorithm (NSGA-II) [8] and among others. All these algorithms have two main purposes: push the solutions toward Pareto-optimal front and maintain diversity among the solutions in the Pareto-optimal front.

Non-dominated sorting genetic algorithm II (NSGA-II) algorithm.

NSGA-II algorithm is a well-known multi-objective evolutionary algorithm (MOEA) that use a fast non-dominated sorting approach with O(MN2) computational complexity (where M is the number of objectives and N is the population size); non-elitism approach; and the need to specify a sharing parameter [8]. The computational complexity of NSGA-II is lesser than other multi-objective evolutionary algorithms. Meanwhile, NSGA-II combines parent and offspring populations to create a mating pool and selecting.

In computer processing, a NSGA-II algorithm assumes that potential solutions of a problem are individuals or phenotypes. Each individual has a set of variables, much like the genes of a chromosome, which can be mutated and altered. Traditionally, these variables are organized by a string of values in binary form. A fitness value, which is always positive, is used to reflect the degree of suitability of the chromosome. Throughout a genetic evolution, chromosomes of the fittest individuals are stochastically selected from current population. The genes of the chromosomes are combined and mixed within the population of offspring in the next generation. A superior chromosome is expected to have a higher chance to produce better offspring in the subsequent generation in nature. This cycle of evolution is repeated until a desired termination criterion is reached. The algorithm is also terminated by the number of evolution cycles.

To facilitate the genetic evolution cycle, the operators of a simulated binary crossover (SBX) and polynomial mutation are required to produce offspring [13]. The crossover selects the genes of a parent chromosome and an offspring chromosome from one generation to the next. This is also known as fitness proportionate selection, where the individual is selected on the basis of fitness. A mixing ratio with a typical value of 0.9 is normally used as the probability of SBX crossover. There are many crossover methods for ordered chromosomes. The mutation operator is applied to each offspring individually after the crossover exercise. It alters one or more genes in a chromosome with a small probability (typically a value of less than 0.1). The aim of the mutation operator is to get a quicker convergence.

2.2. SET OF EXPERIENCE KNOWLEDGE STRUCTURE AND DECISIONAL DNA

From a mechanistic point of view, representing and acquiring knowledge is a computational process, which needs systemic techniques and data structures. It has been shown that knowledge representation (KR) is a medium of translation between human expression and computers [3]. It relies on determining results instead of actions and can facilitate the process of making decisions and recommendations. KR is a set of ontological commitments to answer questions regarding how to interpret the real world.

Set of Experience Knowledge Structure (SOEKS) has been designed in accordance with the fundamentals of the concept of KR [3]. It is a flexible and independent knowledge representation that can handle information and knowledge in differing formats, held within organizations or companies. SOEKS is intended to collect experiences and knowledge from multiple applications that are assembled as formal decision events in order to assist organizations in making precise decisions, predictions, and recommendations. Human DNA carries genetic information coded within combinations of its four elements (nucleic acids). SOEKS is analogous to natural DNA in its ability to extract knowledge and arrange it in combinations of four elements: variables, functions, rules, and constraints. This is an eminently suitable tool for knowledge management tasks, which additionally have been used to collect and store formal decisional events in an explicit manner [14]. Its ontology can be expressed in Extensible Markup Language (XML) or Ontology Web Language (OWL) in order to make it shareable and transportable [15].

SOEKS is composed of variables, functions, constraints, and rules [3]. Functions are made up of interactions between variables, which include dependent variables and a set of input variables. Constraints are another way of representing associations between variables. Although a constraint is a form of function, it has a different purpose. It limits the performance and configuration of a system and restricts the feasible solutions in a decision problem. Lastly, rules are another way to express links between variables. They condition possible relationships that operate on the universe of variables. Essentially, rules use the statements IF–THEN–ELSE to connect conditions with their consequences.

3. DDNA-BASED INTELLIGENT STRUCTURE

This article explores different functionalities in the diagnosis, prognosis, solution, and knowledge macro processes in the DDNA as illustrated in Figure 2 [16]. Decisional DNA, as a domain-independent, flexible and standard knowledge repository, can not only capture and store experiential knowledge in an explicit and formal way, but can also be easily applied to various domains to support decision-making and standard knowledge sharing and communication among these systems[14]. In addition, it introduces a new model offering many advantages. It can gain knowledge from different repositories. It stores different kinds of information in the form of XML for sharing and easy transfer purposes. It not only shares knowledge but assists in decision-making and prediction processes. This structure automatically produces a series of similarities between existing pairs of experiences, and those similarities are commonly computed for prediction purposes. Furthermore, many well-known techniques, such as Multi-objective Evolutionary Algorithms (MOEAs) and Data Mining techniques (DM), can be embedded into this structure for specific purposes. As can be seen in Figure 2, DDNA depicts the process of data workflow.

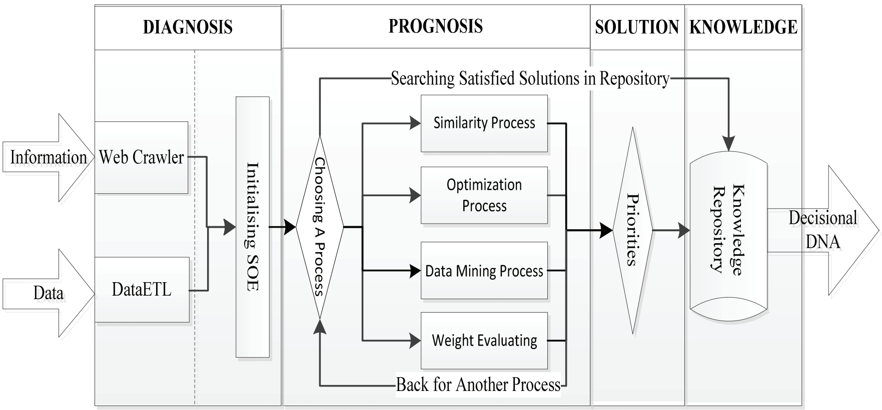


Fig. 1. Architecture of DDNA-Based Intelligent Structure

3.1. DIAGNOSIS MACRO-PROCESS

Organizations pursue the production process in cooperation with other area or other organizations, and exchanging information and knowledge quickly and securely among applications and systems is a progressively important requirement. Because Set of Experience knowledge structure developed as a knowledge representation ease keeping and duplicating knowledge from previous decision events, diagnosis Macro-process is responsible for communicating with the applications in terms of acquiring formal decision events [17]. At this stage, user’s interactions are needed to supply data to the applications to make them produce pre-solutions by following the atomicity and multidisciplinary concepts. Then those pre-solutions are integrated into one unique language and measurement model SOEKS with variables, functions constraints, and rules

Therefore, both a dataETL (Extraction Transformation and Load) component and a web crawler component are introduced in terms of acquiring information from disparate systems and websites. For example, the heritrix can be employed at this stage to gather webpages and filter out redundant information [16]. The collected information is identified and divided to the variables of the SOEKS as well as other information attached to them such as functions, constraints and rules [16]. Alternatively, the dataETL component is responsible for interpreting data from the repositories (e.g., by reading CSV files, which are comma-separated values represented as numbers in plain-text form) to the SOEKS [16].

3.2. PROGNOSIS MACRO-PROCESS

Once the Diagnosis Macro-process has been reached, the prognosis macro process produces a set of proposed solutions or decisions which is achieved by providing several models. Each model provides a scenario which offers measurements for uncertainty, incompleteness or imprecision [17]. For this reason, several search processes shown in figure 2 for improved optimal solution are performed. Those processes, including data mining process, similarity process, weights evaluation process and optimization process and among others, can be implemented either individually or together. In this paper, an evolutionary algorithm NSGA-II is adopted during the homogenization and integration processes. Hence, produced holistic Sets of Experience by using optimization process can be taken into the solution macro-process in order to choose a unique optimal solution.

Optimization process. Fig 2 illustrates the procedure of optimization process of the SOEKS. The optimization process of the SOEKS defines a process oriented to the acquisition of the optimal solutions. Initially, gathered SOE problem instance is evaluated whether it has solutions or not. If the problem already has solutions in knowledge repository, the system will directly index the solutions to find optimal one for the problem. However, if the SOE problem does not have any solution in knowledge repository, then, system will run a multi-objective evolutionary algorithm (MOEA) to seek a set of optimal solutions. In this paper, we adopted a well-known MOEA: NSGA-II algorithm.

Previous sections have illustrated how the DDNA-Based Intelligent Structure gathers information and data from website or data repository to generate SOEKS and DDNA. The next step is to determine whether MOEAs are needed for discovery of optimal solutions. Therefore, the current prototype is to compare the similarity of all of the SOEKS' comprising elements in knowledge repository. If there is no related

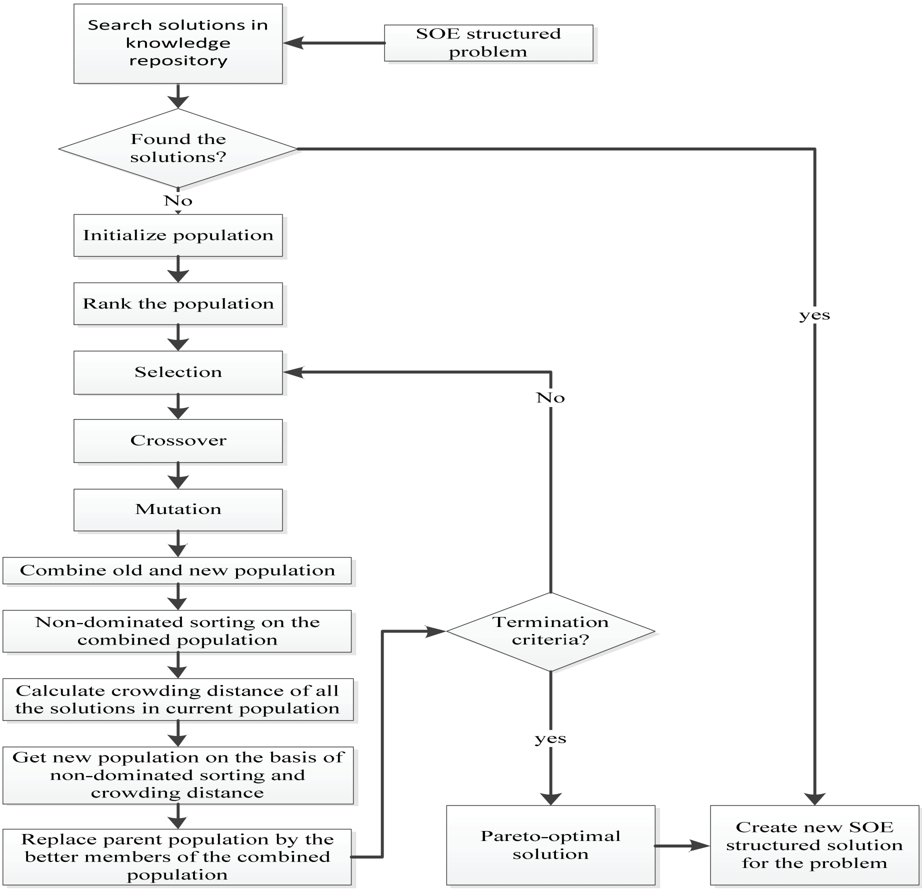


Fig. 2. Optimization Process based on NSGA-II

solution found, the NSGA-II algorithm will be employed and the SOE structured problem extracted from previous layer will be inputted into the NSGA-II algorithm as parameters. The steps involved in the NSGA-II algorithm are shown in Fig. 3. These steps are described below.

Initialization of population. The first population is stochastically generated. Then, each individual is assigned a value by using objective functions.

Rank the population. The population is sorted by using non-dominating method and crowding distance [8]. For multi-objective optimization problem, the dominating concept depicts that if a solution x dominates another solution y, there are two conditions needed. (a), the solution x must be no worse than the solution y in all objectives. (b), the x at least has one objective better than y. the x will not dominate solution y if it violates any of (a) and (b). Therefore, this method is used to repeatedly assign a set of non-dominated solutions. Each solution is assigned a fitness which equal to its non-domination level. Solutions with lesser fitness are the better candidates to be selected for the next generation.

Calculating crowding distance of each solution is the measured to use the biggest cuboid to calculate distance between two neighboring solutions which are in the same non-dominating front in the same objective space. The solution having higher value of crowding distance is probably selected to the next generation.

Selection. Selection is the step of genetic algorithm that usually combines crossover and mutation to breed new population. It is done based on the crowding distance and fitness generated above for a tournament selection. In this paper, we choose binary tournament strategy for our selection method.

Crossover and Mutation. In order to preserve the diversity in solutions of the Pareto-optimal set, crossover and mutation operators are used to mix parents' genetic materials for generating population.

Combine old and new *population*. When a new population has been generated, it will combine old population together in order to implement elitism and the non-dominating sorting on the combined population [8].

Replace the old parent population. Fig 3. shows a replacement strategy of NSGA-II [8]. An old population P_t combine a new population Q_t to form a new population R_t.. The variable t is the current generation. Then, the population R_t is ranked by non-domination. The solutions of the lower ranking fronts such as the front F_1 and F_2 are selected initially to replace parent population. If the parent population still has room for replacement, the solutions having large crowding distance will accommodate in the parent population. These steps are repeated until the termination criteria are contented.

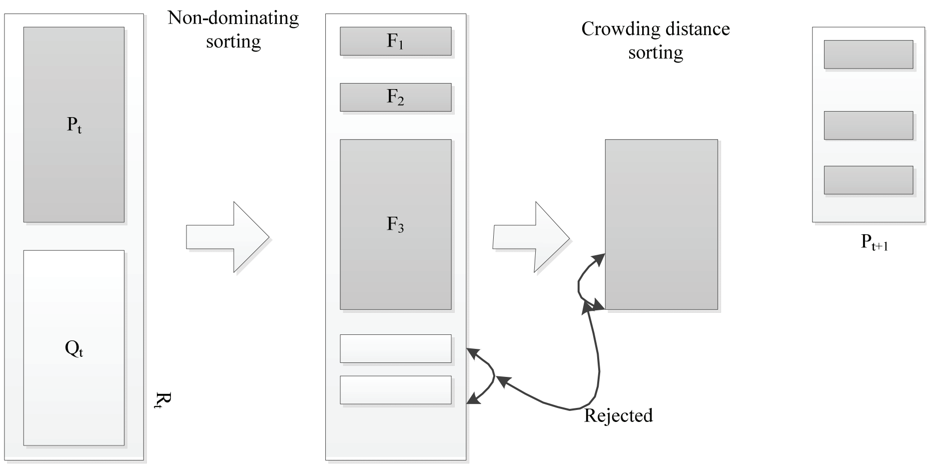


Fig. 3. Optimization population replacement procedure

3.3. SOLUTION MACRO-PROCESS

This macro process filters the solutions offered from previous layer and allows the user to define priorities for choosing the best solution. Finally, it passes the model to the next macro-process.

3.4. KNOWLEDGE REPOSITORY

This repository provides storage and retrieval capabilities for individual, collective, and organizational experience. Providing secure, reliable, location-independent, and fast access to SOEKS and DDNA structures is the main concern of this knowledge repository.

4. CONCLUSION AND FUTURE WORK

This paper introduces prototype of combining Decisional DNA with an evolutionary algorithm. This new prototype can extract information from different sources and convert it into knowledge which can be reused or shared with different systems. In addition, our research represents some small effort to address the above global challenge. The SOEKS and DDNA-based optimization algorithm appear to be a suitable and comprehensive tool for knowledge discovery. The combination of the SOEKS and the evolutionary algorithm not only improves efficiency of finding optimal solution by the evolutionary algorithm but also make new knowledge to be shareable, transportable and easy understandable. In our future efforts we plan to implement our prototype for real world optimization problems.

REFERENCES

[1] COOLEY R.,. MOBASHER B, and SRIVASTAVA J., *Web mining: information and pattern discovery on the World Wide Web*, in Tools with Artificial Intelligence, 1997. Proceedings., Ninth IEEE International Conference on, 1997, pp. 558–567.

[2] FEIGENBAUM E.A. and McCORDUCK P., *The fifth generation : artificial intelligence and Japan's computer challenge to the world*, Reading, Mass.: Addison-Wesley, 1983.

[3] SANIN C., MANCILLA-AMAYA L., SZCZERBICKI E., and CAYFORDHOWELL P., *Application of a Multi-domain Knowledge Structure: The Decisional DNA*, in *Intelligent Systems for Knowledge Management*, Vol. 252, N. Nguyen and E. Szczerbicki, Eds., ed: Springer Berlin / Heidelberg, 2009, pp. 65–86.

[4] PEARL J., *Heuristics: intelligent search strategies for computer problem solving*: Addison-Wesley Longman Publishing Co., Inc., 1984.

[5] GLOVER F., COX L.H., KELLY J.P., and PATIL R., *Exact, heuristic and metaheuristic methods for confidentiality protection by controlled tabular adjustment*, International Journal of Operations Research, 2007.

[6] ADLEMAN L.M., *Molecular computation of solutions to combinatorial problems*, Science-AAAS-Weekly Paper Edition, Vol. 266, 1994, pp. 1021–1023.
[7] TAN K.C., LEE T.H., and KHOR E.F., *Evolutionary algorithms with dynamic population size and local exploration for multiobjective optimization,* Evolutionary Computation, IEEE Transactions on, Vol. 5, 2001, pp. 565–588.
[8] DEB K., PRATAP A., AGARWAL S., and MEYARIVAN T., *A fast and elitist multiobjective genetic algorithm: NSGA-II*, Evolutionary Computation, IEEE Transactions on, Vol. 6, 2002, pp. 182–197.
[9] BANZHAF W. and REEVES C.R., *Foundations of genetic algorithms 5*. San Francisco, Calif.: Morgan Kaufmann Publishers, 1999.
[10] ROCHA M., NEVES J., VELOSO A., FERREIRA E., and ROCHA I., *Evolutionary algorithms for static and dynamic optimization of fed-batch fermentation processes,* Adaptive and Natural Computing Algorithms, 2005, pp. 288–291.
[11] FONSECA C.M. and FLEMING P.J., *Multiobjective genetic algorithms*, in *Genetic Algorithms for Control Systems Engineering, IEE Colloquium on*, 1993, pp. 6/1–6/5.
[12] ERICKSON M., MAYER A., and HORN J., *Multi-objective optimal design of groundwater remediation systems: application of the niched Pareto genetic algorithm (NPGA),* Advances in Water Resources, Vol. 25, 2002, pp. 51–65.
[13] DEB K. and AGRAWAL R.B., *Simulated binary crossover for continuous search space*, Complex Systems, Vol. 9, 1994, pp. 1–34.
[14] SANIN C. and SZCZERBICKI E., *Experience-based Knowledge Representation: SOEKS*, Cybernetics and Systems: An International Journal, Vol. 40, 2009, pp. 99–122.
[15] N.N.T. DUONG T.H., Jo G.S., *Constructing and Mining: A Semantic-Based Academic Social Network*, Journal of Intelligent & Fuzzy Systems, Vol. 21(3), 2010, pp. 197–207.
[16] WANG P., SANIN C., and SZCZERBICKI E., *Introducing The Concept of Decisional DNA-Based Web Content Mining*, Cybernetics and Systems: An International Journal, Vol. 43, 2012, pp. 97–117.
[17] SANIN C., S*mart Knowledge Management System*, Faculty of Engineering and Built Environment, School of Engineering, Faculty of Engineering and Built Environment, School of Engineering, University of Newcastle, Newcastle, Australia, Newcastle, Australia, 2007.

discrete event simulation, modelling, healthcare, regional system

Marek LUBICZ, Jacek ZABAWA*

ADVANCED SIMULATION MODELS OF REGIONAL HEALTHCARE SYSTEMS

We discuss further contributions to the problem of creating a simulation model of a complex regional healthcare system and implementing it using real-world data. The approach is based on modelling patients' pathways comprising diagnostic and treatment processes throughout patients' stay in the system. The general structure of the model and two main modules are outlined, with observations on problems regarding modelling logical processes within the regional system, and pitfalls when modelling large real-world systems. We illustrate the approach with a sample model of a regional system of hospital care in the Lower Silesia Region of Poland for lung cancer patients.

1. INTRODUCTION

Simulation modelling has been for years used to support solving complex decision problems in healthcare systems. In many cases the subjects of modelling are specific healthcare units, such as hospital wards, operation theatres, emergency departments, or ambulatory clinics. One of the most frequent modelling approaches in the field is based on queuing-system or patient-flow perspective, where specific healthcare resources are used for serving streams of patients sequentially travelling through the units of the system. A specific, patient-oriented framework of this type is based on Clinical Pathways, also called Pathways of Care or Clinical Profiles [7], which model the 'path' followed by an ill person through a healthcare unit or clinical process. The path may be defined for single 'episodes' of care (e.g. thyroid surgical treatment, performed at ambulatory and hospital level, from initial endocrinologist consultation to hospital-based surgery and follow up [8]), or for multiple 'episodes' of care, constituting a multi-level long-standing sequence of diagnosis-treatment-follow-up episodes (e.g. for lung cancer (LC) patients, undergoing multiple procedures in a number of

* Faculty of Computer Science and Management, Wroclaw University of Technology, Wybrzeże Wyspiańskiego 27, 50-370 Wroclaw, Poland.

healthcare premises [6]). In the latter case, single ‘episodes of care’ (e.g. radical surgery plus a series of chemotherapy or radiotherapy) can be standardized and modelled in the same way as for the first case, but the repetitions of the ‘episodes’ are difficult for modelling, as they greatly depend on primary disease progression, co-morbidities, physical and mental condition.

In addition to patient-centred perspective, when deciding on the best use of available resources, medical technologies and selecting the most efficient clinical pathway (best care for an individual patient), healthcare decision makers look at the same problem from system-wide perspective, deciding on the use of scarce resources of healthcare systems for the benefit of a population of patients (greatest good for greatest number). Some well known examples of modelling projects of this type include models for specific care groups (emergency care [2], coronary care [1,3], or thyroid surgery[8]), or investigations of specific formal problems arising when modelling complex healthcare systems (designing simulation experiments to evaluate surgical care policies [9], or comparing conventional and distributed approaches to simulation of complex health systems [5]).

The present authors extend previous results [6] in developing a simulation model of a complex Regional Healthcare System (RHS) and discuss problems associated with implementing the model using real-world data from a regional system of hospital care in the Lower Silesia Region of Poland. To simplify the presentation, the considerations are limited to the category of thoracic surgery (TS) and pulmonary oncology LC patients. The ultimate goal of developing the model is to provide an Intelligent Decision Support Environment, which could assist regional healthcare management decisions, using patient-centred (patient pathways based) and system-wide perspectives.

The paper is organized as follows. In section 2 we outline the assumptions and general structure of the model. In sections 3 and 4 we present selected logical processes associated with two main modules of the simulation model: First_arrivals and First_hospital_stay. In section 5 we discuss specific issues of implementing the model with ExtendSim [4] and conclude on the actual results of the model development.

2. MODEL ASSUMPTIONS AND GENERAL STRUCTURE

The simulation model described here is a generalized version of a first prototype presented in [6]. It was developed using discrete-event-simulation methodology to investigate processes in a regional healthcare system of hospital care, serving population of a number of districts. Due to the specificity of the selected category of patients (hospitalised surgical LC patients), it is assumed that the patients, once in the system, may expect multiple ‘episodes’ of care, each consisting of the first diagnostic hospital stay and a number of consecutive stays of specific type (e.g. surgery followed by chemotherapy), while particular stays may take place in different hospitals. Taking into account patient-centred perspective, the following main modules of the model could be defined:

– module 1: First_arrivals of the new-comers to the hospital care subsystem,

– module 2: First_hospital_stay, when main and subsidiary diagnoses are specified, and treatment pathway (methods and sequences) is defined; this module is further decomposed into:

– submodule 21: an interface between arrivals and admission to hospital,

– submodule 22: A&D/ED (emergency department), where decisions on patient admission or transfer to another hospital are made (on the basis of patient status as well as available resources: beds, financial limits),

– submodule 23: Hospital_Wards, modelling diagnostics and initial treatment,

– module 3: First_pathway, consisting of all hospital stays defined by the treatment pathway except the first (diagnostic) hospitalization,

– module 4: Next_pathways, which include: return to district when the first pathway is finished, and consecutive 'next_arrivals' and 'further_pathways', as needed according to patient status (e.g. reoccurrence or co-morbidities).

In addition, taking into account the regional healthcare perspective, we could define functional modules, which interfere with all patient-centred modules:

– population demography and epidemiology module (which determines parameters of the main first arrival and subsequent arrivals processes),

– healthcare resources module (hospital network, wards, beds, budgets, etc.),

– regional healthcare finance module (which in particular determines available hospital budgets and allocation or payment procedures),

– public health and policy module (e.g. health threats, inequalities in access).

3. MODELLING INPUT PROCESSES (FIRST_ARRIVAL)

Each would-be patient enters the system according to a dynamic random Poisson process with time-varying and district-dependent parameters (equal the reverses of the values for each district and particular date, as presented on Fig. 1); inter-arrival times are defined for particular year, day of the week, district of residence, and are recalculated as frequencies for each day of the simulation period. The resulting hierarchical model, as implemented in ExtendSim, is presented on Fig. 2.

Enter a schedule of arrival times

DB	_Create Time	Value	Value	Value 3	Value 4	Value 5	Value 6	
1	1/1/2006 0:00	1	2006	0,0112179487179	0,0080128205128	0,005608974359	0,0040064102564	0,0
2	2/1/2006 0:00	2	2006	0,0495283018868	0,0448113207547	0,0369496855346	0,0149371069182	0,0
3	3/1/2006 0:00	3	2006	0,0328525641026	0,0528846153846	0,0328525641026	0,0120192307692	0,0
4	4/1/2006 0:00	4	2006	0,0344551282051	0,0536858974359	0,0264423076923	0,0088141025641	0,0

Fig. 1. Schedule of inter-arrival times distribution (Create Block) for specific districts

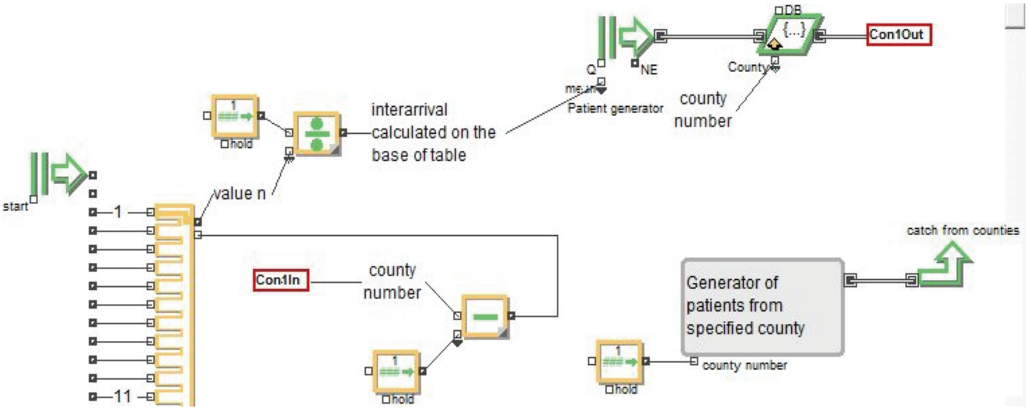


Fig. 2. Hierarchical block (lower right) and its structure for stream of patients from a district

Having generated an object from a particular district (a newcomer to the system on a particular day), the attributes of the object, namely: gender, age group, main disease (ICD-10 group), clinical priority (standard, urgent or critically ill), and number of hospital for the first admission, are determined, using input data database (Fig. 3), and dynamic procedures of selecting and using multi-attribute frequency distributions. The sample source input data, presented on Fig. 4-5, have been determined in a Data Mining project [6], which comprised analyses of the relationships between specific attributes, using Feature Selection and Variable Screening module of Statsoft STATISTICA 10 software (in the case study there are 29 districts, 36 hospitals with over 15k beds, serving approx. 5.5k newcomers per year, for the first time admitted to a hospital).

