Jerzy Mikulski (Ed.)

Communications in Computer and Information Science 329

Telematics in the Transport Environment

12th International Conference on Transport Systems Telematics, TST 2012
Katowice-Ustroń, Poland, October 2012
Selected Papers

Springer

Communications in Computer and Information Science 329

Jerzy Mikulski (Ed.)

Telematics in the Transport Environment

12th International Conference
on Transport Systems Telematics, TST 2012
Katowice-Ustroń, Poland, October 10–13, 2012
Selected Papers

Springer

Volume Editor

Jerzy Mikulski
Silesian University of Technology
Faculty of Transport
Krasińskiego 8
40-019 Katowice, Poland
E-mail: jerzy.mikulski@polsl.pl

ISSN 1865-0929 e-ISSN 1865-0937
ISBN 978-3-642-34049-9 e-ISBN 978-3-642-34050-5
DOI 10.1007/978-3-642-34050-5
Springer Heidelberg Dordrecht London New York

Library of Congress Control Number: 2012948416

CR Subject Classification (1998): H.4.2, I.6.3-7, I.2.6, I.2.10, K.4.3, J.1, J.2, C.4, H.2.7

Typesetting: Camera-ready by author, data conversion by Scientific Publishing Services, Chennai, India

Printed on acid-free paper

Springer is part of Springer Science+Business Media (www.springer.com)

Preface

Transport telematics are the systems using the information and communication technologies in the area of infrastructure and of means of transport and its participants. An intelligent transport covers systems that allow, through the data transmission and analysis, to influence the behaviour of road users and the action of technical elements in means of transport or along the traffic route. Intelligent transport systems - in accordance with the European Directive - are used for the transport management informatisation. The research shows that the use of telematics can significantly increase the efficiency of the transport system, the road safety and the environmental protection.

Solutions for intelligent transport systems are undergoing rapid technological development, and at the same time they are at the stage of intensive dissemination. The design and development of telematic transport systems is a new integrated approach to solving the transportation problems. Transport telematics offers large opportunities to strengthen the positive characteristics of transport (accessibility, mobility), while simultaneously minimizing its negative impacts (environmental pollution, energy consumption, accidents, infrastructure costs) without incurring large investment expenses.

The implementation of traffic management systems faces many problems. The principal one is the lack of an ITS architecture and standards, and also helpful descriptions of best practices are missing. Paradoxically, this situation could create the opportunity to quickly overcome the delays through the construction of entirely new transport management systems, instead of modernising only the existing ones. Transport institutions more and more eagerly reach for IT solutions because they help to manage the transport efficiently, what combined with good examples for sure will result in new good solutions.

The development, implementation and maintenance of intelligent transport systems is a complex interdisciplinary technical, organizational and financial project. We would like to introduce to you a unique opportunity that is the participation in a meeting of scientists and experienced persons dealing every day with issues of transport telematics. This book provides an overview of solutions being developed in the field of intelligent transportation systems, and includes theoretical and case studies in the countries of conference participants.

October 2012 Jerzy Mikulski

Organization

Organizers

- Department of Railway Engineering, Faculty of Transport, Silesian University of Technology, Poland
- Polish Academy of Sciences, Transport Committee, Poland
- Polish Association of Transport Telematics, Poland

Co-organizers

- Faculty of Transport and Electrical Engineering, Radom University of Technology, Poland
- Faculty of Navigation, Gdynia Maritime University, Poland
- Faculty of Economics, University of Economics in Katowice, Poland
- Faculty of Social and Technical Sciences, Silesian School of Management in Katowice, Poland
- Faculty of Transport, Warsaw University of Technology
- Regional Centre of Road Traffic, Katowice, Poland
- Motor Transport Institute in Warsaw, Poland
- Railway Institute, Warsaw, Poland
- Civil Aviation Personnel Education Centre of Central and Eastern Europe, Poland

Organizing Committee

Chair: J. Młyńczak
Members:
A. Białoń
P. Gorczyca
K. Kędziora
J. Łukasik
S. Surma
Secretary: R. Skowrońska

Conference Took Place under the Patronage of

- Minister of Transport, Construction and Maritime Economy
- Rector of Silesian University of Technology
- Silesian Voivode
- Marshal of the Silesian Voivodeship
- Bogdan Marcinkiewicz, Member of European Parliament

Scientific Program Committee

J. Mikulski – chairman	Silesian University of Technology, Polish Association of Transport, Telematics, Poland
E. Szychta - vice-chairman	Technical University of Radom, Poland
A. Weintrit - vice-chairman	Gdynia Maritime University, Poland
K. Banaszek	Polish Air Navigation Services Agency
R. Bańczyk	Regional Centre of Road Traffic, Poland
A. Bujak	Wroclaw School of Banking, Poland
M. Bukljaš Skočibušić	University of Zagreb, Croatia
K. Chwesiuk	Maritime University of Szczecin, Poland
M. Dado	University of Zilina, Republic of Slovakia
J. Dyduch	Technical University of Radom, Polish Academy of Sciences, Transport Committee, Poland
A. Fellner	Silesian University of Technology, Poland
W. Filipowicz	Gdynia Maritime University, Poland
M. Franeková	University of Zilina, Republic of Slovakia
V. Gavriluk	Dnipropetrovsk National University of Railway Transport, Ukraine
M. Holec	Gdynia Maritime University, Poland
A. Janota	University of Zilina, Republic of Slovakia
J. Januszewski	Gdynia Maritime University, Poland
Z. Jóźwiak	Maritime University of Szczecin, Poland
A. Kalašová	University of Zilina, Republic of Slovakia
D. Kevicky	University of Zilina, Republic of Slovakia
J. Kisilowski	Technical University of Radom, Poland
J. Klamka	Polish Academy of Sciences, Katowice Branch, Poland
B. Kos	University of Economics in Katowice, Poland
S. Krawiec	Silasian University of Technology, Poland
O. Krettek	RWTH Aachen, Germany
R. Krystek	Motor Transport Institute, Warsaw, Poland
J. Lewitowicz	Air Force Institute of Technology, Warsaw, Poland
A. Lewiński	Technical University of Radom, Poland
M. Luft	Technical University of Radom, Poland
B. Łazarz	Silesian University of Technology, Poland
Z. Łukasik	Technical University of Radom, Poland
J. Modelski	Warsaw University of Technology, Poland
G. Nowacki	Motor Transport Institute, Warsaw, Poland
S. Oszczak	University of Warmia and Mazury in Olsztyn, Poland
Z. Pietrzykowski	Maritime University of Szczecin, Poland
B. Piotrowski	Silesian Inspectorate of Road Transport, Katowice, Poland

Referees

Table of Contents

A Decision Support System for Real-Time Evacuation Management and Rescue Team Planning during Hazardous Events in Public Infrastructures

Iraklis Tsekourakis, Christos Orlis, Dimosthenis Ioannidis, and Dimitrios Tzovaras

Information Technologies Institute, Center for Research and Technology Hellas
Thermi – Thessaloniki, 57001, Greece
{tsekour,orlis,djoannid,tzovaras}@iti.gr

Abstract. The emergency evacuation from crowded public transport terminals and critical infrastructures worldwide is a matter of great concern. Numerous research efforts exist, attempting to develop systems to abide by and ensure optimal evacuation guidance from disaster areas. Existing safety systems fail to guide effectively the most vulnerable travelers and to take into consideration the mobility impairment of each individual. In this context, current work elaborates on the issue by proposing a robust evacuation mechanism, incorporated in a Decision Support System. The innovative solution is to provide group-wise optimal routes to the exits, along with personalized routing, taking into account user capabilities and preferences. As a further step, the rescue teams are provided with a plan including prioritized targets with trapped travelers. The paper provides a thorough analysis of existing evacuation models and mechanisms, presents the DSS architecture and embedded algorithms, and discusses its performance.

