The HUNGARIAN JOURNAL OF INDUSTRY AND CHEMISTRY (HJIC)
formerly (until 2012) the Hungarian Journal of Industrial Chemistry

The HJIC is an international periodical that focuses on results of fundamental and applied research in the fields of

- Biotechnology
- Chemical Engineering Science
- Chemical Processes
- Energetics
- Environmental Chemistry
- Environmental Engineering & Technology
- Industrial Management
- Materials Science
- Mechanical Engineering
- Mechatronics
- Process & Systems Engineering
- Recycling

in the form of original papers, technical reports, reviews, communications, and conference proceedings written in English. HJIC is abstracted by Chemical Abstracts, indexed by the Hungarian Scientific Bibliography Database (MTMT), and archived by the REAL-J Repository of the Hungarian Academy of Sciences.

TABLE OF CONTENTS

EDITORIAL PREFACE

Systems and control theory is a constantly developing scientific area and a dominant driving force behind key industries and engineering fields, e.g. process engineering, automotive engineering, bioengineering, the energy industry, etc. The aim of the current issue (Volume 44, Number 2) is to overview the actual research topics pursued by some selected Ph.D. students.

The papers presented in this thematic issue were selected from contributions at the 14th International Ph.D. Workshop on Systems and Control Engineering held on August 25, 2016. The objective of the Conference was to establish an international forum for young researchers and young engineers from industry and academia. The meeting provided opportunities for the participants to present and discuss the latest results and up-to-date applications in systems and control engineering.

This issue represents the entire spectrum of systems and control engineering as follows:

- process modeling and analysis;
- control (traditional, intelligent, adaptive, etc.);
- system identification and signal processing;
- electrical transmission systems and Smart Grids;
- bioengineering;
- traffic control;
- reaction kinetic networks;
- artificial intelligence;
- soft computing (neural, genetic, fuzzy algorithms, etc.);
- software issues (parallel computing, distributed and network computing, data; visualization);
- decision making (decision support, data mining);
- applications.

The organizers are grateful to the authors for their contributions.

The tradition of the International Ph.D. Workshop on Systems and Control Engineering continues. You are invited to participate at the 15th International Ph.D. Workshop in Veszprém in 2018!

ATTILA MAGYAR

University of Pannonia, Veszprém, HUNGARY

Guest Editor

ROUTE PLANNING BASED ON URBAN MOBILITY MANAGEMENT

Tamás Mátrai,[1]* János Tóth,[1] and Márton Tamás Horváth[2]

[1] Department of Transport Technology and Economics, Budapest University of Technology and Economics, Műegyetem rkp. 3, Budapest, 1111, HUNGARY
[2] Department of Control for Transportation and Vehicle Systems, Budapest University of Technology and Economics, Műegyetem rkp. 3, Budapest, 1111, HUNGARY

Mobility management centres play a significant role in urban transport, taking into account several factors that have an effect on the flow of vehicles. In the present paper a mobility management centre equipped with necessary information and information technologies for travellers is presented that can provide route plans. A route guidance methodology is described that combines current transportation demands with the results of the traditional four-step model.

Keywords: mobility management, route planning, urban mobility, demand management

1. Introduction

The mobility management centre (MMC) as an organization helps to realize and manage mobility demand. The centre controls both private and public transport modes, as well as the provision of information for individual users with the aim of optimizing the entire transport network. The main goal is to provide safe and quick transportation for everyone. During the optimization process the system satisfies the mobility demands with the best available transport modes and routes, the 4R-s principles can be used (Reduce, Retime, Reroute, Revise mode) [1]. *Fig.1* shows the main tasks of the mobility management centre.

Usually in an urban environment, private transport users are supported by the traffic lights from the mobility management centres. The centres can control all traffic lights in the city based on network level principles. Additionally, dynamic parking management can be developed, where the free spaces are displayed on variable message signs and in a further developed version individuals can be guided towards free spaces. Private users can be informed about the actual transport situation and weather conditions *via* various channels. Nowadays the most used channels are Internet-based mobile applications, but there are also radio-based solutions like RDS-TMC services or roadside infrastructure like variable message signs.

If the individual users expect a response and guidance from the MMC for their route choices and do not decide their routes individually, then these centres can optimize the demand on the level of transport network system. In this case they can manage the

demand to avoid congestion – the overuse of certain routes – during peak hours, in other words, they can distribute the demand in terms of a network evenly. Currently these solutions have not yet been developed enough, due to the lack of communication among centres and users. Currently only a very small proportion of vehicles receive dynamic information from the centre. It should be mentioned here that the slow but steady shift towards shared vehicles and on-demand modes has become an emerging trend, especially in conjunction with public transport. Ideally, these services would be integrated together to provide a sustainable and equally convenient alternative to private cars for door-to-door mobility. The realisation of such

Private Transport	Public transport
• Traffic management (traffic lights) • Parking management • Information provision (actual traffic situation, weather) • Route planning • Traffic forecast	• Fleet management • Vehicle priority • Information provision (actual traffic situation, schedules, timetables) • Route planning

Other services
• Park and Ride information • Bike sharing • Car sharing • Demand responsive transport

Figure 1. Tasks of the mobility management centre.

*Correspondence: tamas.matrai@mail.bme.hu

an alternative has set in motion the development of a new mobility concept known as the Mobility-as-a-Service (MaaS), which is a user-centric, digital and intelligent mobility distribution model in which the major transport needs of users are met *via* a single platform and are offered by a single service provider [2]. There is a strong connection between short-term traffic forecasts and route planning. Users would like to receive route suggestions not only based on the current situation but also the forecasted one. Either they are preparing the route on their own device or they are obtaining a plan from the centre. Therefore, an efficient and acceptable level of prediction is essential.

Automatic vehicle location (AVL) is one of the key elements in the accurate running of public transport systems in urban areas. As soon as the dispatchers identify any deviation from the predefined schedule, they initiate certain actions based on a pre-defined rulebook. One action from this rulebook can be to provide priority for those vehicles that are late, this priority provision can be achieved *via* certain junctions which are equipped with the necessary functions. The traffic lights at these junctions can change their signal plan dynamically in order to provide a green signal to the late public transport vehicles.

Other main task of the MMC is the provision of information concerning both static timetables and current traffic situations of public transport, which can be achieved through different channels. Different information is necessary on different channels before a journey commences (e.g. Internet-based route planning services, mobile applications) and during the trip (e.g. at the public transport stop, on the vehicle). The users are expecting to get not only information, but also suggestions especially from the integrated services (e.g. connections, local public transport, railway services, coach services) based on their individual preferences [3]. The most advanced version of this approach is a dynamic multimodal journey planner which provides the ability to adjust to a variety of anomalies (i.e. network disruptions, high capacity usage, etc.) concerning the transport network and evaluates the most efficient routes to get from A to B allowing for efficient traffic management [2]. It has to be mentioned here, that a feedback loop is essential; in order to continuously fine-tune the algorithm in the background, it needs to be determined if the users are really heeding the advice of the route planner.

In functional urban areas, such as agglomerations with high levels of connectivity between urban centres and hinterland, one of the main goals is to promote the shift in mode of the commuters from private cars to more sustainable modes of transport. Therefore, the provision of information with regards to other services such as P+R facilities and their saturation is vital. In this regard, the transfers should be seamless and the information beforehand should be sufficient to help make the right decision. There are different options to ease the level of inconvenience of the transfers, e.g. visual aids of the location of the P+R, information regarding available free spaces, guidance concerning
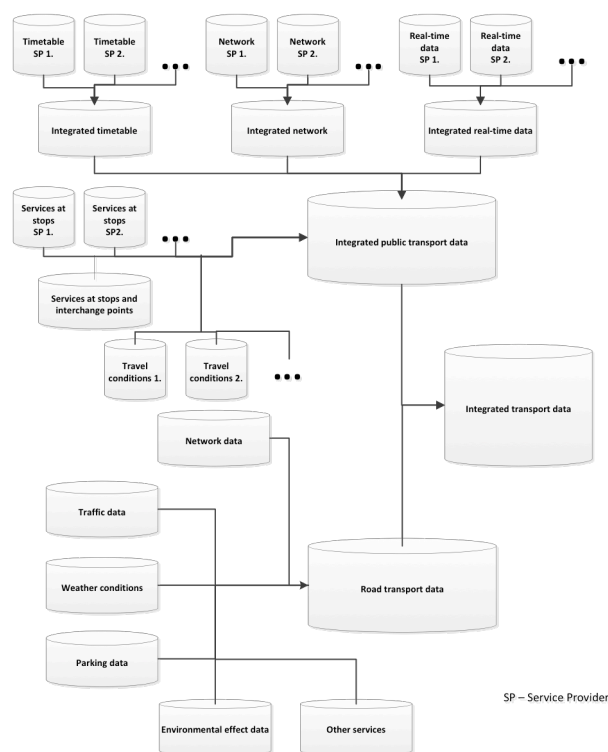
Figure 2. Schematic overview of an integrated transport database [4] (SP = service provider).

public transport services at the parking facilities, and on-the-spot information about public transport schedule.

2. Database of Mobility Management Centres

All the aforementioned tasks are feasible if the MMC is based on complex data collection. From these data, the MMC can create the information necessary to manage mobility demands and give information to travellers. *Fig.2* shows such a possible database.

There are several public transport providers with their own data, which must be integrated into a joint database. Thus, the timetable, network, real-time data, travel conditions and services at stops or interchange points are available. Connecting timetables is particularly important when trips are satisfied by several transport modes. The intermodal hub cannot function without it. The public transport network itself is an integrated network because more public transport companies use the same routes, stops, or stations. Dynamism of data is essential for real-time services (e.g. disturbances on a network). The accurate information is sent according to the current position and status of vehicles. Vehicles equipped with a tracking system are necessary for fleet management. Tariff is an integral part of travel conditions; the passengers must be informed about prices (e.g. is there a transport association selling common tickets) and fares for different passenger groups, etc.

The data concerning fares and travel conditions remain unchanged for long periods of time; they are static or semi-dynamic data. Information systems often
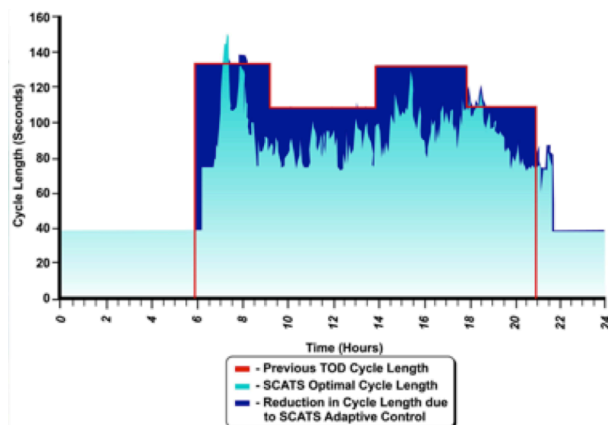
Figure 3. Changing cycle length in the SCATS system (Source: www.scats.com.au).
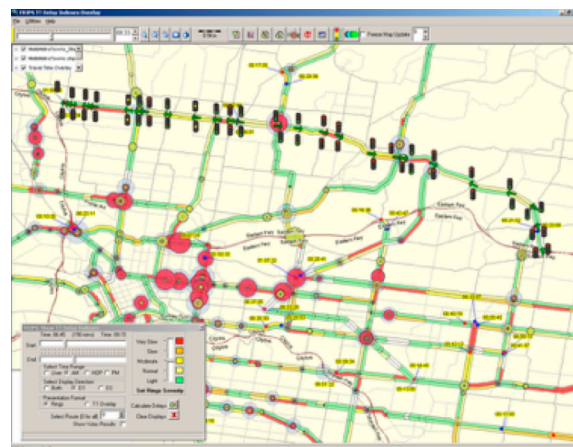
Figure 5. Variable message signs for parking as part of the HelUTC system.

lack such services at stops or stations. It could be helpful for passengers to spend time whilst waiting constructively.

The database of private transport consists of network data, traffic data (presenting the actual situation on the network and also including historical data), weather conditions, parking data, and data regarding environmental effects, and other services. The network data are static, but sometimes include semi-dynamic or dynamic data (e.g. short- or long-term road construction). Traffic data shows accidents, traffic jams, traffic volumes, etc. on the network in real-time and historically. The historical data are very important for traffic forecasting. Numerous methods are based on the comparative examination of present and past situations. Parking data can be static (location and capacity of parking area) or dynamic (parking occupation). Advanced parking systems help to find a free parking place. P+R plays an outstanding role in encouraging drivers to change their modes transport. More and more travellers are sensitive to the environment. They are easily encouraged to use environmentally friendly modes of transport. Therefore, the importance of the data of environmental effects is more accentuated. Data regarding other services includes all the data concerning available services during trips.

2.1. Mobility Management Centre Solutions

According to the needs of modern society, online planning software plays an increasingly important role in optimizing the flow of traffic, and utilizing the capacity of dynamic network traffic models that are strongly connected with real-time traffic management

Figure 4. Estimating the expected travel time using SCATS.

systems. Several models developed for a city may not easily be applied to other cities. The rapid development of IT equipment and the efforts to standardise protocol have facilitated the development of design software for traffic management purposes. The main aim of such software is to support decision made by dispatchers based on actual traffic situations.

2.1.1. Sydney Coordinated Adaptive Traffic System (SCATS)

SCATS is a fully adaptive urban traffic control system that optimises traffic flow. The basic goal is to coordinate traffic lights at junctions in order to improve traffic flow and to reduce travel and waiting times. The main advantage of the system is its ability to adapt to changes in traffic conditions with the modification of the programming of traffic lights.

Fig.3 shows the daily changes of the cycle length of traffic light. The red borderline represents the previous period (before SCATS), the light blue area is the cycle length optimized by SCATS, and the dark blue area is the time saved.

Different travel demands have different goals:
• low traffic demand: minimising the number of stops;
• normal traffic demand: minimising the waiting time;
• heavy traffic: maximising the throughput capacity.

The system can prioritise public transport. In cooperation with TRIPS (traffic modelling) software, traffic forecasting is also feasible (*Fig.4*).

2.1.2. Helsinki – HelUTC, Finland

This mobility management system in Helsinki controls traffic lights, parking and public transport management systems. The concept of controlling traffic lights involves choosing a suitable cycle length according to traffic conditions. There are 10-12 different programmes for each junction. The coordinated intersections have identical cycle lengths. The system can prioritise public transport.

Variable message signs are applied to show parking information. It is very simple: are there any free spaces (*Fig.5*). The public transport management
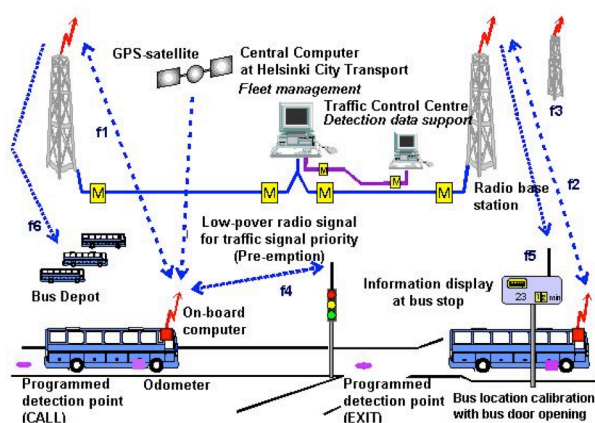
Figure 6. Diagram of the public transport management system in Helsinki.

Figure 7. Overview of the 5T system.

system provides passengers with dynamic information at stops, in vehicles and pre-trips *via* an Internet home page or mobile application. *Fig*.6 shows a diagram for the system.

2.1.3. 5T system in Turin, Italy

The name of the system is "Telematic Technologies for Transport and Traffic in Torino". *Fig*.7 shows an overview of the system. The main goals of the system are to change the travel behaviour of passengers, increase the utilisation of public transport, and optimise travel time, energy consumption, and emissions. Subsystems of 5T consist of traffic control, public transport management including vehicle priority, environmental control, information for travellers (*via* Variable Message Signs, the Internet, mobile applications), and parking management.

Every subsystem has its own management centre, but on the city management level all the information required for coordinating the subsystems is concentrated. The general strategy aims to avoid the overloading of network. The centre consists of seven modules as follows:

- status monitoring module (treating subsystem information);
- origin-destination estimation module;
- modal split module (demand estimation for private and public transport);
- traffic assignment module;
- management module (cooperation with sub-systems);
- operator interface module (creating a connection between the operator and system);
- database module (handling and storing the data of the operating system).

In order to decrease travel time public transport has priority at junctions, and individual drivers are informed about the status of the network and the shortest route for their trip. All the public transport vehicles equipped with AVL (Automatic Vehicle Location) to give real-time information at stops and on vehicles, *via* the Internet and mobile applications.
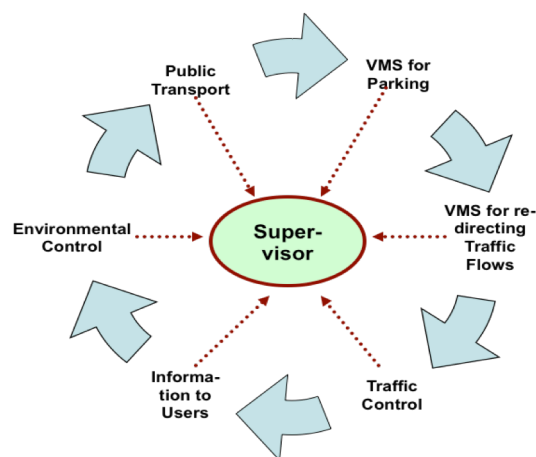
The parking management subsystem uses VMS to show the number of free spaces. The subsystem considers historical data to account for the filling up process of the parking area. The goal is to provide information about the supposed status of the parking lot in the near future, i.e. when the vehicle arrives at the entrance.

2.1.4. Split Cycle Offset Optimisation Technique (SCOOT)

The SCOOT system was developed by the Transport Research Laboratory (TRL) in the UK in order to manage and control traffic lights in cities by taking into account real traffic flows. This system is used in several cities around the globe such as London, Santiago and Beijing. Since SCOOT is designed to update automatically to compensate for the effect of incidents. SCOOT achieved an average reduction of delays of about 25%, when compared with up-to-date fixed-time plans in Nijmegen, The Netherlands [5].

The system also offers additional opportunities:

- providing priority to buses as part of the PT system;
- providing gate effects;
- incident detection;
- online density measurements;
- traffic flow estimation.

All static data, e.g. the physical parameters of the transport network and traffic lights are stored in the SCOOT database.

As an adaptive system, SCOOT requires good quality traffic data, so that it can respond to changes in flow. The system should calculate traffic volumes in every section of the controlled network. Detectors are normally required at every link. Their location is important and they are usually positioned at the upstream end of the approach link. Inductive loops are normally used, but other methods are also available. For the schematics of data collection and processing see *Fig*.8.

When vehicles pass the detector, SCOOT receives the information and converts the data into its internal units in order to construct "Cyclic flow profiles" for
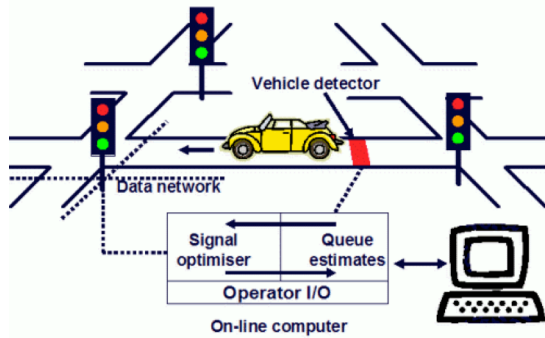
Figure 8. Data collection and analysis in SCOOT [6].

each link. The sample profile shown in the diagram is colour coded green and red according to the state of the traffic lights when vehicles arrive at the stop line at normal cruise speed. Vehicles are modelled down the link at cruise speed and join the back of the queue (if present). Whilst green, vehicles depart from the stop line at the validated saturation flow rate.

The data from the model is then used by SCOOT in three optimisers, which are continuously adapting to three key traffic control parameters - the number of green lights for each approach (Split), the time between adjacent signals (Offset) and the time allowed for all approaches to a signalised intersection (Cycle time). These three optimisers are used to continuously adapt these parameters to all intersections in the SCOOT controlled area, minimising wasted green time at intersections and reducing stops and delays by synchronising adjacent sets of signals. This means that signal timings evolve as the traffic situation changes without the inconvenience of changing fixed time plans of more traditional urban traffic control systems (see *Fig.9*).

The combination of these optimisers provides the necessary solution to deal with the fluctuation in different traffic scenarios, and maintain the stability of different interventions. The Split module is optimizing every status change, Offset module optimizes the duration of every phase of the signalling system, while the Cycle time module optimizes the length of the signal cycle in a certain region of the network.

Providing bus priority happens on three levels, namely (i) equal priority for all buses, (ii) different level of priority for punctual and delayed buses, and (iii) extra priority. Priority provision and intervention are independent from the detection of the buses. In order to detect PT vehicles, the cities usually develop Automatic Vehicle Location systems (AVL), in some cities Selective Vehicle Detection is applied.

Vehicle location monitoring is conducted at the AVL centre. These location data are then incorporated into the selective priority algorithm. As it can be programmed to calculate the length of delays, the vehicle with the longest delay should receive the highest priority. It is possible to provide local priorities under the supervision of SCOOT. A schematic of an example system in London can be seen in *Fig.10* [6].
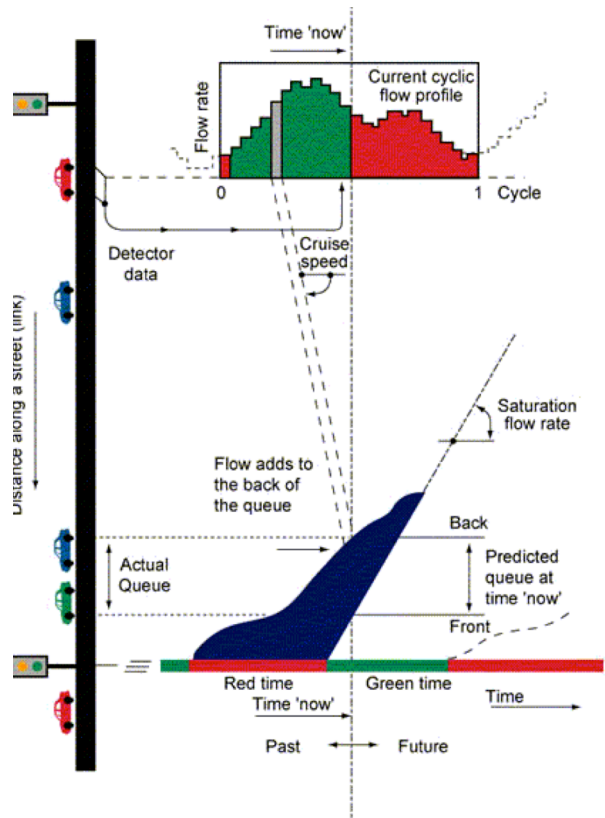
Figure 9. Illustration of the operation of the SCOOT model [6].
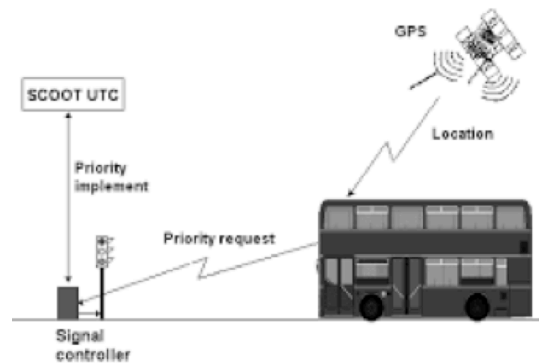
Figure 10. Schematic of the London Bus SCOOT system [6].

2.1.5. Mobility Management in Budapest

In Budapest the following mobility management tasks are in service:
- Public transport management and control (FUTÁR)
  - Fleet management
  - Vehicle priority
  - Provision of information
  - Route planning
- Private transport management and control
  - Traffic management
  - Provision of information
  - Route planning
  - Police monitoring centre
- Other services
  - Bike sharing
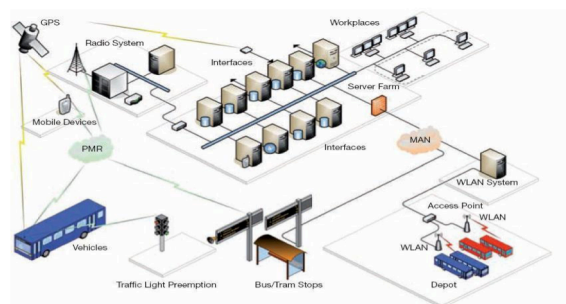
Figure 11. Illustration of the operation of the FUTÁR model.

These tasks are centralized in the same building, but there are three different management systems for the main tasks. They are connected, but there is no common supervisory system. The information flows, although not seamlessly, among them but traffic incidents are managed independently. The aim of project FUTÁR was to completely renew the provision of information and traffic management system at BKV Zrt. This renewal only concentrated on the surface transport network: trams, buses, and trolley buses. The project consisted of four main steps.

The first one was the development of the dispatcher centre and the main central server system. This consists of 32 dispatcher workplaces and all the connecting software and hardware developments. Every dispatcher can view the location of all vehicles in the network in real time and are able to intervene in order to maintain schedules especially when some incidents occur.

The second step was the developments of in-vehicle unit. Budapest has one of the largest surface transport vehicle fleets in Europe. Therefore, equipping these vehicles with the necessary on-board units, communication devices, and vehicle location systems was a long and demanding task. This system has been installed in 2295 vehicles and some of them have also received new displays (2024).

The third step was the deployment of roadside equipment. All over the capital 257 new displays have been deployed. These displays contain information concerning the route number of the next vehicle, while it terminate and the waiting time for this vehicle. Additionally, these displays can provide information remotely through their loudspeakers for visually impaired persons.

The last step was the introduction of the radio system. This is an integrated part of the network, which connects the dispatchers to the drivers. Moreover, the system is capable of providing priority for public transport vehicles at 30 different intersections. *Fig.11* shows a schematic of the FUTÁR system.

Budapest also has a private transport management centre, which in 2012 started to use a new ITraffic version 1.6 developed by Siemens. The new system replaced the old and outdated MIGRA Central system, which was installed in 2002. One of the main reasons for the changes was that support for the operation system required by MIGRA was no longer provided. This meant that the replacement of hardware element

was hardly possible. Of course, not only operating systems, hardware, or communication possibilities but also traffic management systems have evolved over the past 10 years.

The new version of the centre possesses several new features. The communication systems have been replaced with a modern IP-based system: Open Communication Interface for Road Traffic Control Systems (OCIT) and CANTO. This wireless technology has also been integrated into the system (GPRS at intersections), which is able to use any network topology or data transfer media. All the control devices which use the old BEFA 12, 15 or 16 communication systems can be supported by the new system, either as a sub-centre or as an outstation access point.

The software has also been upgraded. It has received a new clear GUI, some statistical analysis tools and further developed MOTION Logistics. Some new variable message signs have also been integrated into this system.

3. A Pilot System in Budapest

The aim of the pilot system presented in this paper is to demonstrate how a Mobility Management Centre equipped with the necessary information and information systems can provide tailor-made multimodal route advice to individual users by not only taking into account the average state of the transport network, but incorporating the current demand, too.

This solution can be more elaborate than already existing private services such as Waze application, for the following reasons:

- this pilot system uses O-D matrices;
- the continuous monitoring of the network is not required; therefore, the number of loop detectors, floating cars or other devices can be reduced;
- it provides real multimodal alternatives before trips in order to choose the best transport mode for the current trip;
- it uses already existing services; therefore, it does not require any major investment;
- it is a module-based system; therefore, easily configurable to match local needs.

The distribution of the demand on the network cannot be estimated without a transport model, which assigns the demand to the available transport network. *Fig.12* shows the proposed concept of this pilot system.

The main concept behind this approach is to use the conventional, static demand matrices for the first iteration of the network load, but later on these matrices will be fine-tuned based on current demands. During the route planning phase, actual road traffic disturbances, e.g. road construction, accident; dynamic information from the AVL system; and the current availability of other services like bikes at the nearest bike sharing station, are taken into account. Based on the combination of this information and predefined user preferences, the system provides a tailor-made route for the user that is compared with at least two alternatives.

The system will use time intervals of 15 minutes to calculate the average situation, based on pre-defined static demand matrices (SDM) for each 15 minute-long time interval of the day, starting from midnight. The current demand, which is the sum of the number of queries concerning a specific route between distinct points of the network, is recorded in an actual demand matrix (ADM). In order to obtain the current state of the network the sum of the SDM and ADM is necessary. However, SDM already indirectly contains the demands of the ADM. However, only as a static estimation from which the real state, represented by the ADM, can differ. Therefore, the ADM cannot be added directly to the SDM, so the system will use a complex method to create a dynamic demand matrix (DDM), which correctly incorporates the actual demand. The static demand matrix is reduced proportionally by the corner sum of the actual demand matrix. Then the ADM and SDM are added together resulting in the new matrix (DDM) for the next 15 minute period. Route suggestions will be provided based on the current state of the transport network.

4. Discussion

As described in the Methodology chapter, the concept of the pilot project has already been elaborated. The existing matrix of the Budapest Macroscopic Transport Model [5] will be used as a daily static matrix and create matrices for each 15 minute interval of the day. Demand matrices will be created for every mode individually (4 freight categories, car, bike, public transport). However, only car and bike matrices will be updated with the current demand.

The averaging method described in the previous chapter is valid until the sum of the actual demand matrix corner sum reaches 30% of the static demand. Over this scale, the averaging might yield false results and a different method will need to be applied. An important assumption is that the sum of the corner of the average daily matrix is fixed, in other words, each day the same number of trips is conducted on the network.

A public transport model will not be developed in this pilot case, since an online journey planner exists which is based on real-time data and can make queries. Alternatively, the existing interface of the FUTÁR system will be used to provide public transport route suggestions. For car and bike users, a search for the shortest route will be based on the results of traffic assignments obtained by VISUM software. Disturbance to the road network will be added as changes in link or node capacity occur. In this case, a simple approach can be formulated, but in the end it represents the real disruption to the network.

A description of the entire procedure is shown in *Fig.13*. The entire procedure consists of two main phases: the preparation and elaboration periods. The preparation period can be done any time before the system starts. The elaboration period can be divided into 15 minute intervals that are represented in the figure by
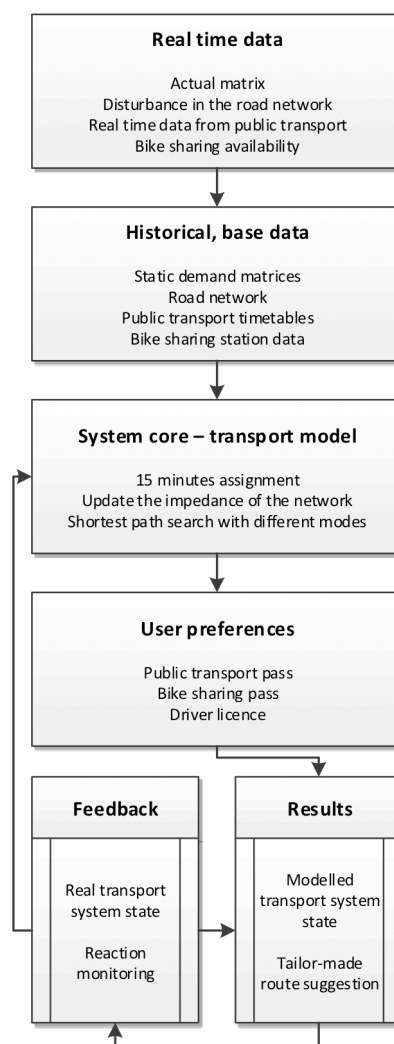
Figure 12. Concept overview of the pilot system introduced in Budapest.

Time Periods of between 1 and T. Time Period 0 is an initialising period that differs slightly from the others, since there are no measurements beforehand. The indexing of time periods and time are the following: τ represents the time when a routing request arrives in the system, $(t-1)$ and (t) are the start and end times of T, T is the time period when τ occurs. During the preparation period the 15 minute-long matrices are created for each mode and the basic static equilibrium assignment STA[PREP] has to be run. These assignments run for roughly 45 minutes for each Time Period on an average PC using VISUM. The results of these assignments are the basic loads of the network.

Time Period 0 is the initial period, there is no routing request that is filled before that period. Therefore, disruptions only exist for the average state. The routing can be drawn from this state and the requests are collected in an actual demand matrix. In this period a traffic assignment STA[0] runs which incorporates the traffic disturbance reported at the beginning (period -1) of the period and the static demand matrix for the next period (period 1). This assignment usually runs for 8-10 minutes on an average PC running VISUM.
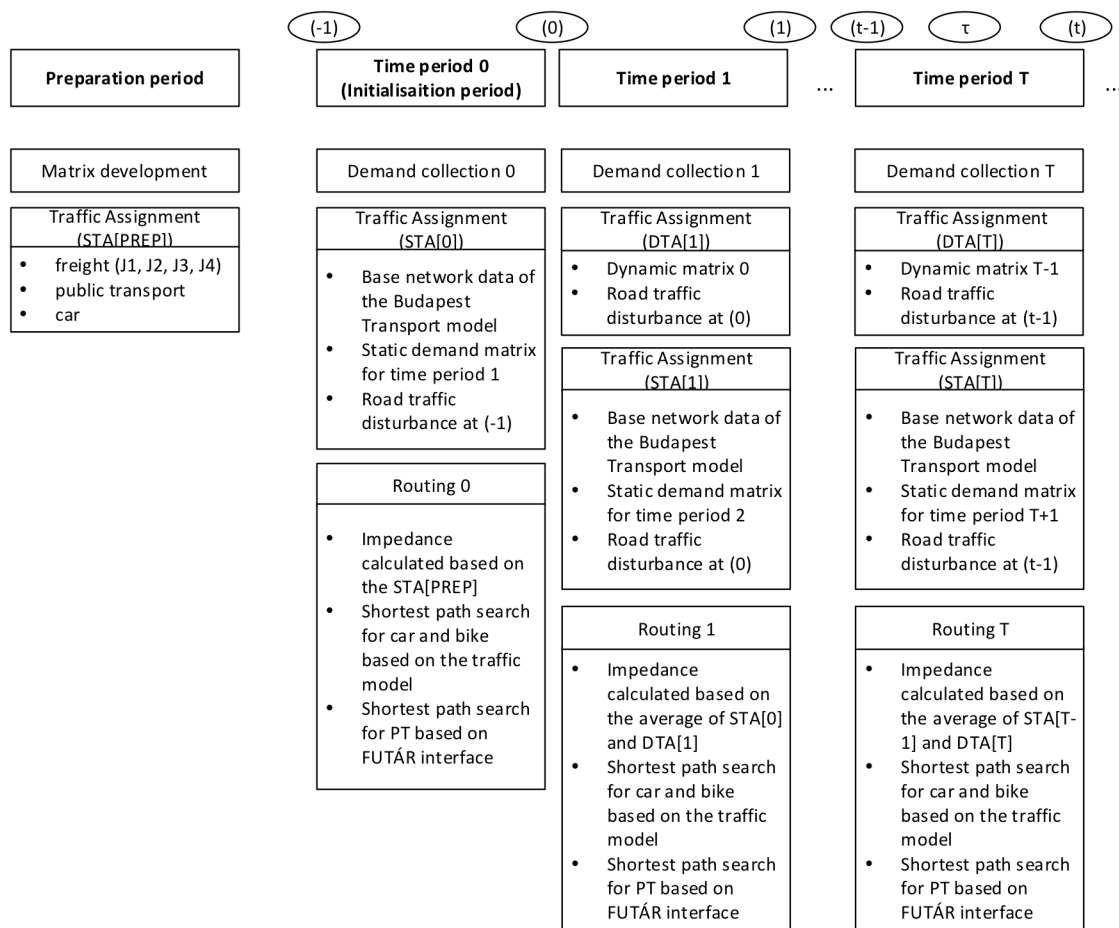
Figure 13. List of processes of the pilot system introduced in Budapest.

Time Period T is the current period, when current routing requests arrive and are collected in an actual demand matrix. An All-or-Nothing assignment DTA[T] is run using the dynamic matrix T-1 and the corrected capacity based on road traffic disturbances in (t-1). This assignment usually runs for less than 5 seconds. STA[T] is the same type of assignment as STA[0], but with the relevant data. The routing algorithm uses an average level of disturbance from STA[T-1] and DTA[T] to determine the shortest route for car and bike users. Time period *T* can represent any subsequent period as well, where the procedure is identical.

5. Conclusion

This paper introduces the concept of mobility management centres. It provides an overview of the tasks, responsibilities and opportunities of mobility management centres together with some international best practices. Later an innovative pilot system is described in detail that will be elaborated in Budapest.

Certainly, as this is a pilot system, it has several limitations and points for further research have been identified. Some of the limitations arise from the required run times of the transport model, e.g. a shorter time interval than 15 minutes cannot be used. From the description of the process, it is evident that it includes averaging in several instances.

This system assumes that all the planned journeys will be made; however, under real circumstances there are more planned than actual journeys. This can be one area of further research, since a good approach can be used to determine the proportion of planned journeys that are actually made. As suggested, a model like this should provide feedback, but this will not be developed during the pilot study.