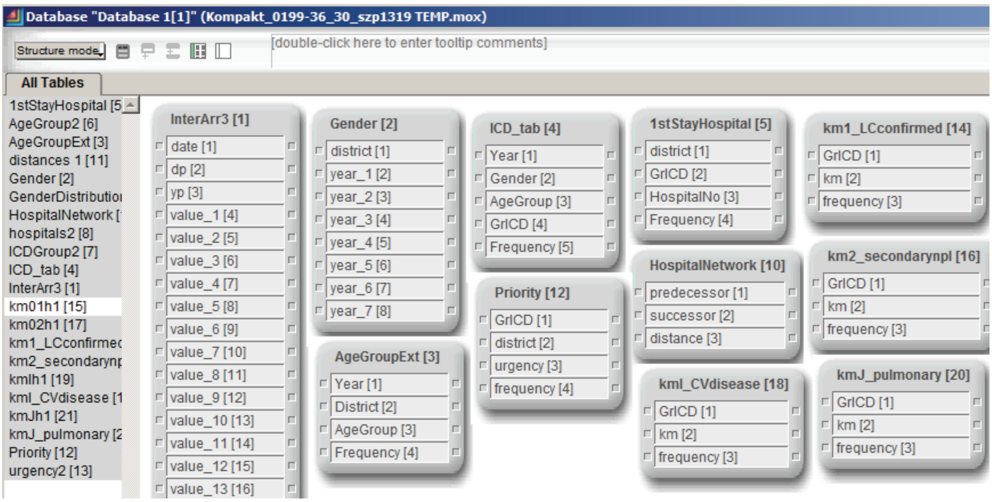


Fig. 3. Structure of the main model database, indicating assumed inter-attribute correlations

(a)

DB	Year	Gender	Age_Group	ICD_Group	Probability
152	2006	0	2	49	0,02247191
153	2006	0	2	50	0,16853933
154	2006	0	2	51	0.
155	2006	0	2	52	0

DB	County_number	ICD_Group	Hospital_Number	Probability
53384	29	27	33	0,120689655
53385	29	27	34	0,103448276
53386	29	27	35	0,465517241
53387	29	27	36	0,034482759

```
(b) real data;
Year = EDDateValue(EDGetCurrentDate(),9);
data=DBDataGetAsNumber(1,2,Year-2004,county);
DBDataSetAsNumber(1,9,2,1,data);
DBDataSetAsNumber(1,9,2,2,1-data);
(c) real Year, data;
integer auxiliary, i;
Year = EDDateValue(EDGetCurrentDate(),9);
auxiliary=(Year-2006)*29*10+county;
for (i = 0; i <10; i++)
{ data=DBDataGetAsNumber(1,3,4,auxiliary+29*i);
DBDataSetAsNumber(1,6,2,i+1,data); }
```

Fig. 4. Sample parts of multivariable empirical distributions (a) as well as ModL language code for database update enabling dynamic generation of gender (b) and age group (c)

DB	county_number	year_1
0	1	0,365758754864
1	2	0,39298245614
2	3	0,431472081218
3	4	0,342465753425
4	5	0,353741496599
5	6	0,427906976744
6	7	0,347457627119
7	8	0,338983050847
8	9	0,305084745763

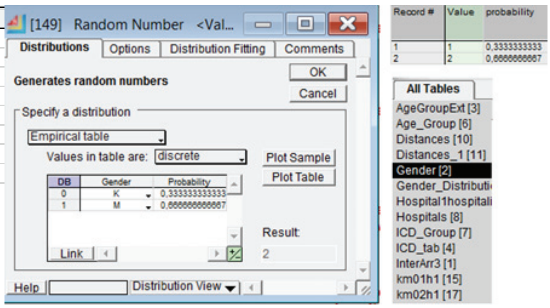


Fig. 5. Linking Table Gender from Database to Empirical Table in Random Number

4. MODELLING INITIAL PATHWAY (A&D AND FIRST HOSPITAL STAY)

The first phase in the initial pathway (submodule 21) is concerned with creating specific object admission attitudes for ordinary (elective) versus urgent and critical (emergency) patients, having been generated in the previous 24 hours of simulation time. It is assumed that emergency patients once generated are sent to the selected hospital, while elective patients have determined time-slots each day (e.g. admission hours 8–12 am). Patients of the latter class should wait until the time-slot, so in the model they must be grouped in a buffer (Fig. 6), which can be described as bunching phase. Technical implementation in ExtendSim involves creating a virtual activity block (lower right on Fig.6), which processes a virtual object generated at the start of the time-slot and activates gate opening signal for Select_item block, which in turn enables all waiting objects to enter the A&D processing buffer.

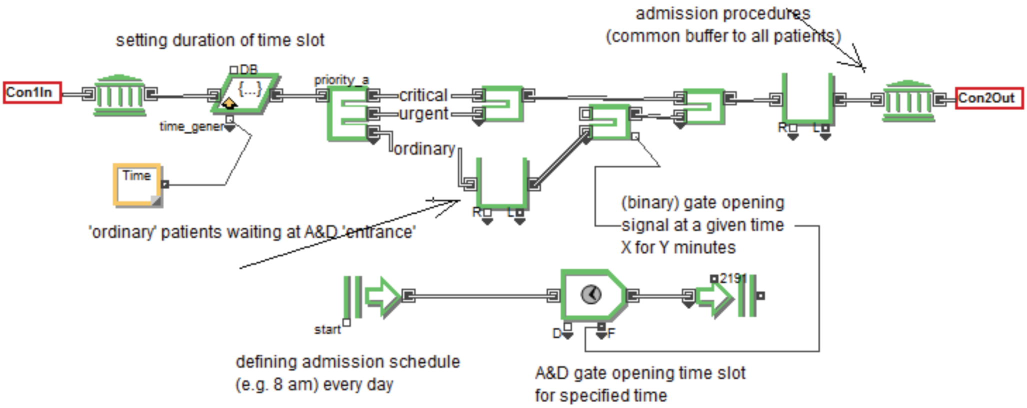


Fig. 6. A simplified diagram for bunching phase (grouping standard patients, waiting for admission hours in a particular hospital)

The most complicated submodule in the current version of the model - sub-module 22 (A&E/ED department at a particular hospital) - is used for making admission decisions for mixed priority newcomers, as well as objects transferred from other hospitals (due to lack of empty beds or monthly limit overspending), and objects assigned to the hospital, coming for subsequent care episodes of their current pathway. Decisions are made in relation to particular categories of resources, and - on the basis of a set of decision rules, each of which sets off a series of calculations or activities. For instance a simplified general rule for urgent (not critical) newcomers may be formulated as:

if an emergency_bed free then assign_scl (subject to cost limit) else if a standard_bed free then assign_scl else if standard extra_bed available then assign_scl else if a standard_bed is to be vacated within h hours then assign_scl else look for an empty bed in other hospitals else assign a (virtual) unlimited bed with immediate preemption, i.e. transfer to a standard bed once it is vacated.

Once the decision rule is fired, it involves starting a number of activities, for instance checking current balance for the hospital (assign_scl) or determining the earliest completion dates for ward service in this or the nearest hospital. Fig. 7 illustrates connection between the A&D (22) and Hospital_wards (23) sub-modules. The latter one models medical procedures for patients admitted to a hospital, using assumptions concerning statistical features of processing times (LOS - length of stay in a hospital) and costs, modelled using results of statistical analysis of the real-world data [6]. In the case study the following assumptions have been made:

– results of diagnostic phase are defined as clinical attributes, depending mainly on the main diagnosis (frequency distributions of attributes km; Fig.1), and generated simultaneously using a dedicated model unit (Fig. 8),

– output distributions for length of stay and costs are fitted as theoretical distributions separately for each combination of priority and all four clinical km attributes; for instance for patients with confirmed LC with metastases, as well as pulmonary and cardiac complications, the LOS is fitted as Lognormal with parameters minimum, mu, sigma equal (0.329, 2.04, 1.15) respectively.

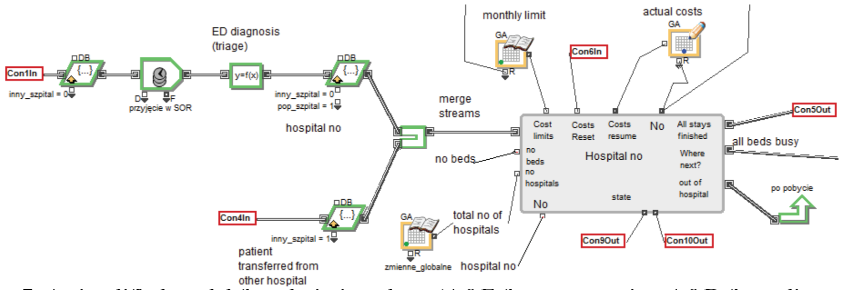


Fig. 7. A simplified model for admission phase (A&E for emergencies, A&D for ordinary)

```
(a) real given_probability;
integer i; auxiliary=ICD_Group;
//km01h1
given_probability=DBDataGetAsNumber(1,14,3,auxiliary);
DBDataSetAsNumber(1,15,2,1,given_probability);
DBDataSetAsNumber(1,15,2,2,1-given_probability);
```

(b)

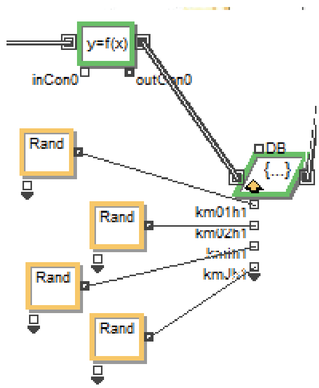


Fig. 8. Sample ModL language code for simultaneous generation of multiple hospital stay attributes (a) and its graphical image in the model (b)

As described above decisions in the A&E module for consecutive objects are made in accordance with the state of the Hospital_wards module with the objects admitted so far. Fig. 9 presents a simplified model of pre-empting (PE) ward service of patients assigned to unlimited extra beds. In this case Activity blocks with PE option have been used, where PE function is initiated according to the priority of the object, using current dynamic information on beds (of different type) in use and current status of the objects on unlimited extra beds; the remaining process time is stored in an attribute and used for completing the service on standard beds once available.

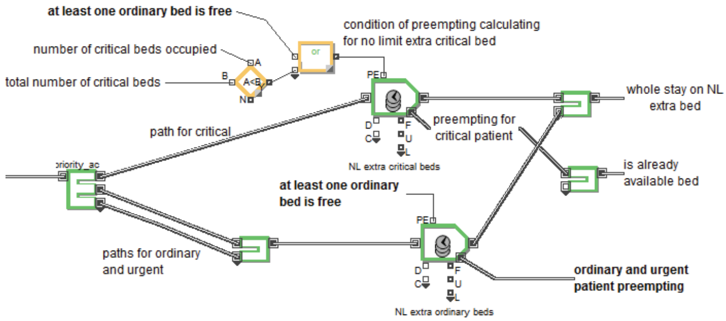


Fig. 9. Sample model of pre-emption unit for unlimited extra beds

5. TECHNICAL IMPLEMENTATION ISSUES AND CONCLUSIONS

The model was implemented in ExtendSim AT 8.0.2 [4], which proved to be a flexible discrete event simulation environment. Nevertheless a number of technical problems had to be solved in this project stage. For a large number of recognizable objects (approx. 150k newcomers and over 0.5 million stays for standard parameters and 5-years simulation period) and realistic simulation parameters required for statistical accuracy (50–100 runs have been used), both the memory considerations and the speed of simulation resulted in designing problem-oriented input and output data management solutions. In particular to enable validation and further statistical analysis of the output data the following solution for output data collection from a large number of runs (15 million of objects, each described with 10 attributes) was developed (Fig. 10)

– determine the number of global objects (that is incremented in subsequent runs) and save the number of objects of the next iteration in a separate, dedicated global array,

– allocate a global array needed to store the attributes of objects (to take care of saving the working memory)

– read the attribute set of the objects and save them to the resulting global array,

– export filled global array to a text (csv) file at the end of all runs.

```
(a) integer arrayIndex, arrayIndex2;
integer i,k;
integer j;
arrayIndex = GAGetIndex("l_w");
GAMultisim(arrayIndex, 1);
j=CurrentSim;
if (j==0)
{ GASetInteger(object_number, arrayIndex, j, 0);
 if (object_number==1)
 { arrayIndex2 = GAGetIndex("wyniki_02");
 GAResizeByIndex(arrayIndex2,0);
 GAResizeByIndex(arrayIndex2,inCon0); }}
else
 {k=GAGetInteger(arrayIndex,j-1,0);
 GASetInteger(object_number+k, arrayIndex, j,0 ); }
```

(b)

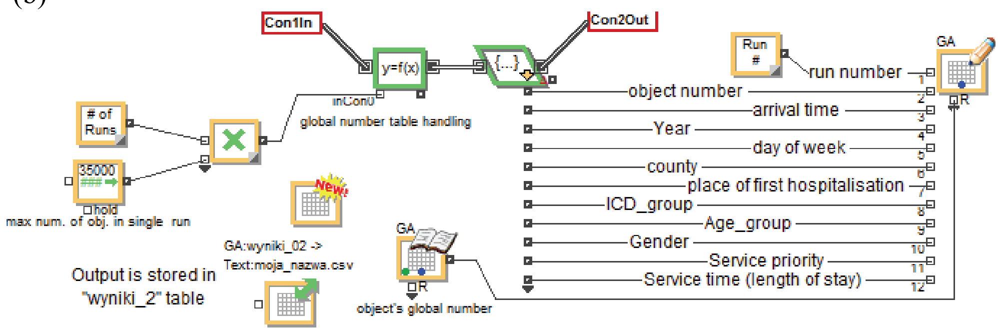


Fig. 10. Sample ModL language code for collecting results of multiple runs (a) and the structure of the corresponding unit of the model (b)

There are also other challenging problems faced during implementation, such as: algorithms for indicating bed vacancies in other hospitals or determining expected times of vacating beds suitable for object of particular priority, dynamic control of monthly limits for particular hospitals and for the whole region. One of the most difficult modelling problems was also selecting the approach for categorizing (data mining phase), defining and implementing conceptual models for region-wide patient pathways (modules 3 and 4).

We conclude with a general observation that in contrast with system dynamic approach, more often applied for modelling large healthcare systems, the application of the discrete-event-based approach enables the modeller to investigate in more details parallel clinical and managerial processes in a regional healthcare system, taking into account stochastic nature of the processes and a great variety of patient categories, resource types and what is very important: detailed costing and payment procedures.

ACKNOWLEDGMENTS

The research by Marek Lubicz was financed by the National Science Centre under the grant N N115 090939 "Models and Decisions in the Health Systems. Application of Operational Research and Information Technologies for Supporting Managerial Decisions in the Health Systems".

REFERENCES

[1] BENSLEY D.C., WATSON P.S., MORRISON G.W., *Pathways of coronary care - a computer-simulation model of the potential for health gain*, IMA Journal of Mathematics Applied in Medicine and Biology, Vol. 12, No. 3–4, 1995, 315–328.

[2] BRAILSFORD S.C., LATTIMER V.A., TARNARAS P., TURNBULL J.C., *Emergency and on-demand health care: modelling a large complex system*. Journal of the Operational Research Society, Vol. 55, No. 1, 2004, 34–42.

[3] COOPER K., DAVIES R., RAFTERY J., RODERICK P., *Use of a coronary heart disease simulation model to evaluate the costs and effectiveness of drugs for the prevention of heart disease*. Journal of the Operational Research Society, Vol. 59, No. 9, 2008, 1173–1181.

[4] Extendsim 8 User Guide; Imagine That Inc. 2010.

[5] KATSALIAKI K., MUSTAFEE N., TAYLOR S.J.E., BRAILSFORD S., *Comparing conventional and distributed approaches to simulation in a complex supply-chain health system*. Journal of the Operational Research Society, Vol. 60, No. 1, 2009, 43–51.

[6] LUBICZ M., MIELCZAREK B., *Simulation Modelling of a Regional Healthcare System – Problems, Framework, Implementation in Arena*. In: Information Systems Architecture and Technology. System Analysis Approach to the Design, Control and Decision Support, J. Świątek et al. (eds). Oficyna Wydawnicza Politechniki Wrocławskiej. Wrocław, 2012, 15–24.

[7] LUBICZ M., RICHARDS B., MIELCZAREK B., LAYCOCK P, SOCHOROVA A., SYPUŁA J., *Modelling clinical and managerial profiles: the value to providers of healthcare*, in: Current perspectives in healthcare computing, Pt. 1., B. Richards (ed), Weybridge, 1998, 353–362.

[8] OZCAN Y.A., TÀNFANI E., TESTI A., *A simulation-based modeling framework to deal with clinical pathways*, Proceedings of the 2011 Winter Simulation Conference, S. Jain, R.R. Creasey, J. Himmelspach, K.P. White, M. Fu (eds.), 2011, 1190–1201.

[9] SOBOLEV B., KURAMOTO L., Cluster-randomized design for simulation-based evaluation of complex healthcare interventions. Journal of Simulation, Vol. 4, No. 1, 2010, 24–33.

financial modeling, discrete event simulation, integrated approach

Jacek ZABAWA*

INTEGRATING SIMULATION AND FINANCIAL STATEMENT CREATING PROCESS

There are many gaps between practice of accounting and simulation modeling. The author had already proposed a new approach to integrate concepts of discrete event simulation modeling and accounting for the generation of financial statements in production processes. The new integrative approach is the multiplication of object representing the discrete events (e.g. sale or purchase) to physical aspect object and book-entry form object. It provides the new abilities to generate and analyze financial ratios. With this approach it is possible to generate not only easy to design items like revenue income statements, cost of goods, incomes, expenses or even activity-based cost but also the balance-sheet with particular items as payables and receivables. This paper presents a business model that uses a new approach and the relationship between supply and demand.

1. LITERATURE REVIEW

Discrete-event simulation (DES), formalized Bernard Zeigler describes the operation of the system as a sequence of events (changes in the system). We often use it to model queuing systems. Key concepts of DES are: event (immediate change in the system), event list (at the time of the simulation events are generated according to a predefined schedule or randomly generated as a consequence of other events), activity-delay (keeping the event / object for a specified period, representing for example duration of the activity that can be deterministic, defined as random or be a function of other variables). Mentioned factors include the type of object / event and the current time in the simulation run.

The models often appears as delay of a different application: it is the delay of the system state changes associated with the state of system resources, e.g. waiting in line to take a position in support (corresponds to the nature of the activity-delay). DES

* Institute of Organisation and Management, Wroclaw University of Technology, Wybrzeże Wyspiańskiego 27, 50-370 Wrocław, Poland, e-mail: jacek.zabawa@pwr.wroc.pl

modeling is currently based on the graphic patterns and building basic mapping relationships between objects and resources and sequence of tasks by context: generators, queues, delays, switches and exits. DES uses mechanisms (often hidden from the user) like the event-based, activity-based, process-based and three-phase approaches [5].

The system dynamics approach is a base for continuous simulation modeling that was developed by Jay W. Forrester. Concepts of system dynamics are visualized by the causal-loop diagrams correspond to the system structure. Knowing the structure characteristics presented in the diagrams (feedbacks and time delays), we can predict the behavior of the system (the trajectories of the variables), but in a computer simulation tools are used more direct record by stock (level) and flow (rate) diagrams that can no longer be seen as a graphical representation of the system of differential equations. If we take into account the frequency of analysis, the continuous simulation calculations are performed regardless of the rate of change of the system state and in the discrete-event system simulation calculations are "increased" to the rapid changes.

In this work the use of the method developed by the author of DES integration and financial reporting will be presented. Foundations and the first use of a new method were presented in the paper [8]. The presented models have been made in the simulation package Extendsim [6]. This tool has been used by the author in a number of previous works such as activity based costing modeling [9]. Modeling in Extendsim is done by placing the blocks on the desktop (i.e., structural elements that make an interpretation the system model – original). Blocks are retrieved from the program libraries (in discrete event simulation: Item and Value) and combined with lines to transfer objects and information and define the model structure. We begin by recalling the basic structures developed in Extendsim that are useful in this approach, and then discuss the assumptions of a simulation model of a business process that uses a new approach, and finally presents the results of experiments and discussion. After developing the model (with more complex experiments, the model is equipped with a block structure for collecting experimental results) take a decision on its launch and conduct experiments (simulation run time, number of runs, the time units). Possibilities of the modeling environment include building their own units, modifying existing action blocks, grouping blocks in hierarchical structures, display charts, etc.

Problems concerning the use of simulation approach in the analysis of cause and effect, what-if analysis and forecasting the impact of different variants of decisions in the area of financial management in manufacturing companies may be caused by differences in the concepts of accountants and engineers [3], lack of formalization of the relationship between material and financial flows [7] - despite the obvious possibility of treatment of cash management issue as an inventory issue, taking into account uncertainty [1], slightly identified relationships between the cash position and planning in tactical or operational dimension [2]. It can be assumed that the reasons for the lack of creative thinking include: deficiencies in the curriculum for DES training [4], such as focusing on mathematical modeling and the applications of simulation tools.

2. THE NEW METHOD AND ITS APPLICATION

2.1. DESCRIPTION OF INTEGRATED APPROACH

The necessity of developing new methods of integrating DES and financial reporting arises from the lack of the possibility of direct calculation of basic categories such as accounting liabilities and receivables. Their value cannot be determined in the same way as income, expenses and profits. For example (Fig. 1), the income of the simulation models are setting multiplying the number of objects that represent sold for (at constant prices) by the price or adding income unit (the variable costs and prices depending on the type of things).

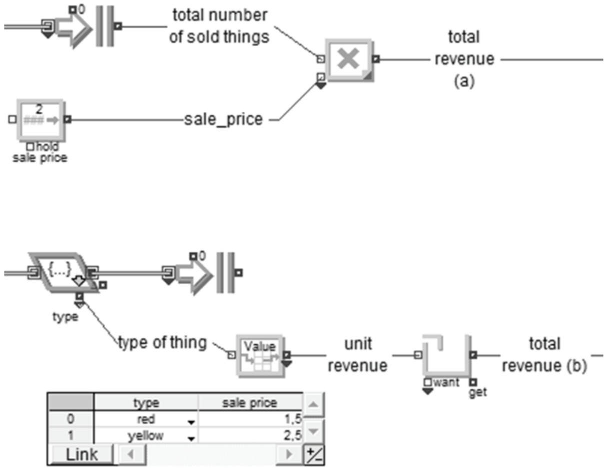


Fig. 1. Calculation of total revenue – two situations

The following submodels presents the essence of the discussed approach: multiplication of the event representation (the object) to the two representations: real objects and accounting type objects. First, we describe a phenomenon called liabilities repayment delay, which involves a relationship (a consequence of the time) cost and total expenditure. We present it in model by a block sustaining object for some time (depending on the identified factors or random). In order to determine the amount of liabilities we need to know the current value of cost and total expenditure. Left side of Fig. 3 shows hierarchical block containing part of the model (service model) for generation of input objects, and for calculating the (total) cost. The core of the method is the multiplication operation (proportion 1:1) of the input object. There are two separate output (Unbatch) objects: one for stream of things and one for accounting record type. It is important to ensure that both facilities still carries the full information about the attributes. This makes it possible to calculate the single and total expenditure on account of liabilities arising from the cost impairment losses incurred (but not yet settled) in obtaining the object to the system.

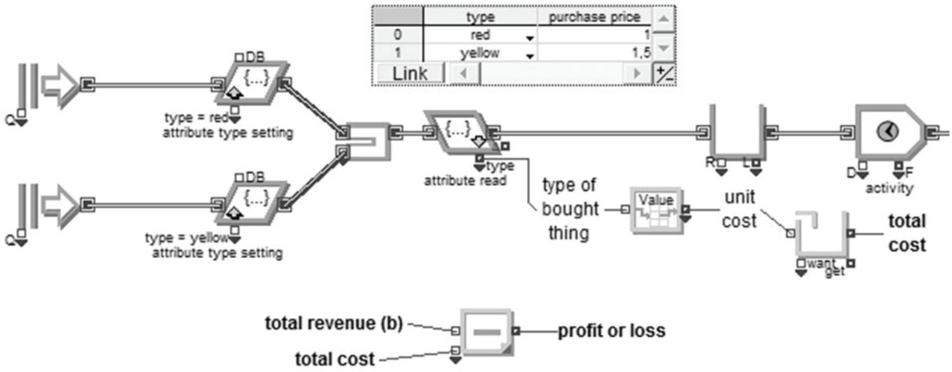


Fig. 2. Calculation of total cost and profit or loss

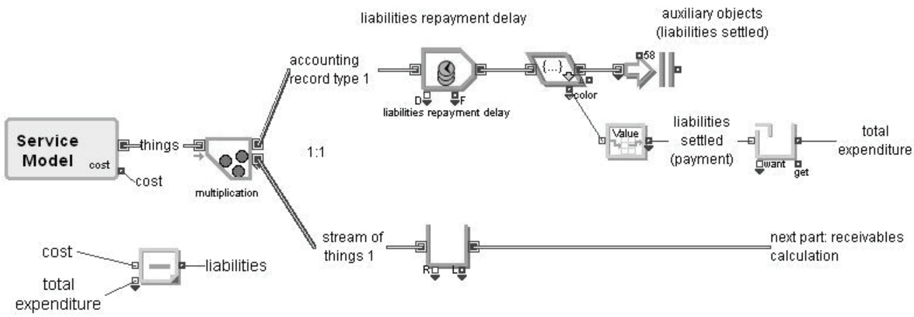


Fig. 3. Model of the liabilities calculation. An example of a new method
– the calculation of the simulated total expenditure and the current status of liabilities. [Source: 8]

We will now discuss economic modeling phenomenon called receivable repayment delay. We apply a new approach in the calculation of the receivables, which binds the relationship between total revenue and total income. This is a consequence of the time. We present it in a block model sustaining business for some time (depending on the identified factors or random). In order to determine the amount of receivables we need to know the current value of the total cost and total income.

The core of the method is the multiplication operation (proportion 1:1) of the input object. There are two separate output (Unbatch) objects: one for stream of things and one for accounting record type Also, at this point it is important that each of these two objects still carries the full information about the attributes. This makes it possible to calculate the single and total income after taking account of receivables arising from the prices (unit revenues) for the sale of property (in the lower part of the model). Accounting record type 2 objects after passing through the block to simulate receivables repayment delay, and after reading the attribute "color" will be sent to the upper branch and disappear. The lower part of the model include stream of the "real" objects.