Keywords: decision support system, emergency evacuation, capacity constraint route planning, rescue team planning.

1 Introduction

Many areas in the world have been and are possible to be affected by severe physical disasters, like earthquake, tsunamis, and hurricanes, whereas fire and floods usually occur as immediate results of these incidents. These disasters have obvious impacts to transport operations and means. And though safety systems are present in every transportation hub, there are still crucial outstanding issues to be addressed and resolved. Public infrastructures such as tunnels and metro stations are paradigms, where travelers are frequently jammed at the exits, even under normal evacuation.

Current work aims to propose an evacuation model based mainly on graph theory with flow analysis. Apart from simple shortest path computations for personalized evacuation guidance, a contraflow configuration algorithm for re-optimization purposes was designed and implemented. Concerning groupwised evacuation, it was decided to apply the Capacity Constraint Route Planner (CCRP) algorithm. The DSS receives real-time data from the localization module and then sends these data to the

J. Mikulski (Ed.): TST 2012, CCIS 329, pp. 1–9, 2012.

crowd and the environmental simulation modules that are responsible for computing the new network situation that is based on the hazard propagation mode. Finally, the DSS provides guidance to the rescue teams by using approximate Deadline-TSP computations. The evacuation plans are also computed in real-time, as imposed by the safety of travelers in an emergency situation. The DSS outputs are filtered and adapted through an intelligent agent framework in order to optimally match each specific user requirements in a personalized manner. The work presented in this paper currently takes place in accordance with the European funded research program SAVE ME [1], and is evaluated in two pilot sites.

The paper is organized as follows: Section 2 discusses state-of-the-art on the theoretical concepts related to evacuation models and algorithms, while Section 3 provides an overview of the DSS architecture, and performs a thorough analysis of the evacuation and rescue plan algorithms and models embedded into the DSS. Section 4 discusses the performance of the DSS in a set of experiments and real-scenario tests. Finally Section 5 proposes future directions and concludes the paper.

2 Evacuation Models

There exists extensive literature related to different kinds of evacuation models. These models are the core of almost every evacuation management system. Various approaches and combinations of them may be applied during the design and the implementation of evacuation management systems. The prominent classification of these models in the literature includes the following categories, the microscopic, the mesoscopic and the macroscopic models. On the one hand, *microscopic* models emphasize the individual movement and reaction of evacuees. On the other hand, *macroscopic* models do not consider individual parameters such as the physical abilities of the evacuees.

First of all, microscopic models were developed to accurately model transportation systems at the individual traveller level. They consider movement of travellers through a network, where the individuals have their own characteristics and respond to the presence of other travellers and to traffic control devices [2, 3]. Available computing power and a requirement for extensive data on roadway geometry and traffic control represents the primary limitation of microscopic models.

Secondly, macroscopic models are mainly used to produce good lower bounds for the evacuation time and do not consider any personal behaviour during emergency situations [4]. Early macroscopic models focused on overall travel demands for a defined evacuation area and examined key measures, such as evacuation clearance times and overall system capacity [5, 6, 7].

Finally, mesoscopic transportation planning tools evolved from a need for a detail level required by microscopic simulation programs in conjunction with the fact that the analysis fidelity did not afford macroscopic models. These models tend to represent the relative flow of vehicles/travellers on a network link, but do not represent individual lanes on the link. Many of the current approaches are based on cellular automata [8].

3 DSS Architecture and Implementation

The DSS implemented in this work, aims to calculate optimal or near optimal routes for every requested evacuation and rescue operation demand by checking for conflicts regarding this route with the appropriate algorithms included in the Personalized Evacuation Module (PEM), the Groupwise Evacuation Module (GEM), and the Rescue Team Planning Module (RTPM), prioritized in the order mentioned above.

Taking into account results taken from the Environmental Simulation Module (ESM), the Crowd Simulation Module (CSM) and the Edge/Arc Affection Module (EAM), the calculation of optimum routes is done by eliminating the combination of them. The combined costs, which are eliminated, are the results of the combination of the environmental, crowd-behavioural and disability (e.g. mobility impairment) related costs. The inputs to the system include the infrastructure network (i.e. a building) and the population distribution with data taken from the Localization Module (LM) in real-time. The calculated routes are then submitted to the Operator Support Module (OpS), which, having a complete view of the network situation, the ongoing and the scheduled routes, it accepts it or rejects it. Moreover, the DSS will calculate alternative routes by using the Replanning Module (RM) in these cases:

- New evacuation plan requested from the OpS.
- The proposed evacuation routes contain segments that have become inaccessible.
- The network situation has changed dramatically during the evacuation cause of a new hazard or the spreading of an existing one.
- There has been a critical change of the evacuees' number.

The above processes continue until the OpS finally allows the application of the proposed evacuation routes. Thus, as it was described above, the DSS considers every evacuation/rescue operation procedure individually. In order to achieve its goal, it combines different types of data in order to achieve high flexibility and adaptation to different kinds of routing problems. Fig. 3 illustrates the general view of the DSS modules and of the peripheral ones.

As is clearly understood from Fig. 1, only a right combination of the existing evacuation models can produce optimal or near-optimal results. The DSS combines all these models in a computational efficient way while in parallel optimal plans are sent to the OpS for approval. Fig. 2 illustrates the combination of microscopic, mesoscopic and macroscopic models that were used by the main parts of the DSS.

The intelligent agent framework module displayed on the left of the DSS in Fig. 3 is the service personalization platform of the DSS. This platform contains a set of agents that enable communication between the DSS and the traveller mobile phone applications. Appropriate agent roles and behaviours are defined targeting the provision of personalized services to end users. The DSS output information is filtered through proper agent behaviours, and as a result, each end user client's application interface is ultimately provided with content tailored to the needs, capabilities and preferences of each respective case. The functionalities of the platform and the agent roles can be seen in Fig. 4.