The validation of the proposed system is important. It can be validated by data from inductive loops and floating cars, but the above mentioned feedback loop would be an asset. There is another validation step, which should be done regarding the capacity reduction of the road traffic disturbance. The current system operates with some educated guesses, but appropriate values can be calculated with actual data. The planned approach for this is to collect incident data from the road network and at the same time collect data from loop detectors in close proximity to the incidents.

The current pilot system is based on a series of static equilibrium assignments, but such a system should use Dynamic User Equilibria or Dynamic Stochastic Assignments. The authors are currently analysing the possibilities of this approach, but it might require greater computing demands as well as more input data, which are currently not available in Budapest. Nevertheless, this is one of the most important research steps.

Acknowledgement

REFERENCES

[1] Juhász, M.; Mátrai, T.; Gál, G.; Kerényi, L.S.: Mobility management survey for the 2012 Olympic games on the basis of published data for "Transport for London", *Közlekedéstudományi Szemle* 2013 **2**, 11–19 (in Hungarian)

[2] Kamargianni, M.; Matyas, M.; Li, W.; Schäfer, A.: Feasibility study for 'Mobility as a Service' concept in London, *UCL Energy Institute, Dept. Transp.* 2015, 1–82 DOI: 10.13140/RG.2.1.3808.1124

[3] Esztergár-Kiss, D.; Csiszár, C.: Evaluation of multimodal journey planners and definition of service levels, *Int. J. Intel. Transp. Syst. Res.* 2015 **13**(3), 154–165 DOI: 10.1007/s13177-014-0093-0

[4] Tóth, J.: ITS international trends, *Proc. XI. Current Issues of Urban Transport: Organisations, Innovation, Quality Service* (Közlekedéstudományi Egyesület, Balatonfenyves, Hungary) 2011 (in Hungarian)

[5] Mátrai, T.; Ábel, M.; Kerényi, L.S.: How can a transport model be integrated to the strategic transport planning approach: A case study from Budapest, *Proc. 2015 Int. Conf. Model. Technol. Intel. Transp. Syst.* pp. 192–199, 2015 DOI: 10.1109/MTITS.2015.7223256

[6] SCOOT Advice Leaflet 1: The "SCOOT" urban traffic control system, available at scoot-utc.com/documents/1_SCOOT-UTC.pdf.

CLOUD-BASED APPLICATION FOR SMART GRID SIMULATION

Szabolcs Tar[1]* and Attila Fodor[2]

[1] Delta-Group Holding Zrt., Szentendrei út 39-53, Budapest, 1033, HUNGARY
[2] Department of Electrical Engineering and Information Systems, University of Pannonia,
   Egyetem út 10, Veszprém, 8200, HUNGARY

This study provides an overview of the content of an R&D project for increasing the efficiency of smart grids to analyse the energy market, consumption, generation, and renewable energy usage. The new energy participants, for example decentralized generation, dynamic consumption, and weather-dependent power plants, are facing challenges. This project offers solutions to these challenges: modelling of distribution networks, forecasts of renewable energy sources distributed generation (RES-DG), micro-grid management and responses to demand in new, complex solutions. A hardware device is being developed for smart grids and security modules to enhance IT security.

Keywords: renewable energy, smart grid, demand response, cyber-security, network assets

1. Introduction

The main objective of our project is to utilize the opportunities in smart grids based on smart grid data and remote control to improve energy and cost efficiency. The service package is composed in a way that our solutions are put into use either as a whole or for the fulfilment of subtasks. The services and hardware devices are being developed like one intelligent cloud-based service.

The rationale of these solutions is supported by several aspects of the energy market. The vastly proliferating renewable energy sources distributed generation producers (RES-DG) are affecting every player in the energy market. The number and built-in power of such power stations are expected to exhibit a rising tendency in the near future according to various energy tenders, see *Tables 1* and *2*.

2. Experimental Data

Grid operators should find a solution for operating grids with more entry points replacing the ones utilizing one directional power flow, where the direction of power flow changes over time and is influenced mainly by the performance of weather-dependent power plants.

Fig.1 shows the electricity consumption of Hungary, which displays both domestic generation and importation, as well as planned and actual values of power transmissions from or to neighbouring countries. The prime cost of the electricity and the volume of the consumption set the actual, base, and peak prices of

Table 1. Number of (pc.) and power generation (kW/MW) of RES-DG producers over recent years in Hungary [1].

	2008		2014		2015*	
	pc.	kW	pc.	MW	pc.	MW
Solar	107	363	8.8k	68	11k	84
Wind	1	10	56	0.5	56	0.5
Hydro	1	16	3	0.06	3	0.06

* based on estimated data

Table 2. Use of renewable energy in Hungary [2].

	Available capacity in mid. 2014 (MW)	Electricity consumption plan for 2020 (GWhr)
Hydroelectric	66	238
under 1MW	6	12
1 – 10MW	22	67
above 10MW	39	158
Geothermal	57	410
Solar	63	81
Wind	750	1545
Biomass	600	3324
Steady	500	2688
Biogas	100	636
Total	1537	5597
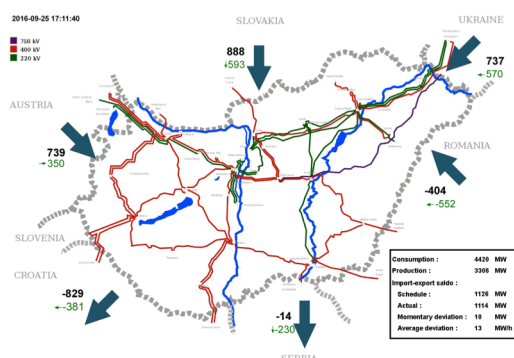
Figure 1. Electricity consumption, export, and import of Hungary as of September 25, 2016 (at 5 p.m.) [1].

*Correspondence: Szabolcs.Tar@Delta.hu
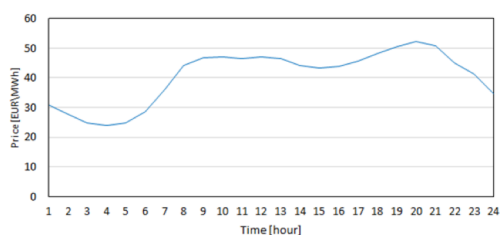
Figure 2. Hourly average price of electrical energy in Hungary (2015).

Figure 3. Consumption on the electricity network of consumers only (power *versus* number of 15 minute sample times).

Figure 4. Consumption of the electricity network of consumers (power *versus* 15 minute long sample times).
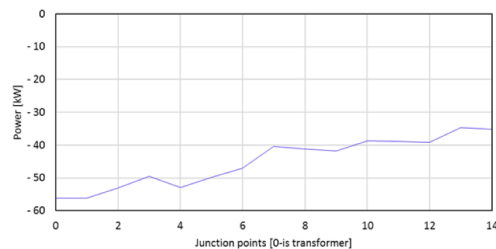
Figure 5. Consumption on the electricity networof - consumers (power *versus* 15 minute long sample times).

electrical energy. *Fig.2* shows the hourly average price of the electrical energy in 2015 in Hungary. As solar and wind renewable energy sources are somewhat uncontrollable, the integration of renewable sources is a big challenge for transmission operators.

As a solution, our group implemented a calculation model of the grid that takes into account parameters of the grid elements and details of network measurements. Details about grids are provided by smart measurement and points of consumption defined in the remote reading circuit. Measurement data are generated over periods of at least 15 minute in length at these measurement points.

3. Simulations

The spread of decentralised energy producers based on *Table 1*, cause time-dependent power flows in grid sections. *Figs.3-5* show the load of a radial grid examined at various times. The lines on the left side of the graphs start from the contact bar of a transformer, the specific points show the power of grid lead-downs. Positive values indicate a flow of power from a transformer towards an endpoint of the grid, while negative values illustrate a flow of power from an endpoint of the grid towards a transformer.

4. Results and Discussion

It is reasonable according to the graphs that the built-in power is insufficient but the trends of consumption and production should be taken into account and a mathematical model created to know the exact operating states as well as for the integration requirement of the new production unit.

In order to be able to model the integration of new RES-DG efficiently and safely over a specific section,

new forecast services need to be impremented. Our service concept is mainly based on mass useage of solar cells in accordance with *Table 1*. External environmental aspects were taken into account that affect the operation of renewable energy source power stations such as clouds, the temperature, humidity, global isolation as well as characteristics of power stations, such as orientation, tilt angle, integrated power, type and efficiency. The sources of data are defined, e.g. online satellite images and analytics, which in a complex fashion, can be used to create forecasts based on factual measurement data.

For microgrid operators, e.g. low- and medium-voltage operators, private grid operators, industrial parks, multiple household buildings, balancing groups, etc., a complex energy service concept was outlined. in which an opportunity for intervention was provided to facilitate a more efficient operation examining the production and consumption habits involved in the given microgrid. If the green energy produced is to be used as widely as possible, it should be identified where and to what extent the available green energy was consumed. By utilizing storage capacities, energy market transactions should be performed that are regarded as new nowadays, e.g. electrical energy is sold and thermal and hydrogen energy sources, vehicle charging, or other services purchased. Our group has proposed a support application service for this. Furthermore, settlement and cost allocation becomes more efficient and clearer based on details for operators. Meters and sensors as well as intervention devices allow microgrid operators to modify the image of a grid proactively based on up-to-date information reducing service outage, which is highly relevant not only for consumers but also producers and trade licence holders too, as an outage location impacting both consumption plan and production plans. An outage location leads to

unplanned resource withdrawal due to the spread of RES-DG power stations. Microgrid providers can initiate interventions to influence the expected production and consumption, providing basic parameters of energy usage. This basic signal can be a plan or request to balance out deviation in the plan. The controller is a decision logics based on the deviation of measurements and basic signals that issues a command within the available limits, which is based on a complex logics by transferring this to the energy use and consumption process in order to achieve the desired result.

The consumer and equipment of the consumer are parts of the energy-use and -production processes (hereinafter referred as to *process*), respectively. This equipment is connected to smart grids among others *via* smart home devices. To allow the consumer to interact, a service shall be developed that is users demand. When intervention in the process does not take place automatically, the consumer is provided with detailed and aggregated information via various solutions in which the result of his/her intervention is displayed as well as the decrease or rise in energy use or costs.

The means of automatic intervention or interaction are dynamic tariffs, which also require smart grids that provide detailed and instantly available data. This allows for the service provider supplying dynamic tariffs to make an offer that reaches the user in a timely manner, which will be either reviewed or accepted by the user or just received as information, as it is possible to intervene remotely and automatically in processes. The goal of our service is to make this information available on a platform based on priorities, irrespective of service providers, ensuring the opportunity to change of energy provider and the access previous data. We provide information to consumers as well as producers by connecting the initiator of an intervention securely *via* processes about which information, states, and backtesting are also generated by us to providers.

To ensure the overall security of smart grids and services included in smart grids, the situation of cyber security and its expectations shall be mapped. Smart grids are regarded as an extended electronic information system, whose endpoints can be accessed and managed remotely. However, they are less verified and have partial physically protection. The main requirements considered are

- Act. CXII of 2011 on the right of information self-determination and information liberty;
- Act. CLXVI of 2012 on the identification, designation and protection of essential systems and facilities;
- Act. L of 2013 on the electronic information security of state- and local government-owned organisations;
- Directive 95/46/EC of the European Parliament and of the Council
- Relevant provisions of energy legislation
- Information methodologies and recommendations

A hardware development has been implemented to enhance information security and a universal communication device has likewise been developed in the framework of hardware development offering a solution that facilitates efficient communication between various devices of smart grids.

5. Conclusion

These services and hardware devices reduce the costs for the participants in the energy sector from generation to consumption via distribution networks, and provide more secure ITC infrastructure. To manage and forecast generation and consumption, new implementations of RES-DG can be easily installed in the current distribution network, without the need for extra investment. The service package is flexible. Anyone can choose whether to use services in the future or logged date. The system will be extended to include other new services.

Such services generate data requirements. Further complex analytics provide useful results for the value chain actors, e.g. to improve the reliability of the distribution network. A source of irregular consumption is unchecked loads on the network, which leads to the overload of built-in elements of infrastructure, premature aging and damage. The simulation engine analyses the data and generate alerts concerning the status of network elements. Before a failure occurs, the system can take various actions to reduce its duration and damage caused in order to expand cloud-based services for industrial consumers to dynamically control consumption for the network balancing. As far as the value chain is concerned, actors provide further bonuses to reduce energy costs.

The devices and services in smart grids, like in telecommunications or IT infrastructure create an opportunity for the further integration of sensors and surveillance camera networks. This integration into the management systems of cities is currently being tested.

Acknowledgement

REFERENCES

[1] Hungarian energy and public utility regulatory authority (MEKH), 2015 www.mekh.hu/nem-engedelykoteles-kiseromuvek-es-haztartasi-meretu-kiseromuvek-adatai-2008-2015-villamos-energia

[2] Hungarian Independent Transmission Operator Company Ltd. (MAVIR) analysis of the Hungarian electricity network, 2015 www.mavir.hu/documents/10258/15461/Forr%C3%A1selemz%C3%A9s_2015.pdf

[3] Hungarian Power Exchange Ltd. (HUPX): www.hupx.hu

SIMULATION OF ELECTRICAL GRID WITH OMNET++ OPEN SOURCE DISCRETE EVENT SYSTEM SIMULATOR

Milán Sőrés and Attila Fodor*

Department of Electrical Engineering and Information Systems, University of Pannonia,
Egyetem út 10., Veszprém, 8200, HUNGARY

The simulation of electrical networks is very important before development and servicing of electrical networks and grids can occur. There are software that can simulate the behaviour of electrical grids under different operating conditions, but these simulation environments cannot be used in a single cloud-based project, because they are not GNU-licensed software products. In this paper, an integrated framework was proposed that models and simulates communication networks. The design and operation of the simulation environment are investigated and a model of electrical components is proposed. After simulation, the simulation results were compared to manual computed results.

Keywords: energy transmission, electrical networks simulation, distributed energy systems

1. Introduction

The simulation of electrical networks is important before network planning, development, servicing, etc. is conducted. There are many planning and simulation software solutions on the market, which can simulate electrical networks and grids, e.g. MATLAB, EPLAN, WSCAD.

The problem with commercial simulation software is that such software packages cannot be ported to a new system. For example, if the development of a cloud-based electrical network and grid simulation system is required, standard simulation software products and methods cannot be used. To solve this problem, a simulation engine was used, which possesses a GNU licence. The OMNeT++ Discrete Event Simulator (DES) [1] was chosen. Mets *et alia* [2] have previously used the OMNeT++ simulator environment.

OMNeT++ is an extensible, modular, component-based C++ simulation library and framework, primarily for building network simulators. OMNeT++ is not supported directly by the simulation of an electrical grid. To solve this incompetency, a model of the most important electrical components was constructed and as a result, the general simulation engine can be used for cloud-based electrical grid simulation. Two methods were implemented in our OMNeT++ software, which investigated their mathematical foundations, as well.

2. Simulated Network

According to load distribution the electrical distribution system can be classified as:

- Fed at one end with one load at the other end;
- Fed at one end with more loads;
- Fed at both ends;
- Radial;
- Ring;
- The distribution system.

To test the OMNeT++ and model of the developed components, a network architecture was chosen, which is convenient and compatible with normal methods for planning of electrical networks. A system has been simulated, which is fed at both ends and consists of three loads (*Fig.1*).

Of course, this OMNeT++ simulation project can simulate any distribution systems, though exact values will only be calculated in this case. To simplify our calculations, the wire parameters shown in *Table 1* were

Table 1. Wire parameters

ϱ	$1.85 \cdot 10^{-8}$ Ωm
q	1 mm^2
l_1	27.027 m
l_2	54.027 m
l_3	27.027 m
l_4	27.027 m
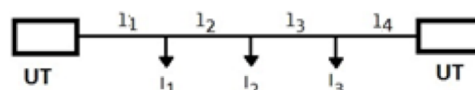
Figure 1. Distribution system fed at both ends with three loads.

*Correspondence: fodor.attila@virt.uni-pannon.hu

considered.

The resistances of all sections of wire are computable using *Eq.(1)*. The calculated resistances of the sections of wire are $R_1 = R_3 = R_4 = 1\ \Omega$ and $R_2 = 2\ \Omega$, which were used in the OMNeT++ simulation.

$$R = 2\,\varrho\,l\,/\,q \tag{1}$$

The selected resistance values are far from the resistance of wire used in real distribution systems, but the calculations are simplified and the illustration of the results more obvious. The values of current loads for each load were $I_1 = 10$ A, $I_2 = 16$ A, and $I_3 = 5$ A.

2.1. Classical Method of Calculating the Voltage Drop

First, the voltage drop of a distribution system was simulated, in order to calculate the voltage drops using the classical method. Our system consists of three loads and it is fed from both ends.

The current of the first and second feeding points should be I_I and I_{II}, respectively. By applying Kirchhoff's Law, *Eq.(2)* is defined as:

$$I_I + I_{II} = I_1 + I_2 + I_3\ . \tag{2}$$

The total length of the wire is calculated as:

$$\Sigma\,l = l_1 + l_2 + l_3 + l_4\ . \tag{3}$$

If I_I and I_{II} are known and the load currents are subtracted from one of them, a load that is fed by both ends is identified [4, 5]. Afterwards, the electrical network can be separated to obtain two networks fed at one end. The method of calculating, for example, I_I is as follows:

$$I_I\,\Sigma\,l = I_3\,l_1 + I_2\,(l_2 + l_3) + I_1\,(l_4 + l_3 + l_2)\ . \tag{4}$$

Using *Eqs.(2)* and *(3)*, the results of the calculations of the total current, currents of feeding points, and the length of the wire are $I_1 + I_2 + I_3 = 31$ A, $I_I = 15.4$ A and $I_{II} = 15.6$ A, and $\Sigma\,l = 135$ m, respectively. From I_I and I_{II}, it can be determined that the voltage drops accotding to Ohm's Law. The voltages of the loads are $U_1 = 214.6$ V, $U_2 = 203.8$ V, and $U_3 = 214.4$ V.

2.2. A Method of Calculating the Voltage Drop Based on the Node-Potentials

The previously presented method can be easily used to calculate networks consisting of topology fed at one end as well as at both ends. However, our electrical grids are obviously not that simple, see *Fig.2* or they can even be more complex. Furthermore, in this kind of method implemented using OMNeT++ the presence of a small solar plant on a rooftop is hard to handle.
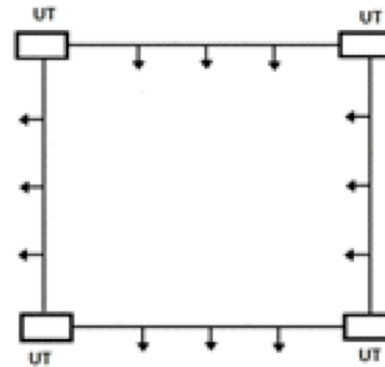
Figure 2. A more complex electrical network than shown in *Fig.1*.
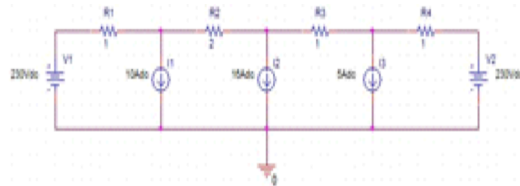
Figure 3. The network fed at both ends modelled with electrical elements.
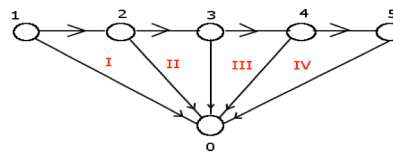
Figure 4. Directed graph of the modelled circuit.

Another method of calculating the voltage drops and currents was identified. Using the node-potential method, any parameter of an electrical network can be calculated. To apply such a method, the feeding points with voltage sources, the loads with current sources, and the wires with resistances were modelled (*Fig.3*).

With this method the feeding points are modelled with voltage sources exhibiting constant voltages, which results in a crucial consequence for more complex networks. If another feeding point is added to the system, some part of it or even the whole network will be parallel to the new feeding point, as it is directly connected to the ground. That would make the investigation of the system and the handling of complex grids easier. Of course this is only a theoretical method with many limitations and conditions, but it can be a good basis on which to start our investigation, plus the method can be developed. Later new elements, both linear and non-linear, can be added to the network. Although the focus of this paper was linear time-invariant systems.

The electrical circuit can be transformed into a directed graph, where the direction of the edges is the same as the direction of the current in calculations. In this case our network possesses six potentials. Two potentials of the feeding points, three potentials of the current sources and the ground potential (*Fig. 4*).

The index of the nodes is identical to the index of potentials. If we apply Kirchhoff's First Law to all nodes, six equations and in ordinary cases five unknown variable potentials from 1 to 5 (the ground potential is 0

V) will results so the linear equation system can be solved. They can be arranged them into a vector $\boldsymbol{\Phi}$.

$$\boldsymbol{\Phi} = \begin{bmatrix} \Phi_0 \\ \Phi_1 \\ \vdots \\ \Phi_5 \end{bmatrix} \quad (5)$$

From the potential, voltages can be calculated from *Eq.(6)*.

$$U = \Phi_i - \Phi_{i+1} \quad (6)$$

Obviously the voltages can be arranged into a vector $\mathbf{U}$, similarly to vector $\boldsymbol{\Phi}$. From the resistances, a resistance matrix $\mathbf{R}$ or conductance matrix $\mathbf{G}$ can be can created. For the node-potential method, a special matrix was used. Once again, the central concept is Kirchhoff's First Law. The first row of the matrix contains the conductances associated with the first node. The direction of the voltage (likewise the direction of the current) will determine their sign.

$$\mathbf{G} = \begin{bmatrix} -G_1 & G_2 & 0 & \dots & 0 \\ G_1 & -(G_1 + G_2) & G_2 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & -G_4 \end{bmatrix} \quad (7)$$

The potential vector can be reduced by omitting Φ_0 thus yielding a reduced potential vector $\boldsymbol{\Phi_r}$. Of course the current matrix labeled $\mathbf{I}$ can be used.

$$\mathbf{I} = \begin{bmatrix} I_I \\ I_1 \\ \vdots \\ I_{II} \end{bmatrix} \quad (8)$$

Ohm's Law helps to calculate the unknown values of the network based on *Eq.(9)*.

$$\mathbf{I} = \mathbf{G} \times \boldsymbol{\Phi}_r \quad (9)$$

In the method of node-potentials, the potentials are considered to be unknown, while the other parameters are given. As a result, matrix $\boldsymbol{\Phi_r}$ contains all parameters that should be calculated. In the present situation the two potentials of the feeding points are considered to be known, 230 V and the ground potential is 0 V. Each current value of the loads is known, but the currents of the feeding points are unknown. Therefore, the matrices should be modified so that all unknown variables will be present in one matrix, $\mathbf{X}$, and all known parameters in another one, $\mathbf{C}$, a constant matrix. The $\mathbf{G}$ matrix must be used as well and denoted by $\mathbf{G_m}$.

$$\boldsymbol{X} = \begin{bmatrix} \Phi_2 \\ \Phi_3 \\ \Phi_4 \\ I_I \\ I_{II} \end{bmatrix} \quad (10)$$

$$\boldsymbol{C} = \begin{bmatrix} G_1 \Phi_1 \\ -G_1 \Phi_1 + I_1 \\ I_2 \\ -G_4 \Phi_5 + I_3 \\ G_4 \Phi_5 \end{bmatrix} \quad (11)$$

$$\boldsymbol{G_m} =$$
$$\begin{bmatrix} G_1 & 0 & 0 & 1 & 0 \\ -(G_1 + G_2) & G_2 & 0 & 0 & 0 \\ G_2 & -(G_2 + G_3) & G_3 & 0 & 0 \\ 0 & G_3 & -(G_3 + G_4) & 0 & 0 \\ 0 & 0 & G_4 & 0 & 1 \end{bmatrix} \quad (12)$$

By arranging the three matrices into one equation, we get

$$\mathbf{C} = \mathbf{G}_m \times \mathbf{X} \quad (13)$$

Eq.(13) was solved using Gauss-Jordan elimination method and the same results presented earlier for $I_1 + I_2 + I_3$, I_I and I_{II}, Σl, U_1, U_2, and U_3 were obtained. To make use of the Gauss-Jordan elimination method, the modified conductance matrix with the constant matrix had to be extended. With the help of a newly created matrix $\mathbf{G_e}$, the elimination process yielded the values for vector $\mathbf{X}$ directly.

The extended conductance matrix ($\mathbf{G_e}$) is as follows:

$$\begin{bmatrix} G_1 & 0 & 0 & 1 & 0 & G_1 \Phi_1 \\ -(G_1 + G_2) & G_2 & 0 & 0 & 0 & -G_1 \Phi_1 + I_1 \\ G_2 & -(G_2 + G_3) & G_3 & 0 & 0 & I_2 \\ 0 & G_3 & -(G_3 + G_4) & 0 & 0 & -G_4 \Phi_5 + I_3 \\ 0 & 0 & G_4 & 0 & 1 & G_4 \Phi_5 \end{bmatrix}$$
$$(14)$$

2.3. Comparison of the Two Methods

Although both calculations yield exactly the same results, in our opinion the second one is preferred. The structure of the matrices indicates that it can be applied to even more complex grids, e.g. in radial topology. Another advantage of the node-potential-based method is that the direction of the load currents is irrelevant. For example, even a small solar plant on a rooftop can be simulated.

3. Simulation with OMNeT++

The OMNeT++ 4.x Integrated Development Environment is based on the Eclipse platform, which has been extended with new editors, views, wizards, and additional functionality. Although OMNeT++ is not a network simulator in itself, it has gained widespread recognition as a network simulation platform in the scientific community as well as in industrial settings, and has built up a large community of users.

The most common area of application of OMNeT++ is the simulation of telecommunication networks. The simulator itself is message-based, so our electrical distribution system had to "communicate" *via*
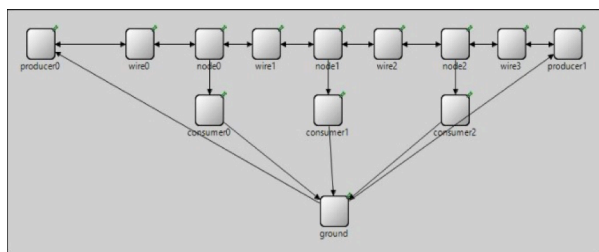
Figure 5. The schematic of the network using OMNET++.

Table 2. Summary of modules used in OMNET++ simulations.

Parts of the system	Module names
Feeding point	Producer
Load	Consumer
Connections	Wire
Node	Node
Ground point	Ground

messages, which is rather unusual in terms of physics or electrical engineering. OMNeT++ provides component architecture for models. Components (modules) are programmed in C++, then assembled into larger components and models using a high-level language (NED).

OMNeT++ simulation comprises three different files, which are (i) Source code (CPP), (ii) Network file (NED), and (iii) Configuration file (INI). In OMNeT++ modules are defined, which can communicate *via* messages. Both modules and messages are special C++ objects. In the source code, we can specialize our modules, and define their tasks precisely. The network file contains the actual topology of the network system, and the position and names of the used modules as well as their connections. With the aid of the configuration file, other network settings can be specified, though none were used in our simulations.

All modules have four functions, which can be defined by users. In our simulation, only two of them were used; the *initialize()* and *handleMessage()* functions. In the initialize function the initial parameters of the modules can be set, in the *handleMessage()* function the action in case of an incoming message can be declared. With these functions, completely different modules can be developed.

The main idea was to show the current or energy flow *via* messages in our system. It worked quite well in the case of a system fed at one end. Although because of the message and handling of messages inclinations of OMNeT++ in systems fed at both ends, the energy flow was not so easy to show.

3.1. Implementing the Distribution System in OMNeT++

The distribution system shown in *Fig.1* had to be implemented. It is obvious that we have at least three different types of modules exist in this network as follows: (i) feeding point, (ii) load, and (iii) connection between them. The implemented electrical distribution network is shown in *Fig.5*.

By analysing the topology more closely, we could identify another module could be identified, connecting two wires and a load, the node. Nodes are important parts not only of grids with their topology, but also of radial grids.

The previously used module is not shown in *Fig.1* but there should be a ground point to make our simulation easier and clearer. This ground point module

is the most important as it carries out the main calculations and connects the loads to the producers. *Table 2* shows the modules and the names used in our source code and files.

3.2. Modules and Messages

The class of messages possesses several variables, including a void-type pointer called *contextPtr*, which is a user-defined pointer. In our simulation, this pointer was used to send data to modules. It points to a class, which contains the used variables such as voltage, current, effective power, reactive power, resistance, etc. Of course, some of the variables are used in only one module (e.g. the resistance in wire module), and some in all of the modules (e.g. voltage, current). In the constructor, all variables were set to zero.

The simulation consists of cycles. Each cycle begins with the producers sending messages, and ends when they receive their messages from ground modules. In simulations, the Event Logging (EV) function can be used to log the parameters or result(s) of the calculations. There are two feeding points in all simulations, though loads are user-defined (N) in pre-processor instruction where both two- and three-point-loaded networks can also be considered.

The first module is called *Producer*. It exhibits a constant voltage value set in pre-processor instruction #define UT 230. In advanced simulations, it should be a user-defined value, and it is not necessarily constant. It could be a function as well, but in this simple simulation, it will remain constant throughout the whole process. At the start of the simulation, only the voltage of the producer is known, but the current and power are unknown parameters until the end of the turn. Thus, these values are set to -1 at the beginning. All three parameters are set in the *initialize()* function. The next stage of the module is handling an incoming message. The incoming message is actually a pointer to the message. At first, in all modules, the values of the pointer were stored as local variables were identified. This is very useful as during simulation some memory allocation problems. The message contains information only about the currents (the voltage remains constant). The current of the producers is set to the value from the message. The power is calculated from the current and voltage values. The module is connected to the ground and one-wire modules. The producer module logs the voltage, current, and power.

The second module simulates the load and referred to as *consumer*. It exhibits a constant current and power factor. Both users are defined. Current values are the same as discussed above in *Section 2*. From these and the voltage from the incoming message, the module calculates the power and reactive power. This block sends the current value to the *ground* block for further calculations. Consumer modules are connected to one node and the ground module, which logs voltage, current, power, reactive power, and the power factor.

In real networks, the connecting wires exhibit resistances as well, causing a voltage drop in the system. In our example, this is desirable, almost 1 Ohm, but our calculations are simple. In the *initialize()* function, the lengths from *Table 1* and R_1 to R_4 resistances from *Section 2* are set. The wire module calculates the voltage drop simply with Ohm's law. Obviously, it shows a useful value from the second cycle, as the current is negative until that turn. Wires are connected to one *producer* and one *node* or between two nodes. Wires log the voltage drop, current, resistance and length.

The Node module is especially useful in radial topology, but also implemented here as well. The aim of this module is to distribute current by applying Kirchhoff's First Law to our network. It works only after the first 'initial stage'. The module is connected to two wires and one consumer module. It shows the value of currents.

The ground module conducts the main calculations of the simulation. It determines the exact current values for each producer module. The calculations are based upon *Eqs.(2)* and *(3)*. The module takes into consideration the length of each wire and the currents from the consumer modules and distributes them. The ground module is connected to all consumer and producer modules in addition to logging the sum of currents and the currents for each producer.

3.3. Simulation

As stated before, this OMNeT++ simulation can be divided into three different stages. At first, the OMNeT++ simulation engines build the network with the user-defined values for each module. This is referred to as stage 0, and there is no logging occurs here. In other simulations, logging is possible here as well, but in this case this opportunity is omitted.

The next stage, or the first cycle, starts when the producers send their first message, and ends when they get their message from the ground module. This part is not necessary in other kinds of simulations. In this part all modules have their own user-defined values, but they do not have any effect on the other modules, e.g. producers exhibit their own voltages, but their current and power are both -1 or the voltage and current of wires are both unknown (-1), as shown in *Figs.6* and *7*.

At this stage, the voltage of the consumers is 230 V, and its powers are calculated with this voltage, shown in *Fig.8*. The first cycle of the simulation is examined in *Fig.9*. Currents are "delivered" between the


```
** Event #1  t=1  enetwork.wire0 (Wire, id=5), on `{}' (cMessage, id=0)
Wire U = -1
Wire I = -1
Wire R = 1
Wire l = 27.027
```

Figure 6. Wire logging during the first cycle.


```
** Event #3  t=2  enetwork.node0 (Node, id=7), on `{}' (cMessage, id=0)
I in = -1 I out = -1
```

Figure 7. Node logging during the first cycle.


```
** Event #10  t=3  enetwork.consumer2 (Consumer, id=13), on `{}' (cMessage, id=13
Consumer voltage: 230
Consumer current: 5
Consumer power: 1115.5
Consumer reactive power: 279.571
Consumer power factor: 0.97
```

Figure 8. Consumer logging during the first cycle.
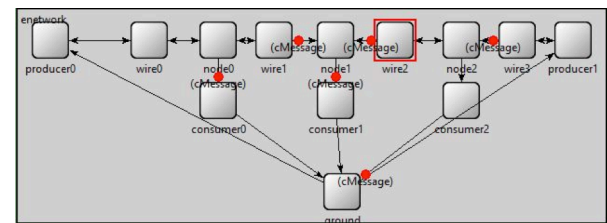
Figure 9. The first cycle of the OMNeT++ simulation.


```
** Event #22  t=23  enetwork.ground (Ground, id=11), on `{}' (cMessage, id=13
I GND = 31
I1 = 15.4002
I2 = 15.5998
```

Figure 10. Current distribution.


```
** Event #23  t=24  enetwork.producer0 (Producer, id=2), on `{}' (cMessage, id=13)
Producer voltage: 230
Producer current: 15.4002
Producer power: 3542.04
```

Figure 11. Voltage, current and power of the first feeding point.


```
** Event #25  t=25  enetwork.wire0 (Wire, id=5), on `{}' (cMessage, id=13)
Wire U = 15.4002
Wire I = 15.4002
Wire R = 1
Wire l = 27.027
** Event #27  t=26  enetwork.node0 (Node, id=7), on `{}' (cMessage, id=13)
I in = 15.4002 I out = 5.40016
** Event #34  t=27  enetwork.consumer2 (Consumer, id=13), on `{}' (cMessage, id=48)
Consumer voltage: 214.4
Consumer current: 5
Consumer power: 1039.84
Consumer reactive power: 260.609
Consumer power factor: 0.97
```

Figure 12. Logging of wire, node, and consumer modules in the second cycle.

modules throughout the whole network system. However, at the end of the initializing stage, the ground module calculates the sum of currents of the consumers and the currents of each feeding point *(Fig.10)*. Thus, the producers receive their currents, and the correct calculated values, e.g. voltages, powers, etc. will be obtained without any unknown parameters *(Fig.11)*.

```
** Event #27  t=26  enetwork.node0 (Node, id=7), on `{}' (cMessage, id=13)
I in = 15.4002 I out = 5.40016

** Event #30  t=26  enetwork.node2 (Node, id=14), on `{}' (cMessage, id=42)
I in = 15.5998 I out = 10.5998
```

Figure 13. Applying Kirchhoff's First Law to the Node modules.

The previously shown modules can be examined and it can be concluded that our voltage and current values are known *(Fig.12)*. It could be interesting to see how our nodes work after the first stage, when applying Kirchhoff's First Law *(Fig.13)*. We have to mention that in the case of node modules I_{out} means the current of the consumer subtracted from I_{in}.

3.4. Validation of the Simulation Results

It has been seen that the results of the OMNeT++ simulation and the results of the classical method of calculating are the same. As an example, the results of the voltage drops are shown in *Fig.14*.

3.5. Topology

All electric systems have unique topologies the position of the components or element(s) and the wires that connect them. By considering a network fed at both ends with only one node, it conatins one node module and two connecting modules between the feeding points and the node. In the case of a similar network with two loads there are two nodes and three connecting channels, as in the previous case. From this point of view the simulated network can be checked, this one contaists of three loads and four channels. If a gred fed at both ends with N loads (where N is a positive integer) exists then the number of wire modules, w (also a positive integer) can be calculated as follows:

$$w = N + 1 \qquad (15)$$

Eq.(15) is only true for this type of topology. For example for a topology in which there is only one feeding point, the number of loads obviously is equal to the number of wires (and nodes). A proven formula can be applied to more complex topology variations. The advantage of these formulae are that they automate the creation of NED files either from another piece of software or implemented from OMNeT++.

4. Summary of Simulation Experiences

It can be concluded that our DES program works properly as shown in *Section 3.4*. Obviously there are some advantages and disadvantages of this method.

4.1. Advantages

First of all it is an open-source platform, thus it can be modified and developed easily. The use of other auxiliary pieces of software can be added, too. As the program includes a graphical viewer, individual
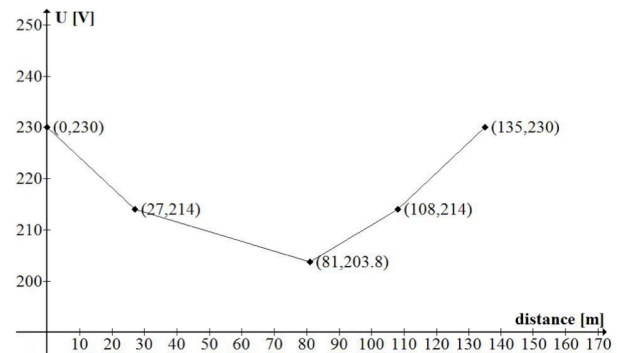
Figure 14. Voltage drop of the loads.

simulation results can be followed. The message direction is the same as the current direction, thus a negative current means failure in the program, or negative values can be defined as unknown parameters. Our system can be observed for values of interest. With this definition the model later may be used in transient analysis to consider the reinitialization cycles each time.