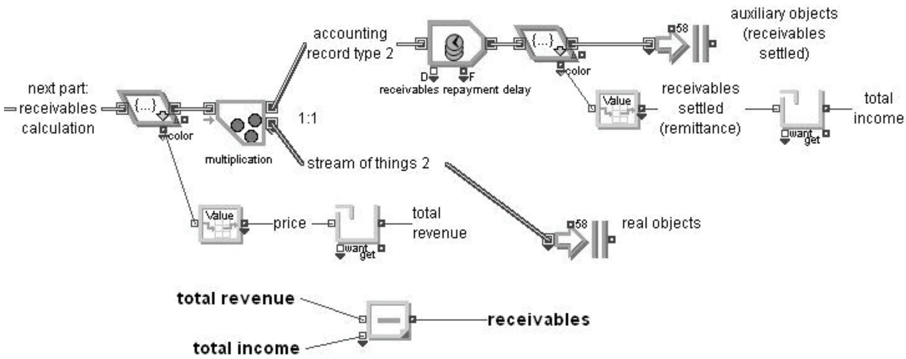


Fig. 4. Model of the receivables calculation. An example of the new method – the calculation of the simulated total income and the current status of receivables. [Source: 8]

2.2. DESCRIPTION OF PROPOSED MODEL

We consider the business model involving the buying and selling of goods of one kind. Goods (products) are purchased from wholesalers, delivered in batches, stored in a warehouse, retrieved from storage when there is a potential buyer (customer) and committed to buyer. The model takes into account the phenomenon of supply and demand, the cost of purchasing the product, the phenomenon of delayed repayment of liabilities from the purchase of the product, cover the shipping costs depending on the size, the cost of storage of the products and their settlement of the recorded profits, sales events and payment of the sold good phenomenon delay of receivables from the sale of the product, the calculation of liabilities and receivables, cash and profit and loss account.

The discussion of the model will begin by defining the function of sell price elasticity of demand in terms of demand per unit of time (time unit). Demand [unit/time unit] = max (0, a [unit/time unit]- b [unit^2 / time unit * currency]*price [currency/unit]) where price >=0. That means, that for sufficiently high prices demand is zero and zero for the price there is a finite volume of demand (equal to a [unit / time unit)].

Demand (namely the intensity of the demand) in the present model is interpreted as follows. The greater the intensity of the demand, the shorter intervals between successive objects - realizations of demand (potential customer). For example, current demand of 100 [unit / time unit] means that the time interval between successive letters of demand is 0.01 time unit. The model assumes that the time intervals between successive letters of demand are realizations of the exponential distribution with parameter 1 / demand. More specifically parameter values and assumed that the coefficient a will be equal to 1200 [unit / time unit] and the parameter b will be equal to 1 [unit ^ 2 / time unit * currency]. This means that the range of reasonable prices and sales volume covers the range from 0 (maximum value of current demand 1200 customers in the [time unit]) to 1200 [currency] (current demand already zero). The model module represents the above assumptions is presented in [Fig. 5].

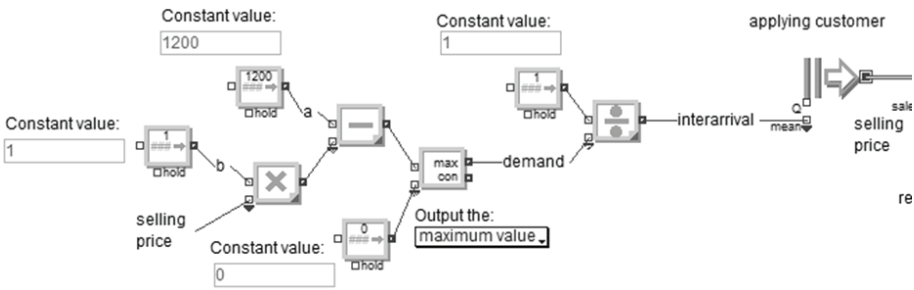


Fig. 5. Calculation of stream objects intensity: interarrival parameter (mean in exponential distribution)

The second discussed module allows you to set the selling price in the course of the experiment: we can declare the price fixed or set it as a constant in the course of time intervals (Fig. 6). Note the interpretation of the margin (for simplicity does not give her in percent) and to periodically repeat the scenario changes the selling price.

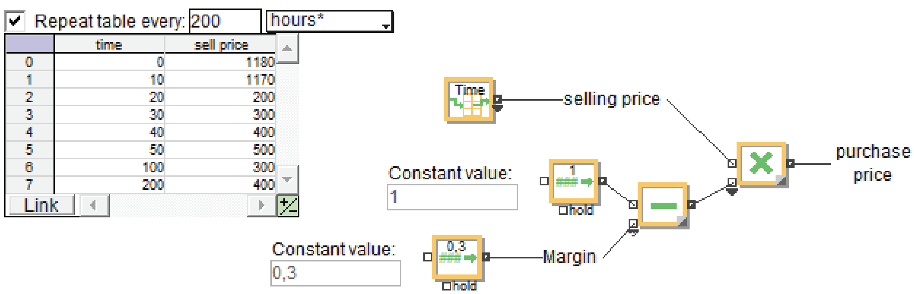


Fig. 6. The establishment of the selling price in the experiment and calculation of purchase price

The third module shows the relationship between the flow properties of demand and supply. Lower left part of Fig. 7 shows the continuation of the module shown in Fig. 5. We give the value of the attribute “sale_price” object representing the client. It will be recalled that the intensity depends on the arrival of customers – as a function of demand - not just from the sales price. The upper left part of the figure shows a portion of the processing of objects from a stream of supply-side (products). It should be noted block “Set” in the attribute “waiting_time” records current time of placement of the product in stock (supply2). The role of the attribute “waiting_time” is to determine the residence time in the store and consistently outstanding waiting cost by calculating the sum of the costs of individual storage products. The issue of financing the costs of storage products will be discussed further below.

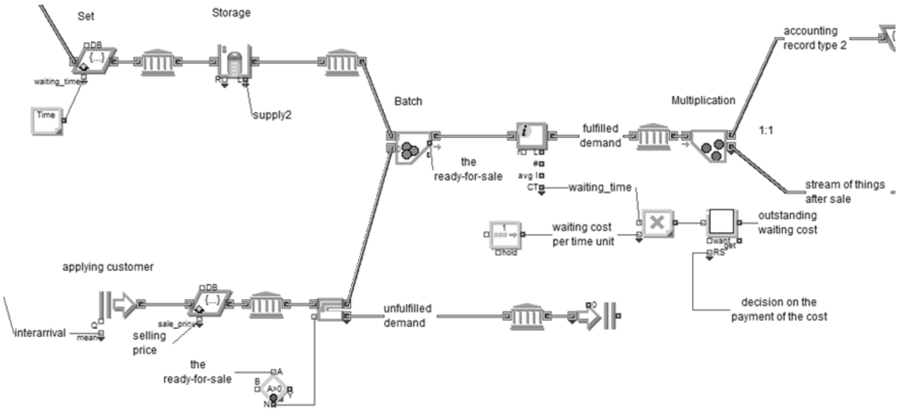


Fig. 7. Part of the model representing relations between demand and the supply

One unit (object) representing the product is transferred from the warehouse to block the “Batch” where the object expects demand (customer) and indicates the status of “the ready-for-sale”. As soon as the object appears, followed by a combination of supply-side object and the demand, the appropriate copy the required attributes and calculating the real waiting_time to stream fulfilled demand. When in the block “Batch” no one object does exist awaiting for demand, the client does not wait for the appearance of new supply-side object, but is classified as an unfulfilled demand (the valuation of lost customers can be considered by the average unit profit on sales).

The process of ordering (buying) products in this model are as follows Fig. 8. Founded “infinite” number of possible to carry supplies from wholesale (9999 units). Deliveries begin if no supply is not active or fit supply (supply) is less than the current demand arising from the sale price. It was assumed that the volume is determined relative to the contract demand volume per unit of time and will be in the range of 0 (blank delivery) and 1 (delivery of demand over the unit of time). The mentioned value is stored in the attribute "volume" in place to control the type of block Unbatch (multiplication). At the time of the contract purchase price is to be read and stored in the corresponding attribute.

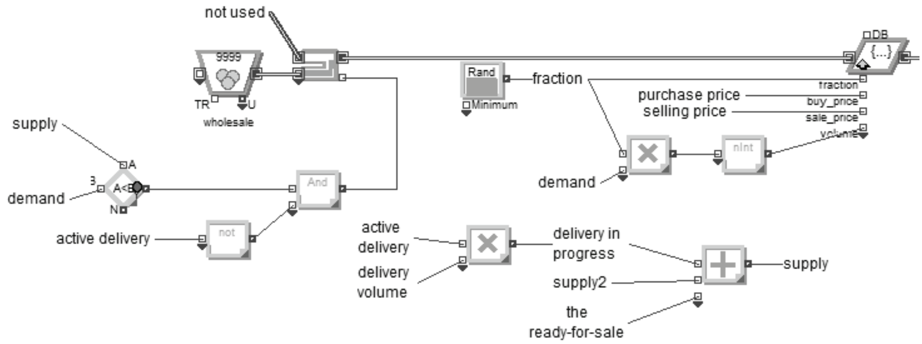


Fig. 8. Module of the delivery starting and determining the size of delivery

The process of the product delivery and unpacking supplies (multiplication) cost calculating is as follows (Fig. 9). Each delivery gets the attribute "costs" necessary to determine the total cost and should take into account the cost of buying all the products that are in it. We include here the cost of delivery that will be proportional to the purchase price of the product, the number of products (delivery Particular value). Shipping costs increase by a factor dependent on the size of delivery and if delivery includes up to 100 units of the product is the delivery cost increase of 100 currency units and more than 100 200 currency. We read here the value of the attribute "costs" for subsequent deliveries and calculate their sum (total cost) using a block-type Holding Tank. Then the object that represents the supply is stopped at the time of delivery (delivery in progress); signal “active delivery” is used in the ordering process module supply (Fig. 8). After the delivery, the block type Unbatch (unpacking supplies) make unpacking supplies, multiplexing the appropriate number of products. It should be noted separation (1:1 multiplication by block Unbatch) input for a “product” of two streams (branches): lower (objects that represent the “real” products) and upper (objects used in the calculation of expenditures) with the use of which is shown in Fig 10. Note the block type of Activity (Liabilities repayment delay), which represents the time between delivery of the products (but after the time of transport, it was also assumed zero time unpacking supplies) to pay for these products by the modeled business. This time the total recorded value of the attribute “costs” is equal to the total expenditures: knowing the total cost and total expenditures we can calculate the current liabilities.

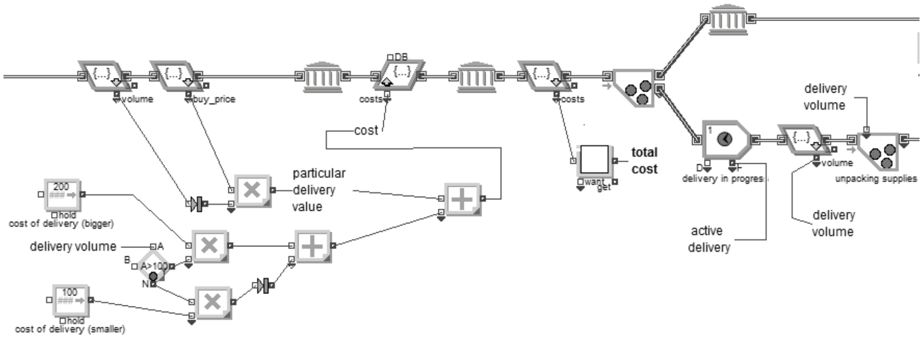


Fig. 9. The module defining the cost of delivery, total cost of products (without cost of storage) and multiplication (unpacking supplies)

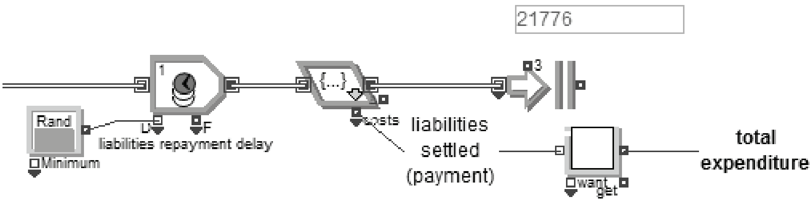


Fig. 10. Module for determining the level of total expenditures

The process of calculating the total revenue and total income is as follows (Fig. 11). The complex object (fulfilled demand) is divided (block Unbatch) into of two objects that follow separate streams: accounting record type 2 and stream of things after sale. For further calculations, less important is the flow of a second – we cannot use it to determine the number of satisfied customers demand side. Read the attribute “sale_price” object type “accounting” (receive unit revenue) and then we sum it in Holding Tank unit and get “total revenue”. It should be noted Activity block type (receivables repayment delay), which represents the time from product sales until you receive cash for these products by the modeled business. Re-read the attribute “sale_price” object type “accounting” (receive unit income), and then we sum it in Holding Tank unit and get “total income”. Because we know the value of total revenue and total income, we can calculate the current receivables. It should be noted that the location of the unit blocks defining the total revenue before or after the block Unbatch has no effect on the results obtained.

The left part of the next module (Fig. 11), constructed of blocks of libraries Value and Plotter is responsible for calculating sought financial statement: Liabilities, receivables and profit. The course of these values may be displayed on the chart. The right side of the module is responsible for the determination of cash and an estimate of storage cost.

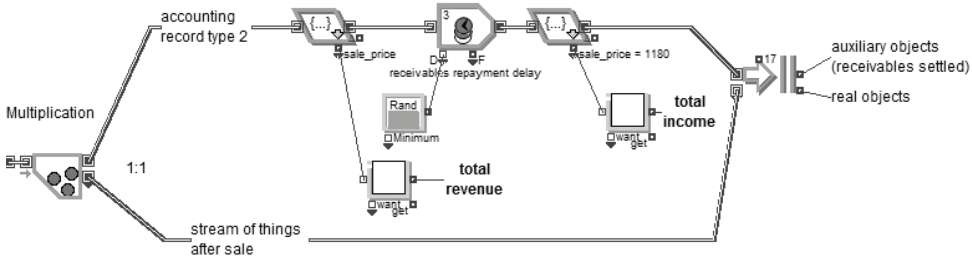


Fig. 11. Module for determining the level of total revenue and total income

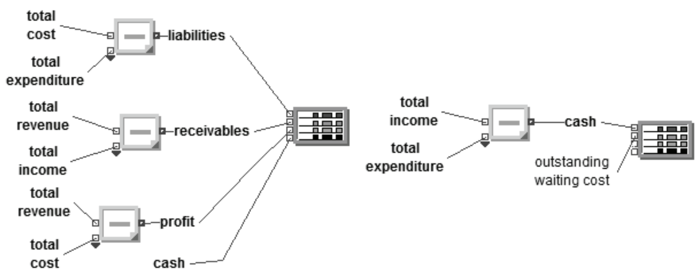


Fig. 12. Module for calculating the amount of liabilities, receivables, profit and payment of the storage cost decision making.

3. SAMPLE EXPERIMENT AND CONCLUSIONS

The experiment was performed five times. It was assumed that the selling price would be fixed in a given course, sales price of 0, 200, 600, 800, 1200 monetary units at each subsequent run, the demand per unit of time will be settled by delivery (fraction = 1). Duration of experiment: 24 time units. The main flow diagram for the results of the selling price of 200 is shown in Fig. 13, and the main results are shown in Table 1. From the chart we find that the receivables are much lower than liabilities. We see higher cash levels than profits and growth trend of cash and profit.

First experiment was repeated by adding the possibility of a random fraction values of the satisfy demand (this means that only a part of the demand per unit of time will be met): fraction between 0 and 1, the uniform distribution.

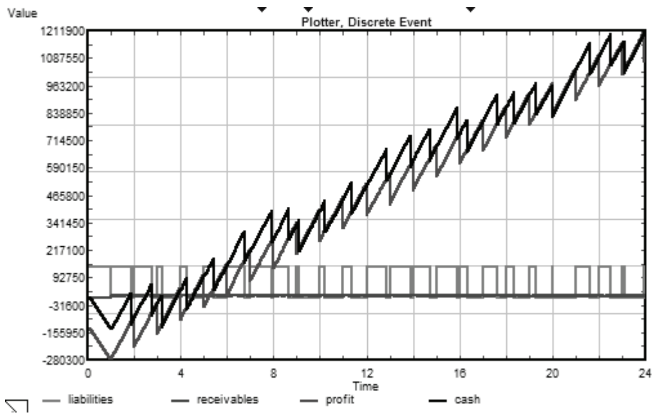


Fig. 13. Chart of the most important results for the sell price = 200.

Table 1. Experiment 1. The values of the principal results in successive runs

Sell price	0	200	600	800	1200
Liabilities	0	$1,4 * 10^5$	$2,5* 10^5$	$2,2 * 10^5$	0
Receivables	0	$1,1*10^4$	0	0	0
Profit/loss	-4100	$1* 10^6$	$2 * 10^6$	$1,8 * 10^6$	0
Cash	-4100	$1,1* 10^6$	$2,2 * 10^6$	$2 *10^6$	0
Total revenue	0	$4,6* 10^6$	$8,3 * 10^6$	$7,4 * 10^6$	0
Total expenditure	4100	$3,3* 10^6$	$6 * 10^6$	$5,4 * 10^6$	0
Total income	0	$4,5* 10^6$	$8,3 * 10^6$	$7,4 * 10^6$	0
Total cost	4100	$3,5* 10^6$	$6,3 * 10^6$	$5,6 * 10^6$	0
Outstanding waiting cost	28000	$1,4 * 10^4$	$4,1*10^3$	$1,8 *10^3$	0

It should be noted that the selection of the most favorable (cash, profit) variant in both experiments did not change. Moreover, significantly decreased the volume Liabilities (sometimes even to zero). As expected, in the absence of demand (sell price = 1200), all results are equal to zero, whereas the sell price = 1200 only losses are recorded below zero and cash and total costs and total expenditures.

3.3. SUMMARY

This paper presents the second application of a new method for simulation and integration of financial reporting during model of business taking into account the phenomenon of supply and demand. The method can also be used in more complex conditions, for example taking into account the phenomenon of prepayments (not shown in this paper). It seems that it would be interesting to apply the method in the construction of the balance sheet and also on the model current account the company, which is needed to pay off liabilities and repaid receivables. Interesting results can also apply the method to other simulation packages.

REFERENCES

[1] BAUMOL W.J., *The transaction demand for cash: an inventory approach*, The Quarterly Journal of Economics, Vol. 66, No. 4, 1952, 545–556.

[2] FENIES P., TCHERNEV N., *From advanced planning system to advanced budgeting system: the next step in supply chain management software,* In: Pathways to Supply Chain Excellence, Groznik A., Xiong Y. (Eds.), InTech, 2012.

[3] GARDNER L.L., GRANT M.E., ROLSTON L.J., *Converting simulation data to comparative income statements*, Proceedings of the 1992 Winter Simulation Conference, Swain J.J., Goldsman D., Crain R.C., Wilson J.R. (Eds.).

[4] MOLNAR I., *Simulation: body of knowledge,* In: Handbook of Research on Discrete Event Simulation Environments: Technologies and Applications, Abu-Taieh, E. M., El-Sheikh A. (Eds.), Information Science Reference, 2009.

[5] PIDD M. *Computer simulation in management science*, Wiley & Sons, 1998

[6] STRICKLAND J.S., *Discrete event simulation using ExtendSim 8*, Lulu 2011.

[7] VIDAL C.J., GOETSCHLACKX M., *A global supply chain model with transfer pricing and transportation cost allocation*, European Journal of Operational Research, Vol. 129, No. 1, 2001, 134–158.

[8] ZABAWA J., *Integrating Discrete Event Simulation and Financial Reporting*, Information systems architecture and technology: the use of IT models for organization management, Oficyna Wydawnicza Politechniki Wrocławskiej, Wrocław 2012.

[9] ZABAWA J., *Zasady rachunku kosztów działań (activity-based costing) na przykładzie symulacyjnego modelu systemu obsługi*, Symulacja systemów gospodarczych, Oficyna Wydawnicza Politechniki Wrocławskiej, Wrocław 2003.

PART 3

MODEL BASED CONTROL AND DECISION SUPPORT

eHealth, sport training, mathematical modeling, optimization

Krzysztof BRZOSTOWSKI*, Adam STĘPIEŃ**, Andrzej DOMAŃSKI*

MODEL BASED ALGORITHM TO PLAN THE LONG-TERM TRAINING FOR SPORT PERFORMANCE OPTIMIZATION

In the paper problem of planning long-term training for sport performance optimization is considered. After short introduction to the problem and survey on solutions in the field of sport training support, the system to wireless sensing physiological data and their processing is presented. Designed system was presented in the details: architecture and the main elements of the proposed approach results of experimentations are presented and discussed. In the final part of this work model-based algorithm to build plan of physical training is given. Our algorithm to design plan of training is based on Banister's model describing relationship between input training (workload) and output training (physical performance). In order to verify proposed approach experimental studies have been performed.

1. INTRODUCTION

Combination of wireless sensor network technologies with system analysis techniques and modern machine learning paves the way for useful and powerful tools. Such tools can be put into practice in various areas: industry, healthcare, sport, emergency management, entertainment etc. Recently, this area of application of wireless technologies and system analysis techniques is intensively investigated by researchers. Moreover, many systems based on these technologies and techniques are built by commercial manufactures. Intensification of research in this area is related to rapid development of segment of mobile phones market [1, 4, 6, 11].

In this paper we formulate and solve the problem for one of the mentioned above area: sport science. Our aim is to design an algorithm for system to support both recre-

* Wrocław University of Technology, The Institute of Computer Science, Wybrzeże Wyspiańskiego 27, 50-370 Wrocław, Poland.

** University School of Physical Education in Wrocław, Faculty of Biomechanics, al. I.J. Paderewskiego 35, 51-612 Wrocław, Poland.

ational and professional training. Proposed system fits in worldwide trend of merging modern wireless technologies with system analysis and machine learning techniques.

Usually such systems use wireless technologies to transfer data from sensing devices. Sensing devices (or simple sensors) usually operates with Bluetooth or ZigBee interfaces. They are popular because combining them with mobile devices such as smartphone helps to design and built system that are comfortable to use and not bothersome. Depending on used sensing devices different networks are built for designed system. The main are called BAN (Body Area Network) or PAN (Personal Area Network).

Sensed data can be processed locally utilized resources of mobile devices or remotely taken advantage of servers or server farms. It is clear that data processing algorithms must be conformed with computational unit. It means that different versions of the same data processing algorithm are designed and implemented both mobile device and remote server.

In this paper we proposed an algorithm to be applied for designed system to support sport training. This algorithm is intended for planning long-term physiological training for both recreational and professional athletes.

This paper is organizing as follows: after short introduction, review of the contemporary solutions in the area of sport training is given. Then architecture of proposed system is presented with placing emphasis on data processing modules. Next section is devoted to proposed algorithm to plan long-term physiological training. At the end results of experimentations, discussion and final conclusions are given.

2. STATE-OF-THE-ART ANALYSIS

As it was mentioned various commercial and research teams work on systems to support both recreational and professional athletes. The results of their efforts are architectures for such systems, “lightweight” and “heavyweight” data processing and supporting decision-making algorithms and even new sensors. On the market we can find solutions with different functionalities, more or less sophisticated. Some of them used built-in sensors, some of them can be connected with sensors through wireless interfaces such as Bluetooth. Most of them uses simple methods to data processing (e.g. simple statistics). But we can find systems based on more sophisticated methods in which, for example, mathematical models are applied [8].

One of the most popular manufacturer is Polar. This company offers either wearable devices for athletes and complex solutions for coaches [9]. Devices from the first group are composed of wrist worn watch and chest worn heart rate sensor. Depending on aims of the training, Polar provides equipment for fitness improvement or for sport performance maximization. These devices incorporate elements of motivational feedback i.e. generate beeps every time when certain amount of calories is burnt. The second group of Polar's products are designed for team sports. Besides heart rate monitors Polar offers software to monitor fitness capabilities of team. It is extremely useful in order to eliminate injuries and prevent overtraining. Moreover the software provides tools to optimize training intensities.

Suunto is another large company that offers wearable devices to support physical training both for amateurs and professionals [10]. Among others the company provides devices that generate personalized training plan. Based on results of user's training monitoring these devices are capable to make recommendations for training volume i.e. the frequency, duration and intensity. Moreover proposed plan can be changed in order to adjust to user's current capabilities i.e. when user's activity level decrease.

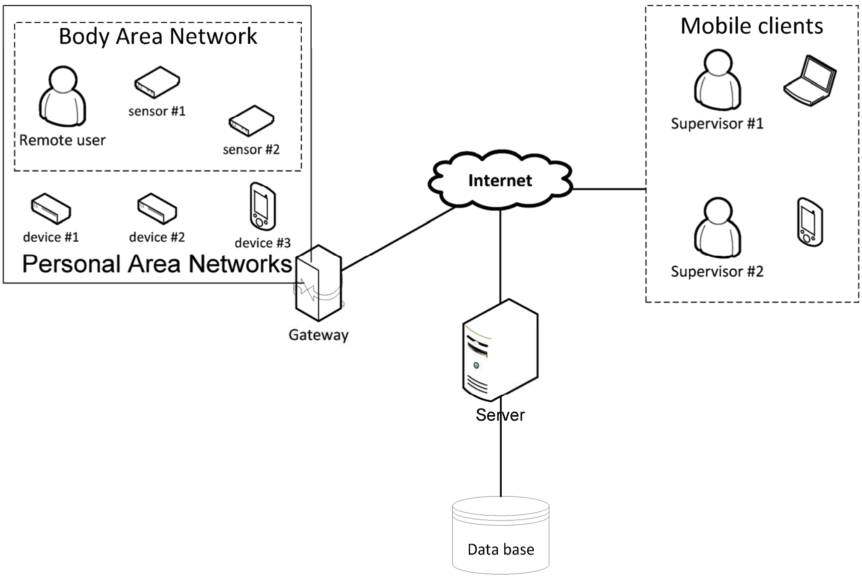


Fig. 1. Architecture of the system

MOPET is developing under research project [2]. It is wearable system to supervise a physical fitness activity. This system has features typical for systems offered commercially by Polar or Suunto. Additionally acquired data are used to build user's model. This model is periodically updated in order to produce exact prediction of user's performance. The results of model based prediction is apply to provide safety and healthy advices adapted to the user. MOPET has also implemented mechanisms that motivate the user by graphical and audio suggestions. For example in order to provide audio feedback system tells the user current speed and recommend increasing or decreasing current speed. Similar project are described for example in [5, 7, 12].

3. ARCHITECTURE OF THE SYSTEM

Proposed architecture of the system (see Fig. 1) is typical three-tier architecture in which following modules are distinguished: data acquisition module (BAN and/or PAN), data processing and supporting decision making module and presentation module (user interface).

In the first group we design wireless sensor networks that are composed of sensing nodes. These nodes are placed on athlete body or in his/her surroundings. In the first case we have BAN whereas in the second case we have PAN [1].

Sensed data can be pre-processed on personal server i.e. mobile device (nowadays in most cases we use smart phones or cellular phones) and then results of computation are presented to the user on the same mobile devices or transmitted to the remote computational server through Internet. It is usually happens when the cost of computations is high and it is not reasonable to continue calculations on mobile phone. We design our system by the means that mobile device can be applied only to the computational tasks which are simple and they are not demanding complex calculations.

The second module is computational module which is used to manage data processing functions. These functions characterize by heavy computational costs and it is not possible to use them on mobile device. Hence they are execute on remote servers and the results of computations are sent to the user. It is worth stressing that methods designed to process data on remote server are "heavy" in opposite to the algorithm executed on mobile device which are "light". Apart from methods of data processing we design some algorithms to support making decision and implement mathematical models.