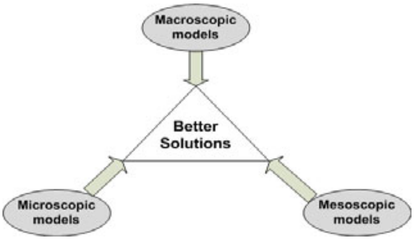


Fig. 1. Evacuation models

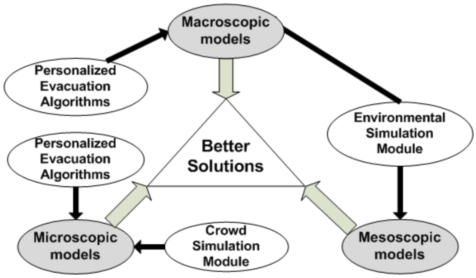


Fig. 2. DSS embedded evacuation models

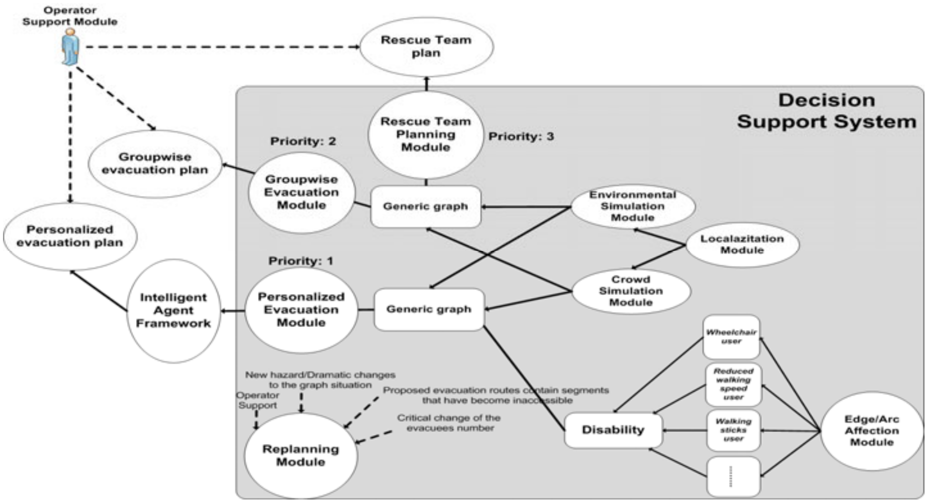


Fig. 3. DSS overall architecture

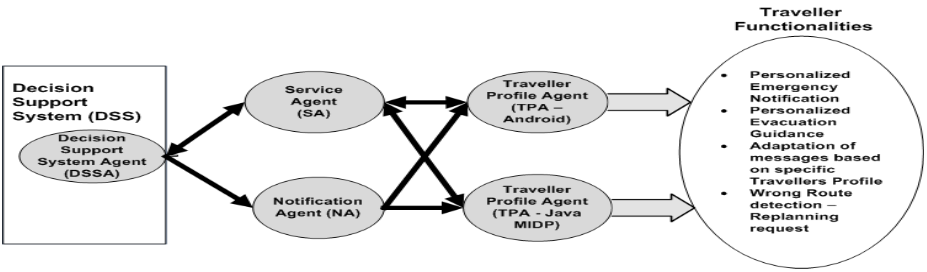


Fig. 4. Intelligent Agent Framework

3.1 DSS Embedded Algorithms

The most striking aspect regarding evacuation management systems is the methodology on how a graph-based model is created, given a specific infrastructure. Graphs are commonly used to model the topological structure of internetworks in order to study problems ranging from routing to resource reservation. Area segments, such as rooms, corridors, stairs and lobbies, are modelled as nodes. The connections between adjacent segments are represented through arcs/edges. Thus, a static network representation $G = (N,E)$ is created.

Groupwised Evacuation Algorithms. MRCCP (Multiple-Route Capacity Constraint Planner) and CCRP (Capacity Constrained Route Planner) [9] are the most adaptable procedures to this problem not only because they compute near-optimal evacuation plans but also because their worst-case time complexity are very good for hard real-time applications. CCRP is the algorithmic procedure that has been selected for the purposes of the SAVE ME project. This is because CCRP has a better worst-case time complexity in contrast with MRCCP which also fits very well to evacuation scenarios but its scalability to large size networks is unsatisfactory.

Personalized Evacuation Algorithms. Personalized routing is the most important functionality of the DSS even if concerns a smaller number of people. The system must follow again the above mentioned algorithmic techniques with some basic changes. In case of a small number of disabled travellers the routes will be provided by shortest path computations, otherwise CCRP is the most effective way to provide the evacuation routes. One major difference between the personalized and the groupwised evacuation procedure is that evacuation routes may differ between people with different distance from a hazard area and people with different disabilities (e.g. a wheelchair user can only use the elevator and not the stairs). In case of more than one person with same special characteristics in the same place, the above mentioned groupwised techniques are applied in a personalized manner in order to reduce the computational effort. Generally speaking, an unassisted manual wheelchair user is almost 45% slower than an average person while an assisted is almost as fast as an average person [10]. Of course, wheelchair users cannot pass through the stairs. Results taken from [10] concerning different kind of disabilities used extensively from the DSS for the personalized evacuation module enhancing in that way the accuracy of the model by applying the special characteristics of each individual group. For example, if the edge type describes stairs then the new edge weight *weight(i,j)’* for the disabled travellers is computed as can be seen in (1),(2) and (3). According to [10] the values of the following parameters were set to: a=1.45, c=2.18 and b was set a very high value, since wheelchair users cannot use the stairs.

$$\text{Reduced speed travelers: weight(i,j)'= weight(i,j)}*a; \quad (1)$$

$$\text{Wheelchair users: weight(i,j)'= weight(i,j)}*b; \quad (2)$$

$$\text{Travelers that use walking sticks: weight(i,j)'= weight(i,j)}*c; \quad (3)$$

Rescue Team Planning Approaches. The Rescue Team planning is a very difficult procedure because the system has to combine different types of information and to decide an optimal scheduling plan for the rescuers. The main objective of this procedure is to secure the safety of the rescuers and of the trapped travellers. To this matter, there exist techniques based on fleet management and vehicle routing problem for collecting hazardous materials and scheduling vehicles under many constraints ([11], [12] and [13]). However, rescuers cannot be considered as vehicles with constant capacities. And this is because rescue teams can be divided through time cause of an urgent need (e.g. a rescuer may decide to help a wheelchair user to exit the building).