4.2. Disadvantages

On the other hand, as an open-source platform it is still in the development phase. Sometimes the program gave us crash reports during the development phase. The message direction is the same as the current direction; however, according to the message sending system of OMNeT++ extra message(s) can be obtained.

Another issue is the network description file (NED). Each time our network system is modified, the whole code needs to be modified, including the connection definition part, as well as the gate declaration. Topology statements are omitted. Although the time of the simulation is user-defined, all the steps are to be followed the simulation may be rather lengthy. In this case, the simulation of another type of electrical distribution system was attempted, e.g. for a system fed at only one end, a totally new program has to be written. This program can be solved only if a formula is created for each different type of topology. Real electrical grids may be more complex than the situations our DES can handle at the present.

5. Conclusion

The models of the electrical components have been developed in this paper based on engineering principles that are able to describe the behaviour of an electrical grid. It seems that the OMNeT++ discrete event simulator is suitable for the simulation of electrical grids. At first we need to fix some problems mentioned in *Section 4.2*.

A future task would be to implement new electrical network components, e.g. photovoltaic power plants, wind turbines, and different loads, etc. to the simulation environment and subsequently the OMNeT++ would be suitable to simulate smart grid networks.

SYMBOLS

DES	Discrete Event Simulatoion
EV	Event Logger
I_1, I_2, I_3	currents of the loads
I_I, I_{II}	currents of the feeding points
l_1, l_2, l_3, l_4	length of the wires connecting loads and feeding points
N	number of loads
w	number of wires
q	diameter of the wire
U_1, U_2, U_3	voltages of the loads
U_T	voltage of the feeding point
ϱ	resistivity
NED	Network Description File
INI	Initialization file
Φ	potential
$\mathbf{\Phi}$	potential vector
$\mathbf{\Phi_r}$	reduced potential vector
$\mathbf{R}$	resistance matrix
$\mathbf{G}$	conductance matrix
$\mathbf{G_m}$	modified conductance matrix
$\mathbf{G_e}$	extended conductance matrix
$\boldsymbol{I}$	current vector
$\mathbf{X}$	vector of variables
$\mathbf{C}$	vector of constants

Acknowledgement

REFERENCES

[1] OMNeT++ Discrete Event Simulator omnetpp.org
[2] Mets, K.; Verschueren, T.; Develder, C.; Vandoorn, T.L.; Vandevelde, L.: Integrated simulation of power and communication networks for smart grid applications, *Proc. IEEE 16th Int. Workshop, Computer Aided Modeling and Design of Communication Links and Networks (CAMAD),* pp. 61-65, 2011
[3] Niemi, R.; Lund, P.D.: Decentralized electricity system sizing and placement in distribution networks, *Applied Energy* 2010 **87**(6), 1865-1869 DOI: 10.1016/j.apenergy.2009.11.002
[4] Gonen, T.: Electrical power transmission system engineering: analysis and design (CRC Press, Boca Raton, FL USA) 2011
[5] Jamniczky, Á.: Electric engines, (University of Veszprém Press, Veszprém, Hungary) pp. 31–49, 1994 (in Hungarian)

MOBILE DATA SYNCHRONIZATION METHODS

MIKLÓS PÁL AND GÁBOR LÁNER

Capture IT Solutions and Consulting, Záhony u. 7 Building B, Budapest, 1031, HUNGARY

The study introduces and compares the performance of a classical and two innovative mobile data synchronization methods. A customized test environment will be created for every selected method researched. Performance, stability, and other measurement results will be produced from these environments, which will be the major outcome of the study.

Keywords: mobile synchronization, comparison, performance measurement, data optimization, slow network connections

1. Introduction

The main goal of Mobile Workforce Management software is to make fieldwork easier and more efficient. This requires all necessary data to be available and in sync with the server databases. This is why the synchronization module is one of the most important parts of Mobile Workforce Management software. During the fieldwork, there are various data connection conditions that the application needs to be conformed with. There are several data objects, e.g. task, user, client, etc., and multipart objects, e.g. document files, images to be synchronized. The usability of an application highly depends on the efficiency of this synchronization.

The main problem that synchronization needs to solve is to transfer data between the devices and the data source. The solution begins with the data source and through the communication channel ends with the saved data on the device. Three potential implementation methods will be introduced and compared in this article.

2. Experimental

All three considered solutions have different technical backgrounds. The first one is a SOAP (Simple Object Access Protocol) web service implementation based on XML communication. This is the classical way to transfer data between two different platforms. The protocol originally was designed for Microsoft in 1998. The XML-based web service is a widespread solution because it is platform independent and built on industry-wide standards. It was one of the first ways to build service-oriented, modular architectures using smaller

applications is by the communication of web services instead of robust, monolithic systems.

As the technology was mainly used in enterprise environments, and has been in use for more than 15 years, naturally it consists of an antiquated approach compared to modern, lightweight services. While the technology is based on standard HTTP protocols, the requests and responses travel in objects called envelopes. All of the envelopes have a header and a body part, where the body contains the actual payload, the data that is the main reason of communication. The format of the envelopes is strict, and furthermore, there are several encoding, formatting, and parsing standards that have been created since the birth of the protocol. Unfortunately, not all implementations are compatible, it is easy to create a server that cannot digest the client's request, while both endpoints use valid but different soap formats, even though the raw format of the envelope is an easily readable XML file.

The strict format has both benefits and inconveniences. When the service at the server is ready, it is easy to generate a unique description XML, called WSDL (Web Services Description Language). This file not only helps the development of the client side, but there are several tools available that can generate almost all of the client-side code, that can be used to connect to the service. Overall, it is a strict, old fashioned, but really reliable way to approach services in the modern mobile world.

The second tested solution is the RESTful (Representational State Transfer) service in the architecture of microservices. REST is also known as RESTful architectural design, and was represented in 2000 by Roy Thomas Fielding in his dissertation at the University of California, Irvine [1]. REST has become the main architectural design for web and mobile development over the last few years. According to 'ProgrammableWeb', 69% of the newly created APIs were using REST while only 22% were using SOAP in 2014.
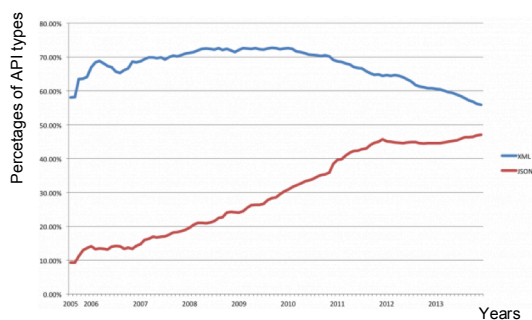
Figure 1. Percentage of APIs added to the ProgrammableWeb directory based on communication types over the years [2].

Table 1. Data structure used in tests with the field data types and typical contents.

Column	Type	Example
ID	numeric	
LA reference	numeric	476483
promoter	character	Vultron
street	character	STONEGATE ROAD
locality	character	MEANWOOD
works type	character	STANDARD
easting	numeric	428804
northing	numeric	437215
location	character	OPP STAINBECK AVENUE VULTRON DUCTING FOR MAINS CABLE TO
description	character	DISPLAY IN BUS SHELTER
works start date	date	19/03/2008
works end date	date	08/06/2013

There are many advantages of this architectural design. The first is its solid performance, due to the high level approach of the solution. It typically communicates over HTTP (Hypertext Transfer Protocol) using HTTP verbs like GET, POST and PUT. REST supports more message formats, e.g. XML, CSV (comma-separated values), JSON (JavaScript Object Notation), etc. The primary communication format is JSON, which is structured text data type. This message format requires significantly less metadata than the XML format, thus greatly increasing the network efficiency as the valuable data can fit in smaller network packages. The spread of the JSON message format is shown in *Fig.1*. The diagram shows the percentage of JSON *versus* XML message formats used by APIs in the 'Programmable Web Directory' between 2005 and 2013 [2]. Another great advantage of RESTful is its simplicity. It is easy to implement and maintain due to its structure. It clearly separates the client and server implementations.

Today's trends point in a direction where developers need to create highly availabile applications exhibiting high level of scalability that are ready to run in cloud environments. Microservice architecture is a method of developing software applications as a suite of independently deployable, small, modular services in which each service runs a unique process and communicates through a well-defined, lightweight mechanism to serve a business goal. The most well-known microservice architecture users are Netflix and Amazon. Applications based on this architecture are easy to understand and modify because of the independent parts. Instead of a robust application which contains all functions, logics and millions lines of code, there are many separated services with a focused function. These applications or application modules are able to run on multiple copies of multiple machines which makes them highly scalable, available and capable of running in cloud environments.

Patterns in programming are reusable solutions to a problem occurring in a particular context. In the world of microservice architecture there are many patterns available to choose from. From the aspect of deployment there are two main patterns:

- *Multiple services per host* - There is one physical server with all services installed on it.

- *Single service per host* - In this case, there is a standalone host for each service. The host could be a virtual machine or a container.

The communication methods between the clients and servers are described with the API gateway communication pattern. In this pattern the gateway is a service discovery between the client and server. This service is the single entry point. From the aspect of the database there are two main patterns. The shared database pattern uses only one database for all services. The database per service pattern uses a standalone database for every single service.

Finally, the third solution is a distributed NoSQL implementation of mobile data synchronization. Early versions of NoSQL databases have existed since the 1960s but the technology started to spread only in the twenty-first century. NoSQL in other words means non-relational database. One of the main benefits of these databases is the simplicity in design, because they store data in a key-value structure. The other main benefit is the horizontal scalability with the support of clustered environments and cloud infrastructure. These types of databases are mainly used in big data environments. All three applications have the same functionalities.

During the experimental three demonstration applications were created. One separated environment for each featured solution. All the applications have three main modules with the same functionalities: (i) database at the backend side to store test data and (ii) synchronization module to transfer data between the backend and mobile application.

Mobile applications exist with the capability to connect to the synchronization module and synchronize data to the mobile device. Additionally, there are some status checking and logging functionalities on this side. Every mobile application was created with its own mobile database to store synchronized records. The data source of solutions was tested, which contains up to 50 thousand historical roadwork items of data from 2011 until 2016. The structure and an example record of the database are shown in *Table 1*. Test data is stored in a single table with a sequence number as a primary key.

The architecture of SOAP web service implementation is shown in *Fig.2*. On the database side
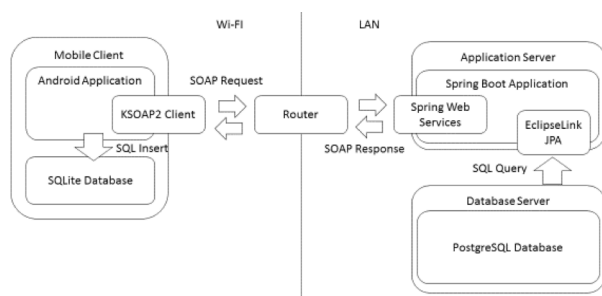
Figure 2. The SOAP based test application's logical architecture. The communication between the client and server component is based on XML / SOAP messages.

Figure 3. The REST based test application's logical architecture. The communication between the client and server component is based on JSON / REST messages.

there is a PostgreSQL server (version: "PostgreSQL 9.0.3, compiled by Visual C++ build 1500, 64-bit") installed. For the experimental, a separated database was created with pg_default tablespace and UTF-8 encoding. Inside the database test, tables were created in a public scheme.

The application server is a J2EE web application implemented with the SpringBoot framework, which builds a standalone runnable jar application that includes a WAR web application and also grants an embedded Apache Tomcat application server. This solution provides a monolithic architecture, which is widely used in the enterprise environment. The core framework of the application server is SpringBoot (version 1.3.6.-RELEASE), where the embedded Tomcat server version is 8.0.36. The web service itself is provided by Spring-WS (version 2.3.0.-RELEASE). The *getRoadworkListRequest* web service provides the main query about synchronization logic. This service performs a *select * from roadworks* query through the persistence layer and returns the whole list of the currently stored roadworks. The response is the XML representation of the data table presented in *Table 1*. As a persistence layer the application uses the EclipselLink JPA provider version 2.5.0.

The mobile application of SOAP implementation was built for the Android SDK version 24.0.0. Android does not offer any built-in library to handle SOAP calls. There are several third-party libraries to fill the gap, but one could not be found that could be a fully satisfactory solution to our problem. In the tests, the kSOAP2 (version 3.6.1) implementation was used that also has some very uncomfortable limitation, but during the test it was working reliably. The synchronized data is saved into the SQLite database in the mobile device. SQLite is the built-in Android database that offers a relational database with functionalities to access and store data.

The architecture of the RESTful service implementation is shown in *Fig.3*. This architecture is a typical microservice architecture where the service itself implements the synchronization functions. The mobile application implements the mobile-device functions like synchronization calls and status reports. A shared database pattern was used by sync service, which means the database used by synchronization service is a database used by other services, too. The same PostgreSQL database was used at the database level in the REST implementation as used before for the SOAP

test application. Thus, the database version and database configuration were the same.

The service is implemented as a standalone Java application supported by Jetty (version 9.2.1.v20140609). Jetty provides a lightweight embeddable web server and it has support for REST APIs of Web Socket. These features make Jetty ideal to use in microservice architecture. Sync Service provides a REST API for mobile clients to conduct synchronization. The main API is the GET /rest/sync/roadworks HTTP/1.1; Content-Type: application/json service that responds with the whole list of roadworks as a JSON content type. The roadworks data is accessed with a full table selected from a database over an EclipseLink persistence layer. The same JPA provider (EclipseLink version 2.5.0) is used for this service as it is for the SOAP application.

The mobile application of the REST implementation was also built for the Android SDK version 24.0.0. However, Android has its own HTTP client provider, in this article Android-async-http (version: 1.4.9) was used for asynchronous HTTP client functionalities at the mobile application level. This is a well-featured and widely used library by top developers like Instagram and Pinterest [3]. The SQLite database was used to store synchronized data in the mobile device as well as in SOAP implementation. It was experienced during development that implementing REST API calls in the mobile environments is relatively easy to perform since it is a widely supported method of communication. Couchbase was used to build the NoSQL database because it offers a complete solution with a server-side database (Couchbase server), synchronization gateway and mobile-side database (Couchbase Lite). The architecture of this solution is shown in *Fig.4*.

The data layer is a Couchbase server (version 4.0.0-4051 Community Edition (build-4051)). The database has a single server node configured. The server node in Couchbase represents an instance of the database. In our test only one instance was run. In production environments, more instances are necessary to improve server availability. The node contains the physical data representation objects, called buckets. A new bucket was configured for the article as a Couchbase bucket type with 200 MB of memory allocated per node. The optimization of disk I/O

operations was set to default, which means the disk I/O priority is low for this bucket. In this article this is an issue, because there are no other buckets in use. The auto-compaction settings are also set to default which means auto-compaction should run if the fragmentation is above 30%.

The Sync Gateway (version 1.2.1 was installed) is located at the server side as a standalone application. This module implements database read / write functions and solution specific APIs to transfer data to and from mobile devices. It has a built-in versioning logic, which adds revision information to the documents stored in the server of the database and handles synchronization metadata like synchronization cycles and user data. The Sync Gateway is configured. It was created to setup the gateway to sync every document type without any user authorization. In this case every connected device synchronizes every document without restrictions between the client and server.

The mobile application of NoSQL implementation is similar to the previous solutions built for Android SDK version 24.0.0. The most significant difference here is the mobile database, which is Couchbase Lite (version 1.3.0). This is a mobile database created for Couchbase and Sync Gateway. It contains the mobile database engine, the mobile database handler and the synchronization interface implementation. The synchronization supports both push and pull requests with version checking so only modified documents are transferred during a call.

2.1. Samples and Measurements

Four types of measurements were performed during the experiment: speed test between server and devices for different numbers of datasets (small 1-10, medium 1,000-10,000 and a large number of records up to 50,000). Speed tests were performed during data transfer to and from the devices and with mixed directions. Speed and stability tests were performed using text and binary data types with a high amount of data to transfer. During the tests, all data packages and sizes of the packages were monitored, as well as the performance of mobile applications, synchronization gateways and databases.

3. Results and Analysis

3.1. Experiments

During the analysis an attempt was made to provide constant conditions with the following hardware infrastructure. An Asus K53S notebook with Intel® Core™ i7-2630QM CPU, 8 GB RAM and HDD WDC WD7500BPVT-80HXZ was used as the server to run the database, application server, synchronization service and synchronization gateway. A Samsung Galaxy SM-G935F (S7 Edge) smart phone with Android 6.0.1 (build number MMB29K.G935FXXS1APG2) was used as the mobile environment.
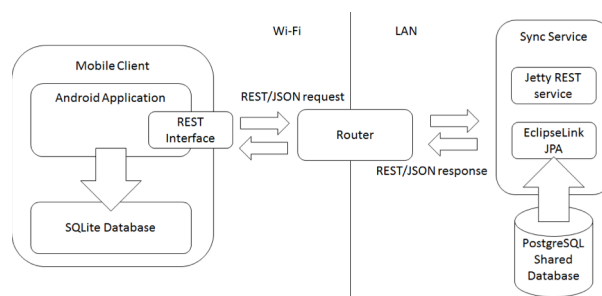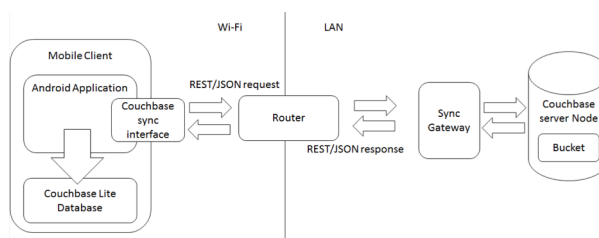

Figure 4. The CouchBase based test application's logical architecture. The communication between the client and server component is based on the database standard synchronization gateway.

Every test was run on the same local network. The network used a 100 Mb/s WIFI router. The server was connected to the router with a local area network (LAN) cable, the mobile device was connected *via* a WIFI network. This way the network speed during the experiment was constant.

During the experiment, six test rounds were run on the three different solutions. The amount of data was raised in every step from 1 row to 50,000 rows. One test round with 100,000 records was also planned, but the tests revealed the limitation of the mobile hardware, for around 30 MB of data, the response could not be parsed in one batch, mainly due to the lack of memory. To achieve realistic conclusions from the measurements, every step was repeated three times. Overall, a total of 54 tests were run.

Measurement results were collected using several methods. In SOAP and REST implementations, most of the information was collected from the mobile platform. Both applications were provided with a logger module that provided log entries in every main step of the synchronization. These steps were the following:

- synchronization initialized
- synchronization started (request was sent from the mobile device to the server)
- synchronization finished (the response came back from the server)
- parse start (when the mobile application started to process the response)
- parse done (when processing finished and all records from the response were saved in the database of the mobile client)

In the case of the NoSQL solution, the monitoring was a bit different because there was no way to write a custom logger module for the built-in processes. Fortunately, the Couchbase Sync Gateway provides a fine-grained log where nearly all equivalent steps can be found that we redefined for the previous tests.

The performance and mobile database monitoring was the other main part of the analysis. This part was the same for all solutions including NoSQL. The performance was monitored continuously with an Android debug tool while the database was monitored from the application with a status screen. It was expected that the SOAP Web Service implementation would be significantly slower than the RESTful and NoSQL solutions mainly because of the larger data packets transferred in XML format.

Figure 5. The synchronization time required for each test application with different record counts, in seconds.

Figure 6. The synchronization time required for each test application with different record counts, in seconds, on logarithmic scale.

Figure 7. The synchronization time required for the SOAP and REST based test applications with different record counts, in seconds, without data parsing and persisting.

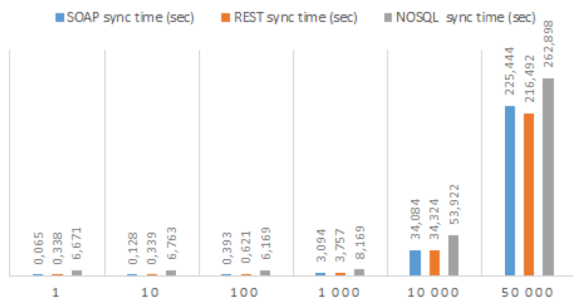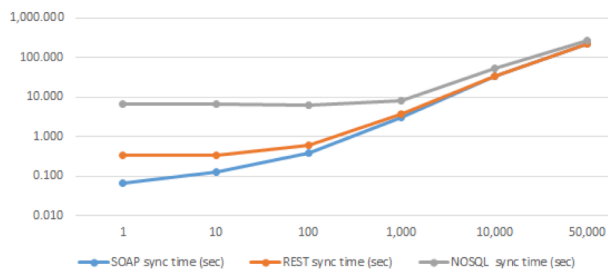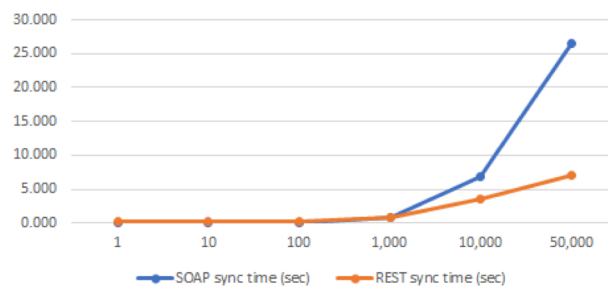The measured results refuted these expectations. As *Fig.5* and *Table 2* show, the average sync times are not just nearly the same but with smaller data amount the SOAP is even faster than the REST. In this result, the NoSQL lags behind the other two solutions, but the sync time gets closer as the amount of data increases. The reason for this difference could be the additional versioning features of Couchbase. Furthermore, NoSQL is built for working with high amounts of data.

The synchronized data amount on the horizontal axis is increasing nearly logarithmically because of this, a logarithmic view of this result set (*Fig.6*) could yield a better understanding. *Fig.6* shows the key point is at 1,000 rows. Here is the point where all solutions start to converge into each other. From this point, increase in the sync time becomes more directly proportional to the increased in the data amount. By taking into consideration the result, numbers and sync time per row values in *Table 3*, is can be seen that the minimum value using SOAP is at 1,000 records, using REST the minimum value is at 10,000 and using NoSQL the minimum is above 50,000. The results above were

Table 2. The synchronization time required for each test application with different record counts, in seconds.

Number of rows	SOAP sync time (sec)	REST sync time (sec)	NoSQL sync time (sec)
1	0.065	0.338	6.671
10	0.128	0.339	6.763
100	0.393	0.621	6.169
1,000	3.094	3.757	8.169
10,000	34.084	34.324	53.922
50,000	225.444	216.492	262.898

Table 3. The time needed to synchronize one record for each test applications in different package sizes.

Number of rows	SOAP sync time (s/record)	REST sync time (s/record)	NoSQL sync time (s/record)
1	0.06533	0.33833	6.67100
10	0.01280	0.03390	0.67633
100	0.00393	0.00621	0.06169
1,000	0.00309	0.00376	0.00817
10,000	0.00341	0.00343	0.00539
50,000	0.00451	0.00433	0.00526

Table 4. The synchronization time required for the SOAP and REST based test applications with different record counts, in seconds, without data parsing and persisting.

Number of rows	SOAP sync time – without parsing (s)	REST sync time – without parsing (s)
1	0.044	0.278
10	0.090	0.239
100	0.134	0.243
1,000	0.815	0.764
10,000	6.842	3.589
50,000	26.540	7.020

Table 5. The size of the data packages using XML and JSON format, in kilobytes for different record counts.

Number of rows	SOAP XML size (KB)	REST JSON size (KB)
1	0.727	0.272
10	6.200	3.700
100	55.400	32.900
1,000	554.400	331.600
10,000	5,734.000	3,481.600
50,000	29,286.000	17,920.000

calculated using synchronization and data processing. If data processing is skipped, the results change as shown in *Table 4*.

Again, the key point here is the limit where the number of records is 1,000. After that point the synchronization using REST increases much faster. As shown in *Fig.7*, the SOAP sync time rises sharply while the REST sync time rises less rapidly. The size of the messages can only be monitored for the SOAP and REST implementations. The results of these measurements were the same as expected. Due to the strict data format, the XML structure requires larger amount of data transfer packages than JSON, as shown in *Table 5*. The logarithmic diagram in *Fig.8* shows that the increase in size is directly proportional to the number of transferred rows.
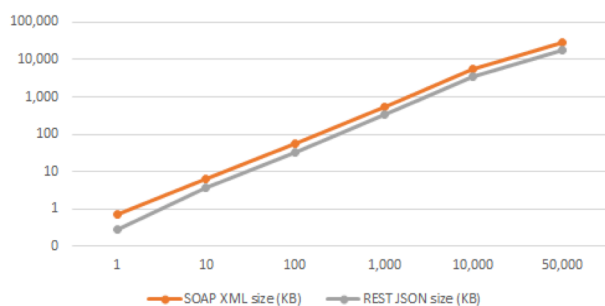
Figure 8. Comparison of the size of the data packages between the XML and JSON format on a logarithmic scale.
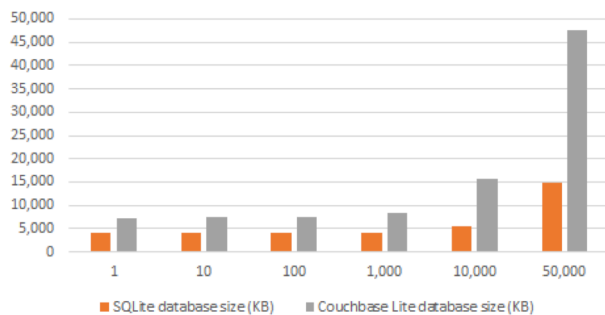
Figure 9. Comparison of the size of the databases between the SQLite and CouchBase databases.

The last measured value is the size of mobile database after data synchronization. The SOAP and REST implementations have the same database size because both of them used an SQLite database with the same data. This is the reason why only SQLite and Couchbase Lite databases were compared in *Table 6*. As shown in *Table 6* and *Fig.9*, the size of the Couchbase Lite database is much bigger than that of SQLite. The difference increases as the amount of data rises.

4. Conclusion

Finally, we need to state that SOAP performed surprisingly well during the experiment performance tests. The biggest limitation with regards to it is the minimal support in mobile development. Because in enterprise companies SOAP is still the most common technique this architecture is still popular.

More measurement data confirmed that there is a common key point where there is only a minimal difference between the selected solutions. This point is around 1,000 records per transaction, which is the point where it does not matter which solution is used. This is

Table 6. The size of the database on the client device using SQLite and CouchBase Lite databases for different record counts.

Number of rows	SQLite database size (KB)	Couchbase Lite database size (KB)
1	3,993.600	7,168.000
10	4,003.840	7,372.800
100	4,044.800	7,536.640
1,000	4,167.680	8,325.120
10,000	5,457.920	15,656.960
50,000	14,704.640	47,462.400

not the optimal point of performance for all solutions but it could be a good compromise.

The best choice is the REST synchronization, if the goal is to quickly implement a customizable, reliable, scalable, and extendable, cloud-ready modern solution. Any amount of data is supported from small datasets to big data solutions. The bottleneck of this solution occurs during data processing.

The best choice is the SOAP synchronization if the goal is to create an enterprise-ready highly secure and auditable solution. This solution is not recommended for big data environments, but up to medium amounts of data, it could offer a real alternative to RESTful service. Working with SOAP has many limitations in mobile development.

NoSQL is the best choice for big data environments where a very large amount of data needs to be processed and there is a limited time for development.

Acknowledgement

REFERENCES

[1] Fielding, R.T.: Architectural styles and the design of network-based software architectures, Ph.D. Dissertation, University of California, Irvine, 2000

[2] DuVander, A.: JSON's eight year convergence with XML, 2013 www.programmableweb.com/news/jsons-eight-year-convergence-xml/2013/12/26

[3] Smith, J.: Android asynchronous HTTP client, 2016 loopj.com/android-async-http

REFRIGERATOR OPTIMAL SCHEDULING TO MINIMISE THE COST OF OPERATION

ROLAND BÁLINT AND ATTILA MAGYAR

Faculty of Information Technology, University of Pannonia, Egyetem u. 10, Veszprém, 8200, HUNGARY

The cost optimal scheduling of a household refrigerator is presented in this work. The fundamental approach is the model predictive control methodology applied to the piecewise affine model of the refrigerator.
The optimisation could not be solved using off-the-shelf tools, e.g. Multi-Parametric Toolbox, so a binary tree-based optimal scheduling algorithm has been developed for this problem.

Keywords: optimal scheduling, model predictive control, operating cost minimization, modelling, day-ahead market

1. Introduction

Environmental awareness is trendy in modern societies. Policy-makers tend to prefer green energy, e.g. solar, wind, etc. in order to decrease carbon footprints. The demand-and-supply-based nature of the electricity market has led to hourly electricity prices in day-ahead markets [1]. Most people seek the cheapest solution by using their appliances during low price periods, however, there are certain devices, e.g. refrigerators work all day long. In this case, the only opportunity lies in the optimal scheduling of the cooling periods. The prescribed inside temperature of the refrigerator makes the problem difficult to solve.

2. Problem Statement

In a day-ahead market, the service provider provides the hourly electricity price for the next day. The problem to be solved is the optimal scheduling of a given refrigerator based on the forthcoming electricity prices. As a constraint, the refrigerator temperature must be kept within given boundary values.

2.1. Energy Price

The service provider provides the electricity price for the next 24 hours (DAM: day-ahead market). *Fig.1* shows the electricity price [2] over a week, each line corresponds to the price of a day.
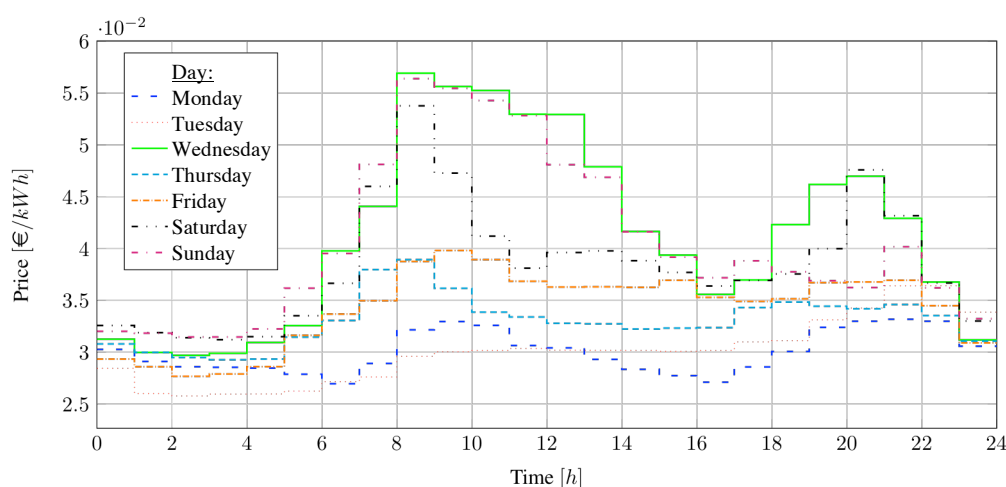

Figure 1. Electricity price of a day-ahead market over a week [2].

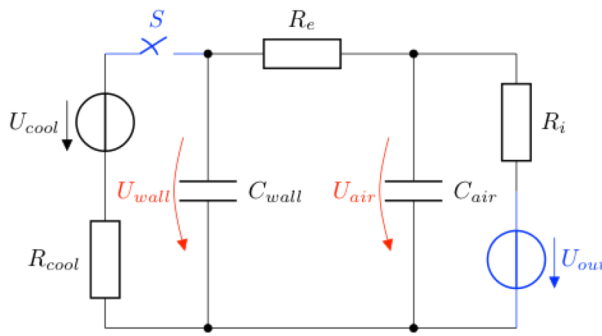*Correspondence: balint.roland@virt.uni-pannon.hu

Figure 2. Electric substitution model of a refrigerator. *S* symbolizes the switch of the devices controlled by a thermostat in an ordinary refrigerator.

2.2. Refrigerator Model

In order to describe the dynamics of the thermal behaviour of the refrigerator, the electric substitution model of *Fig.2* can be used, where the red voltages correspond to the inner air temperature (U_{air}) and the cooled back wall (U_{wall}). The blue elements are the inputs to the system: the outer air temperature (U_{out}) and the on/off switch (*S*). The capacitors demonstrate the heat capacity of the back wall (C_{wall}) and the inner air (C_{air}) while the resistors are the heat transfer coefficients between the components.

2.2.1. Cooling Dynamics

The first case is when the switch *S* is closed, i.e. the refrigerator is cooling. The dynamics of the system can be described by the state-space model of the form outlined in *Eqs.(1)* and *(2)*, where the state and input vectors are given in *Eqs.(3)* and *(4)*.

$$\dot{x} = \mathbf{A}\,x + \mathbf{B}\,u \tag{1}$$

$$y = \mathbf{C}\,x + \mathbf{D}\,u \tag{2}$$

$$x = \begin{bmatrix} T_{air} \\ T_{wall} \end{bmatrix} \tag{3}$$

$$u = \begin{bmatrix} T_{out} \\ S \end{bmatrix} \tag{4}$$

The parameters of the state-space model of the cooling dynamics are given as follows:

$$\mathbf{A}_{on} = \begin{bmatrix} -\dfrac{R_e + R_i}{R_e \cdot R_i \cdot C_{air}} & \dfrac{1}{R_e \cdot C_{air}} \\ \dfrac{1}{R_e \cdot C_{wall}} & -\left(\dfrac{1}{R_e \cdot C_{wall}} + \dfrac{1}{R_{cool} \cdot C_{wall}}\right) \end{bmatrix} \tag{5}$$

$$\mathbf{B}_{on} = \begin{bmatrix} \dfrac{1}{R_i \cdot C_{air}} & 0 \\ 0 & \dfrac{U_{cool}}{R_{cool} \cdot C_{wall}} \end{bmatrix} \tag{6}$$

$$\mathbf{C} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \tag{7}$$

$$\mathbf{D} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} \tag{8}$$

2.2.2. Reheating Dynamics

The second case is when the switch *S* is open, i.e. the refrigerator is reheating to environmental temperature. The system is still linear and time invariant as outlined in *Eqs.(3)* and *(4)*. The parameters of the state-space model are given in *Eqs.(9)-(12)*.

$$\mathbf{A}_{off} = \begin{bmatrix} -\dfrac{R_e + R_i}{R_e \cdot R_i \cdot C_{air}} & \dfrac{1}{R_e \cdot C_{air}} \\ \dfrac{1}{R_e \cdot C_{wall}} & -\dfrac{1}{R_e \cdot C_{wall}} \end{bmatrix} \tag{9}$$

$$\mathbf{B}_{off} = \begin{bmatrix} \dfrac{1}{R_i \cdot C_{air}} & 0 \\ 0 & 0 \end{bmatrix} \tag{10}$$

$$\mathbf{C} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \tag{11}$$

$$\mathbf{D} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} \tag{12}$$

2.3. Control Aim

The aim of this study is to schedule refrigerator's operation to minimise the operational costs with the following assumptions:
- the operational cost is the energy consumption of the refrigerator during the day;
- the price of electricity changes hourly;
- the energy price is known for 24 hours in advance;
- the temperatures must be between the following operating constraints:
  - the inner air temperature should be between 0.1 ºC and 5.5 ºC;
  - the cooled back wall temperature should be between -19 ºC and 7 ºC;
- the outer air temperature is constant;
- the input variable is binary (*On/Off*);
- the models should be chosen based on the value of the input.

The main objective is to minimize the operational costs, defined as

$$cost = \sum_{k=0}^{N-1}(p_k \cdot U_k) \tag{13}$$

where p_k is the price of electricity and U_k is the input at time $T_s \cdot k$ (T_s: sample time).

3. Optimal scheduling

To find the optimal scheduling of a system the model predictive control approach is used. Afterwards a novel heuristic optimisation technique is presented.

3.1. Model Predictive Control of the System

For optimization, the Multi-Parametric Toolbox 3.0 (MPT) [3] is used first, because in this Matlab Toolbox we can find all the necessary options can be found:
- 'Piecewise Affine' (PWA) modelling for linear time invariant (LTI) systems;
- optimal controller design for PWA systems.

The solver uses a cost function as follows:

$$min \sum_{k=0}^{N-1}\left(||\boldsymbol{Q}_{\mathrm{x}}\boldsymbol{x}_{\mathrm{k}}||_{\mathrm{p}} + ||\boldsymbol{Q}_{\mathrm{u}}\boldsymbol{u}_{\mathrm{k}}||_{\mathrm{p}}\right), \qquad (14)$$

where variables $\boldsymbol{x}_{\mathrm{k}}$ and $\boldsymbol{u}_{\mathrm{k}}$ are the state and input vectors at the kth step of the prediction horizon N. The $|| \cdot ||_{\mathrm{p}}$ expression is the standard p vector norm, and $\boldsymbol{Q}_{\mathrm{x}}$ and $\boldsymbol{Q}_{\mathrm{u}}$ are penalty matrices.