The last collection of units is used to present results of data processing and decision making such us visualization (e.g. charts) and reporting (e.g. tables).

4. PLANNING LONG-TERM SPORT TRAINNG PROGRAM

The aim of the sport training is the same for recreational and professional athletes: final score. But there are difference between both groups in physical, physiological, technical and even mental preparation of the athlete. However, in order to reach the best possible score both group of athletes have to improve their skills and train regularly. To this end continuous monitoring of athletes performance is needed. Typical periodization of training is built of training periods called macro-, meso-, and microcycle. In each period there is diagnosis made using specialized tests in order to evaluate current preparation of the athlete. Sometimes during training it is necessary to determine current relationship between workload and physical performance of the athlete but there is limited or no access to laboratory. In this case mobile system to support sport training with feature to measure athlete physical condition is helpful and, in many cases, it is sufficient.

In this paper an algorithm to plan the long-term sport training program is proposed. Our approach based on well-known Banister model of sports performance. Combining this model with system analysis techniques we proposed new method to plan physiological training for recreational and professional athletes.

4.1. MODELLING RELATIONSHIP BETWEEN SPORT TRAINING AND PERFORMANCE

In [3] a model describing relationship between training and physical performance is presented. In the rest of the paper we call this model *Banister's model*. This model is helpful to find question on: how does sport training modify physical performance of athletes throughout the training period? Applying it in computer simulation it is possible to analyze how different intensities and durations of sport training (i.e. volume of the sport training) influence on physical performance of an athlete. Presented model with some parameters describes an athlete and his/her physiological state. In order to analyze influence of sport training on particular athlete it is necessary to determine parameters of Banister's model which characterize the athlete. In this subsection such model is described and procedures to determine parameters of this model is presented.

The model describing relationship between sport training and physical performance is composed of two parts: fitness and fatigue. Fitness (stand for $g(t)$) and fatigue (stand for $h(t)$) are response on workload which we call training input and denoted as $w(t)$. General formula describing relationship between sport training and physical performance is [3]:

$$p(t) = p_0 + \theta_1 g(t) - \theta_2 h(t) \tag{1}$$

where: p_0 initial physical performance of athlete, θ_1 is related to athlete's fitness response to training and θ_2 is related to athlete's fatigue response to training. Functions $g(t)$ and $h(t)$ have similar form:

$$g(N) = \sum_{n=1}^{N-1} w(n) \exp\left(\frac{-(N-n)}{\theta_3} \right) \tag{2}$$

$$h(N) = \sum_{n=1}^{N-1} w(n) \exp\left(\frac{-(N-n)}{\theta_4} \right) \tag{3}$$

where: θ_3 is related to decay effects for fitness and θ_4 is related to decay effect for fatigue. Combining formulas from (1) to (3) the final model describing relationship between sport training and physical performance has form:

$$p(N) = p_0 + \theta_1 \sum_{n=1}^{N-1} w(n) \exp\left(\frac{-(N-n)}{\theta_3}\right) - \theta_2 \sum_{n=1}^{N-1} w(n) \exp\left(\frac{-(N-n)}{\theta_4}\right) \tag{4}$$

In relationship (4) parameters from θ_1 to θ_4 are determined based on measurements. Variable $w(n)$ and $p(n)$ stand for input training (workload) and output training (physical performance) respectively. In our model they have to be measured. In our research we proposed heart rate to estimate workload of sport training. It means that athlete throughout training wearing heart rate monitor to collect such data. It is not perfect but one of the most convenient indicator of workload of training [3]. The second problem is related to measurement of physical performance. In this case we proposed Cooper test-based on method to determine $V_{O_2 \max}$ of the athlete. Utilization of Cooper test is limited to the sports in which endurance is predominate factor. In our case, karate, this procedure is sufficient.

As it was mentioned above parameters (θ_1 - θ_4) of the model (4) have to be estimated in order to build relationship for specific athlete. To this end we have to collect data and then with use of identification procedure determine these parameters. Let us denote measurements for input training sequence (workload) as $W = \{w_m(k)\}_{k=1}^{K}$ and measurements for output performance (physical performance) as $P = \{p_m(k)\}_{k=1}^{K}$. Then, we propose performance index for estimation task in the form:

$$Q(\theta) = \sum_{k=1}^{K} \left(p_m(k) - p(k)\right)^2 \tag{5}$$

Our aim is to find parameters (θ_1 - θ_4) such as minimize defined performance index (5):

$$\theta^* \to Q(\theta^*) = \min_{\theta} Q(\theta) \tag{6}$$

The results of solving estimation task (6) is presented in Figure 2. Model (4) with estimated parameters is applied in the problem of planning program for long term sport training.

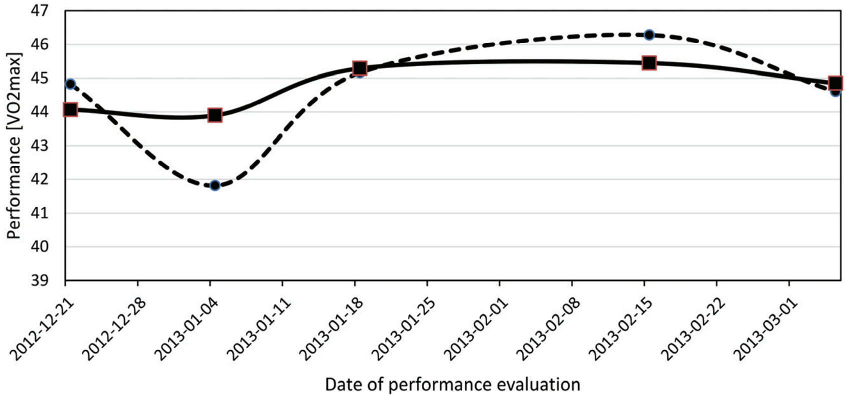


Fig. 2. Physical performance curves: measured (dashed line) and estimated (solid line)

4.2. ALGORITHM TO PLAN PROGRAM FOR LONG-TERM SPORT TRAING

In this section algorithm to determine program for long-term sport training is introduced. Let us rewrite model (4) in compact form:

$$p(n) = \Phi(w(n), \Theta) \tag{7}$$

where: $\Theta = [\theta_1 \quad \theta_2 \quad \theta_3 \quad \theta_4]^T$ is vector of parameters.
Problem of planning program of sport training is as follows:
for a given:
– model $\Phi(w(n), \Theta)$, see (4);
– initial physical performance of athlete p_0;

find: training load $\{w^*(n)\}_{n=1}^{N-1}$ for next training session that maximize physical performance i.e.: $p^*(N)$.

Let us define performance index for formulated optimization task:

$$p^*(N) = Q(w(n), p(n), p_0, \Phi) \tag{8}$$

In order to find optimal training plan following optimization task must be solved:

$$w^*(n) = \arg \max_{w(t) \in W} Q(w(n), p(n), p_{0,} \Phi) \tag{9}$$

where W is a space of possible training plans. Example results obtained solving optimization task (9) is presented in Figure 3.

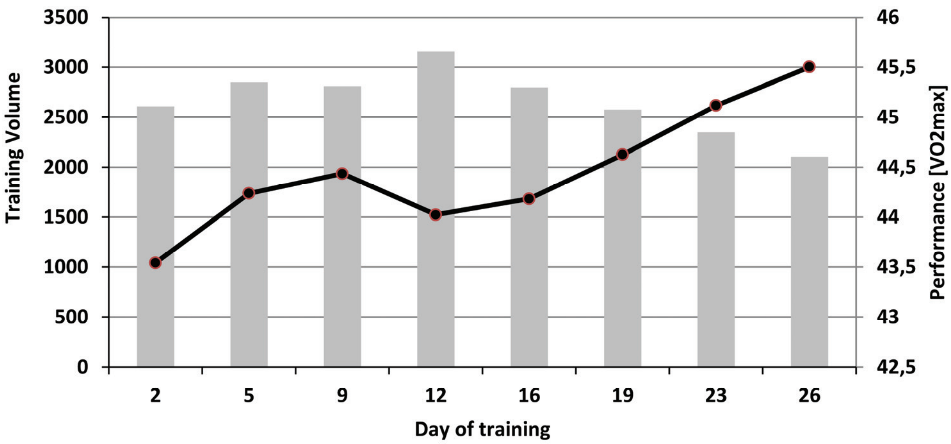


Fig. 3. Relationship between training volume (bars) and performance (solid line)

5. EXPERIMENTAL STUDIES

In this section results of experimentation is presented. In our study one young athlete is involved. In order to collect data throughout training we design application to sense data from heart rate monitor. Collected data are used to estimate workload of training and to performance Cooper's test. Based on this data we determine training plan which afterwards is used by athlete during sport training.

In upper plot of Fig. 4 determined workload (bright bars) and executed workload (dark bars) is presented. Tests are performed in three weeks. It can be seen that in almost all cases executed volume of the training was slightly larger than it is determined. But, as it is shown in Fig. 4 long-term of training was achieved. It is easy to noticed that real (dashed line) and predicted (solid line) curves of physiological performance are different. The reason is that dynamics of physiological performance of athlete is more sophisticated that Banister's model and not all factors are taken in this model into account. However this model – as it can be seen in Fig. 4 and previous ones (Fig. 2 and Fig. 3) – are able to reconstruct of the dynamics of response athlete body on workload.

6. CONCLUSIONS

In the paper an approach to plan optimized program for long-term sport training is presented. Proposed solution based on Banister formula and numerical methods of optimization. In order to verify proposed approach experiments are designed and performed. To collect data during training wearable sensors and mobile system has been built. Sensed data are processed in order to determine athlete-oriented plan of sport training.

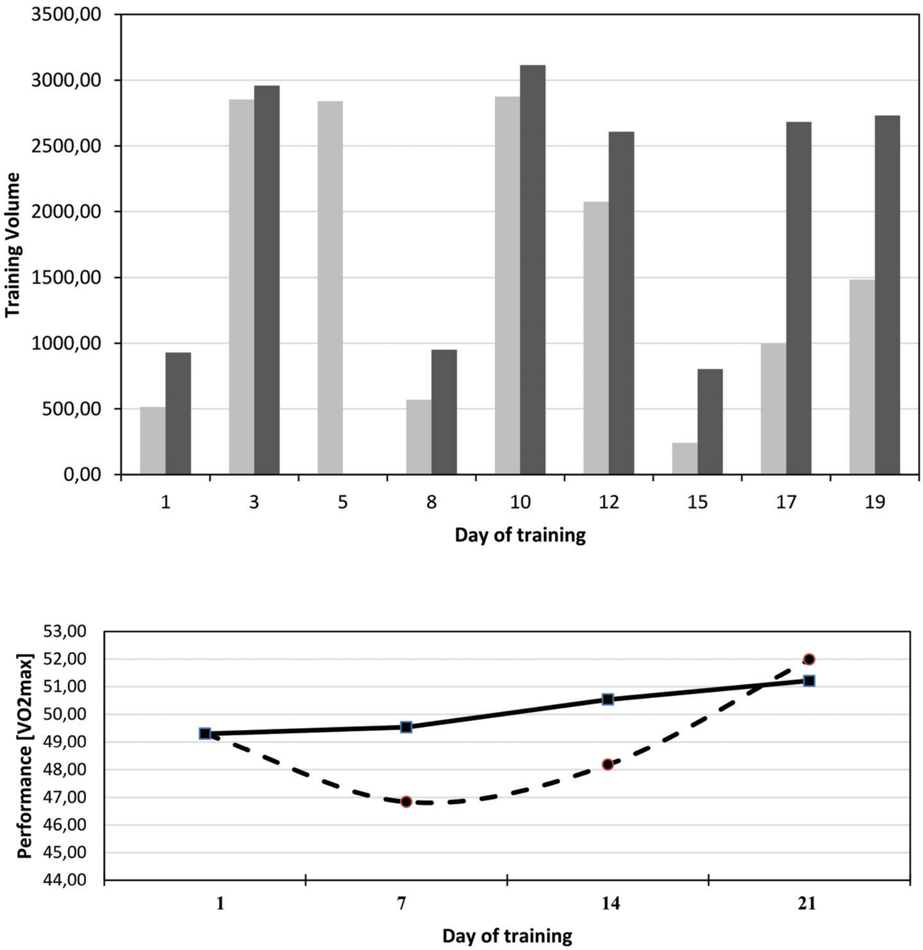


Fig. 4. Planned (light bars) and executed (dark bars) training volume (above) and predicted (solid line) and measured (dashed line) performance (below)

Presented in section 5 results vindicate application of presented approach in described problem. It is must be stressed that the final decision is made by coach or athletes (for example when they train for recreation). Proposed system just recommend and advise further workload of the sport training. But, as it can be seen, resting on recommendations produced by our system give encouraging results.

Future plans are connected with analysis of usefulness more sophisticated models for the problem of modeling relationship between sport training and physical performance. The reason is that Banister's model is linear but we think that such relation is more sophisticated. Finding model which better describe relation between workload and physical performance helps us do build new model-based algorithm to design program of sport training with improved performance.

Proposed approach to support sport training has great commercial potential. Prototype of the system has been tested and will be used as supporting tool in the sport training of young athletes in sport club.

REFERENCES

[1] ALEMDAR H., ERSOY C., *Wireless sensor networks for healthcare: A survey*, Journal Computer Networks: The International Journal of Computer and Telecommunications Networking archive, Vol. 54, Issue 15, 2010, 2688–2710.

[2] BUTTUSSI F., CHITTARO L., *MOPET: A Context-Aware and User-Adaptive Wearable System for Fitness Training*, Artificial Intelligence in Medicine, Vol. 42, 2008, 153 – 163.

[3] CALVERT T.W. at al., *A Systems model on the Effects on Training on Physical Performance*, IEEE Transaction on Systems, Man and Cybernetics, Vol. SMC-6, 1976, 94–102.

[4] CHI-CHUN L.O. et al., *Ubiquitous Healthcare Service System with Context-awareness Capability: Design and Implementation*, Expert Systems with Applications, Vol. 38, Issue 4, 2011, 4416–4436.

[5] HAN D, LEE M, PARK S., *THE-MUSS: Mobile u-health service system*, Journal Computer Methods and Programs in Biomedicine archive, Vol. 97, Issue 2, 2010, 178 – 188.

[6] HERMENS H.J., VOLLENBROEK-HUTTEN M.M.R., *Towards remote monitoring and remotely supervised training*, J Electromyography and Kinesiology, Vol. 18, Number 6, 2008, 908–919.

[7] LIM J.E. et al., *A context-aware fitness guide system for exercise optimization in U-health*, IEEE Trans Inf. Technol. Biomed., Vol. 13, Number 3, 2009, 370–379.

[8] PANTELOPOULOS A., BOURBAKIS N.G., *A Survey on Wearable Sensor-Based Systems for Health Monitoring and Prognosis*, IEEE Transactions on Systems, Man, and Cybernetics, Part C: Applications and Reviews, 2010, 1–12.

[9] POLAR, www.polar.fi

[10] SUUNTO, www.suunto.fi

[11] VALES-ALONSO J. et al., *Ambient Intelligence Systems for Personalized Sport Training*, Sensors, Vol. 10, No. 3, 2010, 2359–2385.

[12] YOUM S. et al., *Development of remote healthcare system for measuring and promoting healthy lifestyle*, Expert Systems with Applications, Vol. 38, Issue 3, 2011, 2828–2834.

mathematical modelling, decision support system, optimization, input-output system

Jarosław DRAPAŁA, Mateusz BARCZAK*

IDENTIFICATION OF THE GLUCOSE-INSULIN COMPLEX SYSTEM

Mathematical model of the Glucose-Insulin system allows to simulate and predict the human body response to meals and insulin intakes. It may also be useful when performing control actions for diabetic patients. In order to make the model valuable for clinical applications, its parameters need to be estimated in such a way, that the model responses resemble responses of a particular person. The problem is that the Glucose-Insulin system is complex, whereas the number of measurement data is low. This makes estimation task ill-conditioned. The contribution of the work is twofold. Mathematical models describing different parts of the Glucose-Insulin system are combined into a single complex model. Moreover, parameter estimation routine based on numerical optimization methods is proposed for this complex system. Simulation study is performed to assess the quality of estimation routine.

* Institute of Informatics, Wrocław University of Technology, Wybrzeże Wyspiańskiego 27, 50-370 Wrocław, jaroslaw.drapala@pwr.wroc.pl, 168050@student.pwr.wroc.pl

oxyfuel combustion, radiation modeling, absorption coefficient, Turów lignite

Jakub DŁUGOSZ*, Maciej ZAWIŚLAK**, Halina PAWLAK-KRUCZEK*

IMPACT OF RADIATION MODELS ON HEAT TRANSFER PREDICTION IN DROP TUBE REACTOR DURING DIFFERENT COMBUSTION SCENARIOS

Energy and heat is necessary for civilization existence and growth. Nowadays energy needed for human communities is provided mainly as electricity in power generation processes from combustion of fossil fuels in power plants. Efforts are being made to eliminate carbon dioxide emissions to atmosphere from combustion processes. Such potential has studied by authors in drop tube reactor $O_2/CO_2/H_2O$ (oxyfuel) combustion technology. Changing of combustion atmosphere from air to $O_2/CO_2/H_2O$ results in different thermo-physical and chemical properties of heat transferring medium which finally affects the heat transfer process in combustion chamber. In oxy-combustion higher than in air combustion concentrations of CO_2 and H_2O (triatomic species playing key role in radiation) may affect especially radiation heat transfer and validated for air combustion radiation models are not suitable. Authors by calculating absorption coefficient studied how different models of radiation (those intended for air combustion and those especially modified for oxyfuel combustion) affected the heat transfer characteristics.

1. CO_2 EMISSIONS REDUCTION IN FOSSIL FUELS POWER PLANTS

Since long ago it has been pointed to reduce the harmful effects of exhaust gases and particulate emissions on the environment from thermal energy generation processes. In recent years, ongoing processes of climate change on earth has been established and linked with the air emissions of anthropogenic carbon dioxide (and other greenhouse gases already known) and the issue of reducing CO_2 emissions have become to

* Institute of Thermal Power and Fluid Mechanics, Faculty of Mechanical and Power Engineering, Wrocław University of Technology, Wybrzeże Wyspiańskiego 27, 50-370 Wrocław, Poland, e-mail: jakub.dlugosz@pwr.wroc.pl

** Institute of Machine Design and Operations, Faculty of Mechanical Engineering, Wrocław University of Technology.

play a key role [1-3]. Further development of the energy sector will be dependent on environmental regulations, which are generally becoming more restrictive. The priority of Polish national policy and the European Union is to improve energy efficiency while reducing carbon dioxide emissions and other major pollutants (SO_x, NO_x and particulates). Almost 90% of the electricity demand in Poland is covered by the 55 power plants and CHP (combined heat and power) power plants, which are fired in 60% by hard coal and 38% by lignite [1]. Coal will be still an attractive future fuel for power generation in Poland and many other countries abundant in coal resources provided the so-called clean coal technologies will develop to further reduce carbon dioxide emissions. Using any new coal technology helping to mitigate climate change is generally more expensive than applying traditional technologies and will have, of course, the inevitable impact on energy prices increase. Due to large carbon content in coal (67–75% in lignite and 78-92% in hard coal [1]) CO_2 emissions abatement from coal combustion may be done only by carbon capture and storage (CCS). CCS involves the capture of carbon dioxide its transport and storage. There are three main lines of capturing action: (a) after the combustion process (called post-combustion capture), (b) combustion in oxygen-enriched atmosphere (called oxyfuel combustion) and (c) before combustion (called pre-combustion). At the moment it is difficult to identify among them the preferred technology. Each of these requires the continued research and demonstration phase.

2. OXYFUEL COMBUSTION OF COAL

The conventional coal fired power unit uses in the combustion process oxygen from atmospheric air (78% vol. N_2). Nitrogen is therefore - as an ordinary ballast - the main component of combustion products in the air environment. Dilution of combustion products with plenty of nitrogen makes it difficult to remove CO_2 from the flue gas. Alternative to the conventional combustion process may be the combustion of the coal with the use of pure oxygen (95% purity is satisfactory and economically optimal), this approach is called **oxyfuel combustion**. The advantage is a completely different composition of the exhaust gases, which contain essentially only carbon dioxide in an amount up to 98% and water vapor and a trace amount of nitrogen and carbon monoxide. Separation of main exhaust gases components is then easy to reach in the normal condenser during cooling process where also carbon dioxide-rich gas stream is easy to capture [1]. By this method can be captured almost 100% of carbon dioxide from the flue (also termed exhaust) gas stream, which may be then liquefied and transported to the storage area. It should be noted that the lack of presence of nitrogen in the combustion process results in less heat rejected in chimney, much lower NO_x emissions, the rise of adiabatic combustion temperature and reduced exhaust gas stream. Power plants operating by this technology are often referred to as a zero-emission power plants or carbon-free power plants.

3. RADIATION HEAT TRANSFER FROM COMBUSTION GASES

The basic element of a large steam boiler is steam-water heat system. The combustion chamber of the boiler is surrounded by the walls of the vertical tubes, in which water circulates. They are so called membrane walls forming a water jacket and receiving heat from exhaust gases by radiation. Hence the term "irradiated boiler". The process of producing steam in a boiler is divided into three phases: preheating of water, evaporation, and superheating. The heat from the flue gas is received in the corresponding heat exchangers: economizer, evaporator and superheater.

Radiation is energy carried by photons and has its principles in electromagnetic theory. Radiative emission is releasing fotons while absorption is capturing photons. Most often radiation can be considered with vibrational and rotational transitions taking place in atoms or molecules (these transitions are not present in monatomic species) in higher temperatures dissociation, ionization and electron transition may play significant role [3,7]. Radiation is specified by radiation intensity.

Radiation intensity from surface area A in direction s (3D direction vector which can also be unambiguously identified by two angles in spherical coordinates) is the amount of energy per unit time per unit solid angle centred around direction s per unit projected surface area (projection of A in plane normal to s) per unit small wavelength interval around the wavelength λ [3]. From this definition it is clear that radiation intensity is spectral property but it can also be regarded as total property if integrated over all spectrum i.e. along $(0,\infty)$ interval. Incoming radiation may be reflected, absorbed or transmitted through the medium. Physical bodies can also be source of radiation (emitters).

While numerically calculating radiation in control volume for its each computational cell one or set of Radiative Transfer Equations (RTEs) must be solved. RTE describes changes of incoming radiation intensity in a medium as composition of three components – attenuation of radiation intensity in a medium due to absorption, emission of a medium and scattering in a medium. Gases are considered as non-scattering (scattering coefficient equals 0) and non-reflecting and for them the RTE takes form [4, 7–9]:

$$\frac{\mathrm{d}I(r,s)}{\mathrm{d}s} = -aI(r,s) + a\mathrm{I}_b \tag{1}$$

where $I(r,s)$ is the radiation intensity, r is the position vector, s is the direction vector, a is absorption coefficient, I_b is blackbody radiation and s is radiation beam length.

To get the idea of absorption coefficient one has to refer to Bouguer-Lambert's law stating that radiation beam travelling through gas medium is attenuated exponentially [4, 7]:

$$\mathrm{dI}_\lambda = -a_\lambda \mathrm{I}_\lambda \mathrm{d}s \tag{2}$$

that after integration gives exponential relation.

Provided the gas is of uniform structure (i.e. well-mixed) we obtain:

$$\mathrm{I}_\lambda(s) = \mathrm{I}_\lambda(0)\, e^{-a_\lambda s} \tag{3}$$

Hence change in radiation intensity equals

$$\mathrm{I}_\lambda(s) - \mathrm{I}_\lambda(0) = \mathrm{I}_\lambda(0)\,(1 - e^{-a_\lambda s}) \tag{4}$$

$1 - e^{-as}$ represents absorptivity and is denoted by α.

Absorption coefficient a give insight how a gas medium is attenuating for radiation passing through it.

Gas radiating properties are closely related with wavelength. Calculations of radiative transfer by line-by-line spectral data integration is most accurate but slows down computations to unacceptable level. To avoid this obstacle WSGG (Weighted Sum of Gray Gases) model (denoted as WSGGM) was created and is very popular in use due to fast computation it provides. Modest showed that WSGGM could be used with any technique of examining solution of RTE [4]. WSGGM approach is to replace the integration of spectral properties with summation over number N_g of gray gases (gray gas is a medium in which absorption coefficient is less than 1 and constant over whole spectrum) to simulate the properties of non-gray gas. Radiative properties of medium in WSGG model are described by means of emissivity ε (ratio that says how close the medium is to ideal emitter – blackbody). WSGG model expresses total emissivity as [4–9]:

$$\varepsilon = \sum_{i=0}^{N_g} w_{\varepsilon,i}(T)\,(1 - e^{-a_i s}), \tag{5}$$

where a_i is the absorption coefficient of i-th imaginary gas, s is mean radiation beam length, $w_{\varepsilon,i}(T)$ are weighting factors, T denotes temperature, N_g represents number of imaginary gases. $\tau_i = e^{-a_i s}$ will be transmissivity of i-th imaginary gas and $\alpha_i = \varepsilon_i = 1 - \tau_i$ is i-th gas absorptivity and emissivity. For models WSGG Smith and WSGG Yin described later in this section $N_g=4$ and $N_g=5$ correspondingly. N_g has nothing in common with number of radiating gases constituting radiating gas mixture; a_i, N_g and $w_{\varepsilon,i}(T)$ in (5) are simply data fitting parameters and are established by correlation with emissivity data generated by narrow band (spectral lines) models. In (5) a_0

equals 0, hence ε_0=1 and since there are no reflections taking place in gas medium 0-th imaginary gas in (5) does not absorb nor emits radiation but it is transparent to all incoming radiation (such gas is called clear). For i=1 to N_g a_i are positive and i-th gas is considered as gray. Total absorption coefficient may be calculated from (5) by applying following formula:

$$a = -\tfrac{1}{s}\ln(1-\varepsilon) \tag{6}$$

Authors used in computer simulations Ansys Fluent Software (version 13.0). This platform has implemented WSGG model for radiation properties calculations but with coefficients in (5) correlated with emissivity data generated by EWBM narrow band model for air combustion conditions (given by Smith and hence will be referred in this paper as WSGGM Smith) [5]. WSGGM Smith has one clear gas and three gray gases. For oxyfuel combustion conditions new model (4 gray gases and one clear gas) and with new coefficients was applied in Ansys Fluent as a C++ code in a form of UDF (User Defined Function). This new model was taken from Yin [6] and its coefficients in (5) were correlated with emissivity data generated by EWBM for oxyfuel gases. Further this model will be called WSGG Yin.

WSGGM Smith is correlated with emissivity data for air conditions and its coefficients were established assuming that mole fraction of CO_2 in gas mixture approaches zero or ratio of mole fraction of CO_2 to mole fraction of H_2O equals 1 or 2 with mole fraction of CO_2 equaling 0,1 for both ratios. Hence WSGGM Smith does not cover most of oxyfuel atmospheres in which mole fraction of CO_2 reaches high values (see Fig. 3, Fig. 4 and section 5).

WSGGM Yin is correlated taking into account 10 different oxyfuel gas radiating cases that cover all oxyfuel conditions but does not consider air cases.