In the suggested approach, the Deadline-Traveling Salesman Problem [14] is solved first. As is known, the TSP can be applied only to complete graphs. However, this is not the case for majority of the public infrastructures. To this matter, generalized shortest path algorithms were applied to create the complete graph that is needed to solve the following problem: *Given a metric space G on n nodes, with a start node r and deadlines D(v) for each node v, the Deadline-TSP problem is to find a path starting at r that visits as many nodes as possible by their deadlines*. The approach described above is presented in the following program code.

```
program Deadline-TSP (Output)
{Pre-process: Create a complete graph G'={S ⊆ N , E'}
for the subset S of nodes with trapped travelers by
using shortest path computations};
P = D-TSP(G');
{P ← Computed path from Deadline-TSP, P1 has the
highest priority in the rescue operation};
begin
For i=1 to NS (NS = nodes with trapped trav.)
While (NumberOfRescuersNeeded(Pi)<
NumberOfRescuersSent)
For each RT in RescuTeams
RescueTeam1=Min(Distance(RT, Pi);
ShortestPath.SendtoRescueTeam1(Pi);
End For
End While
End For
If (!NodeInP.SentRescueTeam())
AllRescueTeams.InformAboutHelpInNode(NodeInP);
End If
end.
```

Contraflow Configuration. Contraflow, or lane reversals, has been discussed as a potential remedy to solve tremendous congestion by increasing outbound evacuation route capacity [15]. With a given evacuation situation using an undirected and capacitated graph with multiple sources and multiple destinations, the problem is to find a reconfigured network by contraflow with the objective of minimizing the total evacuation time. This way, each connection obtains double capacity which means that almost double number of people can now pass through each connection. This has as a result much faster evacuation times and increased safety, especially in stressful situations.

4 Results

In order to test the algorithms in the context of groupwised evacuations, multiple tests were realized. However, the tests are computer-based which means that the input data came from some statistical data regarding the daily movement of the passengers in the Monument metro station in Newcastle. In the final tests of the evacuation algorithms, a percentage of 80% of the graph nodes had initial occupancy. Moreover, this scenario included a huge number of travelers ranging from 150 to 600 and with congestion density ranging from 17% to almost 70%.

The following figures present the overall results regarding the evacuation scenarios for different numbers of travelers. As is obvious from Fig. 5, the evacuation times are higher when a hazardous event occurs. This is normal and the objective of the DSS is to reduce these times as much as possible but always in a heuristic manner. CCRP embedded functionality of the DSS produced all these evacuation plans in real-time and with very good overall egress times as can be seen in Fig. 5. The GEM of the DSS may decide to split some travelers located in a specific node. This may have as a result the splitting of the traveler groups during the evacuation procedure. It is certain that travelers are difficult to follow this kind of guidelines but this ensures their safety and the avoidance of a congestion situation during the evacuation procedure. Fig. 6 presents the number of different groups generated by the GEM of the DSS in contrast with the number of the travelers.

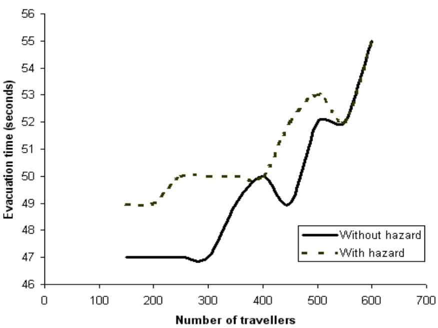


Fig. 5. Evacuation times for different number of travelers

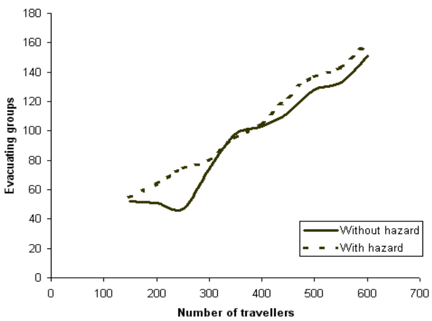


Fig. 6. Number of evacuation groups for different number of traveler

5 Conclusion – Future Work

Current work discusses a Decision Support System, designed and developed for the analysis of evacuation models and systems during hazardous events in public infrastructures. Initially, an overall state of the art analysis on evacuation planning was presented. In this part, different macroscopic, mesoscopic and microscopic models and tools were examined in order to have an overall view of the current research status in this field of study. As a result, it was decided to embed these three models in the DSS functionality. It should be noticed that the DSS delivers a flexible modular architecture, which allows different techniques and implementations to be employed for evacuation planning. The design principles were chosen such as existing components do not have to be changed principally, but only extended by specific functionalities or additional, new applications.

References

1. EU funded SAVE ME SST STREP (FP7-234027), `http://www.save-me.eu/`
2. Braun, A., Musse, S.R., de Oliveira, L.P.L., Bodmann, B.E.J.: Modeling individual behaviors in crowd simulation. In: 16th International Conference on Computer Animation and Social Agents, May 8-9, pp. 143–148 (2003)
3. Yang, L.Z., Zhao, D.L., Li, J., Fang, T.Y.: Simulation of the kin behavior in building occupant evacuation based on Cellular Automaton. Building and Environment 40(3), 411–415 (2005)
4. Hamacher, H., Tjandra, S.: Mathematical modelling of Evacuation Problems: A State of the Art. Pedestrian and Evacuation Dynamics, pp. 227–266 (2002)
5. Chalmet, L.G., Francis, R.L., Saunders, P.B.: Network Models for Building Evacuation. Management Science 28, 86–105 (1982)
6. Montes, Christian: Evacuation of Buildings. M.Sc. Thesis, Department of Mathematics, Universitat Kaiserslautern, Kaiserslautern, Germany (1994)
7. Kostreva, M.M., Wiecek, M.M.: Time Dependency in Multiple Objective Dynamic Programming. Journal of Mathematical Analysis and Applications 173(1), 289–307 (1993)
8. Blue, V.J., Adler, J.L.: Using Cellular Automata Microsimulation to Model Pedestrian Movements. In: Proceedings of the 14th International Symposium on Transportation and Traffic Theory, Jerusalem, Israel (1999)
9. Lu, Q., George, B., Shekhar, S.: Capacity Constrained Routing Algorithms for Evacuation Planning: A Summary of Results. In: Medeiros, C.B., Egenhofer, M., Bertino, E. (eds.) SSTD 2005. LNCS, vol. 3633, pp. 291–307. Springer, Heidelberg (2005)
10. Shi, L., Xie, Q., Cheng, X., Chen, L., Zhou, Y., Zhang, R.: Developing a database for emergency evacuation model. Building and Environment 44(8), 1724–1729 (2009)
11. Wang, F., Yang, M., Yang, R.: Dynamic Fleet Management for Cybercars. In: IEEE Intelligent Transportation Systems Conference, pp. 1246–1250 (2006)
12. Simao, H., Day, J., George, A., Gifford, T., Nienow, J., Powell, W.: An approximate dynamic programming algorithm for large-scale fleet management: A case application. Transport. Science 43(2), 178–197 (2008)

13. Zografos, K., Androutsopoulos, K.: A heuristic algorithm for solving hazardous materials distribution problems. European Journal of Operational Research, New Technologies in Transportation Systems 152(2), 507–519 (2004)
14. Bansal, N., Blum, A., Chawla, S., Meyerson, A.: Approximation Algorithms for Deadline-TSP and Vehicle Routing with Time-Windows. In: Proceedings of the Thirty-Sixth Annual ACM Symposium on Theory of Computing, pp. 166–174 (2004)
15. Kim, S., Shekhar, S., Min, M.: Contraflow Transportation Network Reconfiguration for Evacuation Route Planning. IEEE Transactions on Knowledge and Data Engineering, 1115–1129 (August 2008)

The Analysis of Open Transmission Standards in Railway Control and Management

Andrzej Lewiński[1] ,Tomasz Perzyński[1], and Andrzej Toruń[2]

[1] Kazimierz Pułaski Technical University of Radom
Faculty of Transport and Electrical Engineering
Malczewskiego 29 St, 26-600 Radom, Poland
{a.lewinski,t.perzynski}@pr.radom.pl
[2] Railway Institute
Chłopickiego 50 St, 04-275 Warszawa, Poland
atorun@ikolej.pl

Abstract. The paper deals with problems of delays, queues, interferences and interrupts in open radio transmission standards (GPS, GSM-R) applied in railway control and management systems including Advanced Train Control/Protection\. These systems apply the wireless networks transmitting the signal between different dissipated objects and safety transmission standards require such time related analysis based on mathematical methods (Markov processes).