The MPT toolbox for the PWA system description uses discrete time models, thus initially the continuous time model must be discretised. However, there are two problems to solve:

- the penalty matrices ($\boldsymbol{Q}_{\mathrm{x}}$ and $\boldsymbol{Q}_{\mathrm{u}}$) cannot be changed, so the actual electricity price cannot be used in this form;
- the time information of the system, e.g. system time, is unknown.

To handle these problems, two supplementary state variables have been introduced. The first counts the time and the second stores the actual energy price. If the refrigerator is switched off the energy price is zero, otherwise it provides the actual price value. One pair of (*On/Off*) systems for every price of each hour needs to be created, during the calculation the model parameters and the price are constants. Thus, the choice is based on the state of the binary input and the time value of the state vector. The modified LTI state-space models are the following:

$$\boldsymbol{x} = \begin{bmatrix} T_{\mathrm{air}} \\ T_{\mathrm{wall}} \\ t \\ p \end{bmatrix} \qquad (15)$$

$$\boldsymbol{u} = [\, S \,] \qquad (16)$$

$$\boldsymbol{\Phi}_{\mathrm{on}} = \begin{bmatrix} 0.9998 & 0.0001 & 0 & 0 \\ 0.0004 & 0.9977 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \qquad (17)$$

$$\boldsymbol{\Gamma}_{\mathrm{on}} = \begin{bmatrix} -0.0024 \\ -0.045 \\ 0 \\ p_{\mathrm{h}} \end{bmatrix} \qquad (18)$$

$$\boldsymbol{\Phi}_{\mathrm{off}} = \begin{bmatrix} 0.9998 & 0.0001 & 0 & 0 \\ 0.0010 & 0.9988 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \qquad (19)$$

$$\boldsymbol{\Gamma}_{\mathrm{off}} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ p_{\mathrm{h}} \end{bmatrix} \qquad (20)$$

$$\boldsymbol{f} = \begin{bmatrix} 0.0022 \\ 0.028 \\ 1 \\ 0 \end{bmatrix} \qquad (21)$$

$$\boldsymbol{C} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \qquad (22)$$

Table 1. The simulation time in minutes *versus* horizon size.

Size of horizon, -	8	10	12	13
Calculation time, min	2	8	59	111

$$\boldsymbol{D} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \qquad (23)$$

The hourly changing energy prices are denoted by p_{h} in matrices $\boldsymbol{\Gamma}$, so there are 24 models with two system dynamics (*On* and *Off*) in total. The model actually used is chosen based on the input (*On* and *Off* states) and the third element of the state vector (*t*: define the model over time so the energy price value). Based on the two variables the necessary model is chosen:

- if the refrigerator operates, its value is the actual energy price;
- if the refrigerator is switched off, it is zero.

By the summation of the values of x_4 the operational cost is obtained (in €/kWhr) for the operating period. The piecewise affine model of the system is given by *Eq.(24)*.

$$\boldsymbol{x}_{\mathrm{k+1}} = \begin{cases} \boldsymbol{\Phi}_{\mathrm{on}}\boldsymbol{x}_{\mathrm{k}} + \boldsymbol{\Gamma}_{\mathrm{on}}\boldsymbol{u}_{\mathrm{k}} + \boldsymbol{f}\,, & \mathrm{if}\ [\boldsymbol{S}] = 1 \\ \boldsymbol{\Phi}_{\mathrm{off}}\boldsymbol{x}_{\mathrm{k}} + \boldsymbol{\Gamma}_{\mathrm{off}}\boldsymbol{u}_{\mathrm{k}} + \boldsymbol{f}\,, & \mathrm{if}\ [\boldsymbol{S}] = 0 \end{cases} \qquad (24)$$

The simulation is very sensitive to some design parameters. These parameters are the sampling time (T_{S}), the prediction horizon size, and the cycle number of the simulation horizon. The correlation between horizon size and simulation time is outlined in *Table 1* ($T_{\mathrm{S}} = 5$ min). If the horizon is bigger than 12 – 15, the simulation will end in deadlock. The sample time value cannot be too large because the change in one step would be too high and the algorithm would not be able to function properly. The multiplication of the sample time and the horizon yields to predicted time horizon. If the sample time and the horizon are small, the algorithm runs quickly, but with a small pre-determined time interval and cannot calculate the forthcoming electricity prices. As N_{sim} increases, the running accelerates. In the PWA system, only the systems of the simulated time interval are present. If N_{sim} is small, the number of models is small so the algorithm must choose from fewer systems but in that case, the cycle has to run repeatedly as the new PWA system is defined.

3.2. Heuristic Search Algorithm using Binary Tree Growing

The general optimizer in the MPT is very slow due to several problem-specific assumptions, e.g. binary input, and the horizon size is bounded so another optimisation algorithm is needed that runs faster and uses problem-specific heuristics.
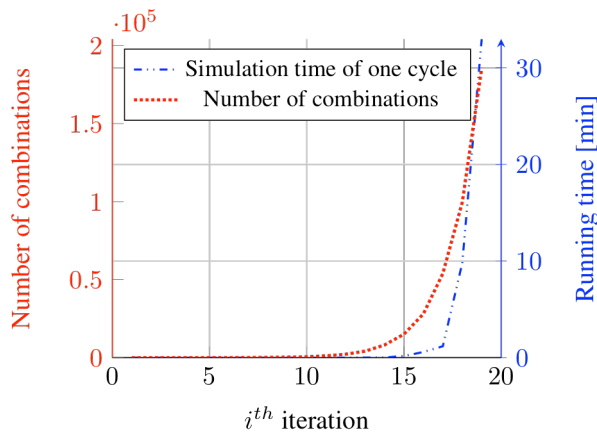
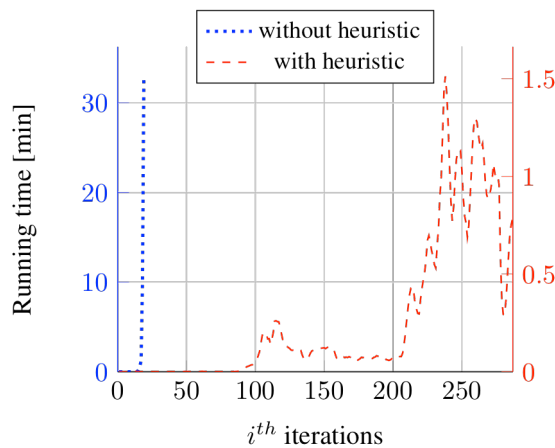*Figure 3*. Running time and number of combinations vs number of cycles.

Figure 4. Without the heuristic algorithm, the simulation time of the 20th cycle is more than 30 minutes, but with the heuristic algorithm, it has a maximum of 1.5 minutes.

3.2.1. Base of Algorithm

In order to find the optimal solution, all input combinations need to be counted over time. Trivially, this number can be reduced by deleting those combinations where the system exceeds the operating constraints. In this case, in all steps, the states are duplicated and 1 or 0 (*On* or *Off*) is added to input vectors so the number of combinations grow exponentially and the simulation time of one cycle grows, too (*Fig*.3). Without additional heuristics, the running time of the 18th loop was as high as the MPT simulation with a horizon size of 13.

Model-related heuristics can be applied to decrease the number of combinations. This rule is simple: between any two input combinations, if the corresponding air and wall temperatures are higher and also the operational cost is greater than or equal to one of the examined combinations then this input combination is not optimal, *i.e.* it can be ignored. The pseudo code of the algorithm is summarized by *Algorithm 1*. With these heuristics, the number of combinations and the running time of the simulation can be reduced. *Fig*.4 shows the running time of the nth cycle with and without this filtering. This figure shows that without the filter, the simulation time of the 20th

cycle is more than 30 minutes but with the filter it is a maximum of 1.5 minutes.

Algorithm 1 Heuristic algorithm

  procedure OPTIMUM SEARCHING
    define:
      $X_0 \leftarrow$ initial state
      bounds $\leftarrow$ bounds of temperatures
      $p_h \leftarrow$ electricity prices of a day
      $h = 1$ (actual hour of day)
      $T_s \leftarrow$ sample time [hour]
      SimTime $\leftarrow$ simulation time
      $cost = 0$
      sys1 $\leftarrow$ reheating DT-LTI system
      sys2 $\leftarrow$ cooling DT-LTI system
      $U_{0,0} = []$
    create:
      $PWA \leftarrow$ PWA system from *sys1* and *sys2*
    loop:
      for $i = 0 : 1 : SimTime/T_s$ **do**
        $U_{i,all} = [U_{i-1,all}; U_{i-1,all}]$ (duplicate U)
        $n = size(U)$ (number of combinations)
        $k \in \{1, 2, ..., n\}$
        if $1 < k <= n/2$ **then**
          $U_{i,k} = [U_{i,k}, 1]$;
        else
          $U_{i,k} = [U_{i,k}, 0]$;
        $X_{i+1,k} = PWA(X_{i,k}, U_{i,k})$
        $cost \leftarrow$ update *cost*
        $k \in \{1, 2, ..., n\}$
        if $X_{i+1,k} \in / bounds$ **then**
          delete $X_{all,k}, X_{all,k}$ and $cost_k$
        $n = size(U)$
        $k,l \in \{1, 2, ..., n\}$
        if $X_{i+1,k} > X_{i+1,l}$ **and** $cost_k > cost_l$ **then**
          delete $X_{all,k}, X_{all,k}, cost_k$
        else
          if $X_{i+1,l} > X_{i+1,k}$ **and** $cost_l > cost_k$ **then**
            delete $X_{all,l}, X_{all,l}, cost_l$
  $n = size(U)$
  $k \in \{1, 2, ..., n\}$
  $opt = min(cost_k)$
  optimal solution $= U_{all,opt}$

3.2.2. Algorithm with Horizon

The algorithm of the MPT uses a horizon to reduce the calculation time for fast online implementation. In the investigated problem, the maximum usable horizon available in MPT is too small. The previous algorithm calculates the optimal solution during a whole day, so the horizon is 24 hour, e.g. if the sample time is 5 minutes then the size of the horizon is 288. This large horizon size causes slow simulations. To decrease this simulation time, a horizon as in the MPT algorithm is defined. This step causes a very large reduction in running time. This time is proportional to the horizon size, but if the horizon is too small the solution would not be optimal. The running times and calculated costs are collected in *Table 2*.

The running times in *Table 2* show that by increasing the horizon size the simulation time increases, too. The operating cost increases, but it is uncertain whether a bigger horizon is better. Based on this, the suggested heuristic search-based optimization algorithm can be used as a basis for a special model predictive scheduling/control algorithm.

4. Results and Discussion

The model predictive control algorithm in the MPT did not work correctly due to the serious limitations connected to the control task. By increasing the horizon size, the simulation becomes slow with maximum horizon size of 13.

The proposed optimization algorithm has been developed to find the optimal scheduling. *Table 2* shows the difference between simulation times and operating costs if the horizon size is changed. *Fig.5* shows the calculated optimal scheduling for Wednesday. *Fig.6* shows the energy prices and the number of input

Table 2. The size of horizon, the simulation time in minutes and the cost of best solution.

Size of horizon	Running time	Cost	Deviation from optimal
none (∞)	75.0	0.20235	0%
24	5.7	0.20427	+0.95%
36	12.6	0.20433	+0.98%
48	16.9	0.20394	+0.79%
60	25.9	0.20413	+0.88%
72	31.0	0.20456	+1.09%

combinations of the ith cycle without and without a horizon size of 4 hours. There is a correlation between the monotonous energy price and the number of input combinations. If the price is increased the number opf combinations increases, too.

Simulations of energy prices for a whole week with and without horizon size of 4 hours ($T_S = 5$ minutes) have also been performed. The results are shown in *Table 3*. The simulation times were reduced by 70-90% while the costs grew only by 1-2%.
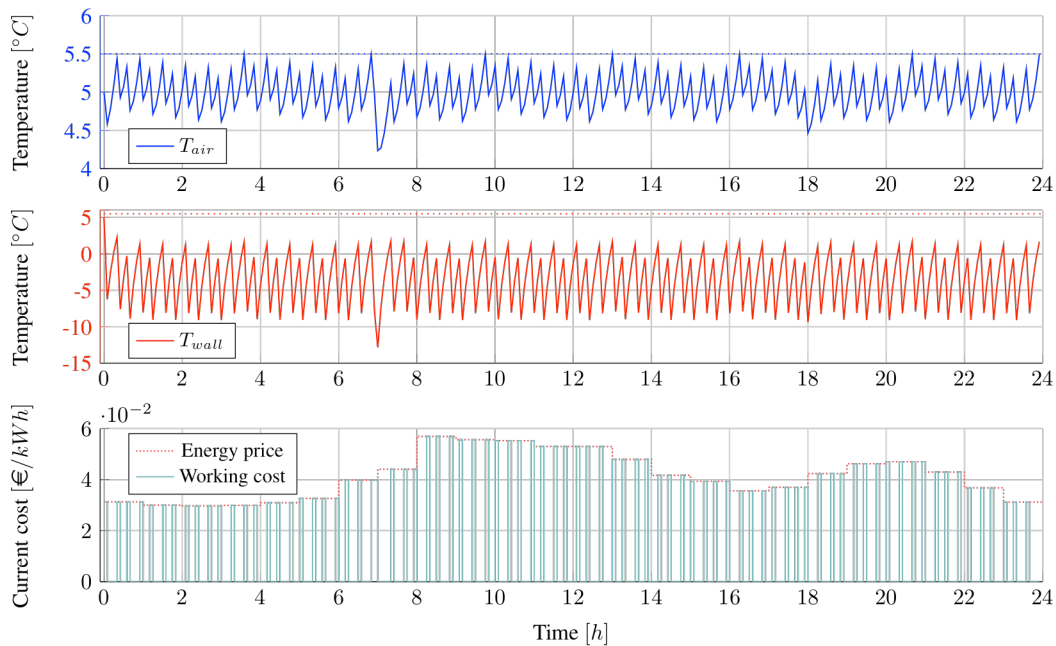
Figure 5. The optimal scheduling for Wednesday's energy prices.
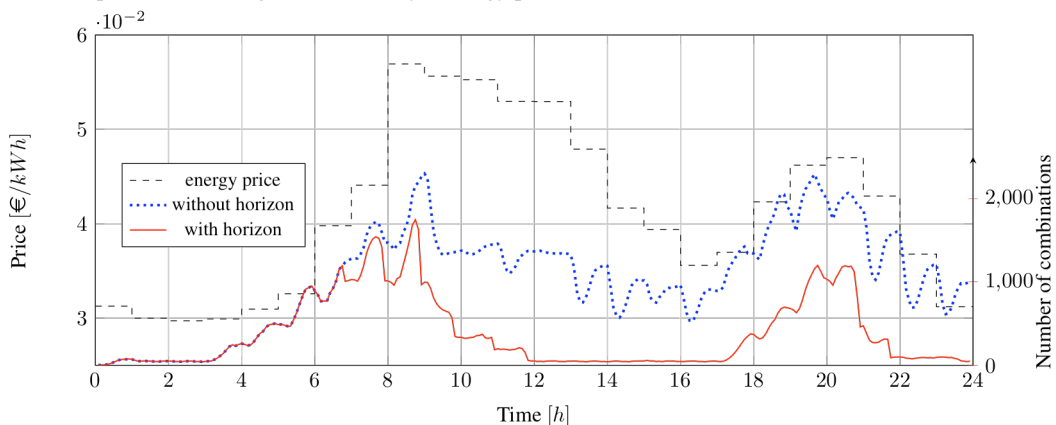
Figure 6. Wednesday energy prices and the number of input combinations of the ith cycle with and without a horizon size of 4 hours.

Table 3. Differences in optimal operating cost and simulation time with and without a horizon.

Day of the week	without horizon		4 hour horizon		difference in percent	
	cost, 1/kW	time, min	cost, 1/kW	time, min	cost, 1/kW	time, min
Monday	0.20235	75	0.20394	16.9	+0.79	-77.47
Tuesday	0.20200	1013.2	0.20650	78.9	+2.23	-92.21
Wednesday	0.27748	234	0.28111	48.8	+1.31	-79.18
Thursday	0.22255	116.1	0.22739	38.3	+2.17	-67.04
Friday	0.23127	267.7	0.23618	32.3	+2.12	-87.93
Saturday	0.25819	222.3	0.26281	45.1	+1.79	-79.70
Sunday	0.27286	100.4	0.27488	18.7	+0.74	-81.34

5. Conclusion

A heuristic search algorithm has been implemented for the optimal scheduling of a refrigerator model. This work was initiated by the fact that the optimization algorithm of the MPT did not work correctly with large horizon sizes for the applied piecewise affine system. The proposed algorithm calculates the optimal input sequence for a whole day with any horizon size but it also works with an infinitely large horizon. The calculation time is fast compared to that of the Matlab MPT solver.

Our own heuristic algorithm is faster without a horizon than the MPT solver with horizon size of 13 hours. I can be run with a larger horizon size, and can yield the absolute optimum without a prediction horizon. The operating cost of the solution is 12 percent less with the aid of a heuristic search. This cost is 25 percent cheaper than the cost of traditional refrigerator control.

The suggested heuristic search-based optimization algorithm can be used as a basis of a special model predictive scheduling/control algorithm. The present work is based on an ideal refrigerator model. As a future step, the validation of the model and the scheduling algorithm in an MPC framework will be investigated. Testing the control scheme in a real situation with random disturbances, e.g. opening the refrigerator's door, etc. will also be performed.

SYMBOLS

A	state matrix
B	input matrix
C	output matrix
C_{air}	heat capacity of the inside of the refrigerator
C_{wall}	heat capacity of the back-wall
DAM	Day-Ahead Market
DT-LTI	Discrete Time Linear Time Invariant
f	constant input or disturbance
Φ_{off}	off state discrete time state matrix
Φ_{on}	on state discrete time state matrix
Γ_{off}	off state discrete time input matrix
Γ_{on}	on state discrete time input matrix
LTI	Linear Time Invariant
MPC	Model Predictive Control
MPT	Multi-Parametric Toolbox

PWA	Piecewise Affine
Q_u	penalty matrix of u vector in MPT
Q_x	penalty matrix of x vector in MPT
R_{cool}	heat resistance of the back-wall
R_e	heat resistance between back-wall and inner air
R_i	insulation of refrigerator
S	binary state switch
T_{air}	inner air temperature of refrigerator
T_{out}	outer room temperature
T_{wall}	cooled back-wall temperature of refrigerator
u	input vector
u_k	input vector value in k^{th} sample moment in DT
U_{air}	voltage corresponding to inner air temperature
U_{cool}	voltage corresponding to minimum reachable back-wall temperature
U_{out}	voltage corresponding to outer air temperature
U_{wall}	voltage corresponding to cooled back-wall temperature
x	state vector
x_k	state vector value in k^{th} sample moment in DT

Acknowledgement

REFERENCES

[1] Huisman, R.; Huurman, C.; Mahieu, R.: Hourly electricity prices in day-ahead markets, *Energy Econ.* 2007 **29**(2), 240–248 DOI 10.1016/j.eneco.2006.08.005

[2] Nord Pool: Elspot prices, www.nordpoolspot.com/Market-data1/Elspot/Area-Prices/ALL1/Hourly, 2016

[3] Herceg, M.; Kvasnica, M.; Jones, C.; Morari, M.: Multi-parametric toolbox 3.0, *Proc. Eur. Control Conf.*, EPFL-CONF-186265 (Zurich, Switzerland) 2013

MODELLING A THREE-PHASE CURRENT SOURCE INVERTER

László Neukirchner* and Attila Magyar

Department of Information Technology, University of Pannonia, Egyetem u. 10, Veszprém, 8200,
HUNGARY

A current source inverter model has been developed in the given paper that is constructed from six LTI models for the different switching modes. The overall model is in a piecewise affine form that supports the use of model predictive control. The model has been verified against engineering expectations and its open-loop performance shows that it is a promising basis of model predictive control structures.

Keywords: model predictive control, current source inverter, space vector modulation, modelling, simulation

1. Introduction

In industrial applications power electronic converters are divided into two major classes, namely voltage source (VSI) and current source inverters (CSI). As a VSI employs modulating the voltage wave through the inverter's legs, its counterpart relies on modulated current waveforms. To operate a VSI the inverter's switching devices should operate in a specified alternating modulation to avoid shorting the inverter's legs, while its current source counterpart could embed this operation into the modulation. Thyristor rectifiers suffer from several drawbacks, e.g. high-input current harmonic distortion or a low power factor. Pulse width modulated current-source inverters, on the other hand, are free from the above-mentioned drawbacks. In such cases, a smoothing LC filter is placed on the AC side and a choke inductor DC side to reduce the harmonic injection resulting from the current modulating PWM operation. In the case of sinusoidal supply voltages, an input power factor with unit value can be achieved if both requirements are met. In conventional control structures, PWM-CSR is usually operated by off-line patterns and look-up tables like in trapezoidal modulation and selective harmonic elimination. The use of these techniques simplifies the calculation of the gate control signals with the special requirements of applied switches like IGBTs. Input filtering, on the other hand, may result in oscillatory transient behaviour. This can be attenuated by adding damping resistive elements, but using them also reduces the overall efficiency. Space vector modulation (SVM) has proven to result in a high voltage gain, reduced switching frequency and low current harmonic distortion, and could be straightforwardly implemented in digital systems [1]. In the second section, the structure of the investigated inverter is presented. In the next section, a detailed open-loop model is presented, and in the fourth section a predictive strategy is chosen for controlling SVM.

2. Inverter Structure

The current source inverter normally requires three-phase inductive and capacitive filtering (L_f and C_f, respectively), improving the load current and voltage waveform's harmonic distortion, and in thermistor appliances assisting the commutation of the switching devices. The legs of the device should include protective diodes, otherwise the absence of a current path introduces a high-voltage spike, causing damage to the switching devices [2]:

- The switching devices used are symmetrical, they do not require freewheeling diodes.
- The current source inverter produces a three-phase pulse-width modulated current instead of a pulse-width modulated voltage. As such, a capacitor needs to be installed on the output for filtering purposes. The high voltage derivative is associated with the VSI.
- The rate at which the DC current rises is limited by the DC choke, allowing sufficient time for the protection circuit to function in the case of a short circuit at the inverter output terminals.
- As a restraint, the DC link current cannot be instantaneously changed, which resulted in a reduction in dynamic performance.
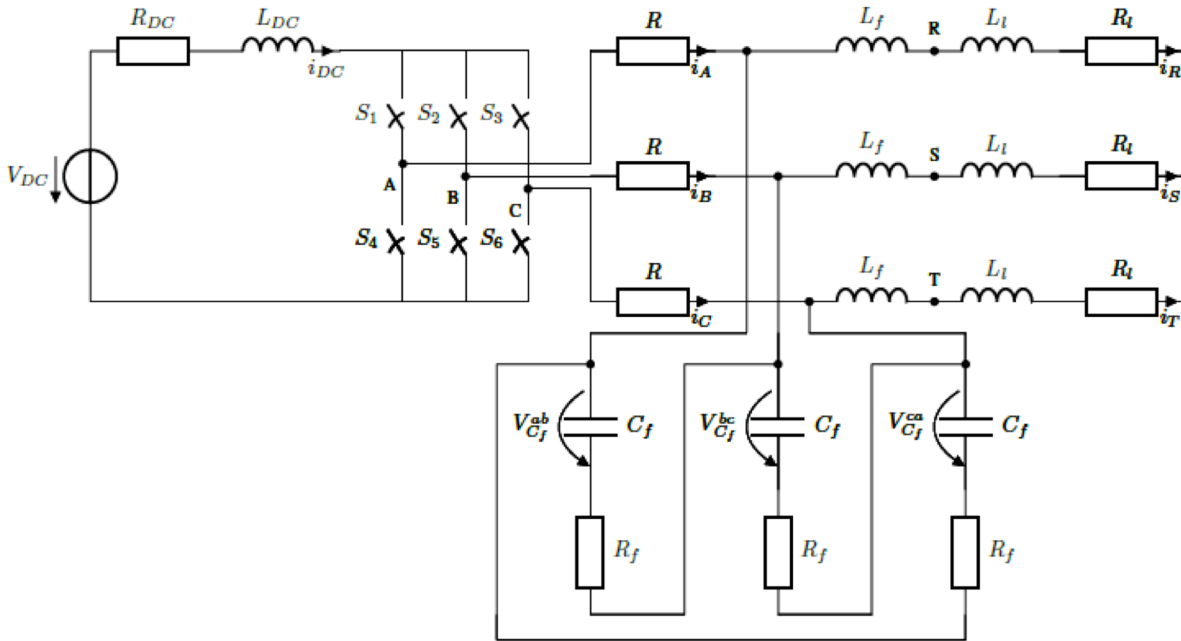
*Correspondence: neukirchner.laszlo@uni-pannon.hu

Figure 1. Current source inverter with an LC filter and RL load.

3. Open-Loop Switched Model with Space Vector Modulation

3.1. Space Vector Formulation

The switched model of the three-phase CSI supplying a three-phase load through an LC filter and RL load is shown in *Fig.1*. The LC filter is a low-pass filter included with the three-phase CSI to decrease the THD by eliminating higher order harmonics. The implemented LC filter includes a resistor R_f in series with a capacitor C_f. This resistor provides a passive damping solution and it was recommended to be within the range of 0.5 to 1 Ω. Moreover, the resistor RDC represents any start-up pre-charging resistance or any resistance in series with the DC voltage source VDC. For constructing the model the following assumptions are made:

- The LC filter and the RL load's circuit components are symmetric.
- Only a continuous conduction mode is assumed.
- Each time, the switch state index s is known.
- The switching devices used for modulation are assumed to be ideal, with zero losses and switching is considered to be instant.
- No external disturbance is assumed.
- No magnetic coupling between the inductances is assumed.

With the above assumptions, the equation system of the inverter, filter and load for all the switching states can be created, as different state-space models [3, 4]. Under the above-mentioned constraints, the three-phase current source inverter has nine switching states as can be seen in *Table 1*.

Table 1. Switching states of SVPWM modulated CSI.

	active switches	i_A	i_B	i_C	Space vector
Zero States	S_1, S_4 S_2, S_5 S_3, S_6	0	0	0	I_0
Active States	S_6, S_1	i_{DC}	$-i_{DC}$	0	I_1
	S_1, S_2	i_{DC}	0	$-i_{DC}$	I_2
	S_2, S_3	0	i_{DC}	$-i_{DC}$	I_3
	S_3, S_4	$-i_{DC}$	i_{DC}	0	I_4
	S_4, S_5	$-i_{DC}$	0	i_{DC}	I_5
	S_5, S_6	0	$-i_{DC}$	i_{DC}	I_6

These switching states can be classified as zero switching states (I_0) and active switching states (I_1, ..., I_6). In addition, there are three zero states which implies vertical conduction of the inverter leg. This implies that the source is shorted, leading to $i_A = i_B = i_C = 0$. This operating mode is often referred to as the bypass operation. Assuming that the operation of the inverter is three-phase balanced then *Eq.(1)* applies

$$i_A + i_B + i_C = 0 \qquad (1)$$

According to *Fig.2*, the five active vectors can be expressed as:

$$I_k = \frac{2}{\sqrt{3}} I_{DC} e^{j\left((k-1)\frac{\pi}{3} - \frac{\pi}{6}\right)}, k = 1, 2, ..., 6 \qquad (2)$$

The active and zero vectors are assumed to be stationary, on the contrary, the current reference vector I_{ref} in *Fig.2* rotates with an angular velocity $\omega = 2\pi f$, and $\theta(t) = \int_0^t \omega(t)dt + \theta(0)$, where f is the fundamental frequency of the inverter output current and θ is the angular displacement. A desired vectorial length and angle, I_{ref} can be calculated from the neighbouring space
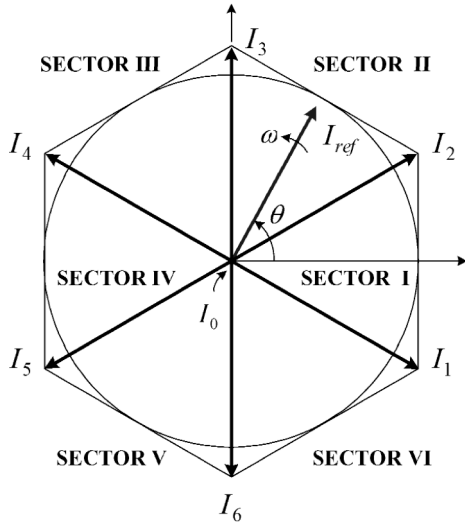
Figure 2. A typical space vector diagram for the CSI. $I_1, ..., I_6$ are the active vectors, I_0 is the zero vector, ω is the angular velocity, θ is the angular displacement and the Cartesian coordinate system represents the imaginary and real axes. The active vectors form a symmetric hexagon with six identical sectors, while the zero vector I_0 lies at the center of the hexagon.

vectors (Fig.3), based on the switching pattern and frequency of the inverter. As I_{ref} passes through the vectorial spaces, different switch sets are applied to formulate the desired output. As a result, when I_{ref} rotates one revolution, the output continues for one cycle based on the fundamental frequency.

3.2. Calculating Dwell-Time

The dwell time for stationary vectors represents the time that space vectors spend in a sector during T_S. The basic method used for dwell-time calculations is the so-called ampere-second balancing principle. The main idea behind it is that the product of the sampling period T_S and the reference vector I_{ref} is equal to the sum of the current vectors multiplied by the time interval of chosen space vectors. With small enough T_S, I_{ref} can be assumed as constant over a period. With this notion, the reference current is approximated from the neighbouring vectors corresponding to the actual sector and from the zero vector for adjusting the length.

The ampere-second balancing principle is given by

$$
\begin{aligned}
I_{ref} T_S &= I_1 T_1 + I_2 T_2 + I_0 T_0 \\
T_S &= T_1 + T_2 + T_0
\end{aligned}
\quad (3)
$$

where T_0, T_1, and T_2 denote the dwell times of the vectors I_0, I_1, and I_2. From Eq.(2), the individual values of space vectors in SECTOR I are defined in Eq.(4):

$$
\begin{aligned}
I_1 T_S &= \frac{2}{\sqrt{3}} I_{DC} e^{-j\frac{\pi}{6}} \\
I_2 T_S &= \frac{2}{\sqrt{3}} I_{DC} e^{j\frac{\pi}{6}} \\
I_0 T_S &= 0
\end{aligned}
\quad (4)
$$
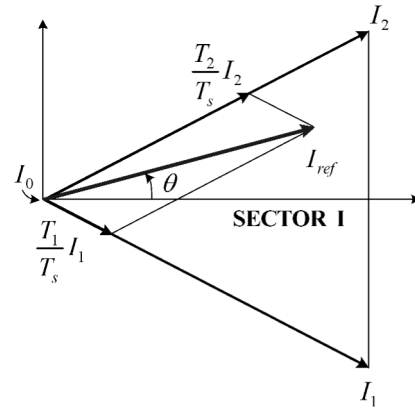
Figure 3. Calculation of I_{ref} in Sector I, where the Cartesian coordinate system represents the imaginary and real axes.

Projecting the resulting complex vector onto the real and imaginary axes leads to:

$$
\begin{aligned}
I_{ref} \cos(\theta) T_S &= I_{DC}(T_1 + T_2) \\
I_{ref} \sin(\theta) T_S &= \frac{1}{\sqrt{3}} I_{DC}(-T_1 + T_2).
\end{aligned}
\quad (5)
$$

Solving Eqs.(3) and (5) yields the dwell times T_0, T_1, and T_2 at any given θ angle as shown in Eq.(6):

$$
\begin{aligned}
T_1 &= m_a \sin\left(\frac{\pi}{6} - \theta\right) T_S \\
T_2 &= m_a \sin\left(\frac{\pi}{6} + \theta\right) T_S \\
T_0 &= T_S - T_1 - T_2 \\
\text{for} &\quad -\frac{\pi}{6} \leq \theta \leq \frac{\pi}{6},
\end{aligned}
\quad (6)
$$

where $m_a = |I_{ref}| / I_{DC}$ is the modulation index in the domain: $m_a \in [0, 1]$.

When I_{ref} is inside of other sectors a multiple of $\pi/3$ is subtracted from θ so that the modified angle σ happens to fall within the range of $-\frac{\pi}{6} \leq \theta \leq \frac{\pi}{6}$ for use in Eq.(7):

$$
\begin{aligned}
\sigma &= \theta - (k-1)\frac{\pi}{3} \\
s &= 1, 2, ..., 6 \\
\text{for} &\quad -\frac{\pi}{6} \leq \theta \leq \frac{\pi}{6},
\end{aligned}
\quad (7)
$$

where s is the sector index. This method save sthe need for implementing unnecessary calculations and code capacities.

3.3. Switching Sequence

As far as designing the switching pattern of the SVM for the CSI a few notions should be specified for switching loss reduction:
- The state transition should involve the least amount of switches (one switch is desired).
- Minimum amount of switches should be involved as I_{ref} transits from the current to the next state.

Table 2. Typical switching sequence for Sector I. The sequence was repeated twice to achieve a smoother output signal. In this case, the sampling period T_S is one 12th of the fundamental time period. The active switch pairs are S_6, S_1 and S_1, S_2 for space vectors I_1 and I_2. The chosen zero state (with switches S_1, S_4) is chosen for I_0 such that the minimum switching loss requirement is met.

Vectors	I_1	I_2	I_0	I_1	I_2	I_0
S_1	X	X	X	X	X	X
S_2		X			X	
S_3						
S_4			X			
S_5						
S_6	X			X		X
dwell	T_1	T_2	T_0	T_1	T_2	T_0
time		T_S			T_S	

The typical sequence can be observed in *Table 2*. I_{ref} is created by I_1, I_2, and I_0 with the mentioned dwell time rule 6. T_s can be divided into three or more segments (T_0, T_1, and T_2) according to the desired THD value of the output current.

From the above, a continuous switching sequence for the SVPWM modulation can be obtained by *Eq.(8)*:

$$\begin{aligned} I_k, I_{k+1}, I_0 \quad & k = 1, 2, \dots, 5 \\ I_k, I_1, I_0 \quad & k = 6 \, , \end{aligned} \quad (8)$$

where k is the sector index.

3.4. Obtaining Open-Loop 'Piecewise Affine' Model

With every state vector I_0, I_1, ..., I_6 a different switch configuration and a different circuit configuration is present at a given time. For every case the circuit's state equations can be written in continuous state-space form:

$$\begin{aligned} \dot{x}(t) &= \mathbf{A}_S x(t) + \mathbf{B}_S u(t) \\ \dot{y}(t) &= \mathbf{C}_S x(t) \\ s &= 0, 1, \dots, 6 \, , \end{aligned} \quad (9)$$

where $\mathbf{A}_S$ is the sth state matrix, $\mathbf{B}_S$ is the sth input matrix, $\mathbf{C}_S$ is the sth output matrix, $x(t)$ and $u(t)$ are space vectors, and s represents the space vector index with $s = 0$ as the zero state and $s = 1, 2, \dots, 6$ as the active states. Since all the equation systems have the same input ($\mathbf{B}_S$) and output ($\mathbf{C}_S$) matrices they will be referred to as $\mathbf{B}$ and $\mathbf{C}$ matrices.

Any linear circuit can be represented by a number of differential equations equal to the number of energy storage units in the circuit. However, in three-phase balanced circuits, the number of equations can be reduced by a third factor because any three-phase quantity can be expressed in terms of the other two phase quantities. Therefore, since currents flowing through L_f and L_l and voltages across C_f are three-phase balanced quantities only four linear differential equations representing the AC side and one linear differential equation representing DC current across L_{DC}

are needed. Hence, linear differential equations for the state variable $x = \left[i_{DC} V_{C_f}^{ab} V_{C_f}^{bc} i_R i_S \right]^T$, input variable $u = V_{DC}$ and output variable $y = [i_R \ i_S]^T$ can be derived using KVL and KCL. Arranging the expressions for the aforementioned state variables and recognizing V_{DC} as the input of the system, yields the following state-space matrices for the zero switching state as shown in *Eqs.(10)-(12)*:

$$\mathbf{A_0} = \begin{bmatrix} -\dfrac{R_{DC}}{L_{DC}} & 0 & 0 & 0 & 0 \\[2mm] 0 & -\dfrac{1}{Q_a} & 0 & -\dfrac{R}{Q_a} & \dfrac{R}{Q_a} \\[2mm] 0 & 0 & -\dfrac{1}{Q_a} & -\dfrac{R}{Q_a} & -\dfrac{2R}{Q_a} \\[2mm] 0 & \dfrac{2R}{Q_c} & \dfrac{R}{Q_c} & \dfrac{Q_b}{Q_a} & 0 \\[2mm] 0 & \dfrac{R}{Q_c} & \dfrac{R}{Q_c} & 0 & \dfrac{Q_b}{Q_a} \end{bmatrix} \quad (10)$$

$$\mathbf{B} = \begin{bmatrix} -\dfrac{1}{L_{DC}} & 0 & 0 & 0 & 0 \end{bmatrix}^T \quad (11)$$

$$\mathbf{C} = \begin{bmatrix} 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}, \quad (12)$$

where $Q_a = C(3R + R_f)$, $Q_b = RR_f + 3RR_l + R_f R_l$ and $Q_c = (3R + R_f)(L_f + L_l)$.