General WSGGM approach is to treat combustion gas mixture as several gray gases and for each of the gray gases one set of RTE is solved. WSGGM Smith and WSGG Yin consider gas mixture as one gray medium for which one set of RTE is solved:

$$\frac{\mathrm{d}I}{I} = -a \cdot \mathrm{d}s \tag{7}$$

4. COMPUTER SIMULATIONS

Important from the standpoint of oxyfuel combustion technology is the phenomenon of radiation due to the presence of large amounts of CO_2 and possibly also H_2O, triatomic species playing a key role in the radiation. Temperature profiles in combustion chamber and radiation heat transfer determines the strategy for optimal heat transfer in heat exchangers. Authors introduced to Ansys Fluent software (version 13.0) 3D geometry of real laboratory reactor (schematically presented on Fig. 1) and conducted

2D simulations (flow across symmetry plane) of the Turów K8 lignite combustion process in air and oxyfuel gas atmospheres (OXY 20 i.e. O_2/CO_2 = 20% / 80% vol. and OXY 30 i.e. O_2/CO_2 = 30% / 70% vol.). In general reactor comprises vertically positioned 2,5 m long ceramic tube of internal diameter 135 mm and 20 mm thick that is surrounded by stack of six heating sections (surface of this reactor zone is termed “Heater” on Fig. 2). Each heating section resembles ring containing from inside spiral electric heater and from outside thermal insulation followed by metallic shielding plate. Both pulverized fuel and inlet gases are supplied to the burner (surfaces named “Inlet 1”, “Inlet 2” and “Burner” on Fig. 2) located at the top of the reactor. Fuel gravitationally falls hence the name drop tube reactor. Both primary gas stream (travelling with coal across surface “Inlet 1” on Fig. 2) and secondary gas stream (moving across surface “Inlet 2” on Fig. 2) may be preheated in electric heaters external to reactor. Separate heating sections stacked along reactor’s length provide possibility of maintaining reactor’s temperature uniform at targeted value up to 1300° C. At the bottom of ceramic tube metallic funnel is situated and this section (which surface is termed “Wall” on Fig. 2) is equipped with optional heating and optional water cooling. Boundary conditions for computations reflected those observed during laboratory tests of the combustion processes in air and oxyfuel atmospheres.

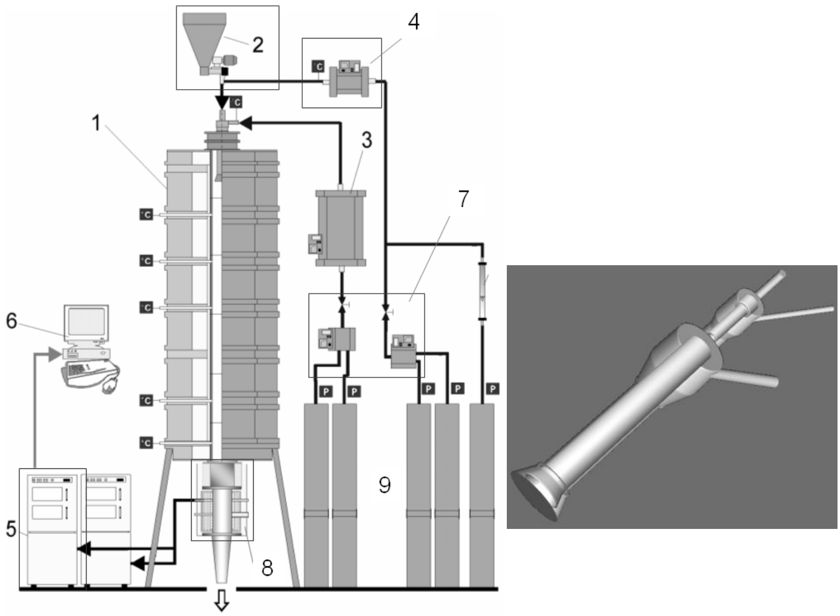


Fig. 1. Schematic of laboratory drop tube reactor: (1 – reactor , 2 – feeder, 3 – secondary gas stream heater, 4 – primary gas stream heater, 5 – Ultramat 23 analyzer and Oxymat analyzer with broadened measurement range of O_2 and CO_2, 6 – controlling unit, data archiving, 7 – O_2/CO_2 gas mixers, 8 – additional heating section, 9 – gas batteries), and geometry of pulverized fuel burner

The calculations assumed the reactor's temperature equalled 1200° C (i.e. the temperature of "Heater" was implemented as 1473 K) and the excess oxygen (designated as λ and indicating how much oxidizer above stoichiometric value was supplied) was taken equal 1,1. Mesh was constructed from rectangular elements, wall function was used (k-ε realizable turbulence model was implemented). For computations no fuel and primary gas preheating was regarded (i.e. temperatures at "Inlet 1" and "Burner" were taken equal 300 K), for secondary gas stream preheating to (i.e. temperature at "Inlet 2") 500K was assumed. *Internal emissivity* property of "Heater" was set equal 0,9. "Wall" surface temperature was taken as 300 K.

"Outlet" is the surface through which flue gases leave reactor and was chosen as Ansys Fluent *pressure outlet*. Discussed in this paper simulations consider oxyfuel combustion variant without flue gas recirculation back to the combustion chamber.

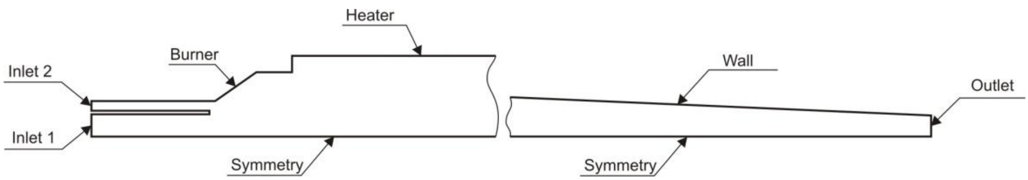


Fig. 2. Two dimensional axis-symmetrical (along Symmetry plane) part of reactor. Labels are names of surfaces

5. RESULTS AND DISCUSSION

Results of obtained simulations are presented in this section. Fig. 3 and Fig. 4 present CO_2 mole fraction along whole reactor's length on all surfaces indicated on Fig. 2 for air and OXY 30 combustion atmospheres correspondingly. Fig. 3 and Fig. 4 both refer to Turów K8 lignite fuel with moisture content 7% (M=7%). Mole fraction of CO_2 is an important parameter because it influences optical thickness and radiation properties of gas mixture that CO_2 constitutes (absorption coefficient and emissivity are among this radiation parameters, see also section 3). It is clearly visible that for OXY 30 lower than in air case CO_2 mole fraction variations are observed and CO_2 mole fraction of exhaust gases in OXY30 case stabilizes after the gases will reach approx. 56% of reactor's length while in air this stabilization takes place after the gases will travel approx. 84% of reactor's length. CO_2 mole fraction in air does not exceed 0,15 while for OXY 30 is generally higher than 0,7.

Absorption coefficients in drop tube reactor along symmetry plane and along heater and wall surfaces (see Fig. 2) for Turów lignite with two moisture levels (7 and 15%) in two combustion atmospheres (air and oxyfuel composed of O_2/CO_2 = 30% / 70% vol.) are shown on Fig. 5 and Fig. 6.

As it was previously discussed in section 3 absorption coefficient may give quantitative information about radiation property of a medium because it is the inverse of mean penetration distance of radiation.

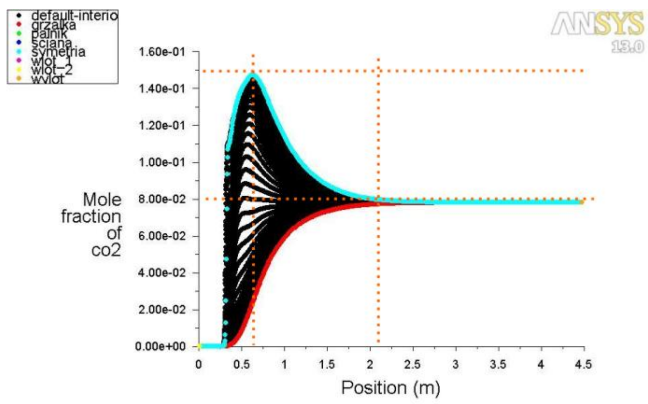


Fig. 3. CO_2 mole fraction along whole reactor's length (all surfaces from Fig. 2 included) for Turów lignite with moisture content equal 7% during combustion in air

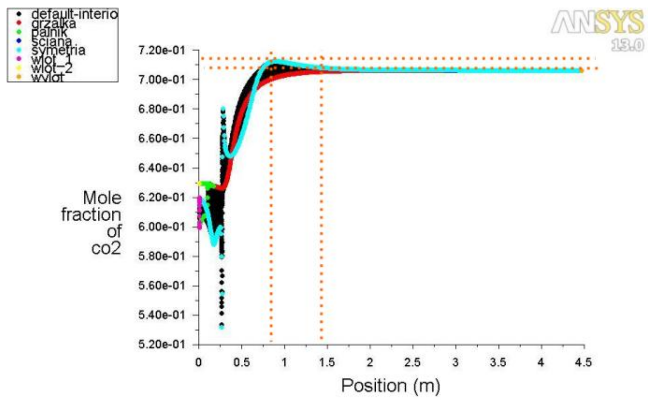


Fig. 4. CO_2 mole fraction along whole reactor's length (all surfaces from Fig. 2 included) for Turów lignite with moisture content equal 7% during combustion in OXY 30 atmosphere

Changing initial moisture level of coal (parameter *M*) from 7 to 15% does not significantly influence absorption coefficient among the same model. Crucial impact on absorption coefficient has combustion atmosphere and applied radiation model. When not suitable for oxyfuel combustion (see section 3) and by default present in Ansys Fluent software WSGG Smith model is applied to OXY30 combustion atmosphere it may lead to very high (above 0,8 1/m) absorption coefficient values which may be misleading for heat transfer calculations. Proper for oxyfuel combustion WSGGM Yin leads to about 8 times lower values (0,1 1/m level).

On the other hand not suitable for air environment WSGG Yin model (see section 3) when applied to air combustion gives lower values of absorption coefficient than dedicated to air conditions WSGG Smith model and shows less magnitude variation along reactor's length. Values from WSGGM Smith in air are maximally four times larger than corresponding values from WSGGM Yin.

When comparing radiation properties of OXY 30 with applied WSGGM Yin and air with applied WSGGM Smith it can be noted that in OXY 30 absorption coefficient is lower than in air. Inverse behaviour is with radiation attenuation which is lower in OXY 30 than in air. Variation of absorption coefficient in air along reactor's length may be attributed to species concentration variations while in OXY 30 high and stable concentration of CO_2 makes radiating gas mixture uniform along the length.

Following emissivity and absorption coefficient calculations radiative heat flux for example between radiative gas mixture and wall may be evaluated using formulas as $q_{rad}=\sigma\cdot\varepsilon_g\cdot T_g^4-\sigma\cdot\alpha_g\cdot T_w^4$, where σ is Stefan-Boltzmann constant and indexes g and w refer to gas mixture and wall respectively.

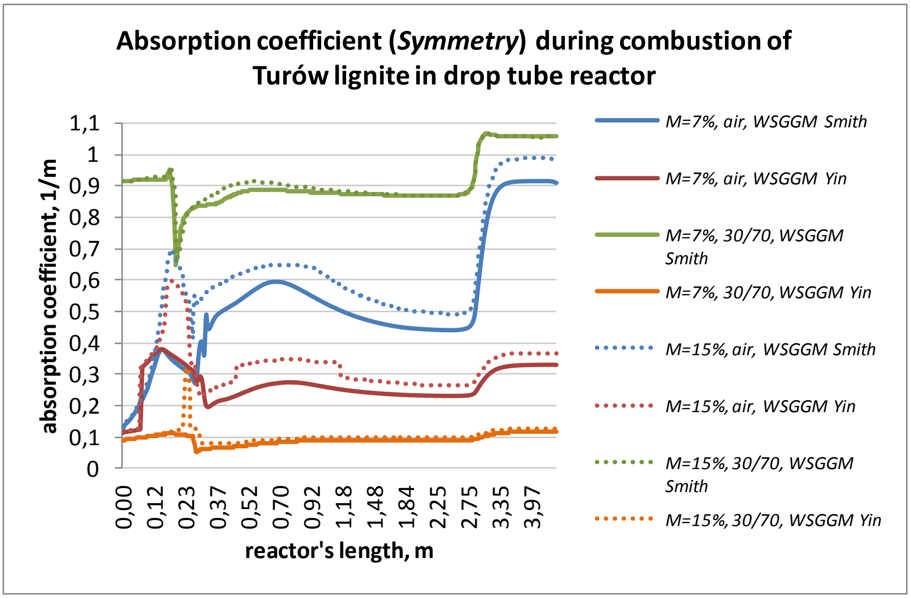


Fig. 5 Absorption coefficient along *Symmetry* plane (see Fig. 2) obtained as a result of simulations. Combustion scenarios of Turów lignite take two variants of moisture content in coal (parameter *M*), two combustion atmospheres (*air* and *OXY 30*) and two radiation models (*WSGGM Smith* and *WSGGM Yin*)

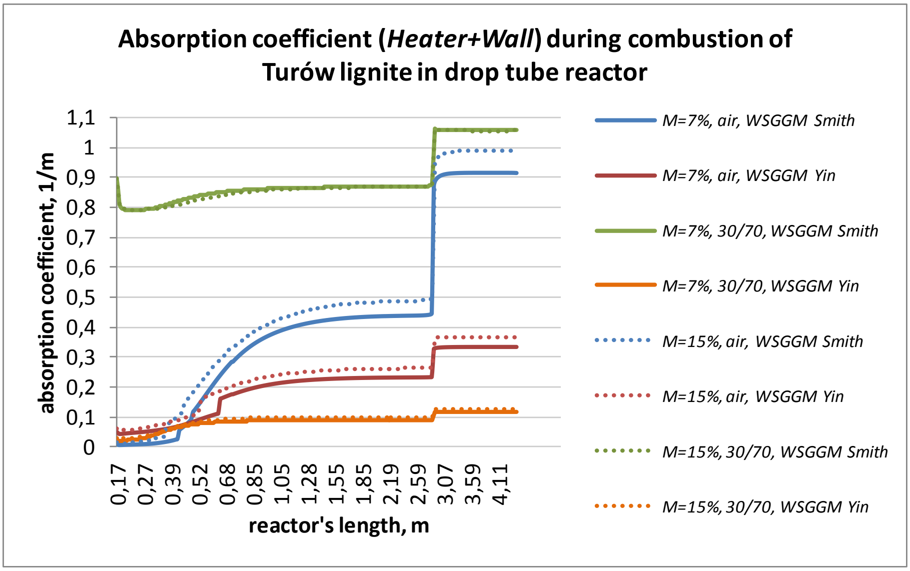


Fig. 6 Absorption coefficient along *Heater* and *Wall* surfaces (see Fig. 2) obtained as a result of simulations. Combustion scenarios of Turów lignite take two variants of moisture content in coal (parameter *M*), two combustion atmospheres (*air* and *OXY 30*) and two radiation models (*WSGGM Smith* and *WSGGM Yin*).

REFERENCES

[1] JABŁOŃSKI J**.,** *Technologie „zero emisji”,* Poznań, Wydawnictwo Politechniki Poznańskiej, 2011, 43–81.

[2] NOWAK W., PRONOBIS M., *Nowe technologie spalania i oczyszczania spalin*, Wydawnictwo Politechniki Śląskiej, Gliwice 2010.

[3] *Wyzwania paliwowe, technologiczne i ekologiczne dla polskiej energetyki*, praca zbiorowa pod red. T. Chmielniaka, Wydawnictwo Politechniki Śląskiej, Gliwice 2010.

[4] SIEGEL R., HOWELL J., *Thermal Radiation Heat Transfer.* 4th ed., Taylor & Francis, New York, London 2002.

[5] SMITH T.F., SHEN Z. F., FRIEDMAN J. N., *Evaluation of Coefficients for the Weighted Sum of Gray Gases Model.*, J. Heat Transfer, Vol. 104, 1982, 602–608.

[6] YIN CH., ROSENDAHL L., KAER S.**.** *Chemistry and radiation in oxy-fuel combustion: A computational fluid dynamics modeling study*, Fuel,Vol. 90, 2011, 2519–2529

[7] KOSTOWSKI E., *Promieniowanie cieplne*, Wydawnictwo Politechniki Śląskiej, Gliwice 2009.

[8] KANGWANPONGPAN T., CORREA DA SILVA R., KRAUTZ H. J., *Prediction of oxy-coal combustion through an optimized weighted sum of gray gases model*, Energy, Vol. 41, 2012, 244–251.

[9] JOHANSSON R., ANDERSSON K., LECKNER BO, THUNMAN H., *Models for gaseous radiative heat transfer applied to oxy-fuel conditions in boilers,* International Journal of Heat and Mass Transfer, Vol. 53, 2010, 220–230.

*flying robots, quadrocopter, control system,
neural network, PID controler*

Michał LOWER*, Wojciech TARNAWSKI*

STABILISATION AND STEERING OF QUADROCOPTER USING NEURAL NETWORK

Quadrocopter is a dynamic system with many variables. PID standard controller is often used to stabilisation of flight of this object. The main disadvantages of this control system is the need of tuning of other parameters arising from the length of arm, different types of propellers with motors and object's weight. The main problem for constructors is the tuning of controller parameters. In this experiment quadrocopter was controlled by neural network. The control system was divided into four subsystems. Each of them is responsible of setting the control values. The neural network was learnt by control system with standard PID controller. This approach is used for checking how neural networks cope with stabilisation of the quadrocopter. The proposed controller was tested in different structure of neural network and different states of flight: in hover, in to forward flight with constant speed, in climbing and in rotation. In all these situations the proposed controller was able to provide foreseeable behavior of the quadrocopter. Simulation results of the neural controller and PID controller working were compared to each other. The mathematical model of quadrocopter and its neural controller were simulated using Matlab Simulink software. The results of this experiment are shown in the paper.

1. INTRODUCTION

UAV-s (Unmanned Aerial Vehicle) have been rapidly developed for a few years by scientists and a commercial company. They are often military construction, which are used for simple recognition tasks. One of the simpliest flight construction in terms of the mechanical structure is a multi-rotor. These constructions have 3-tricopter, 4-quadrocopter, 6-hexacopter etc. motors with special propellers, according to their

* Institute of Computer Engineering, Control and Robotics,Wrocław University of Technology, Poland, e-mail: michal.lower@pwr.wroc.pl, w.tarnawski@pwr.wroc.pl

configuration. Most of them utilize the classical control theory, so they are controlled by proportional integral derivative-PID controller.

PID controller theory and its application in various systems is well described in vast literature. The procedure of tuning should be done before using PID to control the object. Changing the point of work (e.g. weight increase, engines change, etc), the calibration procedure should be done again because the former controller settings do not guarantee the proper work of control system. This is one of the disadvantages of the PID controller thus one is looking for control methods resistant to such changes.

(PID) feedback controller [1, 2, 3, 4]. Scientists try to use more advantage artificial intelligence algorithms for controlling, e.g.fuzzy logic [5, 6, 7, 8, 9] or neural networks controllers [10, 11, 12]. The main advantage of fuzzy logic, in compare to classical method, is the nonlinear character of controller and the ability to develop controller using simple transformation of rules expressed in natural language. The good example of this methodology was demonstrated in our prior publications [5, 6, 7]. In this article controller, which is using to steering and stabilization a quadrocopter is based on neural network. It allows to analyse how this type of algorithm can be cooperative with flying object as a quadrocopetr. In the first stage of the study, which is described in the article, the neural network controller will be tought on the base of the PID control system. In the second stage of the study the learning process is planned to be extended to fuzzy control system.

2. MATHEMATICAL MODEL OF QUADRCOPTER

Quadrcopter is a flying object with four motors with propellers in cross configuration (Fig. 1). Divided into two pairs of propellers in the opposite direction removes the need for a tail rotor. Usually all engines and propellers are identical, so the quadrocopter is the full symmetrical flying object.

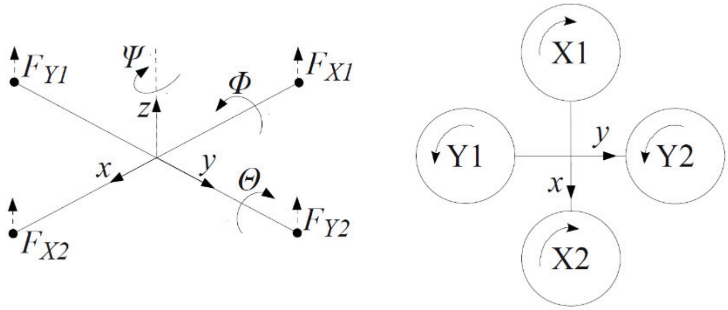


Fig. 1. The coordinate system of quadrocopter

The main effects, acting on a quadrocopter, witch have to be taken into account are:

- gravity efect caused by quadrocopter mass,
- aerodynamic efects of each of the propellers caused by rotating propellers,
- inertial torques of each of the propellers,
- gyroscopic efects of each of the propellers,
- joint aerodynamic efects in all three axes,
- causing linear movement,
- joint inertial torque causing pitch,
- roll and yaw angles changes.

The detailed mathematical model we based on, can be found in [2]. Although, for clarity the main formulas are presented below.

The torque moment around O_y axis:

$$T_x = bl\left(X_1^2 - X_2^2\right) \tag{1}$$

where b is a thrust coefficient, l is the distance between the propeller's axis and the center of mass of the quadrocopter and $X1$, $X2$ are rotation speeds of propellers according to the Fig. 1. As the consequence the angle Θ called pitch can be observed.

The torque moment around O_x axis:

$$T_y = bl\left(Y_1^2 - Y_2^2\right) \tag{2}$$

where $Y1$, $Y2$ are rotation speeds of propellers. As the consequence the angle Φ called roll can be observed.
The join torque around mass center of quadrocopter:

$$T_z = d\left(X_1^2 + X_2^2 - Y_1^2 - Y_2^2\right) \tag{3}$$

where d is so called drag coefficient. As the consequence the angle Ψ called yaw can be observed. The above formulas look quite simple and so they are, thus the quadrocopter attitude can be controlled only via propellers rotation speed changes. But these changes also have an influence on joint thrust force:

$$F_z = b\left(X_1^2 + X_2^2 - Y_1^2 - Y_2^2\right) \tag{4}$$

Combining equations (1) to (4) and the main effects acting on quadrocopter the following system of equations can be derived Eq.(5). It is the simplified model of quadrocopter behavior according to [2, 4, 6, 7] which ignores aerodynamic drags and gyroscopic effects caused by propellers rotation, but this model is good enough to model quadrocopter's behavior in hover state and at low quadrocopter speeds.

$$\begin{cases} \ddot{x} = (\cos\psi \sin\Theta \cos\phi + \sin\psi \sin\phi) F_z / m \\ \ddot{y} = (\sin\psi \sin\Theta \cos\phi - + \cos\psi \sin\phi) F_z / m \\ \ddot{z} = (\cos\Theta \cos\phi) F_z / m - g \\ \ddot{\phi} = [\dot{\Theta}\dot{\psi}(I_y - I_z) + lT_y] / I_x \\ \ddot{\Theta} = [\dot{\phi}\dot{\psi}(I_z - I_x) + lT_x] / I_y \\ \ddot{\psi} = [\dot{\phi}\dot{\psi}(I_x - I_y) + lT_z] / I_z \end{cases} \tag{5}$$

where m is the mass of quadrocopter, g is the gravity acceleration, l is the distance between the rotor and center of quadrocopter, I_x, I_y, I_z are the inertia moment along proper axes, Φ, Θ, Ψ are roll, pitch and yaw angles, respectively.

3. CONTROL SYSTEM

In this article the authors focused on control velocity of moving quadrocopter in two (X,Y). Velocity and rotation in Z axis depend on T_z i F_z, changing this variable has a direct influence on velocity. This dependence is a very simply, so this aspect has been exluded from the article. Otherwise control in X and Y axis can not be steered by simply force changing. Changing RPM for motors, which depend on moving in definite direction, triggers temporary changing of velocity and continuously changing of angular velocity EQ. (5). It has an influence on continuous changing of angle of a quadrocopter, so controller has to check this angle and correct it, otherwise quadrocopter will be rotated around its axis and operator will not control velocity. In this article controller has divided into two block (this is a cascade controller (Fig. 2)). A first block (P) is classic PID with only a proportional term and it responsible for calculating error from line velocity to value (XX'), which is transmissioned to second block (PD/Neural).

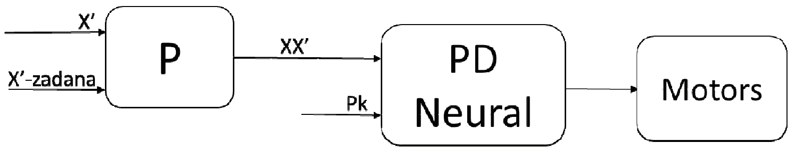


Fig. 2. Dividing the control system

A second block controls a RPM of motors based on its input and internal algorithm, what has directly influence of velocity and rotation quadrocopter. In first test a second block was a classical PID with a proportional and derivative term, schema of this block show on (Fig. 3).

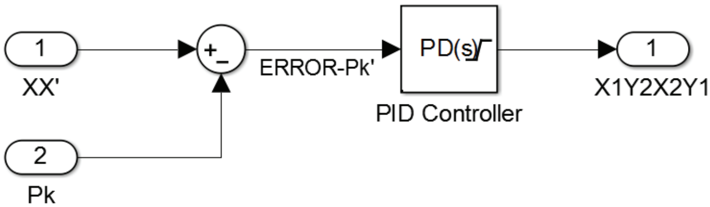


Fig. 3. Basic structure second controller PD

Where:

- XX' – output from first block of control system
- Pk – angle Θ

In second test, second block PD replaced block with neural network (Fig. 4).

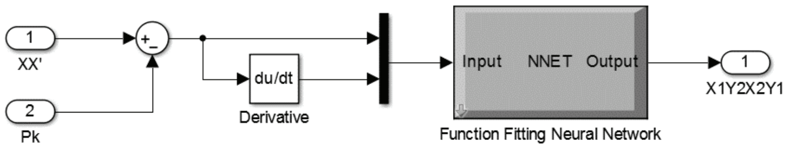


Fig. 4. Basic structure second controller NEURAL

The neural network used as the controller was feed-forward network with two input, sigmoid hidden neurons and one linear output neuron. Number of hidden neurons was selected experimentally and neural network was trained in off-line with Levenberg-Marquardt backpropagation algorithm. Data for training was generated from simulation of quadrocopter when this object was controlled by PID. It is method called "general training" [12].

Input for neural block:

$$\begin{cases} Input_1 = XX' - Pk \\ Input_2 = Input_1 \end{cases} \tag{6}$$

4. SIMULATION

The mathematical model of the quadcopter is implemented for simulation in Matlab with Simulink. In the first step controller P and PD was tuned. Quadrocopter has many dynamic dependent variable, so authors extract potential Matlab's simulation and tuned controller by experimentally. The minimization of integrated square error of one of velocity errors was an objective function. Research was made for all axis (X, Y, Z, R) but in this article the authors show results of simulation with control in two axes (X, Y). Authors chose only those axis, because controlling them makes a lot of difficulties

$$V = [2\ 3\ 0] \tag{7}$$

Input signal was vector with velocity – Eq.(5), which was transformated by transfer function:

$$ln = \frac{1}{xs+1} \tag{8}$$

where:

- x is coefficient transmittance by s – influence on delay step

In simulation used (according to [2]):

- propeller distance – $l = 0.23$ [m],
- quadrocoptertor mass – $m = 0.65$ [kg],
- drag coeff. – $d = 7.5e^{-7}$ [Nms^2],
- thrust coeff. – $b = 3.13e^{-5}$ [Ns^2],
- inertia moment – $I_x = I_y = 7.5e^{-3}$ [kgm^2], $I_z = 1.3e^{-3}$ [kgm^2].