Keywords: Open transmission standards, Safety Transmission in Railway Control and Management Systems, Markov Process and Time Analysis.

1 Introduction

Modern railway management and control systems are based on wireless solutions of transmission. In these new systems the problem of time delay in transmission is very important with respect to functionality and safety. The delay of proper processing of information connected with typical transmission parameters (etc. interferences, interrupts or integrity code faults) is related to number of transmitting devices and time of service in Dispatcher Centre. The open transmission standards connected with public computer networks include the measures and procedures identical for all users. The application of such solutions, recommended by UE railway organizations (UIC, ORE) requires the application of PN-EN 50159-2010 standard [9] (*Railway applications – Communication, signalling and processing systems – Safety-related communication in transmission systems*) including the earlier recommendations of design of safety computer system in rail control an d management PN129-2007 [8].

The new generation of railway control and management systems use now the radio and wireless radio transmission solutions like: GPS and GSM-R [1, 2]. These systems co-operate with existing railway control systems (including fail-safe control) but give more functionality connected with efficient computer monitoring and management. The main rule is related to assuring the same level of safety corresponding to of

J. Mikulski (Ed.): TST 2012, CCIS 329, pp. 10–17, 2012.

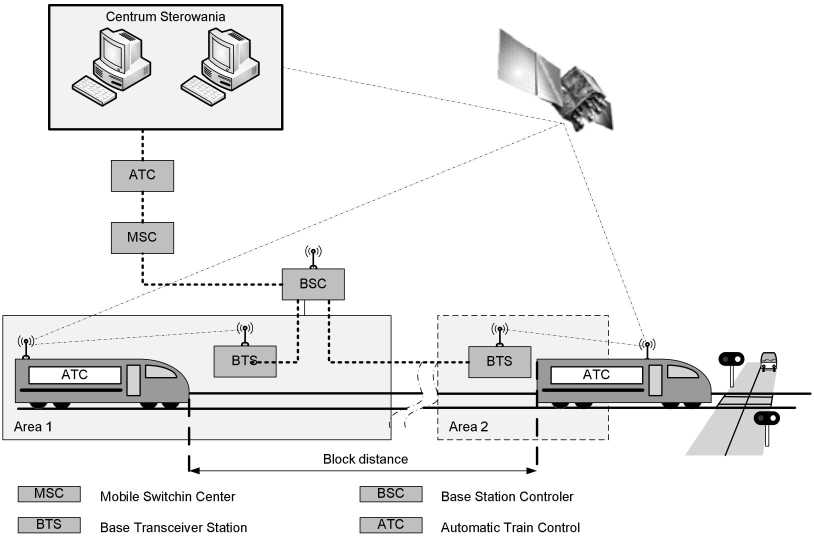


Fig. 1. The open transmission application in Train Management and Control System including the (absolute) Changeable Block Distance Infrastructure

appropriate SIL requirements. The scheme of such co-operation is shown on Fig. 1. Corresponding to above standards the new system - Train control system based on Changeable (absolute) Block Distance (CBD) [5, 6] is elaborated: In the paper the restrictions in application of open transmission standards in this systems are analyzed corresponding to delay of transmission in relation in number of transmitters (equipped trains) and assumed processing time (in control centre).

2 Open Transmission in (Absolute) Changeable Block Distance – Future Railway Control and Management System

2.1 General Characteristics of Open Transmission Systems for CDB Applications

The wireless transmission gives the possibility of dynamic control of the railway systems and cooperation between computers systems [8, 9]. Application the GPS technology to the railway not only enables to supervise the train traffic using the positioning procedure, but may be convenient to synchronize the information between train and dispatcher center,.. In practice there are used the data from four satellites and also information from additional GNSS system (*Global Navigation Satellite System*) which enables to increase the precision of location to 3m. The high precision of measurement of position is necessary. In the real system in order to eliminate the displacement of time there should be used the information also from the fourth satellite. It means that there should be appointed four pseudo–distances and solve the equation:

$$d_i^* = \sqrt{(x_{si} - x_0)^2 + (y_{si} - y_o)^2 + (z_{si} - z_0)^2} + c\Delta\tau \qquad (1)$$

where:

x_{si}, y_{si}, z_{si} - co-ordinate of satellite in rectangular geocentric system

x_0, y_0, z_0 - searching object co-ordinate in rectangular geocentric system

$\Delta\tau$ - displacement object time scale in relation to GPS time scale

c - speed of electromagnetic wave

The GSM-R is well known radio transmission standard applying typical infrastructure from Fig.1: Mobile stations, Base Transceiver Stations, Base Station Centre/ Controller and Automatic Train Centre. for each connection in ATC application the two channel connection (one for speech and second for data) is realized.

The frequencies allocated in Europe to GSM-R is in the 900MHz band : 876 – 880 MHz uplinks (mobile transmit) and 921 – 925 MHz downlinks (base station transmit). For GSM-R we have available 4MHz spectrum witch 19x200kHz frequencies (channels) [11]. From, stand point of safety data transmission GSM-R network must guarantee a high quality of service – data transfer between Radio Block Center and Train (ETCS On-board Unit). Therefore, to ensure that the quality of service parameters required by ETSC, the planning for GSM-R networks must take into account following rules: Minimum field intensity for train speed:

- Lower than 220 km/h : -95dBm,
- Higher than 280 km/h : -92dBm.

2.2 Time Delay in Changeable Block Distance

The proper position of train depends on speed train, because the loss of GPS signal may be connected with distance In the GPS space segment there are atom clocks. Depending on level of precision, in the PPS (*Precise Positioning Service)* time is assumed as ~ 100ns but on SPS level (*Standard Positioning Service*) as ~ 300ns. Such high precision is necessary, because the slightest error can cause the measurement errors in the order 0,5 m.

When the system establishes connection again (finding satellites, synchronization, reading navigation dispatch), there appear considerable delays. In the situation of a lack of information essential to determine actual configuration of satellites, time of delay run into even a dozen or so minutes and is called Time to First Fix (TIFF).