3.5. Open-Loop Performance

As CSI exhibits different LTI regions for different state matrices A_0, A_1, ..., A_6 according to the used space vector I_0, I_1, ..., I_6 in the corresponding dwell time, the piecewise affine approach (PWA) can be used to describe the overall dynamics. The discrete time piecewise affine model can be obtained by discrediting the set of continuous affine models in the form of *Eq.(13)*

$$x_{p+1} = \begin{cases} \Phi_1 x_p + \Gamma u_p \, , & \text{if } \begin{bmatrix} u_1 \\ x_1 \end{bmatrix} \in P_1 \\ \quad \vdots \\ \Phi_S x_p + \Gamma u_p & \text{if } \begin{bmatrix} u_s \\ x_s \end{bmatrix} \in P_s \end{cases} \quad (13)$$

Such systems are composed of s local affine models whose parameters in (Φ_s) change according to which polyhedron P_k contains the state vector.

The simulated waveforms for VDC = 400 V DC link with pre-charge resistance $R_{DC} = 1 \ \Omega$ and choke inductance $L_{DC} = 2e^{-1} \ H$ using the space vector modulation as shown in *Fig.5*, where i_A denotes the PWM current of the inverter and i_R stands for the phase current through the load. The inverter operates at a fundamental frequency of $f = 50$ Hz. The filter capacitor $C_f = 20e^{-6} \ F$, the filter inductance $L_f = 10e^{-3} \ H$ with filter resistance $R_f = 1 \ \Omega$ and $R = 1 \ \Omega$ representing losses. The inverter consists of three-phase balanced inductive load $L_l = 2e^{-3}$ and resistive load $R_l = 5 \ \Omega$. The open-loop performance can be observed in *Fig.4* and the modulation and PWM formulation in *Fig.5*.
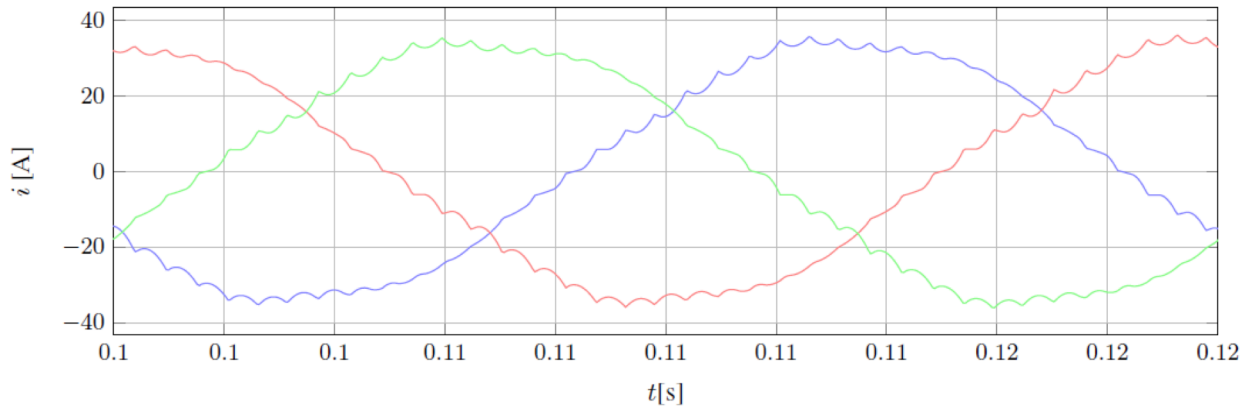
Figure 4. Three-phase sinusoidal waveform of the CSI output currents i_A, i_B, and i_C with a sample period of $T_s = T/36$ and a modulation index of $m_a = 0.9$. The output wave can be smoothened by adding higher filter parameters or further reducing the sampling period.
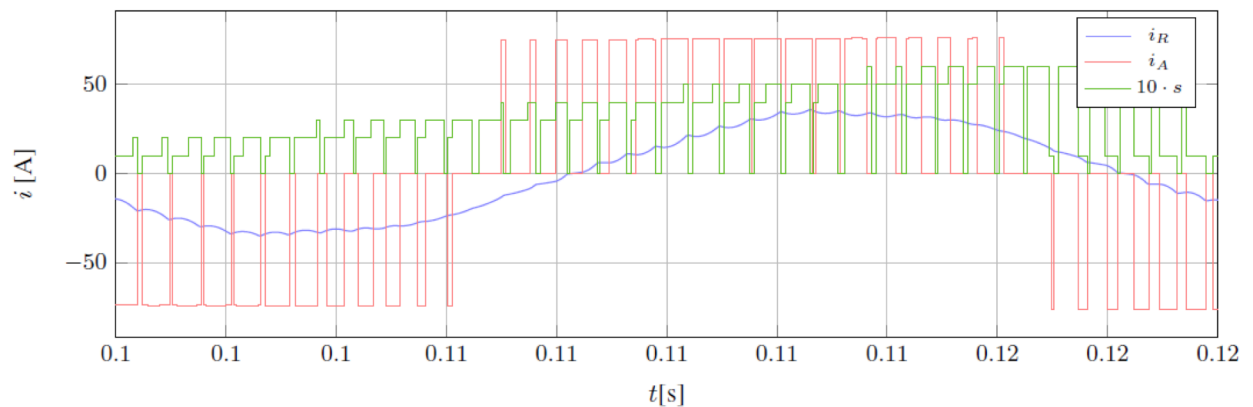

Figure 5. Comparison of the inverter output current i_A and the filtered load current i_R with $T_s = T/36$ and $m_a = 0.9$. The background s's value indicates the corresponding affine state of the model at a given time.

4. MPC assisted SVPWM

The aim of the previous sections was to develop a model for the inverter that describes the switching behaviour of the dynamics. The obtained piecewise affine model gives a good basis for the model predictive control technique that is advantageous from two points of view: the possibility of introducing constraints to the state and input variables and the fact that optimal control problems can be formulated within it.

The aim is to find the optimal space vectors in such a way that the error between the predicted states of the inverter and the chosen reference values are minimized with respect to constraints. Previously, the method of model predictive control was reserved only for applications with relatively slow time scales, like plants and chemical processes. The reason for this is that MPC requires an optimization problem to be solved at each sampling instant in order to obtain the control input. Over the past decade, more and more applications use fast MPC applications, thanks to the advances in computer technology. The optimization problem is often NP hard, and finding a solution is taxing in terms of the calculation time of the control device, thus limiting the rate at which the control could intervene with the application. In particular, model predictive control has

been applied in the field of power electronics, where the control frequency ranges from a few hundred Hz up to a few MHz [5,6]. The main challenge of implementing MPC consists of solving the associated optimization problem in a bounded time.

The predicted values are assumed to be constant after time k. Then, the control problem can be formulated as follows:

$$min \; J(\boldsymbol{x}[k])$$
$$s.t. \; \boldsymbol{x}[k] = \mathbf{x}[k|k]$$
$$J(\boldsymbol{x}[k]) = \sum_{j=1}^{N} ||Q(\boldsymbol{x}(k+j\,|\,k) - \boldsymbol{x}_{\text{ref}}([k]))||^2 \qquad (14)$$

where $\boldsymbol{x}[k+j\,|\,k]$ denotes the prediction of the state at time $k+j$ based on the information available at time k, N is the prediction horizon, Q is the constant weighting. The optimal state-finding algorithm follows:

$$OptimalError = inf$$
for $s = 1 : 7$ **do**
$$[x(1) \quad x(2) \quad \dots \quad x(N)]^{\text{T}} =$$
$$\begin{bmatrix} B_s & 0 & \cdots & 0 \\ A_s & B_s & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ A_s^{N-1}B_s & A_s^{N-2}B_s & \cdots & B_s \end{bmatrix} \cdot \begin{bmatrix} u(0) \\ u(1) \\ \vdots \\ u(N) \end{bmatrix} + \begin{bmatrix} A_s \\ A_s^2 \\ \vdots \\ A_s^N \end{bmatrix} x(0)$$

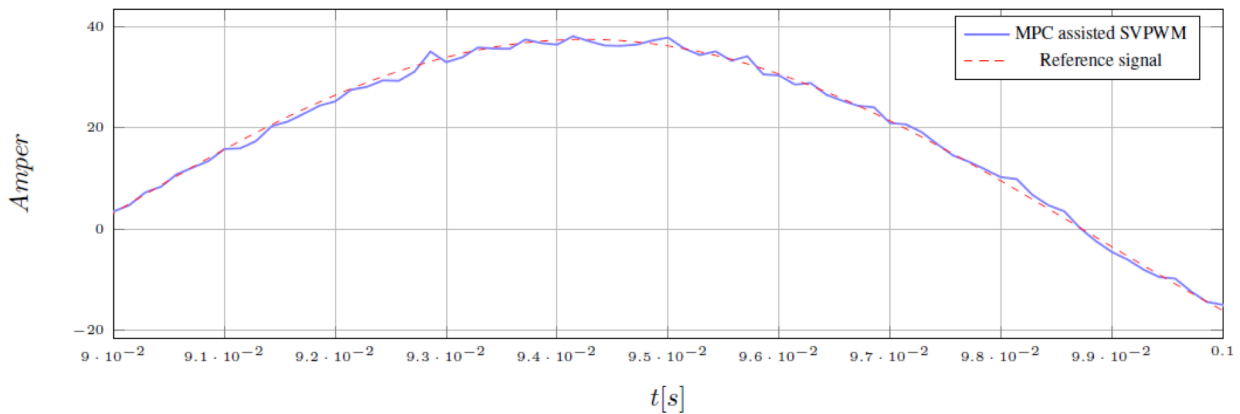*Figure 6.* MPC reference tracking with $f = 50$Hz, $\hat{I} = 37.5$A, and $N = 2$. In order to create the simplest implicit MPC for a finite state PWA environment, every affine state was tested with the according horizon in every iteration.
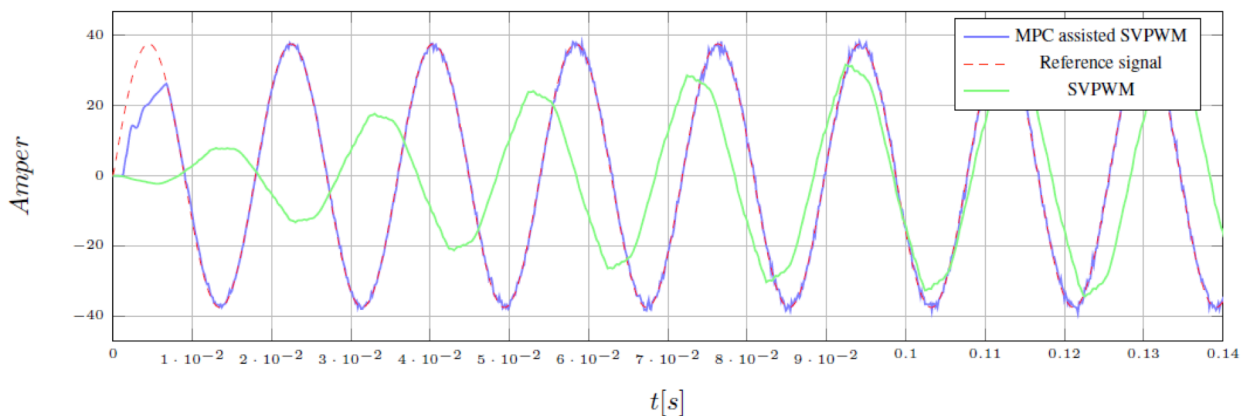

Figure 7. Comparison of classic SVPWM and MPC on the same system. The SVPWM is applied with a modulation index of 0.9, and the MPC reference tracking with $f = 50$Hz, $\hat{I} = 37.5$A, and $N = 2$.

$Error_s = J_s(x)$
if $Error_s < OptimalError$ **then**
     $OptimalError = Error_s$
     $Optimal\_s = s$

where *OptimalError* represents the minimal error value, *Error_s* is the actual error value in each iteration, and *Optimal_s* is the optimal state. The algorithm uses a parallel implicit MPC scheme for every state. Since there are only seven states, with a low horizon the algorithm can be feasibly implemented. When $f = 50$Hz, $\hat{I} = 37.5$A, and $N = 2$, deviations from the optimal trajectory are present during optimization as shown in *Fig.6*, but they are still acceptable in terms of harmonic distortion. The performance observed in *Fig.7* can be compared to the open-loop solution. It can be stated that due to the rigid structure of the open-loop modulation's rigid structure the dynamical properties of the MPC-assisted system exhibit better performance in terms of time response and signal shape.

5. Conclusion

A current source inverter model has been developed in this paper that is constructed from six LTI models for the different switching modes using SVPWM modulation. The obtained overall model is in a piecewise affine form that supports the use of model predictive control in the future. The model has been verified against engineering expectations and its open-loop performance shows that it is a promising basis of further work in terms of identification and control. As the next step, a model predictive control synthesis was performed based on the technology-driven optimality criteria to assist the open-loop SVPWM modulated system. Due to the implicit MPC-assisted modulation the optimal space vector could be found and utilized in each state over an acceptable timescale due to the low horizon. As such a significantly lower time response and better signal shape was achieved.

SYMBOLS

A	state matrix
B	input matrix
C	output matrix
C_f	filter capacitance
f	fundamental frequency
Φ	discrete state matrix
Γ	discrete input matrix
$I_0, I_1, ..., I_6$	space vector
i_A, i_B, i_C	inverter output currents per phase

i_{DC}	DC link current
i_R, i_S, i_T	load currents per phase
$\boldsymbol{I}_{ref}$	reference current value
k	sector index
L_f	filter inductance
L_l	load inductance
LTI	Linear Time Invariant
m_a	modulation index
MPC	Model Predictive Control
v_{cr}	triangular carrier wave for TPWM
v_m	modular wave for TPWM
p	discrete time index
P	space vector containing polyhedron
PWA	'Piecewise Affine' Current Source Inverter (CSI)
PWM	Pulse Width Modulation
R	loss resistance
R_{DC}	DC link start-up pre-charging resistance
R_f	filter resistance
R_l	load resistance
s	space vector index
$S_1, S_2, ..., S_6$	switching devices
σ	modified angle
SHE	Selective Harmonic Elimination
SVPWM	Space Vector Pulse Width Modulation
T	fundamental cycle time
T_k	dwell time of corresponding space vectors
T_s	sampling period
TPWM	Trapezoidal Pulse Width Modulation
θ	angular displacement
$\boldsymbol{u}$	input vector
V_{DC}	DC link voltage source value
$V_{C_f}^{ab}, V_{C_f}^{bc}, V_{C_f}^{ac}$	filter capacitance's voltage
$\boldsymbol{x}$	state vector
$\boldsymbol{y}$	output matrix
ω	angular velocity

Acknowledgement

REFERENCES

[1] Espelage, P.; Nowak, J.M.; Walker, L.H.: Symmetrical GTO current source inverter for wide speed range control of 2300 to 4160 V, 350 to 7000 hp induction motors, *IEEE Ind. Appl. Soc. Conf. (IAS)* 1988 **1**, 302–307 DOI 10.1109/IAS.1988.25079

[2] Wu, B.: High-power converters and AC drives (Wiley-IEEE Press, New Jersey, USA) 2006

[3] Ma, J.; Wu, B.; Rizzo, S.: A space vector modulated CSI-based AC drive for multi-motor applications, *IEEE Trans. Power Electron.*, 1999 **16**(4), 535–544 DOI: 10.1109/APEC.1999.750460

[4] Wiseman, J.; Wu, B.; Castle, G.S.P.: A PWM current source rectifier with active damping for high power medium voltage applications, *IEEE Power Electron. Spec. Conf. (PESC)* 2002 **1**, 1930–1934 DOI: 10.1109/PSEC.2002.1023096

[5] Mariéthoz, S.; Morari, M.: Explicit Model-Predictive control of a PWM inverter with an LCL filter, *IEEE Trans. Ind. Electron.* 2008 **56**(2), 389–399 DOI: 10.1109/TIE.2008.2008793

[6] Mariéthoz, S.; Almér, S.; Baja, M.; Beccuti, A.G.; Patino, D.; Wernrud, A.; Buisson, J.; Cormerais, H.; Geyer, T.; Fujioka, H.; Jonsson, U.T.; Kao, C.Y; Morari, M.; Papafotiou, G.; Rantzer, A.; Riedinger, P.: Comparison of hybrid control techniques for buck and boost DC-DC converters, *IEEE Trans. Control Syst. Technol.* 2010 **18**(5), 1126–1145 DOI: 10.1109/TCST.2009.2035306

COMPUTATIONAL STABILITY ANALYSIS OF LOTKA-VOLTERRA SYSTEMS

PÉTER POLCZ[1,*] AND GÁBOR SZEDERKÉNYI[1,2]

[1] Faculty of Information Technology and Bionics, Pázmány Péter Catholic University, Práter u. 50/A, Budapest, 1083, HUNGARY

[2] Systems and Control Lab, Institute for Computer Science and Control, Hungarian Academy of Sciences, Práter u. 50/A., Budapest, 1083, HUNGARY

This paper concerns the computational stability analysis of locally stable Lotka-Volterra (LV) systems by searching for appropriate Lyapunov functions in a general quadratic form composed of higher order monomial terms. The Lyapunov conditions are ensured through the solution of linear matrix inequalities. The stability region is estimated by determining the level set of the Lyapunov function within a suitable convex domain. The paper includes interesting computational results and discussion on the stability regions of higher (3,4) dimensional LV models as well as on the monomial selection for constructing the Lyapunov functions. Finally, the stability region is estimated of an uncertain 2D LV system with an uncertain interior locally stable equilibrium point.

Keywords: nonlinear systems, Lotka-Volterra models, stability analysis, linear matrix inequalities, Lyapunov function

1. Introduction

Approximating the domain of attraction (DOA) is often a fundamental task in the analysis and control of nonlinear systems. The stability properties of dynamical systems are most often studied using Lyapunov functions. Therefore, extensive literature exists on the computational construction of Lyapunov functions [1].

Due to their advantageous properties and the availability of efficient numerical solvers, the use of linear matrix inequalities (LMI) and semi-definite programming (SDP) techniques has become popular in the field of system and control theory. Important results were announced [2,3] in the context of linear uncertain systems, their stability analysis and control synthesis.

Recently, an optimization-based method for DOA estimation was published [4], where the authors use Finsler's lemma and affine parameter-dependent LMIs to compute rational Lyapunov functions for a wide class of locally asymptotically stable nonlinear systems. Based on these results an improved method was published [5,6], where the transformation of the model to the form required for optimization is done automatically using the linear fraction transformation (LFT) and further automatic model simplification steps, which results in the dimension reduction of the optimization task. As the dimensions of the problem are

reduced, the method is capable of handling more complex models. For example, the DOA is successfully estimated for a 3-dimensional rational uncertain system [6] (bioreactor model with an applied proportional and integral substrate feedback law).

The dynamical descriptive power of Lotka-Volterra systems is so extensive that LV models "have the status of canonical format" within the class of smooth nonlinear dynamical systems [7]. Besides modelling biological/ecological environments, they are widely used in other scientific fields like neural networks [8] or in economics, where the Goodwin-Lotka-Volterra models are applied for modelling the predator-prey mechanism of the technological substitution [9,10]. In order to model the correlation between the employment rate and the share of wages of the working population, the authors of Ref. [11] used a stochastic extension of the Goodwin model.

Some important results that make LV models even more attractive are the existing techniques used to represent a general nonlinear system as a multidimensional LV model [12]. The analysis of stability and behaviour of LV systems is extensive in the literature [13]. Plank has shown [14] that *N*-dimensional LV systems are Hamiltonian if they fulfil certain algebraic properties (Theorem 3.1 in Ref. [14]). He demonstrated that, when using an appropriate Poisson structure, one can obtain the Hamiltonian function of the system, which is a key object in determining the system's DOA. Furthermore, candidates of Hamiltonian function were defined in (Section IV in Ref. [14]). It is important to note that LV systems are

---

kinetic in the sense that they can be formally described as chemical reaction networks with mass action kinetics.

In this paper, the local stability properties of LV models are analysed by applying the underlying method presented [6]. The stability region is estimated on a few locally stable 2-, 3-, and 4-dimensional benchmark LV systems. The DOA of an uncertain 2D system is also estimated by two concentric regions.

The main motivation of our work was to evaluate the applicability of the approach [4] on a general polynomial system class consisting of low-degree monomials, and to study the limits of the method as the number of dimensions of the state space increase.

2. Background

In this section, the basic notions and results on which our computational method is based are presented.

2.1. System Class, Lyapunov Functions, and the Domain of Attraction.

Our computational method can handle general nonlinear systems of the form

$$\dot{x}(t) = f(x(t), \delta(t))$$
$$x(t) \in \mathbb{R}^N, x_0 \in \mathcal{X}, \delta(t) \in \mathcal{D}, \dot{\delta}(t) \in \breve{\mathcal{D}} \qquad (1)$$

where $\mathcal{X} \subset \mathbb{R}^N$ and $\mathcal{D}, \breve{\mathcal{D}} \subset \mathbb{R}^D$ are given polytopes, x is the state vector function with its initial condition $x_0 = x(0)$, and δ is a smooth, bounded vector function of uncertain parameters with a bounded time derivative. The applied method is presented in detail elsewhere [15]. From now on, the time arguments of x and δ will be suppressed as is commonly done in the literature. It is assumed that function $f: \mathbb{R}^N \times \mathcal{D} \mapsto \mathbb{R}^N$ of Eq.(1) is a well-defined smooth *rational mapping*, with the property $f(0, \delta) = 0$ for all $\delta \in \mathcal{D}$. Additionally, it is assumed that $x^* = 0 \in \mathbb{R}^N$ is a locally asymptotically stable equilibrium point of Eq.(1) for all $\delta \in \mathcal{D}$. The set of all initial conditions, from which the solutions converge to x^*, is called the *domain of attraction* (DOA). Furthermore, it is assumed that function $f(x, \delta)$ can be written in a so-called quasi-LPV form $f(x, \delta) =$ function $A(x, \delta) x$

The aim is to identify an appropriate rational Lyapunov function $V(x, \delta)$, which satisfies the following conditions:

$$v_l(x) \le V(x, \delta) \le u_x(x), \forall (x, \delta) \in \mathcal{X} \times \mathcal{D}$$
$$\dot{V}(x, \delta, \dot{\delta}) \le -v_d(x) \, \forall (x, \delta, \dot{\delta}) \in \mathcal{X} \times \mathcal{D} \times \breve{\mathcal{D}} \qquad (2)$$

where v_l, u_x, v_d are continuous positive functions on $\mathcal{X}$. Due to conditions of Eq.(2) any closed level set of the Lyapunov function contained entirely in $\mathcal{X}$ bounds an invariant region of the state space. Our objective is to find a Lyapunov function, which fulfils the conditions of Eq.(2) and to determine its maximal invariant level set completely inside $\mathcal{X}$.

2.2. Dynamical System Representation

A Lyapunov function can be computed [4] in the form shown in Eq.(3).

$$V(x, \delta) = \pi_b^T(x, \delta) P \pi_b(x, \delta), \pi_b = \begin{bmatrix} x \\ \pi \end{bmatrix} \qquad (3)$$

where $P \in \mathbb{R}^{m \times m}$ is a *constant symmetric* matrix, not necessarily positive definite, and $\pi: \mathbb{R}^N \times \mathcal{D} \mapsto \mathbb{R}^p$ is mapping, in which each element is a monomial in (x, δ), or a smooth rational function with a monomial numerator. The arguments of π and π_b will be suppressed below.

Applications of the linear fractional transformation (LFT) and further algebraic steps have been proposed [5,6] to transform the system equation $\dot{x} = f(x, \delta) = A(x, \delta)x$ into the desired differential-algebraic representation that was introduced in the same references:

$$\dot{x} = A(x, \delta)x = Ax + B\pi \quad x_0 \in \mathcal{X}$$
$$0 = \mathcal{N}_{\pi_b}(x, \delta)\pi_b \qquad \delta \in \mathcal{D}, \dot{\delta} \in \breve{\mathcal{D}} \qquad (4)$$

where $A \in \mathbb{R}^{n \times n}$ and $B \in \mathbb{R}^{n \times p}$ are constant matrices, and $\mathcal{N}_{\pi_b}(x, \delta) \in \mathbb{R}^{q \times m}$ is an *affine* matrix function in (x, δ) also known as an "annihilator". These transformations also result in the dimension reduction of the optimization problem compared to the results presented in Ref. [4]. The representation in Eq.(4) separates the linear part of the system (x) from its nonlinear part (π).

3. The SDP Problem for Estimating the DOA

In this section, how to construct the semi-definite optimization task (SDP) is presented, which will uniquely characterize an appropriate Lyapunov function and its maximal invariant level set ε_α as an estimate of the true domain of attraction. It is assumed that polytopes $\mathcal{X}, \mathcal{Y}$, and $\mathcal{D}$ are already given.

Using the model representation of Eq.(4) and a Lyapunov function candidate of the form in Eq.(3), sufficient LMIs for the Lyapunov conditions in Eq.(2) can be formulated. According to Finsler's lemma (Lemma 2.1 in Ref. [4]), if real-valued matrices $L_b \in \mathbb{R}^{m \times q}$ and $L_a \in \mathbb{R}^{(n+2p+n^2+np) \times (2q+n^2+nq)}$ exist

$$\forall (x, \delta) \in \vartheta(\mathcal{X} \times \mathcal{D}):$$
$$P + L_b \mathcal{N}_{\pi_b}(x, \delta) + \mathcal{N}_{\pi_b}^T(x, \delta) L_b^T > 0 \qquad (5)$$

$$\forall (x, \delta, \dot{\delta}) \in \vartheta(\mathcal{X} \times \mathcal{D} \times \breve{\mathcal{D}}):$$
$$P_a + P_a^T + L_a \mathcal{N}_{\pi_a}(x, \delta, \dot{\delta}) + \mathcal{N}_{\pi_a}^T(x, \delta, \dot{\delta}) < 0 \qquad (6)$$

then the conditions of Eq.(2) are satisfied for every $(x, \delta, \dot{\delta}) \in \mathcal{X} \times \mathcal{D} \times \breve{\mathcal{D}}$. Variables P_a and $\mathcal{N}_{\pi_a}(x, \delta, \dot{\delta})$ are
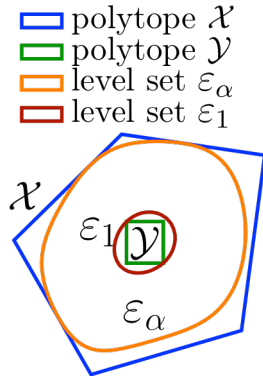
Figure 1. Illustration of the conditions regarding $\mathcal{X}$ and $\mathcal{Y}$. Inside polytope $\mathcal{X}$ of the Lyapunov conditions in *Eq.(2)* are required. The α-level set ε_α of the Lyapunov function should lie inside polytope $\mathcal{X}$ (hence it is invariant). The 1-level set ε_1 should be around polytope $\mathcal{Y}$. This condition ensures that the problem has a bounded solution.
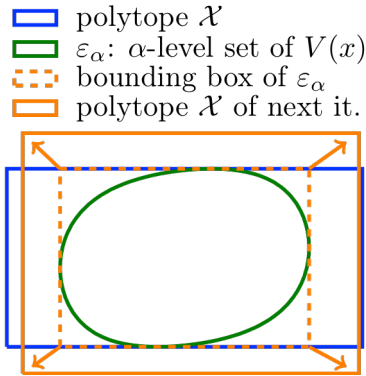

Figure 2. Polytope $\mathcal{X}$ is evaluated through iterations considering the bounding box of the obtained α-level set.

defined in *Eqs.(3.16)* and *(3.22)* of Ref. [15], respectively. $\vartheta(\cdot)$ denotes the corner points (vertices) of a given polytope.

Since the Lyapunov conditions are satisfied within $\mathcal{X}$, an attempt is made to identify the maximal α-level set

$$\varepsilon_\alpha = \{x \in \mathcal{X} \,|\, V(x) = \alpha, 1 \leq \alpha\},$$

which lies inside $\mathcal{X}$. That level set will be *invariant*, in the sense that every trajectory entering this region will never leave it.

$\alpha \geq 1$ is defined, as a free variable of the optimization task, and ε_α is constrained inside $\mathcal{X}$. One can observe that by maximizing α an unbounded feasible solution is obtained, as the function $V(x, \delta)$ can be scaled arbitrarily. Therefore, as *Fig.1* illustrates, an auxiliary polytope $\mathcal{Y}$ is defined around the locally stable origin inside $\mathcal{X}$, constraining the 1-level set $\varepsilon_1 = \{x \in \mathcal{X} \,|\, V(x) = 1\}$ to be around $\mathcal{Y}$.

According to Finsler's lemma, for every $k = \overline{1, M_\mathcal{X}}$ and $l = \overline{1, M_\mathcal{Y}}$, if real-valued matrices $L_{C_k}, \bar{L}_{C_l} \in \mathbb{R}^{m \times q}$, $M_{C_k}, \bar{M}_{C_l} \in \mathbb{R}^{(m+1) \times n}$ exit

Table 1. Number and dimensions of parameter *independent* LMIs of the optimization task.

	Dimension of the LMIs	No. of LMIs
Eq.(5)	m	$M_\mathcal{X} \cdot M_\mathcal{D}$
Eq.(6)	$n + 2p + n^2 + np$	$M_\mathcal{X} \cdot M_\mathcal{D} \cdot M_{\breve{\mathcal{D}}}$
Eq.(7)	$m + 1$	$2M_\mathcal{X} \cdot M_\mathcal{D}$
Eq.(8)	$m + 1$	$2M_\mathcal{Y} \cdot M_\mathcal{D}$

Table 2. Free variables of the optimization task and the number of (scalar) symbolic decision variables they introduce into the optimization task.

Matrix variable	number of (scalar) independent decision variables appearing in the SDP
α	1
P	$\frac{1}{2}\, m(m+1)$
L_b	mq
L_a	$(n + 2p + n^2 + np)(2q + n^2 + nq)$
L_{C_k} and $\bar{L}_{C_l}$	$(M_\mathcal{X} + M_\mathcal{Y}) \times mq$
M_{C_k} and $\bar{M}_{C_l}$	$(M_\mathcal{X} + M_\mathcal{Y}) \times (m + 1)n$

The dimensional parameters are the following:

n	number of state variables (size of x)
P	number of elements in π
m	number of elements in π_b $(n + p)$
q	number of rows in annihilator $\mathcal{N}_{\pi_b}(x, \delta)$
$M_\mathcal{X}$	number of corner points of $\mathcal{X}$
$M_\mathcal{Y}$	number of corner points of $\mathcal{Y}$
$M_\mathcal{D}$	number of corner points of $\mathcal{D}$
$M_{\breve{\mathcal{D}}}$	number of corner points of $\breve{\mathcal{D}}$

$$Q_k^\mathrm{T} P_{C_k}^{(\alpha)}(x, \delta) Q_k \geq 0 \quad \forall (x, \delta) \in \vartheta\left(\mathcal{F}_k^\mathcal{X} \times \mathcal{D}\right) \quad (7)$$

$$\bar{Q}_k^\mathrm{T} \bar{P}_{C_k}^{(1)}(x, \delta) \bar{Q}_k \leq 0 \quad \forall (x, \delta) \in \vartheta\left(\mathcal{F}_k^\mathcal{Y} \times \mathcal{D}\right) \quad (8)$$

then ε_α is inside $\mathcal{X}$ and ε_1 is around $\mathcal{Y}$. Variables $Q_k, \bar{Q}_k$, $\bar{P}_{C_k}^{(1)}$, and $P_{C_k}^{(\alpha)}$ are defined in *Eqs.(3.34)* and *(3.36)* of Ref. [15], $\mathcal{F}_k^\mathcal{X}$ denotes the kth facet of $\mathcal{X}$, furthermore, $M_\mathcal{X}$ denotes the number of facets of $\mathcal{X}$.

The LMI conditions of *Eqs.(5)-(8)* are affine parameter-dependent LMIs, which can be computationally handled by checking their feasibility at the corner points of the polytopic region, on which the parameters $(x, \delta, \dot{\delta})$ are defined. Depending on the number of the corner points, a given number of parameter *independent* LMI conditions (*Table 1*) is obtained.

The SDP task can be summarized as follows. In order to find the maximal invariant level set ε_α of the Lyapunov function, one should maximize α, under the following conditions:

Eq.(5)	$V(x, \delta)$ is positive on $\mathcal{X} \times \mathcal{D}$
Eq.(6)	$\dot{V}(x, \delta, \dot{\delta})$ is negative on $\mathcal{X} \times \mathcal{D} \times \breve{\mathcal{D}}$
Eq.(7)	ε_α lies inside $\mathcal{X}$
Eq.(8)	$\mathcal{Y}$ is inside ε_1

The number of LMIs and their dimensions are given in *Table 1*, furthermore, the free variables of the optimization task are listed in *Table 2*.
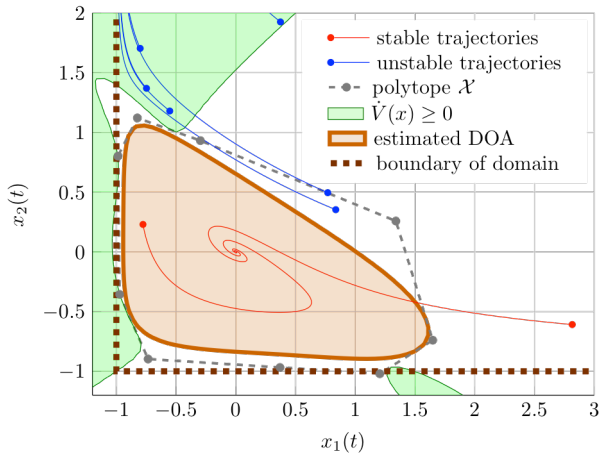
Figure 3. Phase diagram of the 2D locally stable system. Red trajectories converge at the origin, the blue ones do not converge at the origin. The estimated DOA and the used polytope $\mathcal{X}$ are illustrated by the orange-filled region and by the grey-dashed polytope, respectively. The green area highlights the subset of the state space where the Lyapunov function features a positive time-derivative.

3.1. Finding the Most Appropriate Outer Polytope

In the case of 2D systems, polytope $\mathcal{X}$ is evaluated manually through iterations. When having a higher-dimensional system ($n \geq 3$), the iterative procedure [6] is applied that starts from an initial polytope $\mathcal{X}^{(0)}$, then a new polytope $\mathcal{X}^{(1)}$ is defined by enlarging the axis aligned bounding box of the obtained α-level set $\varepsilon_\alpha^{(0)}$. During the iterations, $\mathcal{X}^{(i)}$ is constrained to be a rectangular polytope. *Fig.2* illustrates the operation of a single iteration step.

4. Lotka-Volterra Systems

The N-dimensional LV equation has the form

$$\dot{\bar{x}} = diag(\bar{x} + b), \quad A \in \mathbb{R}^{n \times n}, b \in \mathbb{R}^n, \qquad (9)$$

where $diag(a), a \in \mathbb{R}^n$ denotes an $n \times n$ square diagonal matrix with the $x_1, \ldots, x_n$ on the main diagonal. The system is translated into its, by assumption, unique interior equilibrium point $x^* = -A^{-1}b$ by introducing the centred state vector $x = \bar{x} - x^*$. Then, an autonomous nonlinear system of the form $\dot{x} = \mathcal{A}(x)x$ is obtained, where the matrix function $\mathcal{A}$ can be expressed as: $\mathcal{A}(x) = diag(x + x^*)A$. By applying the LFT and the further algebraic model transformation steps, a model is obtained in the representation of *Eq.(4)*, where the entries of $\pi(x)$ are second order monomials of the state variables.
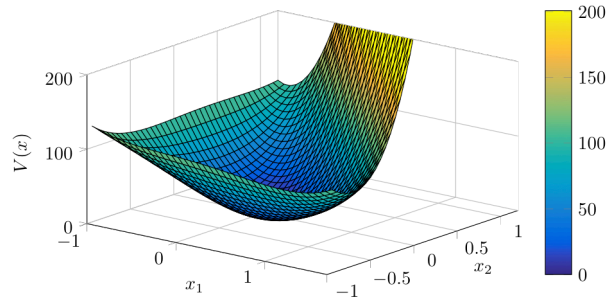

Figure 4. The obtained Lyapunov function in the case of the 2D LV system.

5. Numerical Results

In this section, the applicability of the approach presented above is illustrated through different locally stable Lotka-Volterra models. The results presented here have been computed in a MATLAB environment. For symbolic computations, MATLAB'S built-in Symbolic Math Toolbox was used based on MuPAD. For linear fractional transformations (LFT), the Enhanced LFR-toolbox is used [16, 17]. To model and solve semi-definite optimization (SDP) problems, Mosek solver with YALMIP was used [18].