In the first test a coefficient x was beeing changed and the quality of control of PD controller and neural controller was observed.

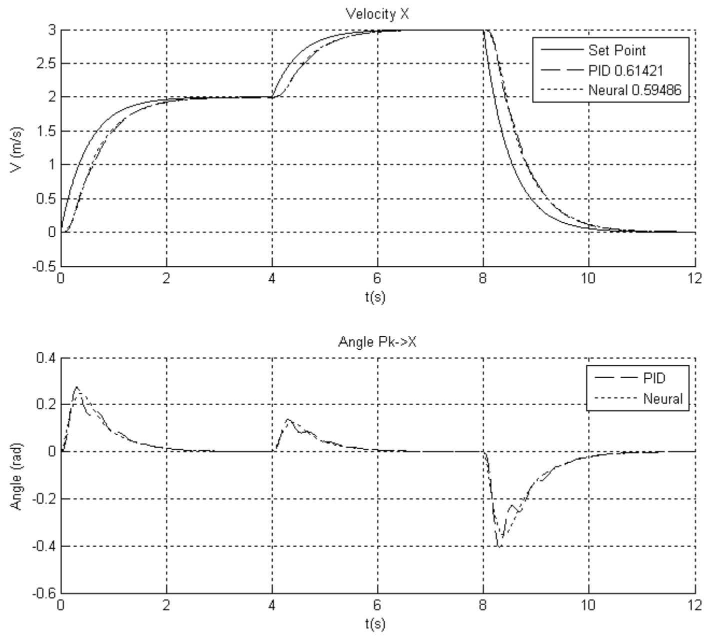


Fig. 5. Input is transfer function: 1/(0.5s + 1)

In the Fig. 5 the simulations results have been shown for the first input signal. The upper graph presents behavior of quadrocopter, when velocity is changing. Object controlled by neural controller has smaller error than object controlled by PID. The lower graph presents Pk-angle changing, the course on the graph for neural controller is more stable, what minimizes stepping in output signal. This is a very positive property for all physical object because oscillations in the drive systems have a negative impact on their lifetime and additionally they can strongly destabilize the flight of quadrocopter.

Fig. 6 shows the simulation results where the coefficient x was reduced. The deterioration of control can be observed for both controllers. Error for neural controller is larger than error for PID and it can be observed in 15–18 second. In spite of this, there are Pk parameter oscillations in PID controller. On the other hand, steps for changing of the angle Pk for PID controller are larger than in previous experiment. Similarly to previous tests, such oscillations can strongly destabilize the flight of real quadrocopter. Finally the authors hold the view that the result of neural control is better for real quadrocopter.

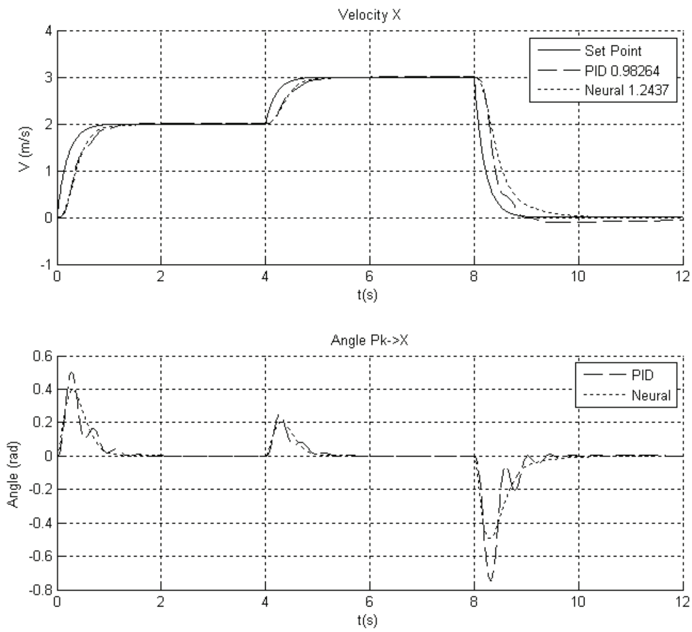


Fig. 6. Input is transfer function: 1/(0.2s + 1)

The next experiment, where the coefficient P in P block has been changed is shown in Fig. and Fig. . Both upper graphs show that neural controller and PID controller have a very similar error. However, similarly to previous tests one can observe that there are oscillations in PID controller which can be the reason of destabilization of the real quadrocopter. The results of neural control are much better for real quadrocopter then results of PID controller. Fig. 8 presents the results of tests for the parameter P increasing to 1. This test definitely shows that neural control is better. The PID controller can not stabilize the flight parameters at all.

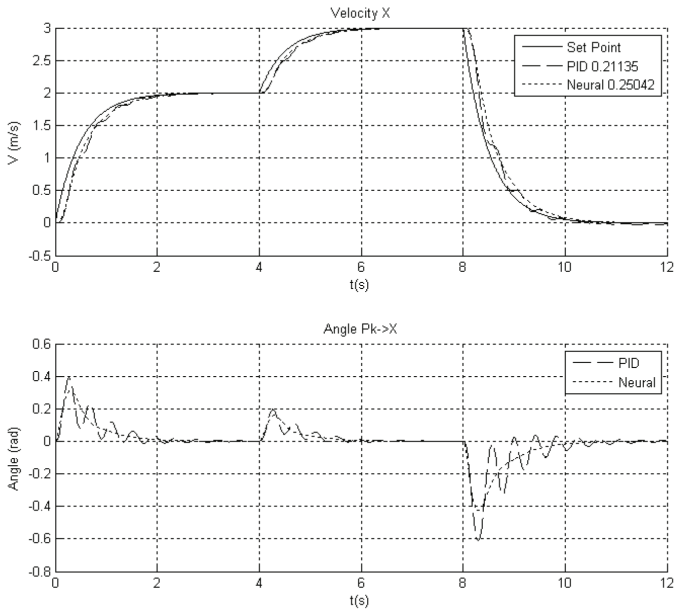


Fig. 7. Output control system when P= 0.7

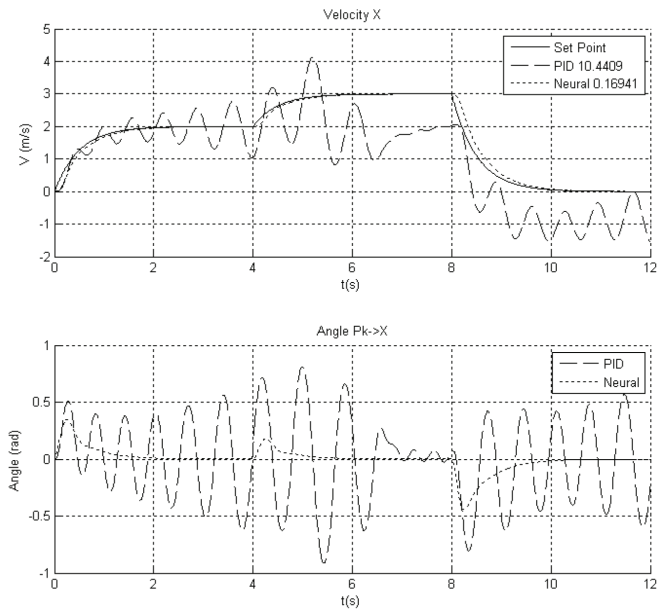


Fig. 8. Output control system when P= 1

5. CONCLUSION

The simulation tests presented in the article have shown that neural controller can be used for stabilization of flight of quadrocopter. It can be expected that the neural controller gives better results of flight control than PID controller. The presented examples show that stabilization of PID controller often reaches the set point with an oscillation. Such oscillation is not observed in the neural controller. That property is very important in the control of real quadrocopter. The simulation results of neural control of quadrocopter are so promising that it is worth to continue this research. Tests of another neural network configurations are considered to be the next step of research. It is planned to expand the space of learning data. The authors want to teach the neural network on the base of PID controller with different parameters which are changing according to the work point. Additionally the authors are going to use fuzzy controller to teach the neural network. In the last stage of research it is necessary to make tests with the real control system of quadrocopter.

REFERENCES

[1] ARAMA, B., BARISSI, S., HOUSHANGI, N.: *Control of an unmanned coaxial heli-copter using hybrid fuzzy-PID controllers*. In: Electrical and Computer Engineering (CCECE), 2011 24th Canadian Conference on. pp. 001064 {001068 (2011).

[2] BOUABDALLAH, S.: *Design and Control of Quadrotors with Application to Au-tonomous Flying,* In: Master's thesis, Swiss Federal Institute of Technology (2007).

[3] LI J., LI Y., *Dynamic analysis and PID control for a quadrotor,* In: Mechatronics and Automation (ICMA), 2011 International Conference on. pp. 573–578 (2011).

[4] HOFFMANN, G.M., HUANG, H., WASL, S.L., CLAIRE, E.: *Quadrotor helicopter flight dynamics and control: Theory and experiment*. In: In Proc. of the AIAA Guidance, Navigation, and Control Conference (2007).

[5] LOWER M., KROL D., SZLACHETKO B., *Building the fuzzy control system based on the pilot knowledge* In: Khosla, R., Howlett, R., Jain, L. (eds.) Knowledge-Based Intelligent Information and Engineering Systems, Lecture Notes in Computer Science, Vol. 3683, pp. 179–179. Springer Berlin / Heidelberg (2005).

[6] SZLACHETKO B., LOWER, M., *Stabilisation and steering of quadrocopters using fuzzy logic regulators*, In: Rutkowski L., Korytkowski M., Scherer R., Tadeusiewicz R., Zadeh L., Zurada J. (eds.) Artfcial Intelligence and Soft Computing, Lecture Notes in Computer Science, Vol. 7267, pp. 691{698. Springer Berlin / Heidelberg (2012).

[7] SZLACHETKO B., LOWER, M., *On Quadrotor Navigation Using Fuzzy Logic Regulators,* In: Computational Collective Intelligence. Technologies and Applications Lecture Notes in Computer Science Volume 7653, pp. 210–219, (2012).

[8] RAZA, S.A., GUEAIEB, W.: *Fuzzy logic based quadrotor fight controller*. In: ICINCO-ICSO'09, pp. 105{112 (2009).

[9] SANTOS, M., LOPEZ, V., MORATA, F.: *Intelligent fuzzy controller of a quadrotor*. In: Intelligent Systems and Knowledge Engineering (ISKE), 2010 International Conference on. pp. 141 {146 (2010).

[10] Yohannes ., José de Gea F., Schwendner J., Kirchner F., On Applying *Neuroevolutionary Methods to Complex Robotic,* In: Tasks New Horizons in Evolutionary Robotics: Extended Contributions from the 2009 EvoDeRob Workshop, Pages 85–108, o.A., 2011.

[11] A. Aswani, P. Bouffard, and C. Tomlin (2012), *Extensions of learning-based model predictive control for real-time application to a quadrotor helicopter*, In: Proceedings of the American Control Conference: 4661–4666.

[12] M. Norgaard, O. Ravn, N. K. Poulsen, L. K. Hansen: *Neural Networks for Modelling and Control of Dynamic Systems*, Springer-Verlag, London, 2000.

fuzzy decision tables, fuzzy flow graphs,
aircraft control

Alicja MIESZKOWICZ-ROLKA, Leszek ROLKA*

FUZZY FLOW GRAPHS IN ANALYSIS OF THE PILOT'S CONTROL ACTIONS

This paper discusses the problem of analysis of the pilot's control actions. The presented approach is based on the fuzzy flow graphs idea. The concept of interactive information systems is utilized to represent the aircraft-pilot influences. The issue of selecting the appropriate attributes is considered. The final decision table with fuzzy attributes is obtained, which is analyzed with the help of fuzzy flow graphs. In consequence, fuzzy decision rules are generated which can be viewed as a part of the pilot's control actions model.

1. INTRODUCTION

One of the main symptoms of the human intelligence is the ability of processing, understanding and storing information about the environment. Various soft-computing methods were proposed to describe the human operator's control actions; in the case of modeling the control behavior of a human pilot, experts systems, neural networks, fuzzy sets, and Petri nets were applied [1]. Human operator, who observes and controls a plant (a part of an environment), performs activity which influences the changes of the state of the controlled object that are perceived by the operator. This process is continued until a desired state of the controlled plant is reached. For modeling these interactions, we can utilize interactive information systems (interactive decision tables) [5], which are suitable to represent agent-environment connections that appear in the process of the agent's perception and action. The rows of a decision table represent elements of a universe, the columns correspond to attributes which characterize the elements of the universe. We consider information systems which describe the process of aircraft control. The pilot perceives the state of the controlled object, inferences and performs activity by using attributes expressed by linguistic values rather than num-

* Rzeszów University of Technology, al. Powstańców Warszawy 8, 35-959 Rzeszów.

bers. Therefore, for description of these processes, we propose to use attributes with fuzzy values. One can generate fuzzy rules from data with the help of different methods, see e.g., [6]. In this paper, we recommend to apply fuzzy flow graphs to this end, our original approach introduced in [2]. It will be used for analysis of data obtained from a pilot's control process. A computer implementation of fuzzy flow graphs is relatively easy. Furthermore, in comparison to other methods, the user is given more insight into the process of rules generation.

2. ATTRIBUTES IN MODELING PILOT-AIRCRAFT INTERACTION

Let us examine the interactions in the process of aircraft control, performed by a human operator. We distinguish two elements:

1. pilot, interacting with the aircraft, who can be viewed as an agent, possessing knowledge, experience and ability to adaptation,
2. aircraft, as the nearest environment (part of the environment); the states of the aircraft have an impact on the pilot and his or her actions.

In order to model the interactions between the operator and the object, we propose to consider information systems on different levels of complexity and abstraction.

In the first stage, the description of the object (aircraft) is based on sensory attributes only. They represent pure information about the state of the aircraft in the form of selected physical quantities. The values of attributes in this stage are real numbers.

In the next stage, the information system represents the point of view of the pilot. In every task of plant control, the operator follows a specified objective. Usually, this is the stabilization of selected control variables on a fixed level. Therefore, we ought to add new attributes, which can express the dynamic character of an object, being a source of prognostic information for the operator, in the form of deviations and changes of variables.

It should be noted that human operators, who control dynamic processes, uses rather linguistic terms than real numbers. During the fuzzyfication process, a crisp value of every attribute is transformed into a degree of membership to a fuzzy set representing linguistic term connected with the attribute.

Next, the obtained information system with fuzzy attributes becomes a basis for creating an interactive decision table DT_F. We assume that a control action of the human operator is preceded by comparison of information about the required state of the plant with the information about the actual state of the plant. Basing on that comparison and using an inner model as knowledge about the behavior of the object, the operator takes a decision and then executes a control action. Hence, the interactive decision table representing the pilot's decision process has the form $DT_F = \langle U, C_F \cup D_F, \{V_a\}_{a \in C \cup D} \rangle$. Every row of this decision table corresponds to an element of a universe U, characterized by two disjoint sets of attributes called condition and decision attributes, respectively, denoted as:

- $C_F = E_F \cup {Ac_F}^{prev}$ is the set of condition attributes with fuzzy values, where E_F is a set of attributes in the form of deviation from required value of flight parameters and the tendency of change of this state variables, ${Ac_F}^{prev}$ is the set of fuzzy attributes, which express the action of the pilot in the previous moment of time; this set represents the connection between an action which has been just done and is being done, because the deflection of control elements has a continuous nature,
- $D_F = {Ac_F}^{cur}$ is the set of fuzzy decision attributes which represent the current action of pilot in the form of qualitative changes of aircraft's control elements.

During the process of creating the decision table DT_F, we ought to take into account the lag caused by the human and the characteristic of the plant. That is why the row of the decision table DT_F, corresponding to the time T_k, is constructed using fuzzy values of condition attributes from C_F for the instant T_k, and the values of decision attributes from D_F usually for the instant T_{k+1}. Summarizing, we assume that the operator perceives the environment at the moment T_k, performs activity at the moment T_{k+1} and observes the effects of that decision at the moment T_{k+2}. The obtained decision table DT_F can be appropriate source for analysis and attempt to build the pilot's decision-action model in tasks of aircraft control.

3. DECISION TABLES WITH FUZZY ATTRIBUTES

As stated in previous sections, we want to use attributes with fuzzy (linguistic) values in interactive information systems. Hence, we need to recall a formal description of a decision table with fuzzy attributes, which we introduced in [2].

Let us consider a finite universe U with N elements: $U = \{x_1, x_2, \ldots, x_N\}$. Any element x of the universe U will be described with the help of fuzzy attributes, which are divided into a subset of n condition attributes $C = \{c_1, c_2, \ldots, c_n\}$, and a subset of m decision attributes $D = \{d_1, d_2, \ldots, d_m\}$.

Every fuzzy attribute possesses a set of linguistic values. We denote by $C_i = \{C_{i1}, C_{i2}, \ldots, C_{in_i}\}$ the family of linguistic values of the condition attribute c_i, and by $D_j = \{D_{j1}, D_{j2}, \ldots, D_{jm_j}\}$ the family of linguistic values of the decision attribute d_j, where n_i and m_j is the number of the linguistic values of the i-th condition and the j-th decision attribute, respectively, $i=1,2,\ldots,n$, and $j=1,2,\ldots,m$.

For every $x \in U$, its membership degrees in all linguistic values of the condition attribute c_i (decision attribute d_j) need to be determined. The value of an attribute for an element $x \in U$ is a fuzzy set on the domain of all linguistic values of that attribute.

We denote by $C_i(x)$ the fuzzy value of the condition attribute c_i for any x, as a fuzzy set on the domain of the linguistic values of c_i

$$C_i(x) = \{\mu_{C_{i1}}(x)/C_{i1}, \mu_{C_{i2}}(x)/C_{i2}, \ldots, \mu_{C_{in_i}}(x)/C_{in_i}\}. \tag{1}$$

$D_j(x)$ is the fuzzy value of the decision attribute d_j for any x, as a fuzzy set on the domain of the linguistic values of d_j

$$D_j(x) = \{\mu_{D_{j1}}(x)/D_{j1}, \mu_{D_{j2}}(x)/D_{j2}, \ldots, \mu_{D_{jm_j}}(x)/D_{jm_j}\}. \tag{2}$$

It is a common property of fuzzy information systems that any $x \in U$ has a non-zero membership to more than one linguistic value of a fuzzy attribute.

Let us assume that for any $x \in U$, all linguistic values $C_i(x)$ and $D_j(x)$ ($i=1,2,\ldots,n, j=1,2,\ldots m$) satisfy the requirements

$$\text{power}(C_i(x)) = \sum_{k=1}^{n_i} \mu_{C_{ik}}(x) = 1, \qquad \text{power}(D_j(x)) = \sum_{k=1}^{m_j} \mu_{D_{jk}}(x) = 1. \tag{3}$$

This restriction is necessary to propose a generalized fuzzy flow graph approach. It is possible to express and examine all possible decision rules obtained by using the Cartesian product of sets of the linguistic values of attributes.

We denote by R_k the k-th decision rule from the set consisting of r possible decision rules ($r = \prod_{i=1}^{n} n_i \prod_{j=1}^{m} m_j$)

$$\begin{aligned} R_k:\quad &\text{IF} && c_1 \text{ is } C_1^k \text{ AND } c_2 \text{ is } C_2^k \ldots \text{AND } c_n \text{ is } C_n^k \\ &\text{THEN} && d_1 \text{ is } D_1^k \text{ AND } d_2 \text{ is } D_2^k \ldots \text{AND } d_m \text{ is } D_m^k \end{aligned} \tag{4}$$

where: $k=1,2,\ldots,r$, $C_i^k \in C_i, i=1,2,\ldots n$, $C_j^k \in D_j, j=1,2,\ldots,m$.

We determine to what degree any $x \in U$, corresponding to a single row of the decision table, confirms particular decision rules. To this end, we calculate the truth value of the decision rule's antecedent and the truth value of the decision rule's consequent. This is done by determining the conjunction of the respective membership degrees of $x \in U$ in the linguistic values of attributes.

The set of those $x \in U$, which confirm a decision rule, is called the support of the decision rule.

In order to determine the confirmation degree of fuzzy decision rules, we apply a fuzzy T-norm operator. By $\mathrm{cd}(x,k)$, we denote the confirmation degree of the k-th decision rule by the element $x \in U$

$$\mathrm{cd}(x,k) = \mathrm{T}(\mathrm{cda}(x,k), \mathrm{cdc}(x,k)), \tag{5}$$

where $\mathrm{cda}(x,k)$ denotes the confirmation degree of the decision rule's antecedent

$$\mathrm{cda}(x,k) = \mathrm{T}(\mu_{V_1^k}(x), \mu_{V_2^k}(x), \ldots, \mu_{V_n^k}(x)), \tag{6}$$

and $\mathrm{cdc}(x,k)$ the confirmation degree of the decision rule's consequent

$$\mathrm{cdc}(x,k) = \mathrm{T}(\mu_{W_1^k}(x), \mu_{W_2^k}(x), \ldots, \mu_{W_m^k}(x)). \tag{7}$$

By using the confirmation degrees (6), (7) and (5), the following fuzzy sets on the domain U are obtained:
the support of the decision rule's antecedent

$$\mathrm{support}(\mathrm{cda}(x,k)) = \{\mathrm{cda}(x_1,k)/x_1, \mathrm{cda}(x_2,k)/x_2, \ldots, \mathrm{cda}(x_N,k)/x_N\}, \tag{8}$$

the support of the decision rule's consequent

$$\mathrm{support}(\mathrm{cdc}(x,k)) = \{\mathrm{cdc}(x_1,k)/x_1, \mathrm{cdc}(x_2,k)/x_2, \ldots, \mathrm{cda}(x_N,k)/x_N\}, \tag{9}$$

and the support of the decision rule R_k, respectively

$$\mathrm{support}(R_k) = \{\mathrm{cd}(x_1,k)/x_1, \mathrm{cd}(x_2,k)/x_2, \ldots, \mathrm{cd}(x_N,k)/x_N\}. \tag{10}$$

The confirmation degrees (6), (7) and (5) will be used to express strength, certainty, and coverage factors of a fuzzy decision rule.

4. FUZZY FLOW GRAPHS

The concept of flow graphs for discovering the statistical properties of crisp decision algorithms was introduced by Pawlak [4]. We proposed a generalized fuzzy flow graph approach [2, 3], suitable for analysis of fuzzy information systems.

A crisp flow graph [4] is given in the form of directed acyclic final graph $G = (\mathcal{N}, \mathcal{B}, \varphi)$, where $\mathcal{N}$ is a set of nodes, $\mathcal{B} \subseteq \mathcal{N} \times \mathcal{N}$ is a set of directed branches, $\varphi : \mathcal{B} \to \mathrm{R}^+$ is a flow function (values in the set of non-negative reals R^+).

For any $(X, Y) \in \mathcal{B}$, X is an input of Y and Y is an output of X. The quantity $\varphi(X, Y)$ is called the throughflow from X to Y.

$I(X)$ and $O(X)$ denote an input and an output of X, respectively. The input $I(G)$ and output $O(G)$ of a graph G are defined by

$$I(G) = \{X \in \mathsf{N} : I(X) = \varnothing\}, \qquad O(G) = \{X \in \mathsf{N} : O(X) = \varnothing\}. \tag{11}$$

Every node $X \in \mathcal{N}$ of a flow graph G is characterized by its inflow and by its outflow

$$\varphi_+(X) = \sum_{Y \in I(X)} \varphi(Y, X), \qquad \varphi_-(X) = \sum_{Y \in O(X)} \varphi(X, Y). \tag{12}$$

For any internal node X, the equality $\varphi_+(X) = \varphi_-(X) = \varphi(X)$ is satisfied. The quantity $\varphi(X)$ is called the flow of the node X.

The flow for the whole graph G is defined by

$$\varphi(G) = \sum_{X \in I(G)} \varphi_-(X) = \sum_{X \in O(G)} \varphi_+(X). \tag{13}$$

By using the flow $\varphi(G)$, the normalized throughflow $\sigma(X, Y)$ and the normalized flow $\sigma(X)$ are determined as follows

$$\sigma(X, Y) = \frac{\varphi(X, Y)}{\varphi(G)}, \qquad \sigma(X) = \frac{\varphi(X)}{\varphi(G)}. \tag{14}$$

For every branch of a flow graph G the certainty and coverage factors are defined as follows

$$\text{cer}(X,Y) = \frac{\sigma(X,Y)}{\sigma(X)}, \qquad \text{cov}(X,Y) = \frac{\sigma(X,Y)}{\sigma(Y)}. \tag{15}$$

The certainty and coverage factors satisfy the following properties

$$\sum_{Y \in O(X)} \text{cer}(X,Y) = 1, \qquad \sum_{X \in I(Y)} \text{cov}(X,Y) = 1. \tag{16}$$

It is possible to express a decision table with fuzzy attributes as a flow graph [2, 3]. We assume that only one decision attribute will be used. Every attribute is represented by a layer of nodes. The nodes of the input and hidden layers correspond to linguistic values of the condition attributes. The output layer nodes correspond to linguistic values of the decision attribute.

We denote by $\widetilde{X}$ a fuzzy set on the universe U, which describes membership degree of particular elements $x \in U$ in the linguistic value represented by X. The membership degrees of all x in the set $\widetilde{X}$ can be found in a respective column of the considered decision table.

The flow $\varphi(X,Y)$ for the branch (X,Y) is equal to power (fuzzy cardinality) of the product of fuzzy sets $\widetilde{X}$ and $\widetilde{Y}$. Let us observe that the T-norm operator prod should be used for determining the product of sets, in order to satisfy the following equation for the input and internal layer nodes

$$\varphi_-(X) = \text{power}(\widetilde{X}) = \sum_{Y \in O(X)} \varphi(X,Y) = \sum_{Y \in O(X)} \text{power}(\widetilde{X} \cap \widetilde{Y}). \tag{17}$$

An analogous equation can be given for the output and internal layer nodes

$$\varphi_+(X) = \text{power}(\widetilde{X}) = \sum_{Y \in I(X)} \varphi(Y,X) = \sum_{Y \in I(X)} \text{power}(\widetilde{X} \cap \widetilde{Y}). \tag{18}$$

The equality $\varphi_+(X) = \varphi_-(X) = \varphi(X)$ is satisfied for any internal node X, when the T-norm operator `prod` is used. The total normalized inflow (outflow) of each layer does depend on the form of T-norm operator used in calculations. In order to satisfy (17) and (18), the total normalized inflow (outflow) of a layer should be equal to 1. This is is fulfilled, when we apply both the properties (3) and the T-norm operator `prod`.

We can merge the input and hidden layers of the flow graph into a single layer, which contains nodes representing all possible combinations of linguistic values of the condition attributes. Let us denote by X^* a node of the resulting layer. The node X^*

corresponds to antecedent of a certain decision rule R_k. Support of the antecedent of the decision rule R_k is determined by using (8).