In GPS system the process of bit synchronization appears every 6s (sometimes there is a need to confirm and repeat it again). When the train drive in tunnel and communication interrupts, renewed connection can take 30s (it is time indispensable for finding a position and to correct time) [3]. For these reasons the delay in GPS system can be estimated:

$$T_{GpS} = T_S + \Delta T \tag{2}$$

TrS – time of synchronization / loss of transmission (6.0 – 30.0 s)

ΔT – time measurement error (0. 000001 – 0.000003 ms)

The GPS delay in train speed function is presented in Table 1).

Table 1. The train distance covered during a loss of transmission with the GPS system

Time	Speed of train			
	60km/h	100km/h	160km/h	250km/h
	16,66m/s	27,77m/s	44,44m/s	69,44m/s
6s	99,96 m	166,62 m	266,64 m	416,64 m
15s	249,9 m	416,55 m	666,6 m	1041,6 m
30s	499,8 m	833,1 m	1333,2 m	2083,2 m

Typical delay data telegram time ($T_{DGSM\text{-}R}$) in the GSM-R (GSM) standard is a sum of the individual partial times: [11, 12]

$$T_{DGSM-R} = T_S + T_R + T_{FTY} + T_{DT} + T_{RC} \tag{3}$$

TrS - conection establish time (2.0 – 10.0 s)
TR – registration at network time delay (30-40s)
TrFT – maximum time of wrong transmission (1 s)
TDT – data transmission time delay (1s)
TRC – recovery time– not interrupted (7s)

3 Delays in Communication Systems with Open Transmission Standards

3.1 Delay Estimation

The assumption of M/M/1 queuing system is an exponential distribution of packet service time and Poisson arrival process, it means that the probability of sending a packet from the sensor node in successive units of time is constant. In Fig. 2 has shown the model of packet transmission what reflects the stationary model of Markov process [4, 7, 10].

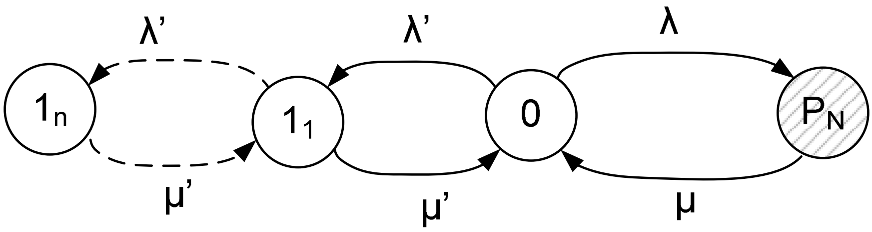


Fig. 2. Model of packet transmission in Markov process

In this model, we distinguish states:

- State 0 – transmission of telegrams without delay (or delay negligibly small),
- State P_N – loss of communication (transmission interrupt, delay, corruption),
- State Pi – state of waiting of telegrams in the queue.

It is thus possible to estimate the queue length N and the average waiting time for service Ts [10]:

$$N = \frac{\lambda}{\mu - \lambda} = \frac{\frac{\lambda}{\mu}}{1 - \frac{\lambda}{\mu}}, \; Ts = N \frac{1}{\lambda} = \frac{1}{\mu - \lambda} = \frac{\frac{1}{\mu}}{1 - \frac{\lambda}{\mu}} \tag{4}$$

where:

- λ – intensity of packets sent per unit time (1/λ – mean time between appearance of the packets),
- μ – the intensity of packet service (1/μ – average service time of packets).

Based on equation (4) the length of the queues and time of delay (waiting for/and service) may be estimated in Train Control System based on Changeable (absolute) Block Distance (CBD) [6].

3.2 Delays in the CBD System

According to the estimates in section 2, assumed that each train in a controlled area communicates with the center at least once per second, and the average service time of telegram (packet) is:

- 1s. (normal transmission –with t_{DT} of service),
- 15s (short loss of GPS signal, normal GSM-R transmission)
- 65s. (15s in GPS and 50 sec. in GSM-R as a worst case of transmission loss in both systems).

Table 2. Values of N i Ts in CBD System

Number of trains per 100sec.	λ	μ=1 $^{-1}$s^{-1}		μ=15 $^{-1}$s^{-1}		μ=65^{-1}s^{-1}	
		N	Ts	N	Ts	N	Ts
2	0,02	0,02	1	0,42	0,21		
5	0,05	0,052	1,05	3	60		
8	0,08	0,087	1,08	Queue grows infinitely due to the inability of service (λ > μ)			
12	0,12	0,13	1,13				
20	0,2	0,25	1,25				

3.3 The Modeling of CBD of System Functionality with Time Parameters

The CDB train control and monitoring system equipped in CDB infrastructure is presented on Fig.3. The following states (corresponding to Markov process) may be assumed:

- 0 – Waiting for service (queue for input trains),
- 1 – Decision about type of service,
- 2 – Service with CDB procedure,
- 21 – Service with fixed block distance procedure (without CDB),
- 31 – Emergency train drive („visibility" drive for first signaling device),

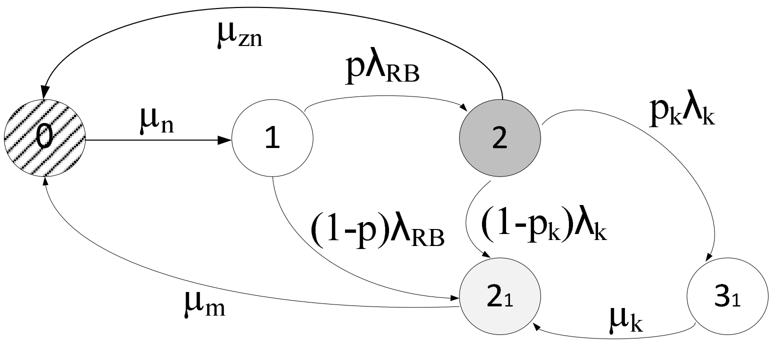


Fig. 3. The functional model of train monitoring system with CDB system

and following transitions between states:

- μ_n – intensity of service in CC (reciprocal to time of waiting for service),
- λ_{RB} – intensity of service in CDB,
- μ_{zm} – intensity of CDB service (reciprocal to time of CDB service),
- μ_m – intensity of CDB service without CDB,

- λ_k – Intensity o radio transmission faults in CC (failure rate),
- μ_k – intensity of service with fixed distance train control without CDB (reciprocal to time of achievement the first , signaling device).

The introduced probabilities require the level of CDB equipment in incoming trains (so called "mixed traffic") :

- p – medium probability (in %) of CDB equipment in incoming trains during day),
- p_k –probability of catastrophic failure of CDB (GPS/GSM-R critical transmission fault).