5.1. 2D Lotka-Volterra System

A locally stable LV system is considered with the model matrix:

$$A = \begin{bmatrix} -2 & -3 \\ 1.4 & 1 \end{bmatrix},$$

with a unique interior equilibrium point $x^* = [1\ 1]^T$. In *Fig.3*, some trajectories of the *centred* system can be observed from different initial conditions. The red trajectories converge at the equilibrium point, the blue ones do not tend to the equilibrium point. After solving the corresponding SDP, the obtained Lyapunov function is $V(x) = \pi_b^T P \pi_b$ (*Fig.4*), where

$$\pi = [x_1^2 \quad x_1 x_2 \quad x_2^2]^T, \pi_b = \begin{bmatrix} x \\ \pi \end{bmatrix} \qquad (10)$$

$$P = \begin{bmatrix} 87.39 & 63.83 & -3.45 & 0 & 0 \\ 63.83 & 165.26 & 67.24 & 100.85 & 28.76 \\ -3.45 & 67.24 & 7.63 & 0.15 & 0 \\ 0 & 100.85 & 0.15 & 38.78 & 39.97 \\ 0 & 28.76 & 0 & 39.97 & 15.53 \end{bmatrix}$$

In *Fig.3*, the shape of the *manually chosen* polytope $\mathcal{X}$, and the obtained invariant α-level set of the Lyapunov function can be seen, which is considered as the estimated DOA (filled orange region). Areas can be observed, where the Lyapunov function's time derivative is positive (green region) can also be observed to be completely outside of $\mathcal{X}$. *Fig.5* illustrates how the value of the Lyapunov function decreases along the trajectories.

ε_α, the invariant α-level set of the Lyapunov function (estimated DOA)

$x^{(i)}(t)$: convergent trajectories, where the larger dots are the initial conditions, the smaller dots are the state values $x^{(i)}(t_k)$ at $t_k \simeq 1, 2, 3$

$v^{(i)}(t) = V(x^{(i)}(t))$, values of the Lyapunov function along the trajectories
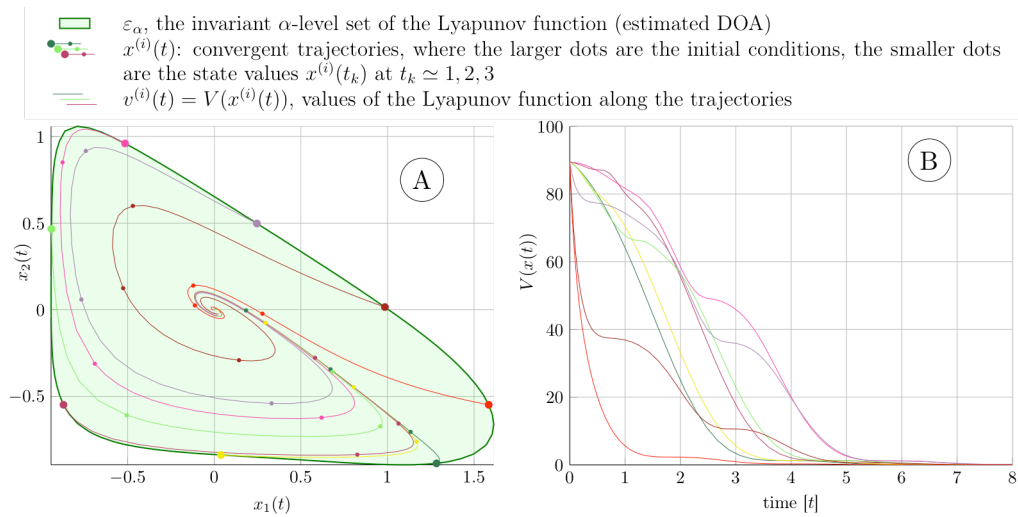
Figure 5. In Panel A, the obtained invariant level set is shown, furthermore some trajectories $x^{(i)}(t)$ of the system can be seen in different colors with initial conditions $x_0^{(i)}$ close to the boundary of the estimated DOA. The values of the Lyapunov function along the trajectories ($v^{(i)}(t)$) are presented in Panel B, respectively.
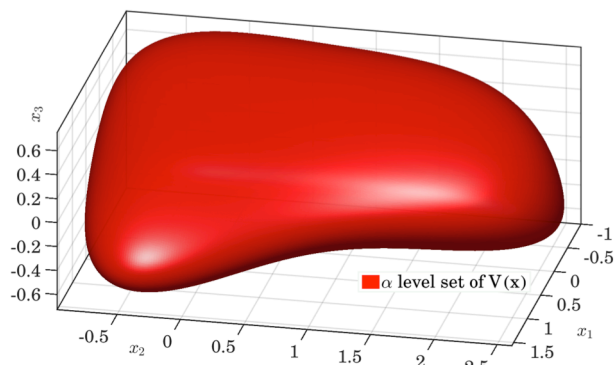
Figure 6. Estimated DOA of the locally stable 3D Lotka-Volterra system (red mesh) and the corresponding final polytope $\mathcal{X}$.

5.2. 3D Lotka-Volterra System

A locally stable 3D LV system has been chosen with the model matrix

$$A = \begin{bmatrix} 0.06 & 0.21 & 0.83 \\ -2.47 & -2.10 & -3.64 \\ 0.06 & 0.47 & -0.45 \end{bmatrix}.$$

Similarly to the 2D example, the equilibrium point is set to $x^* = [\ 1\ 1\ 1\ ]^T$. The monomial set appearing in π_b contains the state variables and every possible 2nd order monomial:

$$\pi^T = [x_1^2 \quad x_1 x_2 \quad x_1 x_3 \quad x_2^2 \quad x_2 x_3 \quad x_3^2] \quad (12)$$

Matrix P is a 9×9 symmetric matrix. *Fig.6* illustrates the invariant α-level set of the obtained Lyapunov function (3D red volume). In *Fig.7*, one can see the cross-sections of the 3D invariant region from three different viewpoints. Some trajectories of the system are shown in red (stable) and blue (unstable).

Figure 7. Cross-sectional view of the 3D estimate along the major planes x_i, x_j and using different values for x_k, where (i, j, k) is a permutation of $(1, 2, 3)$.

Table 3. Number and sizes of the LMIs in the case of *N*-dimensional LV systems.

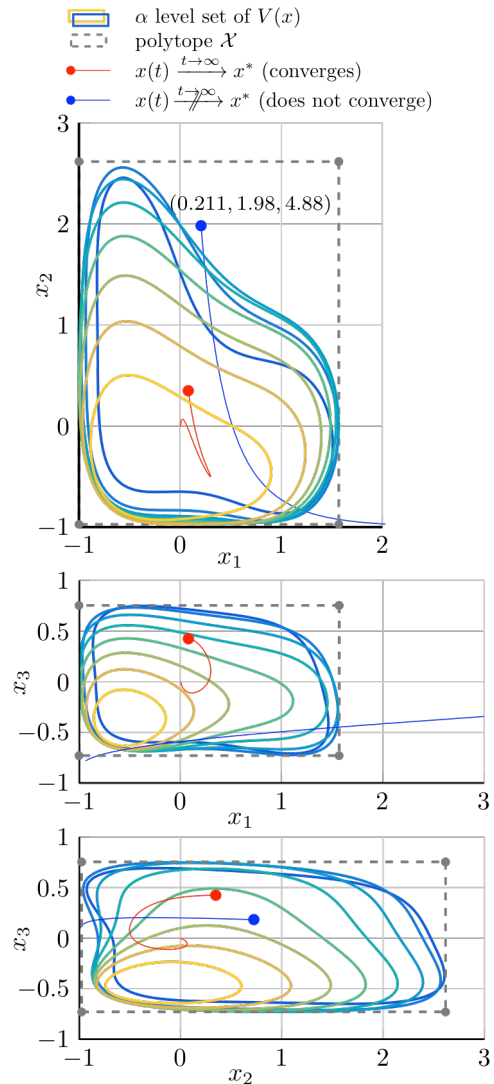| size	number	description
$m_b \times m_b$	2^n	positivity of $V(x)$
$m_a \times m_a$	2^n	negativity of $\dot{V}(x)$
$m_b \times m_b$	$n \cdot 2^n$	$\varepsilon_\alpha \in \mathfrak{Int}(\mathcal{X})$
$m_b \times m_b$	$n \cdot 2^n$	$\mathcal{Y} \in \mathfrak{Int}(\varepsilon_\alpha)$

m_b is the size of the Lyapunov matrix P, $m_a = \frac{1}{2}(n^3 + 5n^2 + 4n)$ is the size of the matrix $P_a(P)$, which appears in the LMI ensuring the negativity of the time derivative of the Lyapunov function.

5.3. 4D Lotka-Volterra System

The method was applied to a locally stable 4D LV system with a unique interior equilibrium point:

$$x^* = [\ 1\ 1\ 1\ 1\ ]^T . \tag{13}$$

The model matrix of the LV system was chosen to be:

$$A = \begin{bmatrix} -4.7126 & -1.5833 & 1.5346 & 2.3230 \\ -9.2461 & -3.1634 & 2.8648 & 2.7796 \\ -14.258 & -4.5477 & 3.6104 & 4.6238 \\ -3.1687 & -0.8016 & 1.2287 & 1.2656 \end{bmatrix} .$$

Due to the fact that model matrix A is a "full" matrix, i.e. there are no zero entries in it, every second-order monomial will appear in the equation of the system. As a consequence, the LFT will produce a model *Eq.(4)*, in which π will contain every possible

second-order monomial. This means that the number of monomials (including the state variables) is $m_b = (n^2 + 3n)/2$ in the case of a "full" model matrix.

In *Table 3*, the number and sizes of the LMIs have been summarised in the case of *N*-dimensional LV systems with a "full" model matrix. The exponential factor 2^n in *Table 3* originates from the rectangular shape of polytope $\mathcal{X}$ possessing 2^n corner points. On the other hand, the feasibility of LMI conditions in *Eqs.(7)* and *(8)* should be checked at each corner point of every facet of the polytope. Furthermore, an *N*-dimensional rectangular polytope has $2n$ facets with 2^{n-1} corner points. As can be seen, a rectangular polytope introduces an exponential increase in the dimension of the given problem. However, *N*-dimensional intervals can be easily handled compared to arbitrarily shaped polytopes defined by (hyper)-triangle meshes. *Fig.8* illustrates the cross sections of the invariant domain along the different pair of axes.

5.4. 2D Uncertain LV System with an Uncertain Equilibrium Point

In this section, the same 2-dimensional system is presented that appeared in Section 5.1 with the same model parameters but with an additional uncertain term $K\delta$, where $K = [1\ 1]^T$. It is assumes, that $\delta \in \mathcal{D} = \left[-0.1, 0.1\right]$ is an unknown *constant* parameter ($\dot{\delta} = 0$). The equation of the uncertain model is the following:

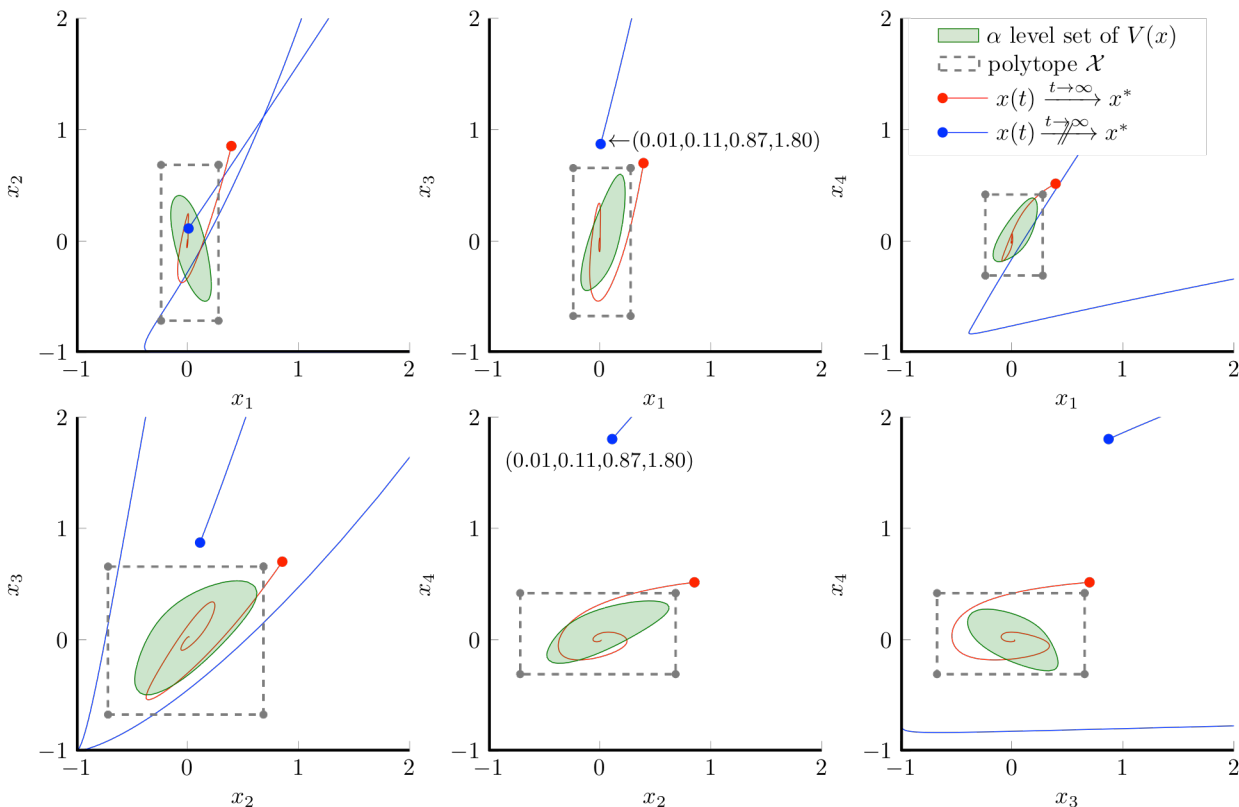$$\dot{x} = diag(\bar{x})(A\bar{x} + K\delta + b) \tag{15}$$

Figure 8. Cross-sectional view of the DOA estimate of the 4D LV system.

The equilibrium point of the uncertain system is an affine function of the uncertain parameter δ:

$$x^*(\delta) = A^{-1}(-K\delta - b) = -A^{-1}K\delta + x_0^* , \quad (16)$$

where x_0^* is the equilibrium point of the system presented in Section 5.1, i.e. when $\delta = 0$. *Eq.(15)* is translated into the following form:

$$\dot{\bar{x}} = diag(\bar{x})A(\bar{x} - x^*(\delta)) . \quad (17)$$

The centred model of the system around the uncertain equilibrium point $x*(\delta)$ can be calculated if the new state vector $x = \bar{x} - x^*(\delta)$ is introduced. The equation of the centred system is the following:

$$\dot{x} = diag\big(x + x^*(\delta)\big)Ax . \quad (18)$$

If the Lyapunov function depends on δ, especially when possessing an uncertain equilibrium point, it is not straightforward to determine an invariant region by considering the α-level set of the Lyapunov function. However, it is possible to compute a region $\mathfrak{I}$, which is "*invariant with respect to*" a larger region $\mathfrak{U}$ ($\mathfrak{I} \subset \mathfrak{U}$), in the sense that every trajectory with an initial condition from $\mathfrak{I}$ will not leave $\mathfrak{U}$.

In order to compute $\mathfrak{I}$ and $\mathfrak{U}$, the α-level set $\bar{\varepsilon}_\alpha(\delta)$ of the obtained Lyapunov function $V(x, \delta)$ is determined in the coordinates system of the *original* system and at the corresponding equilibrium point for every $\delta \in \mathcal{D}$. The overlining in the notation $\bar{\varepsilon}_\alpha(\delta)$ means that this level set is translated into the original coordinates system. As *Fig.9* illustrates, the intersection $\mathfrak{I}$ and the union $\mathfrak{U}$ of the α-level sets obtained for different values of δ have been computed. It can be stated that every trajectory starting from region $\mathfrak{I}$ will not leave region $\mathfrak{U}$.

6. Conclusion

In this work, the Lyapunov function of *N*-dimensional LV systems ($N = 2, 3, 4$) was successfully computed by using the improved optimization-based method [6, 15]. In the case of each deterministic system, an invariant region was given, as the estimated DOA. In the case of an uncertain LV system with an uncertain interior equilibrium point, two regions ($\mathfrak{I} \subset \mathfrak{U}$) were given, which describe the stable regions of the system.

SYMBOLS

$\dot{x}$ denotes the time-derivative of function $x(t)$
$\mathbf{L}^T$ denotes the transpose of a matrix $\mathbf{L}$
$\vartheta(\cdot)$ denotes the corner points of a polytope
$diag(\mathbf{x})$ denotes the diagonal matrix
$x = x(t) \in \mathbb{R}^n$ represents the state variables of a dynamical system

Figure 9. α-level set $\varepsilon_\alpha(\delta_i)$ of the Lyapunov function in the original coordinates system for some values δ_i of the uncertain parameter and the corresponding equilibrium point $x^*(\delta_i)$ of the same color. The red-filled region is the intersection of the level sets for every possible value of δ. The black line shows the contour of the union of the level sets.

$\delta \in \mathbb{R}^n$ vector of uncertain parameters
$\pi \in \mathbb{R}^n$ vector valued function, which represents the set of monomial and rational nonlinear terms to be considered in the Lyapunov function
$\pi \in \mathbb{R}^{m=n+p}$ auxiliary variable denoting $\begin{bmatrix} x \\ \pi \end{bmatrix}$
$V(x)$ denotes the Lyapunov function
$\mathcal{X} \subset \mathbb{R}^n, \mathcal{D}, \tilde{\mathcal{D}} \subset \mathbb{R}^d$ bounded polytopes

Acknowledgement

REFERENCES

[1] Chesi, G.: Domain of Attraction: Analysis and control *via* SOS programming (Springer-Verlag, London, UK) Vol. 415, 2011 DOI: 10.1007/978-0-85729-959-8

[2] Boyd, S.; El Ghaoui, L.; Feron, E., Balakrishnan, V.: Linear matrix inequalities in system and control theory, in Studies in Applied Mathematics (SIAM, Philadelphia, PA, USA) Vol. 15, 1994 DOI: 10.1137/1.9781611970777

[3] Scherer, C.; Weiland, S.: Linear matrix inequalities in control, Lecture Notes (Dutch Institute for Systems and Control, Delft, The Netherlands) 2000 DOI: 10.1201/b10384-28

[4] Trofino, A.; Dezuo, T.J.M.: LMI stability conditions for uncertain rational nonlinear systems, *Int. J. Robust Nonlinear Control* 2013 **24**(18), 3124–3169 DOI: 10.1002/rnc.3047

[5] Polcz, P.; Szederkényi, G.; Péni, T.: An improved method for estimating the domain of attraction of nonlinear systems containing rational functions, *J. Physics*: *Conf. Ser.* 2015 **659**(1), 012038 DOI: 10.1088/1742-6596/659/1/012038

[6] Polcz, P.; Szederkényi, G.; Hangos, K.M.: Computational stability analysis of an uncertain bioreactor model, *Proc. Int. Symp. Stability, Vibration, and Control of Machines and Structures* (SVCS 2016, Budapest, Hungary) pp. 1-12, 2016

[7] Rocha Filho, T.M.; Gléria, I.M.; Figueiredo, A.; Brenig, L.: The Lotka–Volterra canonical format, *Ecol. Mod.* 2005 **183**(1), 95–106 DOI: 10.1016/j.ecolmodel.2004.07.023

[8] Noonburg, V.W.: A neural network modeled by an adaptive Lotka-Volterra system, *SIAM J. Appl. Math.* 1989 **49**(6), 1779–1792 DOI: 10.1137/0149109

[9] Bhargava, S.: Generalized Lotka-Volterra equations and the mechanism of technological substitution, *Technol. Forecast. Soc. Change* 1989 **35**(4), 319–326 DOI: 10.1016/0040-1625(89)90068-1

[10] Goodwin, R.M.: A growth cycle, in Essays in Economic Dynamics (Palgrave, Macmillan, UK) pp. 165–170, 1982 DOI: 10.1007/978-1-349-05504-3

[11] Huu, A.N.; Costa-Lima, B.: Orbits in a stochastic Goodwin–Lotka–Volterra model, *J. Math. Anal. Appl.* 2014 **419**(1), 48–67 DOI: 10.1016/j.jmaa.2014.04.035

[12] Hernández-Bermejo, B.; Fairén, V.: Lotka-Volterra Representation of general nonlinear systems, *Math. Biosci.* 1997 **140**(1), 1–32 DOI: 10.1016/S0025-5564(96)00131-9

[13] Takeuchi, Y.: Global dynamical properties of Lotka-Volterra systems (World Scientific, Singapore) 1996 DOI: 10.1142/9789812830548

[14] Plank, M.: Hamiltonian structures for the *N*-dimensional Lotka–Volterra equations, *J. Math. Phys.* 1995 **36**(7), 3520–3534 DOI: 10.1063/1.530978

[15] Polcz, P.; Szederkényi, G.: An improved method for estimating the domain of attraction of uncertain rational systems using linear matrix inequality conditions and SDP optimization, Technical Report (Faculty of Inform. Techn. and Bionics, Pázmány Péter Catholic Univ., Budapest, Hungary), 2016 users.itk.ppke.hu/~polpe/techreport2016apr.pdf

[16] Hecker, S.; Varga, A.; Magni, J.F.: Enhanced LFR-toolbox for MATLAB, *Proc. IEEE Int. Symp. Computer Aided Control Systems Design* pp. 25–29, 2004 DOI: 10.1109/CACSD.2004.1393845

[17] Magni, J.F.: Linear fractional representation toolbox (Ver. 2.0) for use with MATLAB, 2006 *www.cert.fr/dcsd/idco/perso/~Magni*

[18] Loefberg, J.: YALMIP: a toolbox for modeling and optimization in MATLAB, *Proc. CACSD Conference* (Taipei, Taiwan), 2012 users.isy.liu.se/johanl/yalmip

DIAGNOSIS OF TECHNOLOGICAL SYSTEMS BASED ON THE STRUCTURAL DECOMPOSITION OF THEIR COLOURED PETRI NET MODEL

ANNA I. PÓZNA,[1*] MIKLÓS GERZSON,[1] ADRIEN LEITOLD,[2] AND KATALIN M. HANGOS[1,3]

[1] Department of Electrical Engineering and Information Systems, University of Pannonia, PO Box 158, Veszprém, 8201, HUNGARY
[2] Department of Mathematics, University of Pannonia, PO Box 158, Veszprém, 8201, HUNGARY
[3] Institute for Computer Science and Control, Hungarian Academy of Sciences, PO Box 63, Budapest, 1518, HUNGARY

Diagnosing faults during the operation of a system is an essential task when investigating technological systems. In this paper, a new online fault identification method is proposed which is based on the occurrence graph of the coloured Petri net model of the system. The model is able to simulate the normal and faulty operations of the system given in the form of event lists, so called traces. The diagnosis is based on the search for deviations between the traces of the normal and the actual operations. In the case of complex technological systems, the occurrence graph can contain hundreds of nodes; therefore, the computational effort and searching-time increase significantly. Our proposed structural decomposition method can manage these demands so it has a crucial impact on the practical application of diagnostic processes. The main idea of our method is that the complex systems can be decomposed into technological units. Therefore, the diagnosis can be done by components separately and the diagnostic result of a unit can be used for the diagnosis of the other units connected to it. Because of the structural decomposition, the diagnosis has to be performed on much smaller occurrence graphs but the effect of faults in previous units is taken into account. The proposed method is illustrated by a simple case study.

Keywords: technological system, diagnosis, coloured Petri net model, structural decomposition, qualitative model

1. Introduction

Identifying faults and analysing their consequences are important tasks during the investigation of technological systems. A number of diagnostic methods are known in the literature and the model-based methods are very popular among them [1].

Petri nets and their different extensions (such as coloured, timed or hierarchical nets) are powerful tools for modelling discrete event systems [2]. For example, coloured Petri nets are often used for modelling production lines [3]. It is important that the resultant models describe not only the normal (faultless) operation of the system, but they also take into account different, randomly occurring errors in the system.

In many cases, the normal or faulty operations of technological processes can be characterised by a series of events possessing discrete or qualitative valued variables. In this case, the occurring deviations can be generated by the comparison of the normal and actual events. The occurring faults can be detected and identified based on the observed deviations.

Discrete event systems are usually modelled by automata. In this case, the diagnosis is usually based on the idea of unobservable events [4]. Faults can be modelled as unobservable events, which means only the effects of faults can be noticed. The problem with fault detection is specifying whether any fault has occurred or not in the system. Fault isolation is the problem of identifying which fault has occurred exactly. Since faulty events are unobservable by assumption, the detection and isolation problem must be solved based on the available information of the observed non-faulty events. The diagnosability of discrete event systems was first investigated [5] using the methods of automata theory.

Besides automata, Petri nets are also frequently used for modelling discrete event systems (DES). The structural and mathematical representations of Petri nets both can be used for diagnostic purposes. Methods include various techniques such as analysis of the occurrence graph, marking estimation, linear algebra, integer linear programming, diagnoser nets, or reverse nets.

A simple fault detection method based on the measurement of token quantity is given [6]. It is assumed that the given Petri net is conservative, and any change in the token quantity is caused by faults. If the difference between measured and initial token number

*Correspondence: pozna.anna@virt.uni-pannon.hu

exceeds a predefined threshold then a fault has occurred. Sensor signals are used for token determination instead of modelling the faulty behaviour of the system. The proposed method is very simple and can be used for early fault detection; however, it is not able to isolate faults.

Faults can also be modelled as unobservable transitions in Petri nets. The set of places can be also observable; therefore, the marking of the Petri net has to be estimated. The notion of basis marking [7] (set of markings consistent with the observation) and j-vectors (minimal sequences of unobservable transitions to reach basis markings) are introduced. An online algorithm is developed to detect the occurrence of faults, which uses the basis occurrence graph. The main advantage of the proposed algorithm is that in the case of bounded Petri nets the basis occurrence graph can be computed offline. It reduces the computational effort of the online diagnosis. The basis occurrence graph can be used as an online diagnoser.

Sufficient conditions of diagnosability of faulty transitions are given in the form of a system of inequalities [8]. In this method, the marking of places is observable. Authors introduce the notion of g-markings (markings with negative elements) and unobservable explanations (sequences of unobservable transitions, whose firings can explain the negative elements of a g-marking). After an observed event, the g-marking is updated according to the Petri net equation. When an observed transition fires it removes tokens from its input places and adds tokens to its output places. If this transition is not enabled under the previous g-marking then the removal of tokens causes negative marking. An online fault detection algorithm has also been developed [9], which is based on solving integer linear programming problems and checking the diagnosability conditions.

The integer linear programming approach has also been used [10] to determine if the system behaviour is normal, faulty, or ambiguous. The algorithm has further been improved [11], for a more general situation where different observable transitions can share the same label. Firing times of transitions are also considered, which add more constraints to the ILP problem making the fault detection algorithm more accurate.

Timing characteristics have also been used [12], but with a different meaning: the faults affect the firing speed of the transitions. The fault detection is based on the generation of residuals, which are computed by comparing the markings of observable places with the reference model.

A bottom-up modelling methodology has been proposed using interpreted Petri nets [13]. In the generated model, the faulty and normal states are represented by places. The authors introduce the definition of input-output diagnosability and also give conditions to test this property. The diagnoser model contains the normal behaviour of the system. An online algorithm based on the difference between the system output and the diagnoser model output is developed for detecting faulty markings.

In the case of large systems, the models and associated diagnosers can be extremely large. Furthermore, the diagnostic methods are computationally expensive. Therefore, it is essential to investigate the possibility of distributed diagnosis. The idea of distributed diagnosis is to divide the system into modules or components then make a local diagnoser for each component. The challenging problem is to ascertain the diagnosis state of the whole system from the results of local diagnosers. It usually requires a distributed algorithm and a communication protocol [14].

Coloured Petri nets (CPN) have the advantage of making compact information representations. A CPN diagnoser equivalent to the classical diagnoser has been built [15]. In this approach, places represent different hypotheses and colours represent diagnosis results. The advantage of the CPN diagnoser is the simplified graphical representation. On the other hand, the coloured diagnoser is not necessarily smaller than the classical diagnoser. Decomposition and methods of modular diagnosis of DES are also studied by the authors.

Backward reachability can also be used for diagnosis purposes. If a marking M is reachable from M_0, then M_0 is backward reachable from M in a Petri net. Possible sources of failures for this method can be determined. Backward reachability is extended to coloured Petri nets [16]. Transformation techniques for the inversion of CPN are also presented here.

2. Basic Concepts

A brief description of the basic concepts and notions of our method are given in this section. At first qualitative ranges are introduced to characterise the measured values. After that events, traces, and deviations in a technological system are defined, then the most important parts of the coloured Petri net model and its analysis are introduced. Finally, the structural decomposition-based diagnosis is described in detail.

2.1. Qualitative Range Spaces

In many applications, it is not always necessary to know the exact values of the measured signals. Qualitative models can be used in this case and it is enough to know whether the value of a signal belongs to a specified range space or not. For example, the measurement range of a sensor can be divided into the following range spaces:

$$Q_s = \{\, e_0, 0, L, N, H, e_1 \,\} \tag{1}$$

where $0, L, N, H$ denote the zero, low, normal and high measured value, respectively, while e_0 and e_1 may refer to the extremely low and high values caused by sensor errors. The states of actuators, e.g. valves, switches, etc. can be described similarly. For example, a two-state valve can be represented by
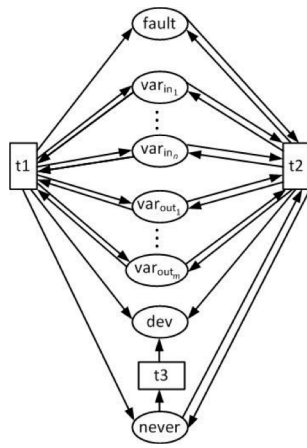
Figure 1. Structure of the Coloured Petri Net Model.

$$Q_v = \{\ op\ ,cl\ \} \tag{2}$$

qualitative range spaces, where *op* and *cl* refer to the open or closed state of the valve.

2.2. Events, Traces, and Deviations

Considering a technological system as a discrete event system, the state of the system can be characterised with the measured values at a given time. The actions in the system, e.g. interactions by operators, modify the values of input and output variables thus the system state changes. An *event* is defined as the arranged (qualitative) input and output values of the system at a given time instance τ:

$$event_\tau = (\tau, in_1, \ldots, in_m, out_1, \ldots, out_n)\ .$$

The course of the system can be described as a sequence of consecutive events, so called *trace*:

$$trace = (event_1, \ldots, event_n).$$

In a technological system the most important types of traces with respect to the diagnosis are the *nominal*, *faulty*, and *characteristic* traces. The nominal trace describes the normal operation of a system. The faulty trace contains the occurring events if a known fault is present while the characteristic trace refers to the actual course of the process. In this paper, it is assumed that only one fault may occur in a process unit of the technological system and this fault evolves before the start of the operation and remains unchanged during the course of the process. If a fault occurs then the trace of the system differs from its nominal trace. As a result, *deviations* between the nominal trace and the current characteristic trace can be defined. In our diagnosis method the following types of deviations are used:

- *never happened*($event_\tau$) (abbreviated as H($event_\tau$)): This type of deviation refers to the events of the nominal trace which ($event_\tau$) never occur in the characteristic trace of the process.

- *chronological deviations*: If an event of the nominal trace ($event_\tau$) happens *later* or *earlier* in the characteristic trace than time point τ, the deviations LAT($event_\tau$) and EAR($event_\tau$) denote them.
- *quantitative deviations*: This type of deviations is used to denote that the ith output value is greater (denoted by $GRE_i(event_\tau)$) or smaller ($SML_i(event_\tau)$) in the characteristic event at time τ than in the nominal event while the input values are identical.

Our diagnosis method is based on the search and comparison of the deviation list on the reachability graph of the CPN model of the technological system.

2.3. Coloured Petri Nets

Coloured Petri nets (CPNs) are extensions of the ordinary or low-level Petri nets. The main differences with respect to ordinary Petri nets are that so-called colours can be assigned to tokens and functions can be assigned to arcs and transitions, too. The detailed formal definition is given in Ref.[17], only the special concepts used in our models are presented here.

- *Places* of the CPN model of the technological system may have three functions. At first input and output variables are represented by places and the colour of the tokens on them denotes the qualitative value of the variable at the current time. On the other hand, places may refer to the occurred fault and the generated deviations. The colours of tokens in these places denote the type of the fault and the occurring deviations, respectively.
- There are three *transitions* in our model, which have different tasks. Transition t_1 is responsible for the generation of faulty or normal operations at the beginning of the process and the initialisation of the variables according to the investigated operational mode. The function of transition t_2 is the timing of the process. It is assumed that the technological process is time-driven and the values of the variables change at the end of the time steps. Therefore, t_2 fires until the end of the process. Transition t_3 is used for the generation of the 'never happened'-type deviations at the end of the process.
- *Arc functions* are assigned to the arcs between places and transitions defining the change in the colours and computing deviations.

The structure of the coloured Petri net for modelling and diagnosing technological systems can be seen in *Fig.1*. Places are represented with ellipses and transitions are represented with rectangles. In a technological system, the consequences of a processing step can be stochastic. For example, the step may be completed in a normal way, or a fault occurs. The probabilistic nature of a transition *t* associated with a processing step can be modelled in a CPN by a fault function, which is built into its guard function. This

fault function returns the logical value true or false with predefined probability, and the token values of the adjacent consequence places of transition t can be controlled by this logical value. This type of transition firing is called a *stochastically fired transition*.

The *occurrence graph* of the CPN is a graph, which contains all of the system states reachable from a given initial state [17]. Assuming that the CPN model of the examined system is given, the occurrence graph can be used for its behavioural analysis. The nodes of the graph refer to system states and the arcs connecting them refer to state transitions, e.g. events. Different paths on the occurrence graph refer to different operational modes of the system and they can be used for analysing the causes and consequences of a system state.

2.4. Diagnosis with Structural Decomposition

The disadvantage of the occurrence graph-based method is the increasing size of the graph as the size, i.e. number of places, of the CPN model increases. In the case of even a simpler technological system containing three or four units, the occurrence graph of its CPN model can contain hundreds of nodes depending on the number of sensors and actuators. The refinement of the qualitative measuring range of sensors or the application of control valves instead of two-state actuators may also cause the growth of occurrence graphs because their branches will be longer. With the growth in the size of the graph, the computational effort and searching-time also increase. This is the reason why the structural decomposition has a crucial impact on the practical application of the diagnostic process.

Generally speaking, complex systems can be decomposed into technological units. By taking advantage of this, the diagnosis can be done by components separately, and the diagnostic result of a unit can be used for the diagnosis of the other units connected to it.

To perform diagnosis with structural decomposition the full trace of the system should be decomposed, too. To do this, first the time instances belonging to the operation of the investigated units have to be selected. Then the variables referring to this unit are picked out from the events belonging to the selected time instances. If the trace is represented in tabular form then specific rows and columns should be selected. Afterwards, time is shifted back in the case of every unit such that the initial time step of the first event should be 1 in every sub-trace. This means that every unit has its own relative time-scale. By applying this decomposing process, the trace describing the operation of the entire system is disintegrated into the event lists referring to the operation of simpler technological units. As a next step the deviation list of the given subsystem is generated by comparing the nominal trace of the subsystem with its characteristic trace. Then the diagnosis is performed using the CPN model and occurrence graph of the subsystem. The task is to compare the generated deviation list with the token distribution of the terminal nodes on the occurrence graph. If the deviation list corresponds with the token distribution of exactly one terminal node then the fault can be determined based on the token of the fault place. If the deviation list matches the token distribution of more than one terminal node then only the set of possibly occurred faults could be determined. If the deviation list cannot be found in the token distribution of any terminal nodes then an unknown failure occurs in the system.

In the case of complex systems, it is necessary to take into account the effect of faults that have occurred in subsystems connected to the diagnosed unit. Therefore, the CPN model of the units has to be modified such that the place of the *fault* has to contain not only the actual operating mode, but the operating modes of previous subsystems, too. To store this information the colour set of this place is extended with an attribute referring to the type of fault and to the place where the fault occurred.

Let us assume that one fault is diagnosed in the first technological unit. This information is added to the *fault* place of the next unit as a previous fault. Then the occurrence graph of this subsystem is generated based on its CPN model where the fault of the previous unit appears as an initial condition. The resultant occurrence graph now contains those states and deviation lists which can occur in this subsystem if the fault of the previous unit is taken into account. The diagnosis is performed on this graph, the possible fault of this unit is determined based on this investigation, and the result is taken into account during the diagnosis of the following unit. In certain cases, the exact type of fault cannot be determined exactly. If the result of the diagnosis of the unit is a set of possible faults then each element of the set is treated separately. This means that the diagnosis of the next unit has to be performed taking into account every one of the possible previous faults. Occurrence graphs of the subsystem are generated according to each fault of the previous subsystem. The result of the diagnosis is the union of the obtained faults.

The main advantage of the described method is the smaller size of the occurrence graphs of subsystems. Therefore, the search requires less time than in the case of the investigation of the entire technological system.

3. Simple Case Studies

A simple case study is presented in this section as a practical illustration of the diagnosis of complex technological systems based on their structural decompositions.