The decision rule R_k is represented by a branch (X^*, Y), where Y denotes a node of the output layer.

Power of the support of the rule R_k, defined by (10), is equal to the flow between the nodes X^* and Y

$$\varphi(X^*, Y) = \text{power}(\text{support}(R_k)). \tag{19}$$

By using the formulae (8), (9) and (10), we can determine, for every decision rule R_k, the certainty factor $\text{cer}(X^*, Y)$, coverage factor $\text{cov}(X^*, Y)$ and strength of the rule $\sigma(X^*, Y)$

$$\text{cer}(X^*, Y) = \text{cer}(R_k) = \frac{\text{power}(\text{support}(R_k))}{\text{power}(\text{support}(\text{cda}(x, k)))}, \tag{20}$$

$$\text{cov}(X^*, Y) = \text{cov}(R_k) = \frac{\text{power}(\text{support}(R_k))}{\text{power}(\text{support}(\text{cdc}(x, k)))}, \tag{21}$$

$$\sigma(X^*, Y) = \text{strength}(R_k) = \frac{\text{power}(\text{support}(R_k))}{\text{card}(U)}. \tag{22}$$

We apply the above notions for calculations in the following example.

5. EXAMPLE

Let us consider the task of stabilization of the aircraft's altitude performed by a pilot during a horizontal flight phase. There are four condition attributes and one decision attribute in the decision system. They have the following interpretation:

c_1 – altitude deviation from the required value: {c_{11} – "large negative", c_{12} – "small negative", c_{13} – "zero", c_{14} – "small positive", c_{15} – "large positive"},

c_2 – rate of climb: {c_{21} – "negative", c_{22} – "zero", c_{23} – "positive"},

c_3 – change of rate of climb: {c_{31} – "negative", c_{32} – "zero", c_{33} – "positive"},

c_4 – change of the rudder deflection angle in the previous moment of time:
$\{c_{41}$ – "negative", c_{42} – "zero", c_{43} – "positive"$\}$,

d_1 – change of the rudder deflection angle: $\{d_{11}$ – "negative (decrease)",
d_{12} – "zero", d_{13} – "positive (increase)"$\}$.

The membership functions of particular linguistic values of the attributes have a typical triangular or trapezoidal shape. Each attribute is represented as a layer of nodes in a fuzzy flow graph. The nodes correspond to particular linguistic values of an attribute. According to the method described in the previous section, we merge layers connected with condition attributes and determine the normalized flow between the resulting and the output (decision) layer. For every decision rule (all combination of linguistic values are considered), the strength, certainty and coverage factors are calculated.

Table 1. Decision table with fuzzy attributes

	C_1					…	C_4			D_1		
	C_{11}	C_{12}	C_{13}	C_{14}	C_{15}	...	C_{41}	C_{42}	C_{43}	D_{11}	D_{12}	D_{13}
x_1	0.00	0.00	1.00	0.00	0.00	...	0.00	1.00	0.00	0.00	1.00	0.00
x_2	0.00	0.00	0.90	0.10	0.00	...	0.00	1.00	0.00	0.00	0.90	0.10
x_3	0.00	0.00	1.00	0.00	0.00	...	0.00	0.90	0.10	0.00	1.00	0.00
x_4	0.00	0.00	0.10	0.90	0.00	...	0.00	1.00	0.00	0.00	0.20	0.80
x_5	0.00	0.00	0.00	1.00	0.00	...	0.00	0.20	0.80	0.20	0.80	0.00
x_6	0.00	0.00	1.00	0.00	0.00	...	0.20	0.80	0.00	0.00	1.00	0.00
x_7	0.00	0.00	0.00	1.00	0.00	...	0.00	1.00	0.00	0.00	0.00	1.00
x_8	0.00	0.00	0.20	0.80	0.00	...	0.00	0.00	1.00	0.10	0.90	0.00
x_9	0.00	0.00	1.00	0.00	0.00	...	0.10	0.90	0.00	0.00	1.00	0.00
x_{10}	0.00	0.00	0.00	0.10	0.90	...	0.00	1.00	0.00	0.00	0.00	1.00

Table 2 contains decision rules obtained from analysis of decision Table 1. Only the rules above a certain level of strength (4%), certainty factor (0.6), and coverage factor (0.1) for branches between resulting and output layer, are taken into account. These decision rules can be treated as a part of the pilot's control action model for the selected flight phase.

Table 2. Decision rules obtained from Table 1

Rules		Strength [%]	Coverage	Certainty
R_1	$C_{13}C_{22}C_{32}C_{42} \rightarrow D_{12}$	41.69	0.613	0.981
R_2	$C_{14}C_{22}C_{31}C_{43} \rightarrow D_{12}$	12.24	0.180	0.850
R_3	$C_{14}C_{23}C_{33}C_{42} \rightarrow D_{13}$	18.10	0.624	0.910
R_4	$C_{15}C_{23}C_{33}C_{42} \rightarrow D_{13}$	8.10	0.279	1.000

6. CONCLUSIONS

We give only one simple example in order to illustrate the presented ideas and to prove the possibility of application of fuzzy flow graphs in the analysis of pilot's control actions. In the case of complex decision tables generated from real data, we created and used dedicated software to facilitate calculations and select decision rules with highest values of strength, certainty and coverage factors. In future research, we plan to apply the presented method to analyze other flight phases and to build a model of pilot's actions, which could be used in a control system. It should be emphasized that our fuzzy flow graph approach is an original generalization, which can be useful in different applications.

REFERENCES

[1] GESTWA M., *Using Fuzzy Control for Modeling the Control Behaviour of a Human Pilot*, In: Fuzzy Controllers, Theory and Applications, Grigorie L. (Ed.), InTech, 2011.

[2] MIESZKOWICZ-ROLKA A., ROLKA L., *Flow graphs and decision tables with fuzzy attributes*, In: Artificial Intelligence and Soft Computing – ICAISC 2006. Lecture Notes in Artificial Intelligence, Rutkowski L. et al. (Eds.), vol. 4029. Berlin, Springer-Verlag, 2006, 268–277.

[3] MIESZKOWICZ-ROLKA A., ROLKA L., *Variable precision fuzzy rough sets model in the analysis of process data*, In: Transactions on Rough Sets VI: Lecture Notes in Computer Science (Journal Subline), Peters J. F. et al. (Eds.), vol. 4374. Berlin, Springer-Verlag, 2007, 191–210.

[4] PAWLAK Z., *Flow graphs and data mining*, In: Transactions on Rough Sets III: Lecture Notes in Computer Science (Journal Subline), Peters J. F. et al. (Eds.), Vol. 3400. Berlin, Springer-Verlag, 2005, 1–36.

[5] SKOWRON A., WASILEWSKI P., *Interactive information systems: Toward perception based computing*, Theoretical Computer Science, Vol. 454, 2012, 240–260.

[6] WANG L., MENDEL J.R., *Generating Fuzzy Rules by learning from examples*, IEEE Transactions on Systems, Man, and Cybernetics, Vol. 22, 1992, 1414–1427.

wind energy, forecasting, data mining

Tomasz POPŁAWSKI*, Piotr SZELĄG*,
Cezary GŁOWIŃSKI**, Łukasz ADAMOWICZ**

FORECAST OF ELECTRIC POWER GENERATED BY THE WIND FARM USING DATA CLUSTERING METHODS

The increasing number of wind farms results in increasing importance of wind power in the power system. It entails the necessity of preparation more and more precise forecasts of wind farms work. However, the characteristic of the source exploited to produce energy makes the task extremely difficult. The increasing amount of data is another problem, since its processing influences the pace of forecast preparation. The article shows the method allowing decreasing the amount of data needed to prepare a forecast, and as a result, shorten the time of calculations. At the same time, satisfactory precision of the forecast is maintained.

1. INTRODUCTION

For a few decades, a substantial increase in producing energy from renewable sources of energy has been observed [1, 2]. This tendency concerns all the renewable sources – biomass, water, wind and solar energy [3].

The growth of new power plants based on renewable sources of energy is possible thanks to the efforts of the international community which aims at the restriction of the use of fossil fuels, which are limited, and supporting the production of energy from unlimited sources which do not influence the environment as negatively as conventional methods. The rules of the E.U. climate and energy package even oblige member states to achieve a specified share of renewable energy in their energy balance until 2020.

The most dynamic increase may be seen in case of wind energy. This tendency is slightly influenced by events connected to the economic recession. Over the last four

* Czestochowa University of Technology, Faculty of Electrical Engineering, Armii Krajowej 17, 42-200 Częstochowa.

** SAS Institute Polska, Gdańska 27/31, 01-633 Warszawa.

years, ca. 40,000 MW new installations per year have been built (see Figure 1). As far as the pace of the process is concerned, Poland does not deviate from this tendency. However, if one looks at the amount of installed capacity (2,500 MW), there is still a lot to do, especially in comparison with the European leader – Germany (31,000 MW). Despite such considerable differences, wind capacity installed in National Power System (NPS) constitutes 6.8%. However, on account of its work characteristics, it makes up only ca. 2.5% the electric energy produced. The amount is increasing every year.

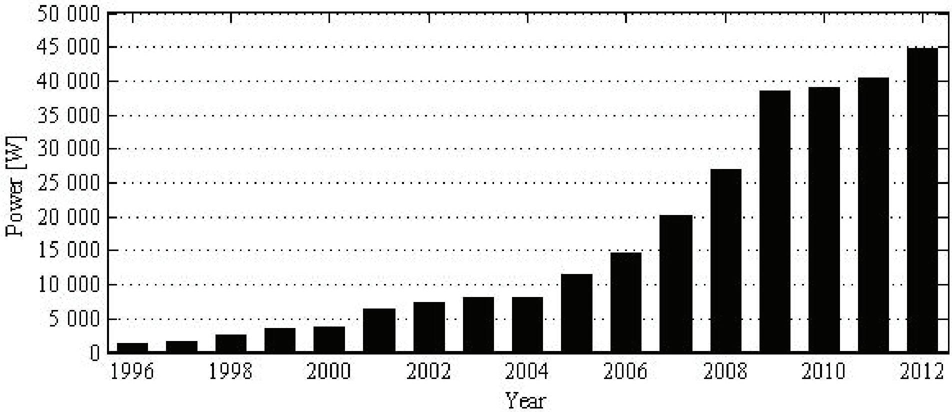


Fig. 1. Global Annual Installed Wind Capacity 1996-2012. Source: [2]

The dynamic development of wind energy sector results in establishing a new path of forecasting connected to working out new prediction tools and methods designed for wind power plants, which are characteristic of instability of the amount of generated power and the difficulties in its prediction [5, 6, 7]. In spite of a few decades devoted to creating and improving new tools, the process cannot be thought of as completed. There are still new problems demanding a solution [4, 6, 8]. One of them, whose importance will increase in the future, is the increase in the amount of data which can be used as input information in order to prepare a forecast of power generated by a wind power plant.

Data transmitted by a wind farm may have very short intervals (there is a possibility of receiving the data every few seconds). SCADA system transmits comprehensive information about the work of such a plant, e.g. wind speed, generated power, atmospheric pressure. Processing of such large amount of data is very difficult and time-consuming. Preparation of a forecast must be preceded by processing and selecting information which substantially influence the precision of a forecast done for a wind power plant use.

As far as economic activity is concerned, a sector called Business Intelligence (BI) is increasing dynamically. The aim of BI is to convert data into information necessary to gain knowledge which could be used in order to raise the company's competitiveness on the market. The BI system generates standard reports or calculates Key Performance Indicators on the basis of which hypotheses can be made. Afterwards, they are verified by performing detailed "sections" of data. Various analytical tools are used in the process (e.g. OLAP, data mining). From National Power System management and the preparation of a comprehensive forecast of power generation by a wind power engineering point of view, a new challenge arises; namely, the challenge of processing large amount of data received from wind power plants. Adding any new installations inevitably leads to the use of tools dedicated to a new direction in data analysis – Big Data.

The tools of one of the major companies in Poland and in the world which specializes in preparing and implementing Big Data and Business Intelligence software were used for and adjusted to this article. Three analytical modules were chosen out of many: SAS Enterprise Guide, SAS Enterprise Miner and SAS Forecast Server. The process of data preparation and performing a forecast is described further in this article.

2. MEASURING DATA PREPARATION

The preparation of a wind generated power forecast will entail the need for processing large amount of data. Because of their location and connection to a network, most of wind power plants are gathered on wind farms. For an owner of such a wind farm a forecast of total power generated is more important that a forecast of power generated by a single wind power plant. That is why it is logical that a wind farm should be the smallest unit taken into account while preparing a forecast.

The authors preparing the forecast for the chosen wind farm had data covering a 4-month period at their disposal. The available readouts were done with 15-minute intervals and contained information on: time and date of a readout, generated power, wind speed, the number of power plants ready to work, and the number of working power plants. Especially, the latter information appeared to be precious and allowed to standardize data thanks to a simple calculation: dividing the amount of generated power by the number of working power plants. As a result, one variable was reduced (the number of working power plants) and the new value specifies power generated by one wind power plant within a given period of time and measured wind speed. This information was used to prepare the curve of power.

A 4-month observation gave circa 11,500 records. In order to prepare a forecast founded on a database of such an extent, standard tools and methods can be used. However, it is worth remembering that it is just a small sample of information available in the power system. Taking this into account, tools prepared for operating large databases were used.

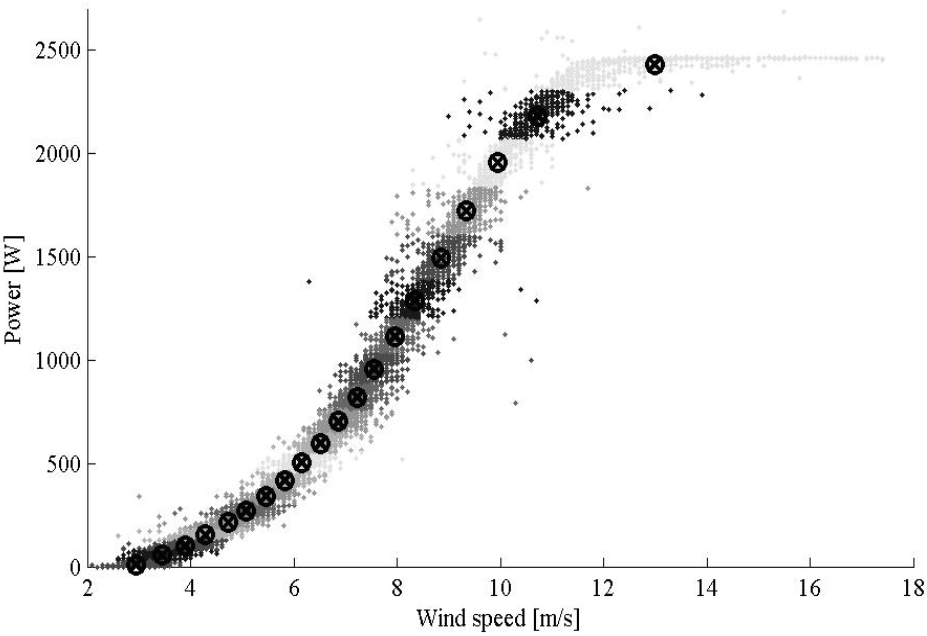


Fig. 2. An exemplary division of source data into clusters indicating chosen centroids

The original set was divided into teaching and testing parts in three to one relation. As a result, the database used for approximation numbered ca. 8,500 records. In order to reduce this amount, a k-means method was used. The method allowed the set of input data to be limited to several dozen of points which were the centroids of individual clusters the original set was divided into (Figure 2). The centres of clusters which were the pairs of coordinates representing wind speed and generated power were approximated in the next stage.

3. POWER FORECAST

The work of a wind power plant characterizes by large fluctuations of the level of power generate. It is a result of a lot of factors, most importantly the speed of the wind flowing over the turbine set. However, other factors cannot be forgotten, temperature and atmospheric pressure among others. They influence air density and, as a result, the amount of generated power. Besides weather conditions, there is a possibility of controlling the power plant by adjusting the nacelle to the direction of the air flow and setting the angle of the blades. The influence of each factor is shown in formula 1 [4].

$$P = c_p \rho D^2 v^3 \tag{1}$$

where:

P – power generated $[W]$

c_p – total efficiency of transformation wind energy into mechanical energy

ρ – air density $\left[\frac{kg}{m^3}\right]$

D – diameter of blade circle $[m]$

v – wind velocity $\left[\frac{m}{s}\right]$

The value of wind speed is given to the power of three. That is why even slight changes in this value may provoke big differences in the amount of generated power. Wind power plants producers also acknowledge the substantial influence of wind speed and, therefore, they characterize an object by presenting the curve of power, which is the dependence of generated power on the wind speed. The use of this very characteristic is one of the methods used to prepare a forecast. Nonetheless, it may be used only at the early stages of power plant work until a sufficient work history of a power plant in a specific place is obtained.

The curve of power given by a producer differs from the curve of power obtained from measuring a specific object. There are a few reasons, for example: specific location, natural obstacles, or the presence of other wind power plants. Therefore, it is justified to create and periodically verify the characteristics of working turbine sets. Having measuring data (wind speed, generated power), one is able to prepare a characteristic which is more precise than the results obtained from the use of a producer's curve. The problem which appears at this stage is the choice of an appropriate approximation function. It is important on account of the previous step, that is limiting the measuring data. Substantial reduction of samples and, as a result, narrowing the area of the approximation may negatively influence some types of functions. The matching error increases if the whole range of power plant work is taken into account, not just the approximated part. Polynomial functions may pose such problems. That is why; it was decided to use a sigmoid function which is shown in the formula 2.

$$\hat{P} = \frac{c}{1 + e^{-a(\hat{v} - b)}} \tag{2}$$

where:

$\hat{P}$ – forecast power $[W]$

$\hat{v}$ – forecast wind speed $\left[\frac{m}{s}\right]$

a, b, c – approximation function coefficients

On the basis of fixed coefficients and a wind speed forecast it is possible to determine a forecast of power generated by a wind power plant.

4. TOOLS AND METHODS OF CALCULATION

4.1. TOOLS USED

Calculations were made with the use of SAS Enterprise Guide software. This application allows preparing data needed for analyses; it contains a large collection of statistical methods, and allows a comprehensive visualization of results with help of various graphs and final reports. Thus, the described process was possible to perform without the use of any other tools. A user's graphic interface, which allows designing the full analysis in the form of a process diagram and simultaneously enabling the access to all the detailed parameters of statistical procedures, is an additional advantage, by dint of which it is possible to quickly modify analyses and perform a lot of iterations of a specific calculation experiment.

Within the framework of this very thesis, cluster analysis methods available within the confines of PROC FASTCLUS procedures and approximation methods based on PROC NLIN nonlinear regression procedure were used. The former procedure allows doing cluster analyses with the use of various methods. For the use of this article, the final analysis was done on the basis of the WARD method. The latter procedure allows the approximation through adjusting the parameters of a specific model of nonlinear regression.

4.2. CALCULATIONS

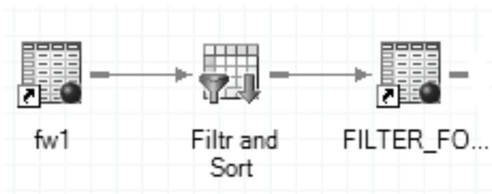


Fig. 3. Process of source data filtrating

The analysis of measuring data from the SCADA system showed the necessity of performing the initial filtration. It resulted from the lack of continuity in saving some parameters. The capabilities of SAS Enterprise Guide application allowed carrying out this operation intuitively and obtaining clean data.

Each of the process elements, from loading of the data by filtration to obtaining the results showed on a graph as a separate component of the process can be modified individually. Figure 3 illustrates the scheme of the task performed.

The next stage (as mentioned before) was to divide the data into teaching and testing parts. The task came down to introducing specific filters (similarly to the previous task). The test data underwent a cluster analysis. The number of groups changed from three to twenty. A creator available within the confines of this function allows choosing data, method of analysis, and the way of presenting the results easily. The results in the form of coordinates of subsequent analyses centroids were shown in a separate table. They were used in the next stage of calculations in order to determine the coefficients of the sigmoid function showed in the formula 2. Nonlinear regression which is needed to do that, is one of the components available within the framework of this task, together with linear and logistic regression, as well as generalized linear models.

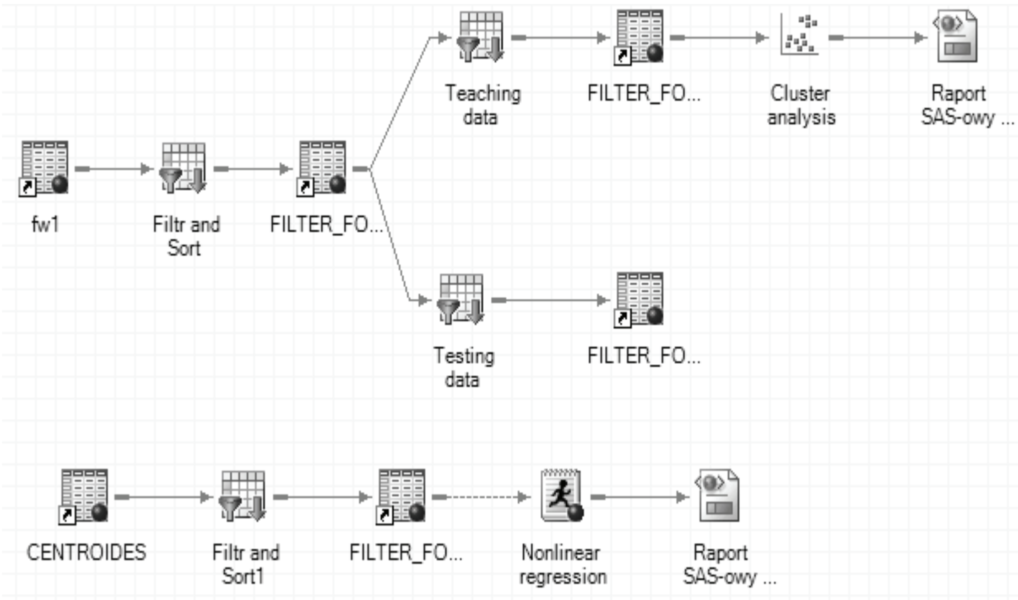


Fig. 4. Process of data division, cluster analysis and nonlinear regression

Filling further information with the use of the creator, the independent variable, the variable explaining the method of calculations, the initial point, and the range of coefficients changeability were specified. The problem appeared on the point of introducing the formula of sigmoid function, since in the collection of dozen models there was no such a model which would suit the authors of the article. The problem was quickly solved. Each element of the process created with the use of the creator has a related program code. A slight modification of one of the elements of the PROC NLIN procedure allowed obtaining the required model. The process was illustrated in the Figure 4.

The authors obtained only measuring data. In order to do a forecast of power generated by a wind farm, a forecast of wind speed is essential. A pseudo-forecast was made which involved upsetting the real readouts by random values. The amount of changes fluctuated from –20% to +20% of the measured value. On the basis of the wind speed forecast and specified coefficients of the function, it was possible to prepare power forecasts.

5. RESULTS

The testing data were used in order to prepare the power forecast and to verify its precision. The teaching data were divided into specified numbers of subsets. The minimal number of groups is three, since it is a number essential to perform the regression. The coefficients of centroids of these groups were used as input data to determine the parameters of a sigmoid function which was a curve of power used to perform a power forecast at the same time. A dozen of curves were chosen, one per each division. Each curve was used to do the forecast. Next, an average error of the forecast was determined for each of the curves individually.

Table 1. Forecast error

	Number of clusters					
Error [%]	3	6	9	12	15	18
average	9,57	7,64	7,68	7,72	7,93	8,06
max.	49,14	54,59	55,34	55,83	59,28	61,21
min.	0,01	0,00	0,00	0,00	0,00	0,00

The error was not determined in a classical way. Wind power engineering worked out its own method of determining the error, taking into account the characteristic of wind power plant work. Namely, the difference between the real and the forecast value is compared with the nominal size of the object for which the forecast is prepared. In this case, it is a wind farm of 20 MW. The value of the error was shown in the Figure 5.

The biggest forecast error is seen if data are divided into three subgroups. The increase in the number of clusters by one causes a drop in the value of the error – over two percentage points. The increase in the number of divisions results in the increase of the value of the error and its stabilization within the range from 8 to 8.2% if the number of clusters is over twenty.

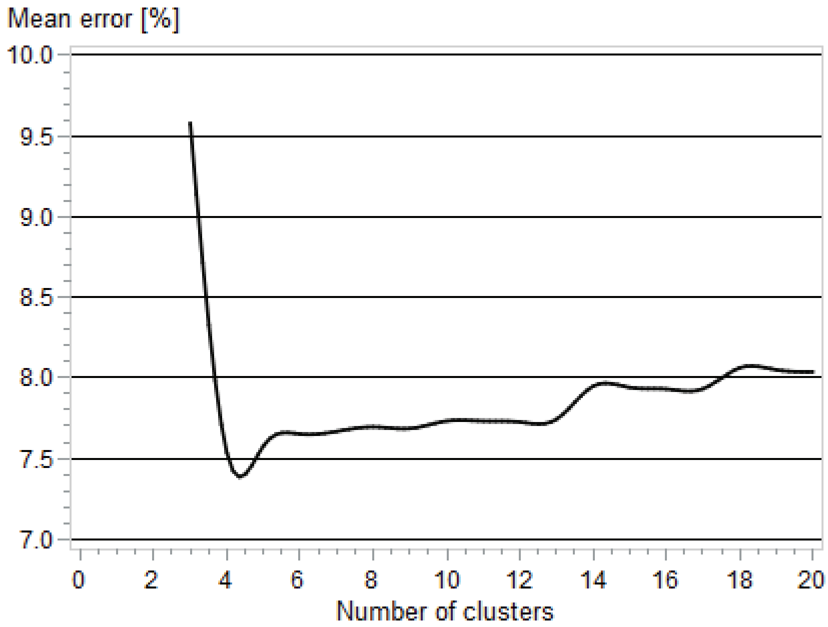


Fig. 5. The scale of the forecast error in the number of the groups function

The maximal error is shaped in the range from 49% to 62% which is shown in the Table 1. It is worth noticing that a maximal error increases together with increasing the number of clusters.

6. CONCLUSIONS

The use of data mining methods is connected to the pace of development of wind power engineering and, hence, the amount of measuring data obtained from wind farms. The application of cluster analysis as a stage preceding the preparation of the proper forecast is to simplify the complexity of calculations and, at the same time, to preserve the properties and the characteristic of the forecast process. The method shown in this article allows, to a great extent, reducing the amount of data needed and, simultaneously, ensures a satisfactory effectiveness measured by a nominal Mean Absolute Percentage Error (nMAPE).

The choice of appropriate tools is also very important. SAS Institute solutions described in the article allow for making necessary calculations intuitively, and for their clear presentation. Mechanisms processing large amount of data are an additional advantage of the presented software.

The obtained forecast errors in the range 7.4 – 8% are comparable with the results showed in the literature. However, the chosen method allows limiting the amount of input data and thus shortens the time of a single forecast which is extremely important especially while preparing forecasts with several to dozen-minute intervals. Therefore, the research direction is justified and it should be continued. Nevertheless, the amount of data used in order to make a forecast obliges to being careful in drawing conclusions. It is necessary to carry out further tests on a bigger sample.