The very important for functional analysis is estimation of boundary probabilities connected with occurrence the state 2 and 21. These values may be evaluated with respect to formula $P_i = P_i(t)\big|\ t \to \infty$.

$$P2 = \frac{\mu_m p \lambda_{RB} \mu_n \mu_k}{\lambda_{RB}\mu_n\mu_k(\mu_{zn} - \mu_{zn}p + \lambda_k) + \mu_m((\lambda_{RB} + \mu_n)(\mu_{zn}\mu_k + \mu_k\lambda_k) + p\lambda_{RB}\mu_n(\mu_k + p_k\lambda_k)} \quad (5)$$

$$P21 = \frac{\lambda_{RB}\mu_n\mu_k(\mu_{zn} - \mu_{zn}p + \lambda_k)}{\lambda_{RB}\mu_n\mu_k(\mu_{zn} - \mu_{zn}p + \lambda_k) + \mu_m((\lambda_{RB} + \mu_n)(\mu_{zn}\mu_k + \mu_k\lambda_k) + p\lambda_{RB}\mu_n(\mu_k + p_k\lambda_k)} \quad (6)$$

Assuming values of coefficients, the calculation of probability P2 and P21 is presented on Table 3. Two cases are analyzed:

- 1 – light traffic (2 trains per. 100s., CBD equipment - 20%),
- 2- heavy traffic (12 trains per. 100s., CBD equipment – 80%).

(These values and others from the Table 3 are assumed arbitrary).

Table 3. Values of P2 and P21

	Values	
	Case 1	**Case 2**
μ_n	1/10s	1/5s
λ_{RB}	0.02s	0.12s
μ_{zm}	1/25s	1/12.5s
μ_m	1/5min	1/2min
λ_k	1/2days	1/10days
μ_k	1/4min	1/2min
p	0.2	0.8
p_k	0.7	0.3
P_2	**0.01**	**0.25**
P_{21}	**0.94**	**0.61**

It's obvious that CBD system is designed for heavy traffic and trains equipped with CBB radio devices. The time parameters have no significant influence for service with CBD procedure.

4 Conclusion

The analysis of CBD system with open radio transition shows, that assuming typical delays related to transmission failures must be restricted with respect to number of transmitters and time of service in the dispatcher center. For CBD system the lost of transmission requires the switch to the classical fixed distance control with insulated rail sections (for normal operation the number of train may be 20). Cause of delay in the GPS and GSM-R networks, the CBD system can not be a main source of safety information of train localisation and must work together with existing speed and distance measurement train systems (odometer). The presented system may be treated as an additional radio transmission overlay into existing infrastructure (disability of it has no influence for fail safe operation of now exploited rail control systems).

References

1. Latré, B., De Mil, P., Moerman, I., Van Dierdonck, N., Dhoedt, B., Demeester, P.: Maximum Throughput and Minimum Delay in IEEE 802.15.4. In: Jia, X., Wu, J., He, Y. (eds.) MSN 2005. LNCS, vol. 3794, pp. 866–876. Springer, Heidelberg (2005)
2. IEEE Std 802.11-1999 Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) specifications
3. Jackowski S., Chrzan M.: Contemporary telecommmunications systems. Technical University of Radom (2008) (in Polish)
4. Lewiński, A., Perzyński, T.: The delay analysis in dissipated railway management and control systems. In: Telematic Transport Systems Conference (2005)
5. Lewiński, A., Toruń, A.: The Changeable Block Distance System Analysis. In: Mikulski, J. (ed.) TST 2010. CCIS, vol. 104, pp. 67–74. Springer, Heidelberg (2010)
6. Perzyński, T.: The problems of safety of computer nets applied in the railway control. PhD dissertation – Technical University of Radom, Faculty of Electric Engineering and Transport, Radom (2009) (in Polish)
7. Standard PN-EN 50129:2007 Railway applications – Communication, signalling and processing systems. Safety related electronic systems
8. PN-EN 50159 – 2010. Railway applications – Communication, signaling and processing systems – Safety-related communication in transmission systems
9. Tanenbaum, A.S.: Computer Networks. Prentice Hall PTR, New Jersey (1996)
10. UIC No. O-2475: ERTMS/GSM-R
11. Quality of Service Test Specification form 11.09.2003 and parameters: GSM-R Interfaces Class 1 Requirements. SUBSET-093 z 11.10.2005
12. Winter, P., et al.: Compendium on ERTMS, UIC, 1st edn. (2009) ISBN 978-3-7771-0396-9

Evaluation of Microsimulated Traffic Light Optimisation Using V2I Technology

Wolfgang Niebel[1], Oliver Bley[2], and Rüdiger Ebendt[1]

[1] German Aerospace Center (DLR), Institute of Transportation Systems
Rutherfordstr. 2, 12489 Berlin, Germany
{wolfgang.niebel,ruediger.ebendt}@dlr.de
[2] Technische Universität Braunschweig, Institute of Transportation and Urban Engineering
Rebenring 31, 38106 Braunschweig, Germany
o.bley@tu-braunschweig.de

Abstract. Within the research project KOLINE a cooperative system for urban road transport is developed. Its traffic related goals are to reduce travel time and fuel consumption as well as noise and pollutant emissions. The herein described evaluation shall determine whether, to which extent and how economically the KOLINE system is able to address these goals. Three differently comprehensive quantifying evaluation procedures are applied to the outputs of a microscopic traffic simulation. Thus not only the ranking of all scenarios, but also comparisons between these procedures become possible.

Keywords: transport economy and policy, cost-benefit-analysis CBA, microscopic simulation output, data analysis.

1 Introduction and Project Aim

Within the German research project KOLINE a cooperative system for urban road transport is developed and tested. In order to reduce travel time and fuel consumption as well as noise and pollutant emissions it uses the communication between vehicles and traffic light infrastructure (V2I) to avoid stops of vehicles at traffic lights and to increase the network capacity. The system description, technical aspects and the specific component for tailback estimation are discussed in detail within two other papers presented at this conference.

This paper concentrates on two aspects of the always recommendable traffic-related evaluation at the end of ITS projects. The first aspect is the in-depth evaluation for several scenarios to determine whether and to which extent the fully productive technical KOLINE system is able to address the above mentioned goals. The thereto applied methods and procedures, the microscopic traffic simulation as single source of evaluation input data, and exemplary outcomes are described.

The second aspect is the investigation of possibilities to produce general project-independent recommendations for future ITS project evaluations. Therefore the available methodology is analyzed and some major terms are defined in the next chapter.

J. Mikulski (Ed.): TST 2012, CCIS 329, pp. 18–25, 2012.

2 Goals, Evaluation Methodology and Procedures

2.1 Goals

Literature lists up to six *global goals* for transport-related projects, which can be grouped and summarized as follows [1, 2]:

1. Mobility (Travel Time)
2. Resource Efficiency (Energy, Money)
3. Environment Friendliness (Noise, Pollutants)
4. Safety
5. Security
6. Customer Satisfaction.