3.1. Description of the Technological Process

Our simple technological example contains three uniform tanks, *TA*, *TB*, and *TC*, which are serially connected as can be seen in *Fig.2*. Each tank has an input valve, an output valve (denoted by Vx, where $x =$ A, B, C, and D), and a continuous level sensor (*lev_x*, x
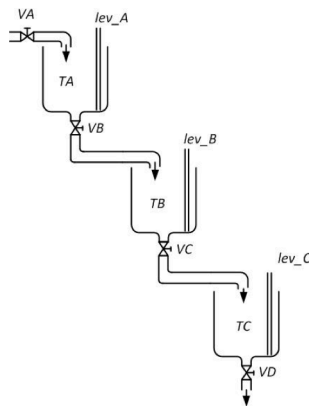
Figure 2. The investigated technological example.

= A, B, and C). The short description of the technological process is as follows: as an initialisation, it is assumed that all the valves are closed. The process starts with the opening of the first valve (*VA*) and then the filling of the first tank (*TA*) starts. The flow of liquid is considered constant so the control of the filling process is based on time. The role of the level sensors is to measure the actual liquid level only. At time step 3, the filling process of tank *TA* completes and its output valve (*VB*) is opened. The second tank *TB* is filled the same way as described for tank *TA*. At time step 5, the *TB* tank is full and its output valve (*VC*) is opened. The filling process of the third tank (*TC*) happens the same way. During the filling of the second and third tanks, the first (*TA*) then the second tank (*TB*) operates as a continuous unit. After the filling of the third tank (*TC*) has completed, the technological systems work in continuous mode.

It is assumed that five possible faults can occur in each tank:

- *2 faults of the level sensor*: negative or positive bias error. In this case, the sensor signal is lower or higher with qualitative unit than the actual value of the level.
- *leaking of the tank*: it is assumed that all of the incoming liquid runs out through the hole, i.e. it is a "big" hole.
- *2 different faults of the output valves VB, VC, and VD*: besides their normal operations, they can stay closed or open only halfway.

For the sake of simplicity, only one of these faults can occur with each tank and all of the faults evolve before the filling process of the first tank *TA* starts.

3.2. CPN Model of the System

For the diagnostic investigation, a coloured Petri net-based (CPN) model of the technological system is developed as follows: the system can be decomposed into three subsystems. Each subsystem represents a tank together with its input and output valves. As can be seen from the technological description of the system, all the subsystems (tanks with their input and output valves and sensors) work in a very similar way, so the structure of their CPN models is identical. The CPN models were
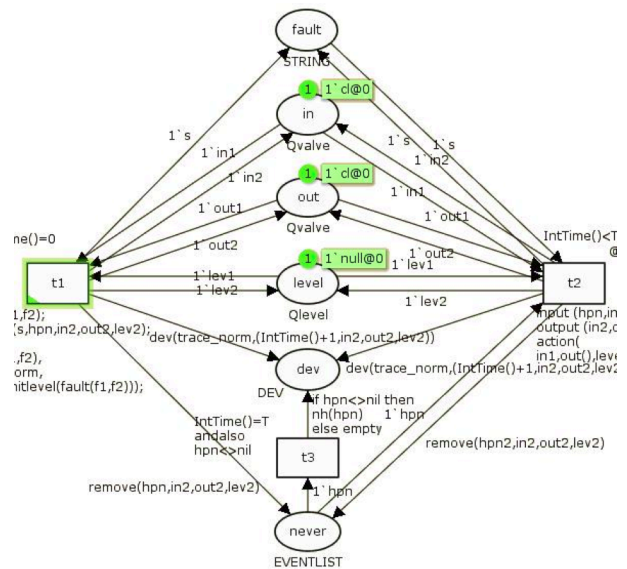

Figure 3. The coloured Petri net model of a tank (for the sake of clarity parts of some inscriptions were omitted).

built using CPN Tools 3.4.0. The CPN model of a subsystem can be seen in *Fig.3*.

The description of the CPN model is as follows: the locations of the Petri nets represent input and output variables, i.e. the state of the input and output valves (denoted by *in* and *out*), and measured level value (*level*), respectively. The qualitative values of these variables are represented by different colour sets.

- colour set of valves: $Q_{valve} = \{cl, op\}$, which represents the closed or open-state of the two-state valve.
- colour set of level sensors: $Q_{level} = \{ e_0, 0, L, N, H, e_1 \}$, where $0, L, N, H$ denotes that the level is zero, low, normal or high, respectively, and e_0, and e_1 indicate that the level is below or above the measurement range, respectively.

Three additional places are needed: one to store the operating mode (place *fault*), one to collect the deviations (place *dev*) and one for the list of events that have not occurred until a given simulation time step (place *never*). The CPN model contains three transitions (t_1, t_2, t_3). t_1 is the initialisation transition, it fires only once at the start of the simulation. It generates an operating mode (normal or faulty) and updates the variables according to the generated operating mode. Afterwards, transition t_2 fires until the end of the process. It updates the values of variables in every simulation step and generates quantitative and chronological deviations except for the '*never happened*'-type. The '*never happened*'-type deviations can be generated after the process has ended. This is done by the firing of transition t_3, which removes the events that remain at place *never* at the end of the simulation and attaches the *NH* guideword to them. The values of variables at a given time step can be read from the trace files. Each trace file contains the list of events that describe the process according to the operating mode. All traces were generated by a MATLAB script.
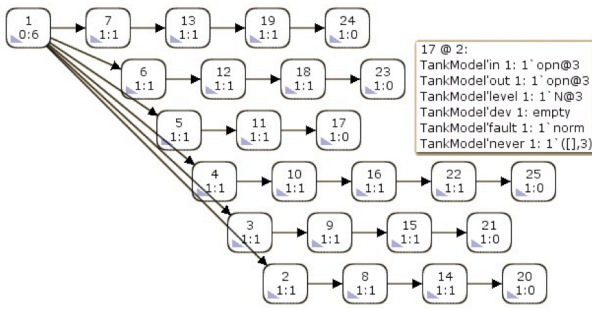
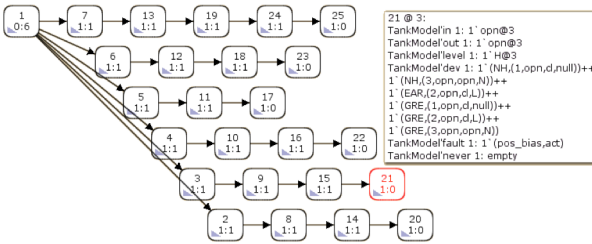*Figure 4*. The occurrence graph of a CPN tank model.

Figure 5. The occurrence graph of the first tank.

3.3. Diagnosis of the System with Structural Decomposition

According to the general description, the technological system consists of three uniform tanks. Because of the same structure and operational mode, all three tanks have the same CPN model. This model can be seen in *Fig.3*. The occurrence graph of the CPN model of a tank can be seen in *Fig.4*. At the time of the generation of this occurrence graph, any fault occurring in a previous technological unit was not taken into account. As can be seen in *Fig.4*, node No. 17 refers to normal operation of the system, while the other terminal nodes (No. 20, 21, 23–25) belong to the five different faulty modes.

The nominal trace of the complex technological system can be seen in *Table 1*. The input variables of the first tank (tank *TA* in *Fig.2*) are the states of valves *VA* and *VB*, while the output variable is the value of level sensor *lev_A*. The variables of the second tank are *VB*, *VC*, and *lev_B*, while *VC*, *VD*, and *lev_C* belong to the third tank, respectively.

The rows belong to time steps 1, 2, and 3, while columns *VA*, *VB*, and *lev_A* compose the trace of the first tank. These cells are framed with a dotted line in *Table 1*. Similarly, rows 3, 4, and 5 as well as columns *VB*, *VC*, and *lev_B* define the trace of the second tank (framed with a continuous line) while rows 5, 6, and 7, along with columns *VC*, *VD*, and *lev_C* give the trace of the third tank (framed with a dashed line). Consider the characteristic trace of the technological system given in *Table 2*.

As a next step, the trace pieces belonging to each individual tank are removed from the characteristic trace of the entire system. The initial time step is shifted to 1 for every unit. The resultant event lists belonging to the three tanks can be seen in *Table 3*.

The diagnostic process is started with the first tank. By comparing the nominal trace of the first tank with the characteristic trace (first column of Table 3) the deviation list is generated. This deviation list is then

Table 1. Decomposition of the nominal trace. Dotted line: first tank, continuous line: second tank, dashed line: third tank.

time	input variables				output variables		
	VA	*VB*	*VC*	*VD*	*lev_A*	*lev_B*	*lev_C*
1	*op*	*cl*	*cl*	*cl*	0	0	0
2	*op*	*cl*	*cl*	*cl*	L	0	0
3	*op*	*op*	*cl*	*cl*	N	0	0
4	*op*	*op*	*cl*	*cl*	N	L	0
5	*op*	*op*	*op*	*cl*	N	N	0
6	*op*	*op*	*op*	*cl*	N	N	L
7	*op*	*op*	*op*	*op*	N	N	N

Table 2. Decomposition of the characteristic trace. Dotted line: first tank, continuous line: second tank, dashed line: third tank.

time	input variables				output variables		
	VA	*VB*	*VC*	*VD*	*lev_A*	*lev_B*	*lev_C*
1	*op*	*cl*	*cl*	*cl*	L	0	0
2	*op*	*cl*	*cl*	*cl*	N	0	0
3	*op*	*op*	*cl*	*cl*	N	0	0
4	*op*	*op*	*cl*	*cl*	N	0	0
5	*op*	*op*	*op*	*cl*	N	0	0
6	*op*	*op*	*op*	*cl*	N	0	0
7	*op*	*op*	*op*	*op*	N	0	0

Table 3. Characteristic traces of the three tanks after decomposition.

TA	TB	TC
$(1, op, cl, L)$	$(1, op, cl, 0)$	$(1, op, cl, 0)$
$(2, op, cl, N)$	$(2, op, cl, 0)$	$(2, op, cl, 0)$
$(3, op, op, N)$	$(3, op, op, 0)$	$(3, op, op, 0)$

searched for among the terminal nodes of the occurrence graph of the first tank. (This occurrence graph can be seen in *Fig.5*). It can be stated that terminal node No. 21 contains the same deviation list and based on the token of the *fault* place the type of fault can be determined: the level sensor exhibits a positive failure bias in the first tank.

The diagnosed fault in the first tank is used during the investigation of the second tank. This fault is added to the place *fault* as a token (*pos_bias*, *prev1*) in the model of the second tank. Then the occurrence graph of the second tank is generated which contains those states that can occur in the second tank if the sensor of the first tank exhibits a positive bias error. The resultant graph can be seen in *Fig.6*.

The deviation list of the second tank is generated by comparing the second column of *Table 3* with the characteristic trace of the second tank. By checking the terminal nodes of the occurrence graph in *Fig.6*, it can be stated that terminal node No. 24 exhibits the same deviation list. This means that the fault of the second tank is leakage and it can be identified unambiguously.

The diagnosed faults of the first and second tanks are added to the model of the third tank in the form of tokens (*pos_bias*, *prev1*) and (*leak*, *prev2*) belonging to the *fault* place. Based on this information the occurrence graph of the third tank is generated in accordance with *Fig.7*. The nodes on this occurrence graph refer to the states if a positive failure bias occurs in the first tank and leak in the second tank.
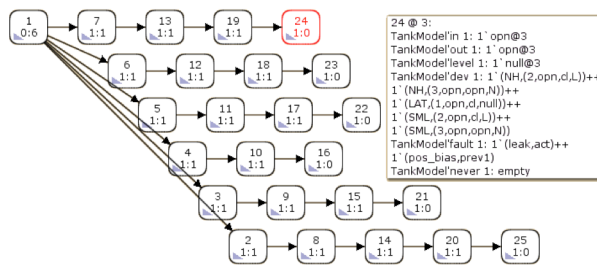
Figure 6. The occurrence graph of the second tank in the case of a positive failure bias in the first tank.
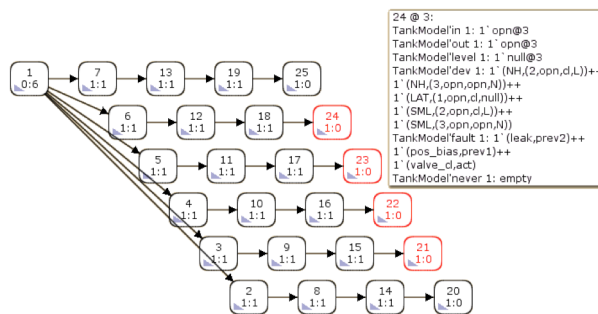

Figure 7. The occurrence graph of the third tank in the case of a positive failure bias in the first tank and leak in the second tank.

The deviation list based on the trace pieces stemming from the filling process of the third tank (see the third column of *Table 3*) is generated and compared to the terminal node of the occurrence graph. It can be stated that terminal nodes No. 21–24 possesses the same deviation list as the deviation list obtained from the characteristic trace of the third tank. This means that the operating mode of the third tank cannot be unambiguously determined, the set of possible operating modes, i.e. normal, leak, or fault, of valves can be defined.

4. Conclusion

A novel method for online fault diagnosis in a technological system is described in this paper. The method is based on the structural decomposition of a complex technological system. The process starts with the modelling of the technological system in the form of coloured Petri nets. For the characterisation of sensor values and actuator states, qualitative value sets are used in the form of coloured tokens. This modelling method allows for the simulation of both normal and known faulty operations of the system. The diagnosis is performed using the occurrence graph of the basic units of the complex system. By generating the deviation list based on the normal and characteristic traces the fault or the set of possible faults can be determined. As a result of the structural decomposition, the diagnosis has to be performed on much smaller occurrence graphs but the effect of faults in previous units are taken into account. Our method reduces the demand of computational efforts and search time. The proposed method was illustrated by a simple case study.

Acknowledgement

REFERENCES

[1] Blanke, M.; Kinnaert, M.; Lunze, J.; Staroswiecki, M.: Diagnosis and fault-tolerant control (Springer-Verlag, Berlin, GERMANY) 2006 DOI: 10.1007/978-1-84628-877-7

[2] Hrúz, B.; Zhou, M.: Modeling and control of discrete-event dynamic systems with Petri nets and other tool (Springer-Verlag, London, U.K.) 2007

[3] Campos, E.J.; Seatzu, C.; Xie, X.: Formal methods in manufacturing (CRC Press Taylor and Francis Group, Boca Raton, USA) 2014 DOI: 10.1201/b16529

[4] Zaytoon, J.; Lafortune, S.: Overview of fault diagnosis methods for discrete event systems, *Ann. Rev. Control* 2013 **37**(2), 308–320 DOI: 10.1016/j.arcontrol.2013.09.009

[5] Sampath, M.; Sengupta, R.; Lafortune, S.; Sinnamohideen, K.; Teneketzis, D.: Diagnosability of discrete-event systems, *IEEE Trans. Automat. Control* 1995 **40**(9), 1555–1575 DOI: 10.1109/9.412626

[6] Prock, J.: A new technique for fault detection using Petri nets, *Automatica* 1991 **27**(2), 239–245 DOI: 10.1016/0005-1098(91)90074-C

[7] Cabasino, M.P.; Giua, A.; Seatzu, C.: Fault detection for discrete event systems using petri nets with unobservable transitions, *Automatica* 2010 **46**(9), 1531–1539 DOI: 10.1016/j.automatica.2010.06.013

[8] Basile, F.; Chiacchio, P.; Tommasi, G.D.: Sufficient conditions for diagnosability of Petri nets, in *Proc. 9th Int. Workshop on Discrete Event Systems, WODES* (Göteborg, SWEDEN) pp. 370–375, 2008 DOI: 10.1109/WODES.2008.4605974

[9] Basile, F.; Chiacchio, P.; Tommasi, G.D.: An efficient approach for online diagnosis of discrete event systems, *IEEE Trans. Automat. Control* 2009 **54**(4), 748–759 DOI: 10.1109/TAC.2009.2014932

[10] Dotoli, M.; Fanti, M.P.; Mangini, A.M.; Ukovich, W.: Online fault detection in discrete event systems by Petri nets and integer linear programming, *Automatica* 2009 **45**(11), 2665–2672 DOI: 10.1016/j.automatica.2009.07.021

[11] Fanti, M.P.; Mangini, A.M.; Ukovich, W.: Fault Detection by Labeled Petri Nets and Time Constraints, *Proc. 3rd Int. Workshop on Dependable Control of Discrete Systems* (DCDS, Saarbrucken, GERMANY) pp. 168–173, 2011 DOI: 10.1109/DCDS.2011.5970336

[12] Lefebvre, D.; Aguayo-Lara, E.: Initial study for observers application to fault detection and isolation with continuous timed petri nets, *IFAC-PapersOnLine* 2015 **48**(7), 97–103 DOI: 10.1016/j.ifacol.2015.06.479

[13] Ramirez-Trevino, A.; Ruiz-Beltran, E.; Rivera-Rangel, I.; Lopez-Mellado, E.: Online fault diagnosis of discrete event systems. A Petri net-based approach, *IEEE Trans. Automation Sci. Engng.* 2007 **4**(1), 31–39 DOI: 10.1109/TASE.2006.872120

[14] Genc, S.; Lafortune, S.: Distributed diagnosis of place-bordered Petri nets, *IEEE Trans. Automation Sci. Engng.* 2007 **4**(2), 206–219 DOI: 10.1109/TASE.2006.879916

[15] Pencolé, Y.; Pichard, R.; Fernbach, P.: Modular fault diagnosis in discrete-event systems with a CPN diagnoser, *IFAC-PapersOnLine* 2015 **48**(21), 470–475 DOI: 10.1016/j.ifacol.2015.09.571

[16] Bouali, M.; Barger, P.; Schon, W.: Backward Reachability of colored Petri nets for systems diagnosis, *Reliability Engng. System Safety* 2012 **99**, 1–14 DOI: 10.1016/j.ress.2011.10.003

[17] Jensen, K.: Coloured Petri nets: Basic concepts, analysis methods and practical use (Springer-Verlag, Berlin, GERMANY) 1997 DOI: 10.1007/978-3-642-60794-3

METAL-CHLORIDE-ACTIVATED EMPTY FRUIT-BUNCH CARBONS FOR RHODAMINE B REMOVAL

MUHAMMAD ABBAS AHMAD ZAINI[1,2]* AND MOHD SHAFIQ MOHD SHAID[1]

[1] Faculty of Chemical & Energy Engineering, Universiti Teknologi Malaysia, Johor Bahru, Johor, 81310, MALAYSIA
[2] Centre of Lipids Engineering & Applied Research, Ibnu Sina Institute for Scientific & Industrial Research, Universiti Teknologi Malaysia, Johor Bahru, Johor, 81310, MALAYSIA

This study aimed to investigate the adsorptive ability of activated carbons derived from empty boil palm fruit-bunch carbons through metal-chloride activation. The derived activated carbons were characterized in terms of yield, pH, surface functional groups, and specific surface area. Rhodamine B dye was used as a pollutant probe to evaluate the performance of activated carbons. Results show that empty, $ZnCl_2$-activated fruit-bunch carbon exhibits a higher surface area of 866 m^2 g^{-1} and a Rhodamine B removal yield of 105 mg g^{-1}. Activation at the same temperature of 600 ºC using the recovered $FeCl_2$ yields an activated carbon with nearly twice the surface area compared to the fresh one. A direct correlation was obtained between the roles of the specific surface area and removal of Rhodamine B. Empty fruit-bunch carbon is a promising adsorbent precursor for colour removal from water.

Keywords: activated carbon, chemical activation, empty fruit-bunch, metal-chloride, Rhodamine B

1. Introduction

Malaysia and Indonesia are the leading producers worldwide of palm oil. Due to the growth in palm oil production and despite the high economic returns, this industry also generates a huge amount of was that has negative implications on the environment. As one of the largest oil palm producers in the world, Malaysia generates abundant empty fruit-bunch (EFB) residues amounting to 12.4 million tonnes annually [1].

At present, only a small quantity of EFB is used as fuel for boilers in the oil palm mills while the remaining large quantity is left to decay in fields or disposed of in landfills. Due to the limited area of landfill sites and other associated environmental implications, the quest for the effective utilization of EFB has become a subject of significant interest. Because EFB is rich in carbon and lignin, it has a great potential to be converted into adsorbent or activated carbon for a variety of purification and environmental purposes [2].

Nasir *et alia* [3] reported the selectivity of methylene blue removal over copper(II) ions by empty fruit-bunch. Yet, the uptake capacity was too small (32.3 mg methylene blue per g of EFB) to warrant large-scale adsorption operation. In a recent related work, Wirasnita *et alia* [4] reported the preparation of $ZnCl_2$-activated empty fruit-bunch carbon with a specific surface area of 86.6 m^2 g^{-1}. However, the surface area obtained is somewhat small to be regarded as that of activated carbon [5]. There are also reports on the use of CO_2 and steam to physically activate the EFB into activated carbons. However, the surface area is often not controllable and the yield is very low because of high activation temperatures (800 to 900 ºC) [6,7].

Our present work aimed to evaluate the adsorptive characteristics of metal-chloride-activated empty fruit-bunch carbons. Zinc(II) chloride and iron(II) chloride were used for the chemical activation of EFB. Attempts were also made to recover the activator for the second activation. Rhodamine B was used as a model pollutant to establish the adsorption data. The activated carbons were characterized and the adsorptive results compared and discussed.

2. Experimental

Oil palm empty fruit-bunch (EFB) was obtained from Sungei Kahang palm oil factory in Johore State, Malaysia. The material was oven-dried at 110ºC overnight to remove moisture. All chemicals used in the preparation of activated carbons and adsorption were of analytical reagent grade.

2.1. Preparation of Activated Carbon

Empty fruit-bunch was loosened and separately modified with zinc chloride and iron(II) chloride in the mass ratio (activator : EFB) of 3:2. Firstly, the activator was dissolved in water, sufficient for the EFB to be immersed. Next, the EFB-activator mixtures were

stirred at 90 °C for 40 minutes. After that, the sample was placed in the oven overnight at 110 °C for impregnation. The impregnated sample was put in a crucible wrapped in aluminium foil, and heated in a furnace for 1.5 h at 300 °C and 550 °C for zinc chloride activation and iron(II) chloride activation, respectively. The selected activation temperatures are half of the boiling points of the activators. The resultant activated carbon was washed with distilled water, and the washed water was used for the second activation using the same impregnation ratio of 3:2 at 550 °C for both activators. Activated carbons were designated as Z1 and F1 for activation using fresh zinc(II) chloride and iron(II) chloride, respectively, and Z2 and F2 when using recovered activators, respectively.

2.2. Characterization of Activated Carbon

The yield of activated carbon was calculated from the mass of the resultant product over that of the dried EFB used. The adsorbent pH was determined by soaking 1 g of adsorbent in 100 cm^3 of distilled water. The pH was measured using a pH meter (HI 8424, Hanna Instruments). The specific surface area of activated carbon was measured using a surface area analyser (Pulse ChemiSorb 2705, Micromeritics) at the temperature of liquid N$_2$, 77 K. Fourier transform infrared spectroscopy (FTIR) (IR Tracer-100, Shimadzu) was used to obtain the peaks of functional groups at specific wave numbers ranging from 400 to 4000 cm^{-1}.

2.3. Adsorption of Rhodamine B

Rhodamine B of commercial purity was utilized without further purification. 500 mg of Rhodamine B powder was weighed using an analytical balance and then dissolved in 1 dm^3 of distilled water in a volumetric flask to make up a stock solution. The dilution of stock solution was needed for preparing the working concentrations for adsorption.

Adsorption was performed according to the bottle-point technique. Briefly, about 0.5 g of activated carbon was added to 50 cm^3 of Rhodamine B solution at varying concentrations. The mixture was allowed to equilibrate on an orbital shaker at 120 rpm and room temperature for 72 h. The residual concentration was measured using a spectrophotometer (Halo VIS-10, Dynamica Scientific Ltd.) at a wavelength of 555 nm (absorption unit = 0.014 × concentration, R^2 = 0.99). The adsorption capacity (mg g^{-1}) was calculated by a simple material balance (*Eq.(1)*), and the adsorption data were analysed by general isotherm models, namely Langmuir (*Eq.(2)*) and Freundlich (*Eq.(3)*). The respective constants were solved by non-linear regression using *Solver* as implemented in MS Excel for the smallest sum-of-squared error (*SSE*) and optimum coefficient of determination (R^2).

$$q_e = \frac{(C_0 - C_e)}{m} V \qquad (1)$$

Table 1. Properties of activated carbons.

activator	type	temp. (°C)	yield (%)	pH	surface area (m^2 g^{-1})
ZnCl$_2$	Fresh (Z1)	300	44.0	3.8	2.64
	Recovered (Z2)	550	41.2	4.9	866
FeCl$_2$	Fresh (F1)	550	46.8	4.6	98.4
	Recovered (F2)	550	37.9	5.3	226

$$q_e = \frac{QbC_e}{1 + bC_e} \qquad (2)$$

$$q_e = K_F C_e^{\frac{1}{n}} \qquad (3)$$

where, q_e (mg g^{-1}) is the adsorption capacity of Rhodamine B, C_o and C_e (mg dm^{-3}) are the initial and equilibrium concentrations, respectively, m (g) is the mass of activated carbon, and V (dm^3) is the volume of Rhodamine B solution. Constant Q (mg g^{-1}) is the maximum monolayer capacity, b (dm^3 g^{-1}) is the adsorption intensity, and K_F ((mg g^{-1})(dm^3 mg^{-1})$^{1/n}$) and $1/n$ are the Freundlich constants related to the adsorption capacity and intensity, respectively.

3. Results and Analysis

3.1. Characteristics of Activated Carbons

Chemical activation using freshly prepared metal-chlorides was performed at a temperature of about half of the boiling point of each activator. The boiling point of ZnCl$_2$ and FeCl$_2$ are 723 and 1023 °C, respectively. This is to allow a sufficient amount of activator to be recovered upon activation as it is commonly understood that chemical activators, e.g. ZnCl$_2$, KOH, etc., are prone to intercalate with the matrix material and/or lost when the activation is done close to the boiling point of the activators. In addition, it enables one to evaluate the effectiveness of activated carbon preparation at the selected temperatures. *Table 1* displays the yield, pH, and specific surface area of metal-chloride-activated empty fruit-bunch carbons.

In general, the yield of activated carbons ranges between 38 and 47%. This indicates the underlying role of metal-chlorides as dehydrating agents to enhance the burning off the carbonaceous material. This is also true for ZnCl$_2$ activation at a temperature of 300 °C. In addition, it signifies that a significant portion of metal-chlorides could be recovered for the subsequent activation. The pH values of activated carbons are in the range of 3.8 to 5.3. The activated carbons are slightly acidic because both metal-chloride activators are Lewis acids. The pH values were found to increase when the recovered activators were used in the activation. This could be attributed to the decreased amount of metal-chlorides that probably could not be fully recovered after the first activation.

From *Table 1*, the developed surface area of Z1 upon activation is undeniably small, even smaller than for raw EFB (28.4 m^2 g^{-1}) [3]. This could be due to the
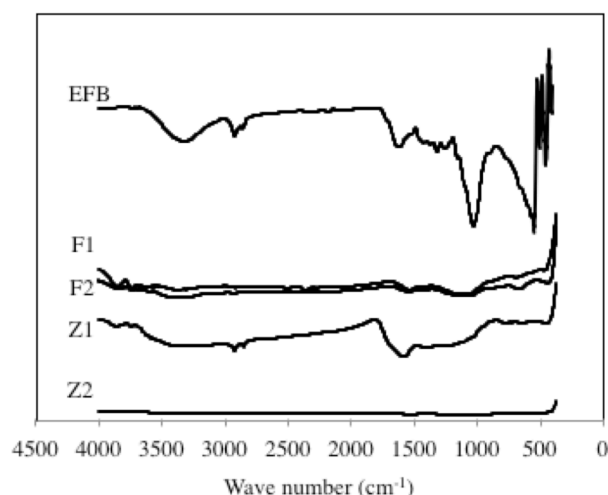
Figure 1. FTIR spectra of EFB and activated carbons.

Table 2. Functional groups used in characterizing samples.

Wave number (cm⁻¹)	Functional group	Type	Sample
3100–3550 [a]	O—H	alcohol	EFB, F1, F2, Z1
1000	C—O		
2900, 2810	C—H	aldehyde	EFB, Z1
1600, 1470	C=C	aromatic	EFB, Z1
1600	C=C	alkene	EFB, Z1, F1, F2
3300	N—H	amine	EFB
1300	C—N		

[a] broad feature

blockade of existing pore channels because of the intercalation of zinc cations within the material matrix. Intercalation of a chemical activator normally creates new pathways for the porous structure when adequate heat is supplied to the impregnated material. This often results in an increase in pore volume, thus increasing the specific surface area. Yet, the activation temperature for Z1 (300 ℃) may not be sufficient to initiate the job, consequently the activator becomes lodged inside the rudimentary pores even though burning off decreases the activated carbon yield. Therefore, Z1 could not be regarded as activated carbon because of its inferior development of surface area.

Activation using the recovered activators was performed at 550 ℃. There is a tremendous increase in the surface area of Z2 using the recovered $ZnCl_2$ from Z1. Although the amount of $ZnCl_2$ in the recovered solution is presumably less than for the fresh one (ratio 3:2), the elevated activation temperature shows a positive effect in increasing the surface area by more than 300 times. This could be related to the fact that more volatiles (nearly 65% weight loss) are liberated from the empty fruit-bunch at 550 ℃ [8]. It is suggested that the liberation of volatiles from the material also contributes to creating the pore pathways. These combined effects bring about the development of activated carbon with a high surface area.

From Table 1, F2 shows a higher surface area than for F1 at the same activation temperature. This could be associated with the optimum impregnation ratio in the preparation of activated carbon. The specific surface area of activated carbon normally increases as the impregnation ratio increases, but decreases when an excessive amount of activator is used. Too much activator may result in the collapse of pore textures, thus decreasing the surface area [9]. In other words, $FeCl_2$ used in the activation of F1 could have already exceeded the optimum impregnation ratio. However, further experimental works of varying impregnation ratios of $FeCl_2$ are needed to establish the optimum surface area of $FeCl_2$-activated empty fruit-bunch carbons. Nevertheless, F2 demonstrates a 3.8 times

lower surface area than Z2. This shows that $ZnCl_2$ is an effective activator for empty oil palm fruit-bunch-based activated carbon. Fig.1 shows the FTIR spectra of EFB and its derived activated carbons. The designated possible functional groups are summarized in Table 2.

The spectroscopy technique measures the absorption of various wavelengths of infrared light by materials of interest to identify specific organic functional groups on the surface of activated carbon. From Fig.1, the EFB displays various peaks that correspond to the presence of functional groups. The broad and strong band at 3300 cm⁻¹ is assigned to the stretching vibration of the (–OH) hydroxyl group. The intensity of the peak decreased by the order of EFB > Z1 > F1 ≈ F2 > Z2. The peak completely disappeared in Z2 probably due to the stronger dehydrating effect of $ZnCl_2$ compared to $FeCl_2$ in activating the EFB at 550 ℃. It also signifies that the activation of Z1 remains incomplete because most of the attributes of EFB spectra remained unchanged. As the EFB is converted into activated carbon, the complicated peaks become simplified indicating the liberation of functional groups in the activated carbons. The absorption peaks at 2930–2850 cm⁻¹ are attributed to the (C–H) stretching vibration of the (–CH₃) group. The peaks between 1400 and 1000 cm⁻¹ are ascribed to (C–O) stretching or (Si–O) of silica containing minerals (ash). The peak at 1026 cm⁻¹ in Z1 is assigned to the out of plane (C–H) bending.

3.2. Adsorption of Rhodamine B

Water polluted with dyes especially from the textile industries has become a subject of great concern because of the disruption to biodiversity and food chains [10]. Basic or cationic dyes are very bright dyes that are water-soluble. In this work, Rhodamine B ($C_{28}H_{31}C1N_2O_3$, mw = 479 g mol⁻¹, solubility in water = 15 g dm³) was chosen as the model dye to evaluate the performance of empty fruit-bunch-based activated carbons. The pH of the Rhodamine B solution was not adjusted, and the values were measured as 5.1±0.2 for all concentrations. At equilibrium, the pH values slightly changed as they are measured to be 5.3±0.1. Fig.2 illustrates the molecular structure of Rhodamine B.
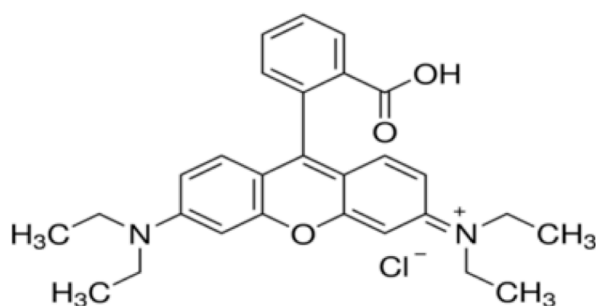
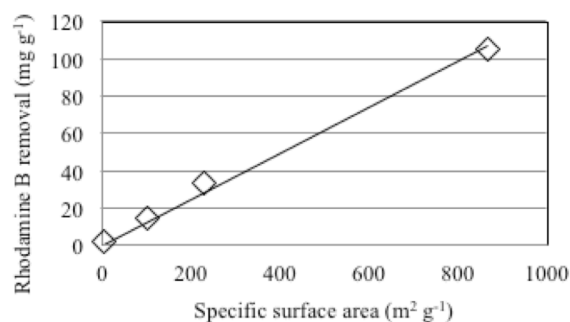*Figure 2*. Chemical structure of cationic Rhodamine B dye.

Figure 4. Correlation between Rhodamine B removal and specific surface area.

Table 3. Parameters of isotherm models.

	Z1	Z2	F1	F2
Q_{exp} (mg g^{-1})	2.33	105	14.8	33.4
Langmuir model				
Q (mg g^{-1})	2.85	97.9	18.5	42.2
b (dm^3 g^{-1})	0.166	2.92	0.248	0.187
R^2	0.907	0.434	0.857	0.967
SSE	0.233	403	23.7	26.3
Freundlich model				
K_F (mg g^{-1})(L mg^{-1})$^{1/n}$	0.560	72.5	4.73	8.86
n	2.10	13.4	2.61	2.33
R^2	0.882	0.707	0.718	0.869
SSE	0.435	209	44.7	94.3

Fig.3 shows the removal of Rhodamine B by EFB-derived materials. From *Fig.3*, the removal of Rhodamine B was found to increase with concentration by the order of Z2 > F2 > F1 > Z1. Activated carbon Z2 demonstrates the highest Rhodamine B removal of 105 mg g^{-1}. This trend is in agreement with the increase in specific surface area of adsorbents. In general, the concentration of dye in the solution provides a driving force for adsorption if the adsorbent or activated carbon possesses abundant active sites. In this case, the active sites are directly associated with the surface area. *Fig.4* displays the correlation between the removal of Rhodamine B and the surface area.

A linear relationship for Rhodamine B removal (mg g^{-1}) = 0.124 × specific surface area (m^2 g^{-1}), R^2 = 0.993 was obtained. A bigger surface area normally provides better interaction probabilities for the Rhodamine B molecules to lodge onto the pore channels. In addition, it is presumed that all types of pore play a dominating role in the adsorption especially mesopores [5,11].
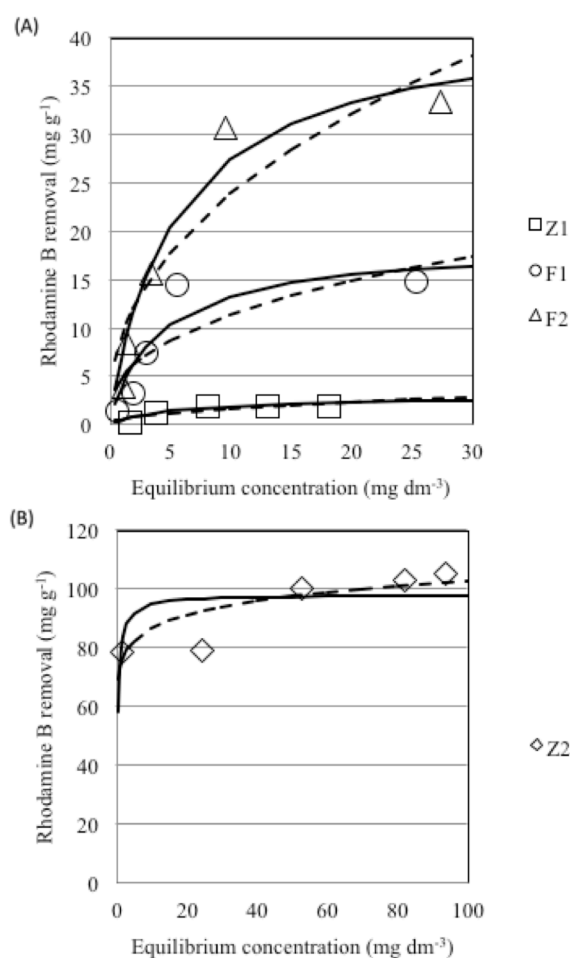

Figure 3. Removal of Rhodamine B by (A) Z1, F1 and F2, and (B) Z2. Lines were predicted by the Langmuir (solid) and Freundlich (dashed) models.