REFERENCES

[1] BRENTON S., MOE G., *Status, plans and technologies for offshore wind turbines in Europe and North America*, Renewable Energy, 2009, Volume 34, No. 3,.646–665.

[2] FRIED L., SAWYER S., SHUKLA S., QIAO L., *Global wind report –Annual market update*. Global Wind Energy Council, 2012.

[3] RODZIEWICZ T., TENETA J., ZAREMBA A., WACŁAWEK M., *Analysis of Solar Energy Resources in Southern Poland for Photovoltaic Applications*. Ecological Chemistry and Engineering S., 2013, 20(1), 177–198.

[4] POPŁAWSKI T., SZELĄG P., *A Fractal Approach Towards Work of a Wind Power Station*. Systems Architecture and Technology. System Analysis Approach to the Design, Control and Decision Support. Eds. Jerzy Świątek, Leszek Borzemski, Adam Grzech, Zofia Wilimowska, 2013, 121–130.

[5] POPŁAWSKI T., SZELĄG P., DĄSAL K., ŁYP J., *Zastosowanie modeli ARMA do przewidywania mocy i energii pozyskiwanej z wiatru*. Polityka Energetyczna, 2010 T.13z.2 385–400.

[6] POPŁAWSKI T., DĄSAL K., ŁYP J., *Problematyka prognozowania mocy i energii pozyskiwanych z wiatru*. Polityka Energetyczna, 2009, T.12z.2 511–523.

[7] KHALID M., SAVKIN A.V., *A Method for Short-Term Wind Power Prediction With Multiple Observation*. Power Systems, IEEE Transactions, 2012, Volume 27, Issue: 2, 579–586.

[8] BARTHELMIE R.J., MURRAY F., PRYOR S.C *The economic benefit of short-term forecasting for wind energy in the UK electricity market.* Energy Policy, 2008, Volume 36, no. 5, 1687–1696.

Anna KIŁYK, Zofia WILIMOWSKA*

VALUE AT RISK DYNAMICS FOR INVESTMENT PORTFOLIO

This work is a presentation of results obtained during the investigation of portfolio risk dynamics. This research, apart from the Value at Risk method uses the following method of determining the dynamics: moving average and moving standard deviation. This approach of calculating the Value at Risk will not only permit observation of changes but also allows an investigation of repeating characteristic patterns.

INTRODUCTION

Operation and continuous development of the company is mainly dependent on investment decisions. You can invest in people, equipments, technology, but all these actions are related with capital, which in future by definition should be returned in different forms (eg. knowledge, reduced production time, new technology etc.). However, on the other hand these actions are associated with taking risks. It is worth noting that many people identify company risk as capital investment risk. On the one hand companies mostly invest in securities to raise capital, but on the other companies are represented by people, who define risk in they own way.

Speaking about measurement methods, which have to 'protect' us from taking actions that might lead to loss, usually we're thinking about:

- Statistical values such as standard deviation or variation,
- Methods like VaR (Value at Risk), CFaR (Cash Flow at Risk), LaR (Liquidity at Risk).

* Institute of Organization and Mangement, Wrocław University of Technology, Wybrzeże Wyspiańskiego 27, 50-370 Wrocław.

Among those approaches the most important is Value at Risk. This method is so highly interesting because of its versatility and ease of use. In this paper authors show research results of value at risk dynamics for an earlier made portfolio.

1. CHARACTERISTIC OF VAR METHOD

Speaking about value at risk (VaR) one usually has in mind the value that can be lost during the investment with specific initial parameters. In this case, the initial parameter will be:

- Time period (time window) – time period in which the investment is analyzed,
- Level of tolerance (confidence level) – which shows how big probability of capital loss can be accepted by the investor. One should keep in mind that the level of tolerance and confidence are directly related, and their relationship can be written as:

$$level\ \ of\ \ tolerance = 1 - confidence\ \ level \tag{1}$$

In this paper authors analyzed an investment portfolio, that's why it's worth to quoting the definition given by P. Best which describes VaR value as *the maximum amount that can be lost through investment in a portfolio during a specified time horizon and with a given confidence level* (Pera 2008, pp.275). This definition shows that value at risk is a function of two variables (confidence interval and time horizon) and both of them are proportional to the VaR. It means that in a situation where one of the parameters increases the whole value VaR increases as well.

The most common method of calculating value at risk (2) is an equation that indirectly expresses VaR as the probability of situation (Jajuga and others 2000):

$$P(W_\tau \le W_\alpha) = \alpha \tag{2}$$

Where:

W_τ – final value of the investment at the end of analyzed period time τ (random value),

W_α – quartile of the distribution for a given level of tolerance,

α – the level of tolerance.

The value of W_α can be expressed as the initial value of the investment W_0 via (3):

$$W_\alpha = W_0 - VaR \tag{3}$$

Further studies will be conducted on the rates of return from investment (R_α), that's why equation (2) can be rewritten as:

$$P(W_\tau \leq R_\alpha) = \alpha \tag{4}$$

Where the rate of return (R_α) is:

$$R_\alpha = \frac{W_\alpha - W_0}{W_0} \tag{5}$$

Combining equation(5) with (3) one finally receives and indicator of risk which is dependent on the rate of return from investment:

$$VaR = -R_\alpha W_0 \tag{6}$$

Presented relation (eq. 6) is directly used in one of the VaR methods, i.e. variance – covariance approach. One of the basic assumptions of this approach is that the returns have a normal distribution. It should be noted that the rates of return from financial markets are dominated normal distribution with fat tails, as is demonstrated by the positive value of kurtosis. It should be also known that this is visible in return distributions with short time intervals. However, presented assumptions won't play the main role in this paper, as it will focus on the study of the changing dynamics.

Based on the assumption of a normal distribution one can write the equation for percentile rates of return (eq. 5) as a function of two variables, i.e. μ and standard deviation δ (risk):

$$R_\alpha = \mu - k\delta \tag{7}$$

Where:

k – constant depending on the confidence level, e.g. for $1-\alpha = 0.95 \Rightarrow k = 1.65$.

Combining equations (6) and (7) one can get a relation, which show the dependence between basic statistical values and risk for the analyzed data:

$$VaR = (k\delta - \mu)W_0 \tag{8}$$

Considering a portfolio consisting of the N assets it should be remembered that the presented equation (8) will be calculated by using average return of the portfolio μ_N

and the total portfolio risk δ_N. Then the VaR equation for the whole portfolio will take the form:

$$VaR = \left(k\delta_N - \mu_N\right)W_0 \tag{9}$$

Because in this paper authors analyzed the dynamics of the VaR, that's why the equation (9) will be modified to form:

$$VaR = \left(k\delta_N - \mu_N\right) \tag{10}$$

This way of calculating VaR gives the possibility to observe percentage changes of value at risk in the analyzed time period.

2. STUDY OF RISK FOR PORTFOLIO INVESTMENTS

This section will present obtained results from the calculation of VaR as well as observations about the dynamics of the investment portfolio which consists of Polish companies from the oil sector. Choosing this particular sector from the stock exchange authors think about less stabile situation on the world oil market, which is caused by unstable situation in the countries supplying this material (tensions in the middle east). With this fact it can be concluded that the uncertainty can affect the risk related to our investment.

This analysis used most companies from the WIG-fuel index:

- PKNORLEN (fuel),
- PGNIG (gas),
- LOTOS (fuel and oil),
- PETROLINV (natural gas and petroleum),
- CPENERGIA (natural gas).

(the other two companies 'MOL' and 'KOV' weren't noted long enough to be analyzed)

All of companies were examined from 04.01.2010 to 31.12.2010, which gives a time horizon of approximately 250 data. In order to analyze the dynamics, time windows were sequentially moved one day later up till 11.03.2011. This step will result in additional data (describing the dynamics) of value at risk. The algorithm can be described as:

1. Calculate VaR for data points [x, x+n]
2. Calculate VaR for data points [x+1, x+n+1]
3. Repeat until calculating for data points [x+y, x+y+1] where x+y+1 is the length of the series.

2.1. STATISTICAL ANALYSIS OF OIL COMPANIES

Conducted statistical research for the investigated companies has shown interesting results. Basic portfolio values such as total return from portfolio and risk are presented in Table 1 and show the expected lack of market stabilization.

Also it’s worth to noting that in case of two companies (PETROLINV and PGNING) there are negative returns and relatively high risk (especially in PETROLINV case).

Table 1. Statistics for the companies investigated

	Return	Risk	Kurtosis
PNORLEN	0,3532	0,0184	0,82
PGNING	-0,0364	0,0145	-0,07
LOTOS	0,1669	0,0195	1,41
PETROLINV	-1,035	0,055	2,99
CPENERGIA	0,0007	0,0303	0,99

In addition, kurtosis indicated that only in one case – PGNING – distribution of returns can be compared to a normal distribution of returns (the value of kurtosis is roughly equal 0) – this is a basic assumption of VaR. In other cases, one can speak of leptokurtosis (value greater then 0), reflecting that the returns are focused around their mean value but tails of the distribution are much bigger (higher probability).

The next table (2) shows the correlation coefficient assigned to pairs of companies.

Table 2. the correlation coefficient for the companies

	PNORLEN	PGNING	LOTOS	PETROLINV	CPENERGIA
PNORLEN	1	0,193	0,797	-0,618	-0,182
PGNING		1	0,183	0,174	0,368
LOTOS			1	-0,358	-0,012
PETROLINV				1	0,613
CPENERGIA					1

What is surprising is the fact that companies belonging to the same economic sector have such a stretched scale of correlation coefficient values. It’s also unexpected that there is no correlation between LOTOS and PETROLINY although they occupy the same sector of the oil industry. One additional point of interest is the negative correlation between PKNORLEN and PETROLINV.

Next stages of research will optimize the discussed portfolio. The method for optimizing the portfolio was shown in the schematic model of work “Multi-stage model for selection of short-term stock portfolio using the DMA method” (Kiłyk, Wilimowska 2010). Using the formulas presented in that paper, one can create the optimal investment portfolio, which consists of the following companies: LOTOS (11,55%) and PKNORLEN (88,45%).

Such portfolio has a positive return $R_P = 0{,}33\%$, which is surprising considering the constant price fluctuations and a prevailing downward trend. Second important information is low risk value $\delta_P = 0{,}017\%$, which despite the unstable economic situation and pessimistic information coming in from around the world suggests the possibility of a profit from this investments.

Study of the risk dynamics will be conducted for two scenarios with different parameter k. The first scenario will analyze the tolerance level equal to 5% and the second will set it to 1%. Initial values of VaR for the analyzed portfolio shows that in both cases their investments lose about 30% ($VaR_{\alpha=0,05} = -0.303$ and $VaR_{\alpha=0,01} = -0.292$).

2.2. ANALYSIS OF VAR FOR THE TOLERANCE LEVEL EQUAL 0.05 (FIRMST SCENARIO)

In the first scenario, one will see how the dynamic of VaR will change, when the level of tolerance is specified at 5%. It means that investor agrees to incur losses (loss understood as a negative return from investment presented in the form of VaR) with probability of 5%. Analyzing the results of the sliding time window one can see the existence of the downward trend with a slope equal to $a = -0.0016$. This means that no matter how long the time investment will be losses will continue to grow.

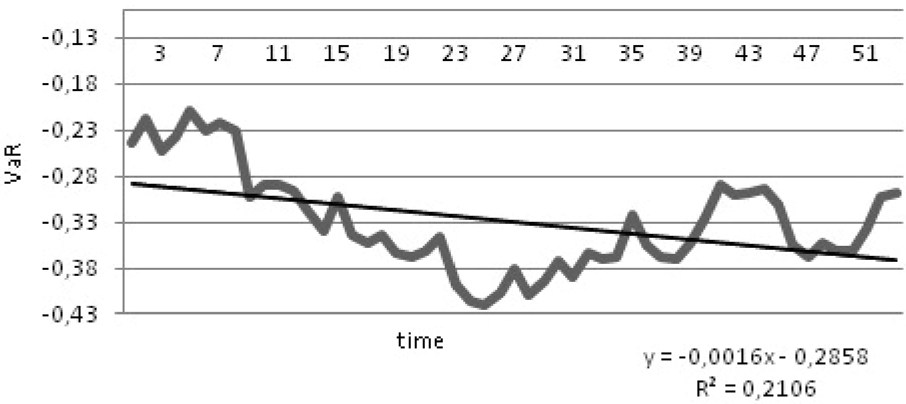


Fig. 1. Changes the VaR values of the fuel portfolio for the first scenario

On the other hand, one should analyze results not with global a approach but with a local approach and break down plot on two smaller interval. The point of breaking on the graph will be at day 24, which may be called the trend breaking point (it change the trend). Obtained statistical values for both parts (presented in Table 3) show that the first part despite a significant standard deviation (large fluctuation) has a slightly smaller loss %. The second interval, as can be seen, it's more stable (also accompanied by an upward trend) that ultimately generates a higher loss.

Table 3. First scenario: VaR statistic for the two time period.

	Mean	Standard deviation
First time period (1–24 day)	-0.303	0.0613
Second time period (25 – 53day)	-0.354	0.0385

The first time period (first 24 days – figure 2), same as in the global approach has a downward trend, but in this case slope is greater ($a = -0.0081$). In addition, it's worth noting the relatively smooth transitions between values and small fluctuations around the trend.

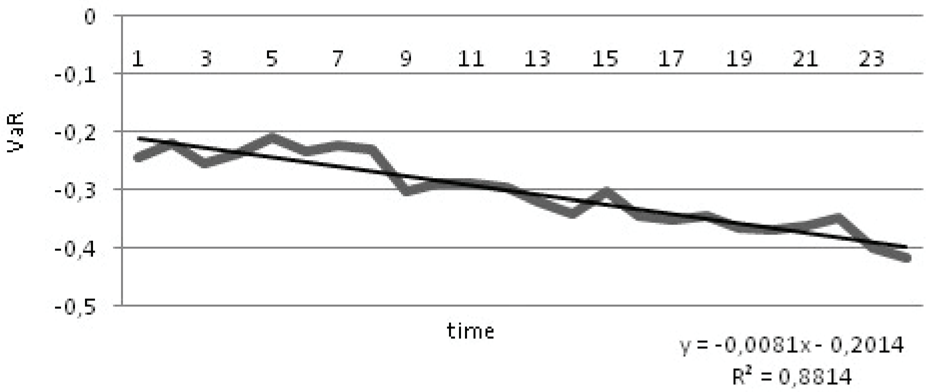


Fig. 2. Changes the VaR values of the fuel portfolio for the first scenario – first 24 days

The second time interval (25–53 days), as can be seen in figure 3, is a fairly strong upward trend (angle $a = 0.0031$). This strong growth is good news for investors and can reduce the risk for this investment.

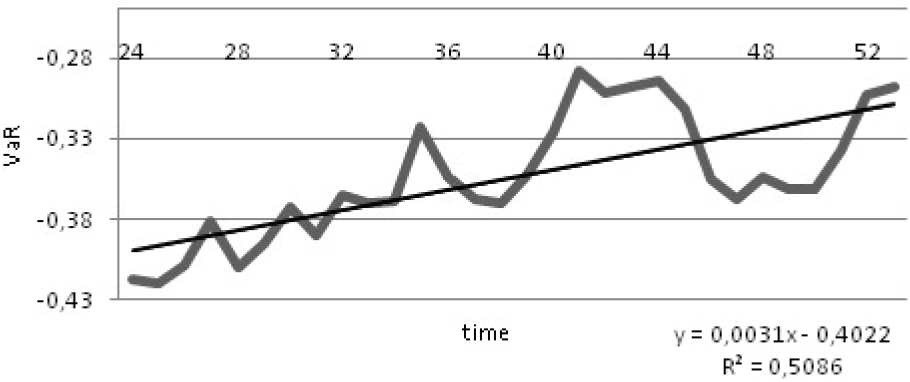


Fig. 3. Changes the VaR values of the fuel portfolio for the first scenario – from 25 to 53 days

2.3. ANALYSIS OF VAR FOR THE TOLERANCE LEVEL EQUAL 0.01 (SECOND SCENARIO)

Similar results were achieved in the second scenario, where the probability of loss on equal to 1% ($k = 2.33$). The global approach (as in first scenario) – figure 4 – can be seen a downward trend with the same angle as in the previous case. However, as has already been show ($VaR_{\alpha=0,01} = -0.292$) the initial value of VaR is lower, and this leads to the fact that following values are relatively smaller (when compared to the first situation with $\alpha = 5\%$).

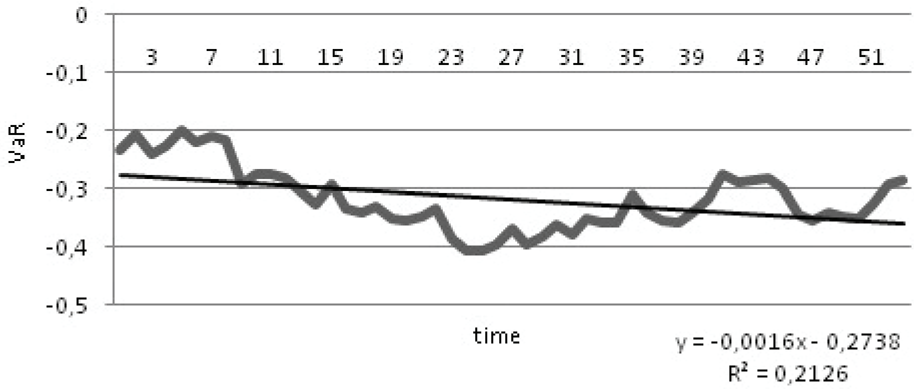


Fig. 4. Changes the VaR values of the fuel portfolio for the second scenario

When turning to the local approach, figure 4 is again divided into two areas where a point of intersection (as before) is on the 24th day. Analyzing the two time windows in statistical terms shows lower values than the same statistics in the previous scenario (practically the same values of the standard deviation for both scenario is a interesting phenomenon).

Table 4. Second scenario: VaR statistic for the two time period

	Mean	Standard deviation
First time period (1-24 day)	-0.291	0.0612
Second time period (25 – 53day)	-0.342	0.0386

Since both scenarios lead to similar graphs (which is caused by the small impact of the k coefficient in the formula 10), trend analysis and accuracy of the linear regression will only be performed for the second scenario. Looking at the graph of the first time interval (Fig. 5) and the parameters describing it (equation trend and parameter R^2) one can notice two things.

Firstly, trend is linear enough to fit nicely with the real data, which confirmed by the high $R^2 = 0.881$ measure. Secondly, looking at the equation of the trend one can observe that the trend's slope is rather small with $a = -0.0081$. The declining trend (described by parameter a) and a fairly strong fit (R^2 value) can be an indicator of the trend's stability.

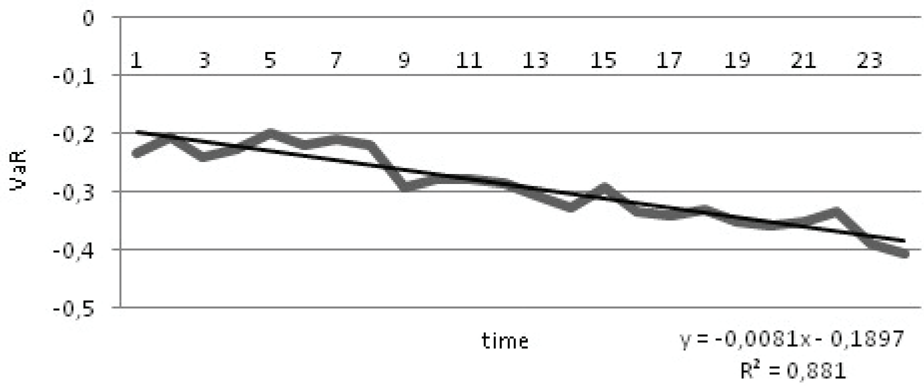


Fig. 5. Changes the VaR values of the fuel portfolio for the second scenario – first 24 days

The second part of the time window, showing an upward trend for the VaR values (Fig. 6), is characterized by larger fluctuation (similar to the first scenario). These variations are reflected in the lower $R^2 = 0.509$. The next parameter, which is the trend slope is small, i.e. $a = 0.0031$. The combining those two pieces of information suggests that in the near future the trend won't change.

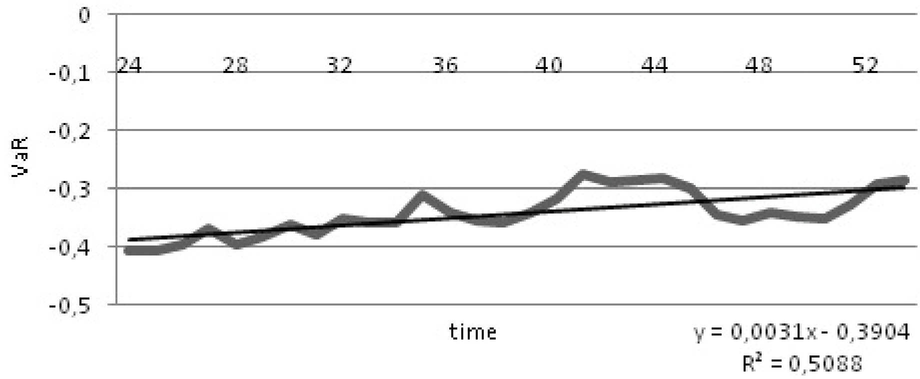


Fig. 6. Changes the VaR values of the fuel portfolio for the second scenario – from 25 to 53 days

3. CONCLSION

The analysis allowed to note that it is possible not only to study changes in the risk for their investments, but also attempt to forecast changes in existing trend. It has been shown that the trend equation and the R^2 measure brings a lot of interesting information.

Research shows large VaR changes for time windows with a length of 53 quotations. These variations confirm the currently unstable situation on the world oil market. Current and upward trend of risk supported by strongly fluctuating VaR signal can be perceived as good sign for the analyzed investment portfolio.

REFERENCES

[1] BEST P., *Wartość narażona na ryzyko, obliczenia i wdrożenia modelu VaR,* Kraków, Oficyna Ekonomiczna, 2000.

[2] JAJUGA K., KUZIAK K., PAPLA D., *Ryzyko rynkowe polskiego rynku akcji – Value AT Risk i inne metody pomiaru. Rynek kapitałowy: skuteczne inwestowanie część 1,* Szczecin, Wydawnictwo Uniwersytetu Szczecińskiego, 2000.

[3] KIŁYK A., WILIMOWSKA Z., *Multi-stage Model for selection of a short-term stock portfolio using the DMA method,* In: Information Systems Architecture and Technology. System Analysis Approach to the Design, Control and Decision Support, Jerzy Świątek (Ed.), Wrocław, Politechnika Wrocławska, 2010, 25–035.

[4] PERA K., *Koncepcja VaR (Value at Risk) w pomiarze ryzyka surowcowego projektu inwestycyjnego,* Gospodarka Surowcami Mineralnymi, tom 24, 2008.

BIBLIOTEKA INFORMATYKI SZKÓŁ WYŻSZYCH

Information Systems Architecture and Technology. Advances in Web-Age Information Systems, pod redakcją Leszka BORZEMSKIEGO, Adama GRZECHA, Jerzego ŚWIĄTKA, Zofii WILIMOWSKIEJ, Wrocław 2009

Information Systems Architecture and Technology. Service Oriented Distributed Systems: Concepts and Infrastructure, pod redakcją Adama GRZECHA, Leszka BORZEMSKIEGO, Jerzego ŚWIĄTKA, Zofii WILIMOWSKIEJ, Wrocław 2009

Information Systems Architecture and Technology. Systems Analysis in Decision Aided Problems, pod redakcją Jerzego ŚWIĄTKA, Leszka BORZEMSKIEGO, Adama GRZECHA, Zofii WILIMOWSKIEJ, Wrocław 2009

Information Systems Architecture and Technology. IT Technologies in Knowledge Oriented Management Process, pod redakcją Zofii WILIMOWSKIEJ, Leszka BORZEMSKIEGO, Adama GRZECHA, Jerzego ŚWIĄTKA, Wrocław 2009

Information Systems Architecture and Technology. New Developments in Web-Age Information Systems, pod redakcją Leszka BORZEMSKIEGO, Adama GRZECHA, Jerzego ŚWIĄTKA, Zofii WILIMOWSKIEJ, Wrocław 2010

Information Systems Architecture and Technology. Networks and Networks Services', pod redakcją Adama GRZECHA, Leszka BORZEMSKIEGO, Jerzego ŚWIĄTKA, Zofii WILIMOWSKIEJ, Wrocław 2010

Information Systems Architecture and Technology. System Analysis Approach to the Design, Control and Decision Support, pod redakcją Jerzego ŚWIĄTKA, Leszka BORZEMSKIEGO, Adama GRZECHA, Zofii WILIMOWSKIEJ, Wrocław 2010

Information Systems Architecture and Technology. IT TModels in Management Process, pod redakcją Zofii WILIMOWSKIEJ, Leszka BORZEMSKIEGO, Adama GRZECHA, Jerzego ŚWIĄTKA, Wrocław 2010

Information Systems Architecture and Technology. Web Information Systems Engineering, Knowledge Discovery and Hybrid Computing, pod redakcją Leszka BORZEMSKIEGO, Adama GRZECHA, Jerzego ŚWIĄTKA, Zofii WILIMOWSKIEJ, Wrocław 2011

Information Systems Architecture and Technology. Service Oriented Networked Systems, pod redakcją Adama GRZECHA, Leszka BORZEMSKIEGO, Jerzego ŚWIĄTKA, Zofii WILIMOWSKIEJ, Wrocław 2011

Information Systems Architecture and Technology. System Analysis Approach to the Design, Control and Decision Support, pod redakcją Jerzego ŚWIĄTKA, Leszka BORZEMSKIEGO, Adama GRZECHA, Zofii WILIMOWSKIEJ, Wrocław 2011

Information Systems Architecture and Technology. Information as the Intangible Assets and Company Value Source, pod redakcją Zofii WILIMOWSKIEJ, Leszka BORZEMSKIEGO, Adama GRZECHA, Jerzego ŚWIĄTKA, Wrocław 2011

Information Systems Architecture and Technology. Web Engineering and High-Performance Computing on Complex Environments, pod redakcją Leszka BORZEMSKIEGO, Adama GRZECHA, Jerzego ŚWIĄTKA, Zofii WILIMOWSKIEJ, Wrocław 2012

Information Systems Architecture and Technology. Networks Design and Analysis, pod redakcją Adama GRZECHA, Leszka BORZEMSKIEGO, Jerzego ŚWIĄTKA, Zofii WILIMOWSKIEJ, Wrocław 2012

Information Systems Architecture and Technology. System Analysis Approach to the Design, Control and Decision Support pod redakcją Jerzego ŚWIĄTKA, Leszka BORZEMSKIEGO, Adama GRZECHA, Zofii WILIMOWSKIEJ, Wrocław 2012

Information Systems Architecture and Technology. The Use of IT Models for Organization Management, pod redakcją Zofii WILIMOWSKIEJ, Leszka BORZEMSKIEGO, Adama GRZECHA, Jerzego ŚWIĄTKA, Wrocław 2012

Wydawnictwa Politechniki Wrocławskiej
są do nabycia w księgarni
plac Grunwaldzki 13, 50-377 Wrocław
budynek D-1 PWr., tel. 71 320 29 35
Prowadzimy sprzedaż wysyłkową
zamawianie.ksiazek@pwr.wroc.pl

ISBN 978-83-7493-802-0