The adsorption data were analysed using the Langmuir and Freundlich models, and the constants are tabulated in *Table 3*. The adsorption data reasonably fitted to both adsorption models with R^2 ranging between 0.71 and 0.97, except for Z2 according to the Langmuir model. The values of Langmuir model capacity (Q) are in close proximity to the experimental data (Q_{exp}). In addition, Z2 shows a higher adsorption affinity (b) for Rhodamine B when compared with the other counterparts. This signifies a higher removal efficiency (~99%) at concentrations below 80 mg dm^{-3}. Similar explanations apply for the Freundlich model. The deviation of the lines of the model from the experimental data is shown in *Fig.3*, which indicates that the removal of Rhodamine B by EFB-based materials is neither a strict monolayer adsorption nor heterogeneous coverage, but could be a blend of the two – monolayer adsorption onto a heterogeneous surface [11].

In a related work, Santhi *et alia* [12] reported a removal capacity of 22.3 mg g^{-1} of Rhodamine B using H$_2$SO$_4$-treated *A. nilotica* leaves. Generally, Z2 shows a higher dye removal capacity compared to this chemically treated natural adsorbent. This is likely due to the well-developed graphitic structure and surface area of Z2 for the effective removal of Rhodamine B.

4. Conclusion

Oil palm empty fruit-bunch was used in the preparation of activated carbons *via* metal-chloride activation. Zinc chloride is a more effective activating agent for activated carbon than iron(II) chloride. Activation at 500 ºC yields empty fruit-bunch-based activated carbon with a surface area of 866 m^2 g^{-1}. A bigger surface area offers greater removal of Rhodamine B dye and a higher adsorption affinity at lower dye concentrations. The maximum removal efficiency was recorded as 105 mg g^{-1}. The mechanism could be described as monolayer Rhodamine B adsorption onto heterogeneous activated carbon. Empty fruit-bunch is a promising alternative to activated carbon precursors for wastewater treatment.

Acknowledgement

REFERENCES

[1] Tanaka, R.; Rosli, W.; Magara, K.; Ikeda, T.; Hosoya, S.: Chlorine-free bleaching of Kraft pulp from oil palm empty fruit-bunches, *Jpn. Agric. Res.* 2004 **38**(4), 275–279 DOI: 10.6090/jarq.38.275

[2] Duan, X.; Peng, J.; Srinivasakannan, C.; Zhang, L.; Xia, H.; Yang, K.; Zhang, Z.: Process optimization for the preparation of activated carbon from *Jatropha* hull using response surface methodology, *Energy Sources, Part A: Recovery Util. Environ. Effect*, 2011 **33**(21), 2005–2017 DOI: 10.1080/15567030903515047

[3] Nasir, N.H.M.; Zaini, M.A.A.; Setapar, S.H.M.; Hassan, H.: Removal of methylene blue and copper(II) by oil palm empty fruit-bunch sorbents, *J. Teknologi* 2015 **74**(7), 107–110 DOI: 10.11113/jt.v74.4707

[4] Wirasnita, R.; Hadibarata, T.; Yusoff, A.R.M.; Lazim, Z.M.: Preparation and characterization of activated carbon from oil palm empty fruit-bunch wastes using zinc chloride, *J. Teknologi* 2015 **74**(11), 77–81 DOI: 10.11113/jt.v74.4876

[5] Ming-Twang, S.; Lin-Zhi, L.; Zaini, M.A.A.; Zhi-Yong, Q.; Pei-Yee, A.Y.: Activated carbon for dyes adsorption in aqueous solution, In: Advances in Environmental Research, Ed.: Daniels. J.A. (Nova Science Publishers, Inc., New York, USA) pp. 217–234, 2015

[6] Hidayu, A.R.; Mohamad, N.F.; Matali, S.; Sharifah, A.S.A.K.: Characterization of activated carbon prepared from oil palm empty fruit-bunch using BET and FT-IR techniques, *Proc. Engng.* 2013 **68**, 379–384 DOI: 10.1016/j.proeng.2013.12.195

[7] Alam, M.Z.; Ameem, E.S.; Muyibi, S.A.; Kabbashi, N.A.: The factors affecting the performance of activated carbon prepared from oil palm empty fruit-bunches for adsorption of phenol, *Chem. Engng. J.* 2009 **155**(1–2), 191–198 DOI: 10.1016/j.cej.2009.07.033

[8] Nyakuma, B.B.; Johari, A.; Ahmad, A.; Abdullah, T.A.T.: Thermogravimetric analysis of the fuel properties of empty fruit-bunch briquettes, *J. Teknologi* 2014 **67**(3), 79–82 DOI: 10.11113/jt.v67.2768

[9] Zaini, M.A.A.; Okayama, R.; Machida, M.: Adsorption of aqueous metal ions on cattle-manure-compost based activated carbons, *J. Hazard Mater.* 2009 **170**(2–3), 1119–1124 DOI: 10.1016/j.jhazmat.2009.05.090

[10] Ming-Twang, S.; Zhi-Yong, Q.; Lin-Zhi, L.; Pei-Yee, A.Y.; Zaini, M.A.A.: Dyes in water: Characteristics, impacts to the environment and human health, and the removal strategies. In: Advances in Chemistry Research, Ed.: Taylor, J.C. (Nova Science Publishers, Inc., New York, USA) Vol. 23, pp. 143–156, 2015

[11] Zaini, M.A.A.; Zakaria, M.; Setapar, S.H.M.; Yunus, M.A.C.: Sludge-adsorbents from palm oil mill effluent for methylene blue removal, *J. Environ. Chem. Engng.* 2013 **1**(4), 1091–1098 DOI: 10.1016/j.jece.2013.08.026

[12] Santhi, T.; Prasad, A.L.; Manonmani, S.: A Comparative study of microwave and chemically treated acacia *Nilotica* leaf as an eco-friendly adsorbent for the removal of Rhodamine B dye from aqueous solution, *Arabian J. Chem.* 2014 **7**(4), 494–503 DOI: 10.1016/j.arabjc.2010.11.008

CHARACTERISATION OF CEMENTS FROM DOMINANTLY VOLCANIC RAW MATERIALS OF THE CARPATHIAN BEND ZONE

TIMEA HALMAGYI,[1] EMILIA MOSONYI,[2] JÓZSEF FAZEKAS,[2]* MARIA SPATARU[1] AND FIRUTA GOGA[2]

[1] S.C. Valdek IMPEX SRL, Sfantu Gheorghe Str. 1 Decembrie 1918, Bl. 15, Sc. E, Et. 3, Ap. 9., Judetul Covasna, ROMANIA
[2] "Babes-Bolyai" University, Kogalniceanu Str. 1, 400082 Cluj-Napoca, ROMANIA

This paper presents the results of laboratory investigations regarding the production of cements from local raw materials, such as limestone from Varghis, gypsum from Nucsoara, basaltic scoria from Racosul de Jos, volcanic tuff from Racosul de Sus, diatomite from Filia, and red mud from Oradea. The raw mixtures, based on modified Bogue calculations, contain limestone, gypsum, and one or two of the above-mentioned materials. The cements resulted from clinker grinding in a laboratory gas furnace at 1260–1300 ºC, with one hour at the peak temperatures, and were characterised for Blaine specific surface area, specific density, and mineral phases. Physico-mechanical properties, such as water content for normal consistency, setting time, soundness, and compressive strength were also determined. Results show that these cements contain belite, ferrite, calcium sulphoaluminate, anhydrite, and some minor compounds.

Keywords: experimental cement, Varghis limestone, Racosul de Sus volcanic tuff, setting time, mechanical properties

1. Introduction

Manufacturing of Portland cement is energy consuming, globally accounting for 2% of primary energy and 5% of industrial energy consumption. Moreover, Portland cement production contributes significantly to greenhouse gas emission in the order of 5% of the global CO_2 emissions due to anthropogenic sources [1]. New civil engineering requirements impose the production of a new type of cement, which is of good quality, environmentally friendly, and requires low energy utilization. Low-energy cement manufacturing is economically and ecologically preferable. These cements could be used in places where high early strength or expansion compensation and also increased durability are required.

In cement chemistry, notation of oxides are abbreviated by their first capital letters: C=CaO, S=SiO_2, A=Al_2O_3, F=Fe_2O_3, C$\overline{S}$=SO_3, and H=H_2O. Low-energy cements comprise those that belong to the CaO-SiO_2-Al_2O_3-Fe_2O_3- C$\overline{S}$ system. They are mainly sulphoaluminate belitic and sulphoferroaluminate belitic cements [2]. These can be produced from both natural raw materials or raw material mixtures, containing by-products or industrial waste, by firing at lower temperatures than for Portland cement clinkers. A large variety of cements was developed in China based on the

composition $C_4A_3\overline{S}$. These were standardized and named as "Third Cement Series" [3]. These cements have special features such as quick setting time, good impermeability, and rapid strength development even at low temperatures.

There are numerous investigations on the laboratory-scale production of sulphoaluminate belite (SAB) cement [4–18]. SAB cement can be classified into three major categories according to the content and proportions of the phase compositions. The tentative naming and some basic properties are summarized below:

1. *Calcium sulphoaluminate-rich belite cement:* They are mostly composed of only two main phases of SAB cement, the major components being $C_4A_3\overline{S}$ (around 55–75 (g/g)%) and C_2S. This type of cement is typically used for applications requiring rapid setting and high early strength [4, 6, 8, 9, 19, 20].

2. *Expansive belite-rich calcium sulphoaluminate cements*: Besides the main components of the SAB cements they contain free lime up to 10 (g/g)% which promotes expansion. This type of cement can be used in restricted areas requiring shrinkage-resistant and self-stressing cements [4, 12, 16, 20, 21].

3. *Non-expansive belite-rich calcium sulphoaluminate cements*: These cements have higher belite, lower calcium sulphoaluminate, and much lower or completely deficient of free lime content than those of commercially produced SAB cements. Industrial by-products with high sulphate content can be used in high

---

*Correspondence: chemiceramic@gmail.com

Table 1. Chemical compositions of the raw materials in (g/g) %.

oxides	Varghis limestone	Oradea red mud	Racosul de Jos basaltic scoria	Racosul de Sus volcanic tuff	Filia diatomite	Nucsoara gypsum
CaO	89.61	12.71	10.38	2.88	0.71	28.31
SiO_2	4.01	8.93	46.10	64.53	92.52	9.83
Al_2O_3	5.16	17.04	18.78	11.80	3.24	2.73
Fe_2O_3	-	48.37	9.93	2.57	2.18	1.07
Na_2O	-	3.68	3.23	1.89	-	0.15
K_2O	-	-	1.69	2.75	0.46	0.55
MgO	-	1.30	7.21	0.45	-	0.39
TiO_2	-	6.80	1.61	0.27	-	0.13
V_2O_5	-	0.19	-	-	-	0.04
P_2O_5	1.22	0.98	1.07	-	-	0.04
SO_3	-	-	-	-	0.89	37.08
Mn_2O_3	-	-	-	-	-	0.02
P.C.	-	-	-	12.40	-	19.90

Table 2. Theoretical mineralogical and chemical compositions of the cements in (g/g) %.

Sample	mineral composition				oxidic compositions				
	C_2S	$C_4A_3\overline{S}$	C_4AF	$C\overline{S}$	CaO	Fe_2O_3	SiO_2	Al_2O_3	SO_3
C1, C2, C3	56	12	22	10	55.15	7.23	19.53	10.63	7.46
C4	50	18	15	17	53.11	4.93	17.44	12.17	12.36
C5	43	28	23	6	51.38	7.56	15.00	18.86	7.20

percentages in production. This type of cement shows high mechanical strength both at early and late ages comparable to ordinary Portland cement and has the potential to replace it [4, 6, 8, 14, 16].

There is no commercial production of the belite-rich calcium sulphoaluminate-type of cement. Mehta [18] produced SAB cements containing no free CaO, but large amounts of belite rich in $CaCO_3$, silicic acid, hydrated alumina, iron oxide, and gypsum. The clinkers were obtained by heating the raw materials in an electric muffle furnace at 1200 ℃ for about 1 hour. Clinkers were ground to a Blaine specific surface area of about 400 m^2 kg^{-1}.

Microstructure examinations of the clinkers showed that $C_4A_3\overline{S}$ appears as cubic crystals whereas belite appears as large rounded grains. The clinkers were very easy to grind due to their brittleness [4]. Sahu *et alia* [12] produced cements of Types 2 and 3 from limestone, fly ash and gypsum at 1200 ℃, for a bearing time of 30 minutes and cooled by fresh air. Kasselouri *et alia* [17] in 1995 obtained cement of Type 3 at 1280 ℃ from limestone, gypsum, bauxite, silica sand, and iron-rich industrial by-products. Cements composed from a mixture of baghouse dust, F-class fly ash, and scrubber sludge, sintered at 1175–1250 ℃ for bearing times of 30, 45, and 60 minutes, and cooled by natural air were obtained according to Arjunan *et alia* [10]. Mixtures composed of limestone, bottom ash, baghouse filter ash, bauxite, and gypsum fired at 1250 and 1300 ℃, for a bearing time of 60 minutes, lead to SAB cement of $C_4A_3\overline{S}$ – ye'elimite, C_2S-larnite (belite), C_4AF-brownmillerite, and $C\overline{S}$- anhydrite compositions [4]. The absence of tricalcium aluminate (C_3A) in these cements indicates that the decomposition of the desired mineral $C_4A_3\overline{S}$ does not take place at these temperatures. The presence of the desired minerals and

the absence of $C_5S_2\overline{S}$ confirmed the formation of the SAB cement at 1250 ℃ and 1300 ℃ after 1 hour.

This article presents our study on the cement series sintered using local raw materials, such as volcanic tuffs from Racosul de Sus, basaltic scoria from Racosul de Jos, and diatomite from Filia (all in the Carpathian orocline), and red mud.

2. Experimental

The raw materials used for the cement experiments carried out as part of this study are natural (Varghis) limestone, Bodoc clay, Nucsoara gypsum, Racosul de Jos basaltic scoria, Racosul de Sus volcanic tuff, Filia diatomite, and artificial industrial waste (Oradea red mud). The selected raw materials were investigated for chemical compositions.

The chemical compositions of the limestone, red mud, diatomite, and basaltic scoria were analysed by SEM/EDAX. The volcanic tuff and gypsum analyses were performed by wet chemical methods, according to the SR EN 192-2 (*Table 1*). The theoretically estimated mineralogical, chemical compositions based on modified Bogue calculations [1, 4, 22] are given in *Table 2*. The compositions of the raw material mixtures are presented in *Table 3*. The raw material mixtures were obtained by grinding them in a laboratory ballmill up to a sieve residue of 90 μm about 12%. Afterwards, the raw material mixtures underwent a process of briquetting and drying followed by firing in a laboratory gas oven for one hour at a constant temperature of 1260–1300 ℃. The inside temperatures of the furnace were estimated with a thermocouple thermometer. Fast cooling was achieved in the air. The obtained cements from grinded clinkers (five hours in a mill, balls:clinkers ratio of 2:1) were tested. Mineralogical

Table 3. The compositions of raw material mixtures in (g/g) %.

Samples	Varghis limestone	Oradea red mud	Racosul de Jos basaltic scoria	Racosul de Sus volcanic tuff	Filia diatomite	Nucsoara gypsum
Cement 1	48.90	-	36.67	-	-	14.43
Cement 2	54.08	-	-	-	13.70	32.22
Cement 3	54.35	-	-	23.37	-	22.28
Cement 4	54.66	0.98	31.53	-	-	12.83
Cement 5	47.45	1.19	20.66	-	-	30.70

Table 4. Mineral compositions in (g/g) % of the cement samples C1-C5 according to the IL-LAB-41 testing method.

Minerals	C1	C2	C3	C4	C5
Belite	24.2	69.1	65.9	52.7	58.4
Ferrite	7.5	13.1	8.5	24.6	20.8
Cubic aluminate	8.5	0.4	0.0	0.2	4.7
Orthorhombic aluminate	20.5	0.0	0.0	1.5	0.0
Free lime	1.4	0.5	1.7	0.6	0.1
Anhydrite	2.1	6.8	3.8	1.5	5.3

Table 5. Physical properties of cement samples C1-C5.

Sample	Blaine specific surface area, cm^2 g^{-1}	Heat of hydration, J g^{-1}	Specific density, g cm^{-3}
C1	7367	29	3.17
C2	8745	123	3.04
C3	7963	32	3.07
C4	6501	44	3.14
C5	7553	110	3.26

Table 6. Physico–mechanical properties of cement samples C1-C5.

sample	water for normal consistency, cm^3	setting time, hour : minutes early	final	soundness, mm	compressive strength, N mm^{-2} 2 days	28 days
C1	135	0:44	1:44	2.75	0.24	0.42
C2	225	0:13	0:35	0.50	4.12	26.67
C3	124	3:35	4:29	3.00	0.23	8.42
C4	185	1:57	> 10 hour	1.00	0.53	0.72
C5	178	0:05	0:15	2.00	1.44	2.28

compositions of cements made by XRD analyses with a Panalytical-Philips Cubix PRO X-ray spectrometer, according to IL LAB 41 proceedings are presented in *Table 4*. Experimental laboratory-produced cements were tested in conditions provided by the Romanian CRH CEMENT S.A. plant in Hoghiz (Brasov county). The Blaine specific surface area, specific density, heat of hydration (*Table 5*), volume of water for normal consistency, setting time, soundness, and compressive strength (*Table 6*) have been determined. Identification of mineral phases formed during the burning of clinkers was carried out by means of the Cubix PRO spectrometer.

3. Results and Discussions

The mineralogical compositions of the experimental cements are summarized in *Table 4*. All cement samples contain belite, ferrite, anhydrite, and many other phases in small amounts. The largest amount of belite found in a cement sample is in the raw mixture containing diatomite (sample C2 with $C_2S = 69.1\%$), followed by the sample prepared with volcanic tuff (sample C3 with $C_2S = 65.9\%$). Soner *et alia* reported [4] that the mineral $C_4A_3\bar{S}$ is stable between 1250–1350 ºC, but probably it can decompose during cooling. Furthermore, it was demonstrated that aluminium could be substituted by iron in the ye'elimite structure forming $C_4A_{3-x}F_x\bar{S}$ [23–26]. The experimental cement samples were characterized with regards to specific density and fineness, representing the Blaine specific surface area
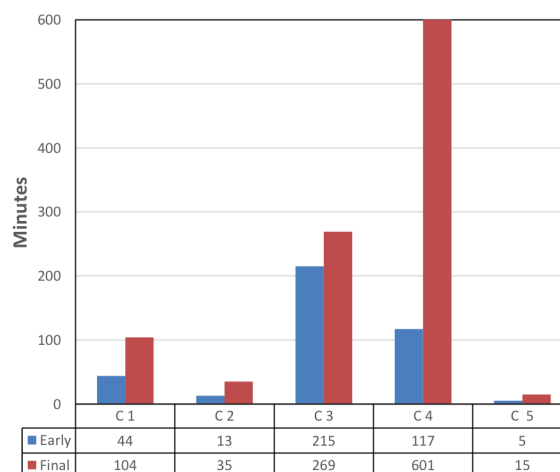

	C 1	C 2	C 3	C 4	C 5
Early	44	13	215	117	5
Final	104	35	269	601	15

Figure 1. Early and final setting times of the cement samples in minutes.

[27]. The heat of hydration according was also determined to the SR EN 196-9/2006 method [28]. The physical properties of experimental cements are presented in *Table 5*.

The experimental Blaine fineness data are characterised by large specific surface areas. The biggest Blaine specific surface area was found in sample C2 (with diatomite) followed of sample C3 sample (with volcanic tuff). These samples have the smallest specific densities in the same order. The binding behaviour of the cements was estimated by measuring the setting time, volume of water for normal consistency, soundness, and compressive strengths (after second and twenty-eighth days). To determine the
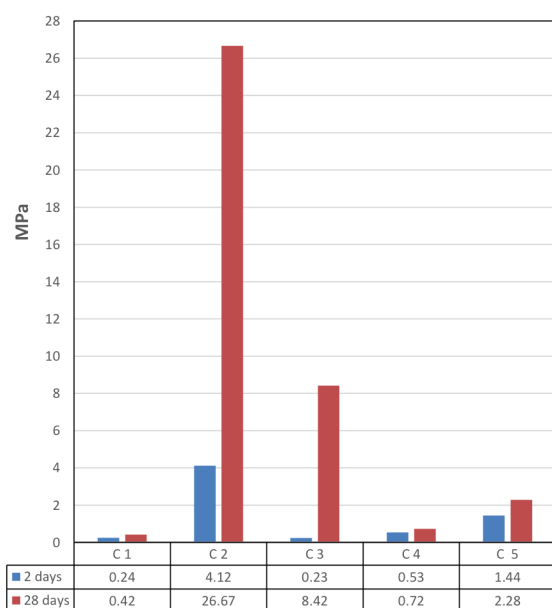
Figure 2. Compressive strength of cement samples after 2 and 28 days in MPa.

setting times, the quantity of water required to form cement paste of standard (normal) consistency was determined previously [29]. The water data for standard consistency are presented in *Table 6.* The setting time for the paste with standard consistency was measured using a Vicat device. The initial and final setting data of the investigated cements are also shown in *Table 6* and *Fig.1*. The compressive strengths have been determined according to the SN EN-1/2006 method [30] and the data are shown in *Fig.2*.

The investigated cements exhibited different mechanical strengths, as a function of their mineral compositions. The strength of SAB cements depends mainly on the mineral ye'elimite ($Ca_4Al_6O_{12}SO_4$ or $C_4A_3\bar{S}$) during the initial minutes up to hours of hydration [31, 32]. Ye'elimite is almost entirely responsible for the hydration reactions at early ages of CSA-type cements [23]. The presence of belite was found to be responsible for compressive strength at late ages. The compositions of raw material mixtures for these experimental cements influenced their mineral contents. Taking into consideration these, the initial strengths are better for the cements containing diatomite (sample C2), volcanic tuff (sample C3), basaltic scoria (sample C4), and red mud (sample C5). A higher content of red mud is favoured over sample C4 (see *Tables 3* and *4*, *Fig.2*).

A good evolution of mechanical strength over time is shown for samples C5 and especially for the C2 cements. C3 cement sample is also notable, but with lower initial strength, which is an important characteristic for a favourable evolution of specific surface area. This is clearly the largest value for cement sample C3. The cements containing basaltic scoria generally developed lower mechanical strengths in comparison to those containing diatomite or volcanic tuff. The cement containing diatomite is noticeable due to its very high initial strength (*Table 6*). For this

cement, the compressive strength, after two days increased, which may be a consequence of the increased specific surface area. In terms of practical applications, the development of cement of good mechanical strength and workability depending on the setting time is of importance. The setting time of investigated cements was decisively influenced by the content of raw material mixtures (see *Table 3*). In the order of C4→C3→C1→C2→C5, the setting time becomes shorter. For the cement samples containing 1.2% red mud, 20.7% basaltic scoria (C5) and 13.7 % diatomite (C2), the setting time is quick. This is a consequence of rapid hydration processes. Because of this, these latter samples can be considered for practical applications only as retarding admixtures or super-plasticizers due to their set-retarding effect. Soundness shows reasonable values, which is required to be less than 10 mm for Portland cement.

4. Conclusion

Based on investigations into cement samples in the laboratory of the Hoghiz plant, it can be concluded that from all raw material mixtures, heated at temperatures of between 1260 and 1300 ºC, resultant clinkers contain more belite, ferrite, and anhydrite. Physico-mechanical properties show good compressive strength at early ages, good soundness, the biggest Blaine specific surface area ($S_{sp} = 8745$ cm^2 g^{-1}) for cement containing limestone, diatomite, and gypsum, as well as belite and anhydrite. The use of the local raw materials from the Carpathian orocline area, e.g. Varghis limestones, Bodoc clays, Nucsoara gypsum, volcanic tuffs from Racosul de Sus, basaltic scoria from Racosul de Jos, and Filia diatomite, facilitated the formation of cements that are more belitical than sulphoaluminate.

Acknowledgement

REFERENCES

[1] Chen, I.A.: Synthesis of Portland cement and calcium sulfoaluminate-belite cement for sustainable development and performance, Ph.D. Dissertation, The University of Texas at Austin, 2009 www.lib.utexas.edu/etd/d/2009/cheni55279

[2] Jun-an, D.; De-dong, L.; Mu-zhen, S.; Jan-mon, W.: A study of mineral composition of cement in the quinary system CaO-SiO$_2$-Al$_2$O$_3$-Fe$_2$O$_3$-SO$_3$, *J. Chin. Ceram. Soc.* 1982 Article No. 4 en.cnki.com.cn/Article_en/CJFDTOTALGXYB198204000

[3] Zhang, L.; Su, M.; Wang, Y.: Development and use of sulpho- and ferro-aluminate cements in China, *Adv. Chem. Res.* 1999 **1**, 15–21 DOI: 10.1680/adcr.1999.11.1.15

[4] Soner, I.: Utilization of fluidized bed combustion ashes as raw material in the production of a special cement, M.Sc. Dissertation, Middle East Technical University, Haziran, Turkey, June 2009 www.osti.gov/scitech/biblio/5635811

[5] Majling, J.; Roy, M.D.: The potential of fly ash for cement manufacture, *Am. Ceram. Soc. Bull.* 1993 **72**(10), 77–81 www.osti.gov/scitech/biblio/5635811

[6] Gartner, E.: Industrially interesting approaches to "Low-CO_2" cements, *Cement Concrete Res.* 2004 **34**, 1489–1498 DOI: 10.1016/j.cemconres.2004.01.021

[7] Quillin, K.: Performance of belite-sulfoaluminate cements, *Cement Concrete Res.* 2001 **31**, 1341–1349 DOI: 10.1016/S0008-8846(01)00543-9

[8] Phair, W.J.: Green chemistry for sustainable cement production and use, *Green Chem.* 2006 **8**, 763–780 DOI: 10.1039/B603997A

[9] Zhou, Q.; Milestone, B.N.; Hayes, M.: An alternative to Portland cement for waste encapsulation - The calcium sulfoaluminate cement system, *J. Hazard. Mat.* 2006 **136**, 120–129 DOI: 10.1016/j.jhazmat.2005.11.038

[10] Arjunan, P.; Silsbee, R.M.; Roy, M.D.: Sulfoaluminate-belite cement from low-calcium fly ash and sulfur-rich and other industrial by-products, *Cement Concrete Res.* 1999 **29**, 1305–1311 DOI: 10.1016/S0008-8846(99)00072-1

[11] Janotka, I.; Krajci, L.; Mojumdar, C.S.: Influence of Portland Cement Addition to Sulphoaluminate - Belite Cement on Hydration and Mechanical Properties of Hardened Mortars, *Ceramics-Silikaty* 2002 **46**(3), 110–116 www.irsm.cas.cz/materialy/cs_content/2002/Janotka_CS_2002_0000.pdf

[12] Sahu, S.; Majling, J.: Preparation of sulphoaluminate belite cement from fly ash, *Cement Concrete Res.* 1994 **24**(6), 1065–1072 DOI: 10.1016/0008-8846(94)90030-2

[13] Janotka, J.; Krajci, L.; Mojumdar, C.S.: Performance of sulphoaluminate-belite cements with high C_4A_3 content, *Ceramics-Silikaty* 2007 **51**(2), 74–81 www.irsm.cas.cz/materialy/cs_content/2007/Janotka_CS_2007_0000.pdf

[14] Adolfsson, D.; Menad, N.; Viggh, E.; Björkman, B.: Steelmaking slags as raw material for sulphoaluminate belite cement, *Adv. Cement Res.* 2007 **19**(4), 147–156 DOI: 10.1680/adcr.2007.19.4.147

[15] Popescu, D.C.; Muntean, M.; Sharp, H.J.: Industrial trial production of law energy belite cement, *Cement Concrete Composites* 2003 **25**, 689–693 DOI: 10.1016/S0958-9465(02)00097-5

[16] Katsioti, M.; Tsakiridis, E.P.; Leonardou-Agatzini, S.; Oustadakis, P.: Examination of the jarosite-alunite precipitate addition in the raw meal for the production of sulfoaluminate cement clinker, *J. Hazard. Mat.* 2006 **131**(1–3), 187–194 DOI: 10.1016/j.jhazmat.2005.09.028

[17] Kasselouri, V.; Tsakiridis, P.; Malami, C.; Georgali, B.; Alexandridou, C.: a study on the hydration products of a non-expansive sulfoaluminate cement, *Cement Concrete Res.* 1995 **25**(8), 1726–1736 DOI: 10.1016/0008-8846(95)00168-9

[18] Mehta, K.P.: Investigations on energy-saving cements, *World Cement Technol.* 1980 **11**(5), 166–177

[19] Zhang, L.; Glasser, P.F.: Hydration of calcium sulfoaluminate cement in less than 24 h, *Adv. Cement Res.* 2002 **14**(4), 141–155 DOI: 10.1680/adcr.2002.14.4.141

[20] Uchikawa, H.: Management strategy in cement technology for the next century Part 3, *World Cement* 1994 **VOL**(11), 47–54

[21] Teoreanu, I.; Muntean, M.: Expansive sulphate aluminate cements, *Cement Concrete Res.* 1983 **13**, 711–720

[22] Dan, E.: The low energy cements durability, manufactured with addition of industrial waste materials, Ph.D. Thesis (University Politehnica Bucharest, Bucharest, ROMANIA) 2004 (in Romanian)

[23] Bullerjahn, F.; Haka, M.B.; Scrivener, K.: Iron solid solutions of ye'elimite-effect on reactivity, *Proc. 19th Int. Building Material Expo* (Weimar, Germany) 2015 researchgate.net/publication/282356671

[24] Cuesta, A.; Álvarez-Pinazo, G.; Sanfélix, S.G.; Peral, I.; Aranda, M.A.G.; De la Torre, A.G.: Hydration mechanisms of two polymorphs of synthetic ye'elimite, *Cement Concrete Res.* 2014 **63**, 127–136 DOI: 10.1016/j.cemconres.2014.05.010

[25] Idrissi, M.; Diour, A.; Damidot, D.; Greneche, J.M.; Alami Talbi, M.; Taibi M.: Characterisation of Iron Inclusion during the Formation of Calcium Sulfoaluminate Phase, *Cement Concrete Res.* 2010 **40**(8), 1314–1319 DOI: 10.1016/j.cemconres.2010.02.009

[26] Pöllmann, H.; Stöber, S.; Schmidt, R.: Synthesis and characterization of sulfoaluminate (belite) cements from industrial residues, *Proc. 1st Int. Conf. on Sulphoaluminate Cement: Mat. Engng. Technol* (Wuhan University of Technology Press, Wuhan, China) pp. 366–384, 2013 www.irsm.cas.cz/materialy/cs_content/2013/Ma_CS_2013_0000.pdf

[27] SR EN Standard 196-6: Method of testing cement - Part 6: Determination of blaine specific surface area and specific density, 2006 tpm.fsv.cvut.cz/student/documents/files/BUM1/Chapter12.pdf

[28] SR EN Standard 196-9: Method of testing cement - Part 9: Determination of the heat of hydration, 2006 tpm.fsv.cvut.cz/student/documents/files/BUM1/Chapter12.pdf

[29] SR EN Standard 196-3: Method of testing cement - Part 3: Determination of setting time and le chatelier soundness, 2006 tpm.fsv.cvut.cz/student/documents/files/BUM1/Chapter12.pdf

[30] SR EN Standard 196-1: Method of Testing Cement - Part 1: Determination of Strength, 2006

[31] Zhang, L.; Su, M.; Wang, Y.: Development and use of sulfo- and ferro-aluminate cements in China, *Adv. Cement Res.* 1999 **1**, 15–21 DOI: 10.1680/adcr.1999.11.1.15

[32] Winnefeld, F.; Barlag, S.: Calorimetric and thermogravimetric study on the influence of calcium sulfate on the hydration of ye'elimite, *J. Therm. Anal. Calorimet.* 2009 **101**(3), 949–957 DOI: 10.1007/s10973-009-0582-6

Guide for Authors

1. Electronic copies of manuscripts should be uploaded to the Author's Page at **hjic.mk.uni-pannon.hu** or emailed directly to **hjic@almos.uni-pannon.hu**.
   – Please visit the Journal's website to download a **Word or Latex template** and follow closely the suggested layout for the text, figures, tables, and references. Manuscripts without the recommended structure and formatting for publication will be returned without a review.
   – The use of the template is required in order to avoid imposing any article processing charges (APCs), page fees, or article submission charges.
2. In a brief letter attached to each manuscripts, **authors must declare that their work is original** and has not previously been published elsewhere. The editorial board requests an originality evaluation score of less than 10% according to CrossCheck/Ithenticate excluding quotes, references, and expressions consisting of three words. If originality of a given manuscript is between 10% and 20% the authors are kindly requested to rewrite the overlapping sections. Nothing above 20% similarity will be accepted for publication and will not be reviewed.
3. For the purposes of correspondence, the authors are asked to provide their current **address**, **telephone number**, **professional website address (if applicable) and regularly checked e-mail address**.
4. **Authors are requested to adhere to the rules** outlined below and follow the examples:
   – The layout of manuscripts should be on A4 page (21 cm × 29.7 cm) consisting of double columns, single-spaced (approximately 45 characters per line, 55 lines per page), and numbered pages. **The text of the papers should be concise, and not exceed twelve pages**; even with Tables and Figures the manuscript should not exceed a total of twelve pages. These limits may be extended for reviews and invited articles for themed issues.
   – The first page should contain the title, the full name(s) (no abbreviations please) of the author(s) in the sequence to be published, the name(s) and address(es) of the institution(s) where the work was completed, and a brief summary of the article (maximum 15 lines). The title of the paper should not exceed 15 words. At least three keywords must be specified, but not more than seven.
   – The corresponding author or authors should be clearly marked.
   – The article should include the following sections: Introduction, Experimental or Theoretical Methods, Results and Analysis, Discussion, and Conclusion. A Conclusion section is mandatory with at least a paragraph presenting the main outcome(s) of the given study, not just repeating the abstract or summarizing the study.
   – **Tables and Figures must be inserted into the text** as close as possible to their first mention and flushed with the top margin as much as possible. Drawings must be of high contrast and consist of continuous curves. Please attach the source of the graphics of the figure/scheme if available in at least 300 dpi in resolution. Characters and labels should be sufficiently large to allow for reproduction. Decimal points must be used and not commas.
   – Note that the electronic deposition of the published version of the manuscript will contain colored graphics; however, the printed version will only be reproduced in black and white unless otherwise requested for a fee.
   – The Tables and Figures should not exceed one page of A4, (**maximum width of Figures: portrait 8 cm, landscape 16.5 cm**). If diagrams are presented, only the explanation and dimensions of the abscissa, ordinate and marking numbers are required, further explanatory texts can be given in the caption. Please remove any excessive white space from around the figures.
   – Equations should always be stand-alone, i.e. occupy an extra line and be labelled with Arabic numerals in parentheses in the right-hand margin. Efforts to avoid equations running into a second column are appreciated.
   – A summary of the symbols used must be included at the end of the manuscript after the Conclusion section under the heading "SYMBOLS". Symbols and abbreviations that represent variables, constants, quantities, properties, etc. must be defined in the text at their first occurrence.
   – **Only SI (System International d'Unites) units are to be used.** If data with non-SI units are reported, they should be in parentheses following the corresponding data in SI units.
5. References should be numbered throughout the text and written in square brackets and not in the superscript position. The references are also to be listed at the end of the paper. *Abbreviations of journal titles should conform to international standards, as accepted by Chemical Abstracts*. The style and punctuation of the references should conform to the ACS Style Guide. As a brief reference please consult the following examples:
   a) Journal articles: Debye, P.; Hückel, E.: The theory of electrolytes I. Lowering of freezing point and related phenomena, *Physik. Z.* 1923 **24**(1), 185–206
   b) Books: Verwey, E.J.W.; Overbeek, J.T.G.: Stability of Lyophobic Colloids (Elsevier, Amsterdam) 1948
   c) Book Chapters: Gunn, A.M.; Winnard, D.A.; Hunt, D.T.E. : Trace metal speciation in sediment and soils, in *Metal Speciation: Theory, analysis and application*, Eds.: Kramer, J.R.; Allen, H.E. (Lewis, Boca Raton, FL, USA) 1988
   d) Patents: U.S. Pat. 3,425,442 (1984)
   e) Published lectures, conference proceedings: Hih, S.; Hhu, C.; Leech, W.J.: A Discrete Method to Heat Conduction Problems, *Proc. 5th Int. Heat Transfer Conf., Tokyo, Japan*, pp. 2.4, 1975. Citations of conference presentations are strongly discouraged and should be avoided unless the talk was made available electronically.
   f) Papers that are unpublished, but have been submitted to a journal may be cited with the journal's name, followed by "submitted for publication" or "in press". This will be accepted if the author uploads the submitted manuscript as 'editorial' material. No reference to "unpublished work" or "personal communication" will be accepted.
   g) All references must be provided with a digitial object identifier (DOI). Please provide the editorial office with a rationale for any citations without DOI numbers and explain why these references are necessary.
6. Authors receive galley proofs, which are to be corrected and returned as soon as possible, but no later than 1 week after receipt. No new material may be inserted into the text whilst being proofread.
7. **Authors who have less experience of written English are urged to have their manuscripts checked by scientists who are working in the Author's respective field and proficient in English**.

Editorial Board of the
Hungarian Journal of Industry and Chemistry