While as the first three global goals are explicitly addressed by the KOLINE project, the latter three ones are nevertheless regarded in other ways:

4. It is investigated whether safety estimations could be gained from the simulation's single car trajectories.
5. A secure implementation of the developed technology is considered as an a-priori condition from the point of evaluation and therefore assumed as being fulfilled.
6. Customer satisfaction expresses itself amongst others in the degree of acceptance of automatic driving manoeuvres, e.g., the automatic vehicle speed control. This degree is inserted as scenario simulation input parameter rather than calculated as output.

The global goals are broken down into more detailed and project tailored *sub-goals* or *criteria*. These criteria need to be operative by *indicators*, meaning a suitable performance figure which always includes a denominator has to be defined. In a final step the necessary physical *measures* for each indicator are determined [3a]. Table 1 shows the global goals, their criteria and physical measures (without denominator) in KOLINE.

2.2 Evaluation Methodology

The generated simulation output values of the measures need to be interpreted to understand their impact. This interpretation process bases on some generic concepts - called *methods* or *techniques* - which can be carried out through different specific formalized *procedures*. The evaluation methodology for transport-related projects in general comprises several methods, which can be of either describing or quantifying nature. The latter ones further divide into non-monetizing and monetizing approaches. Representatives of the first approach are the Multi-Criteria (Utility) Analysis (MCA) and mono- or multi-criteria based qualitative assessments (QA), whilst the most popular monetizing method is a Cost-Benefit-Analysis (CBA). Legal binding CBA procedures with detailed execution directives and values are, e.g., New Approach to Appraisal (NATA) in the UK, Bundesverkehrswegeplan (BVWP) in Germany [4], and System Informatyczny Monitoringu i Kontroli Finansowej Funduszy Strukturalnych i Funduszu Spójności (SIMIK) in Poland.

The inventory of possible evaluation methods and procedures especially for ITS projects yields a broad range of results, but also shows that neither a common approach nor a dedicated or even compulsory procedure yet exist. Examples stretch from simple descriptive [5] or multi-dimensional quantifying reports [6] and specific Multi-Criteria-Analysis [7] up to recommendations for [3b] or application of Cost-Benefit-Analyses [8].

Since no clear preference or advantage for one of these approaches evolves, a bunch of three different procedures representing three methods are chosen. Thus not only the evaluation of all scenarios and their impact, but also the assessment whether these differently comprehensive methods are fit for the given task and which one is the best becomes possible.

2.3 Evaluation Procedures

The applied and - where necessary - modified procedures can all be counted as quantifying ones. Namely they are the compulsory determination of the Level of Service (LOS) according to the German HBS [9], the Performance Index (PI) calculation derived from TRANSYT [10], and the Cost-Benefit-Analysis for German road infrastructure projects "EWS" [11]. Each procedure makes use of a different number and range of the above stated criteria to determine the respective result (Table 1). Due to the procedures' fixed definitions these results are generally comparable with outcomes of other projects.

Table 1. Global Goals, their Criteria and Measures, and Procedures in KOLINE

Global Goals	Criteria	Measures	Procedure		
			HBS	**PI**	**EWS**
Mobility	Travel Time	s; min; h			x
	Delay Time / Stop Time		x	x	x
	Number or Percentage of Stops	n; %	x	x	
Environment Friendliness	Pollutant Emissions (No_x, CO, HC, PA)	g; t			x
	Climate Gas CO_2	g; t			x
	Noise Emissions	db(A)			x
	Fuel Consumption	l			x
Resource Efficiency	Building / Acquisition Costs	€			x
	Operating + Maintenance Costs	€			x
	Occupancy Rate	n; %		x	x

HBS. The Level of Service (LOS) according to the German HBS is stated separately for each accessing lane and traffic mode of a single junction and often uses the peak hour as its time denominator. The sole criterion for signal actuated intersections is the average waiting time. On signal actuated coordinated sections it is the percentage of unstopped vehicles. Table 2 gives an overview of the threshold values.

Table 2. Threshold values for Level of Service (HBS)

Level of Service	Average waiting time *w* [s]				Percentage of unstopped vehicles [%]
	public transport	bikes	pedestrians*	motor vehicles (uncoordinated)	motor vehicles (coordinated)
A	≤ 5	≤ 15	≤ 15	≤ 20	≥ 95
B	≤ 15	≤ 25	≤ 20	≤ 35	≥ 85
C	≤ 25	≤ 35	≤ 25	≤ 50	≥ 75
D	≤ 40	≤ 45	≤ 30	≤ 70	≥ 65
E	≤ 60	≤ 60	≤ 35	≤ 100	≥ 50
F	> 60	> 60	> 35	> 100	< 50

* supplementary 5 s when crossing a divisional island

PI. The bi-criteria Performance Index synthesises the waiting time w and the number of stops h of all traffic modes z and all access sections i of an intersection (Eq. 1). The weight G_h is assumed to be 60, since the emissions of a start-up after a stop equal 60 seconds idling. In difference to the HBS not vehicles but passengers P are calculated with, thus incorporating the occupancy rate. The car occupancy rate in Braunschweig is about 1.23 people per car. The rate for public transport buses is taken from detailed passenger survey data and differs between 1 passenger in the evening and more than 90 passengers in the morning peak. While in [10] only stops of motor vehicles and buses are included in the second addend of Eq. 1, KOLINE extends this to cyclists due to comfort aspects. Normally the peak hour is used as time denominator, whileas the areal denominator can comprise a single or multiple junctions.

$$PI = \frac{\left(\sum_i \sum_z w_{i,z} \times P_{i,z} + G_h \sum_i \sum_z h_{i,z} \times P_{i,z}\right)}{\sum_i \sum_z P_{i,z}} . \quad (1)$$

EWS. This comprehensive Cost-Benefit-Analysis is based on eight benefit criteria of which six are used in KOLINE. The benefit is defined as the difference between the criterion's values of the base scenario and a comparison scenario. A project is worth to be realized when the benefits are greater than the costs.

3 IT Setup

Although the technical feasibility and security of the system are demonstrated with several research vehicles on a public road section, a realtime all vehicles comprising field operational test is impossible. Thus the assessment is solely based on the outcomes of a microscopic traffic simulation, which is described below.

3.1 Microscopic Traffic Simulation

The test area is an urban sub-network in Braunschweig, Germany, with three signalized intersections at a total distance of 700 m. The network is part of an arterial road of its inner city traffic with an average daily traffic of up to 36.220 vehicles. It is thus important for commuters, but also for other traffic, e.g., six public transport bus lines pass the network in several directions. All streets running east-west and the streets running north-south at its central intersection have two lanes per direction with some additional turning lanes at signalised intersections. The remaining streets have one lane per direction. The speed limit is 50 km/h.

The representation of the network as modelled with TSS software AIMSUN 6.1.5 [12] is shown in Fig. 1. The numbers of the nodes correspond to the numbers of real traffic signal systems as assigned by the Braunschweig authority in charge of signalisation. Detectors are located on each lane about 20 m in front of traffic signals. However, not all turning flows are detected directly because of mixed lanes.

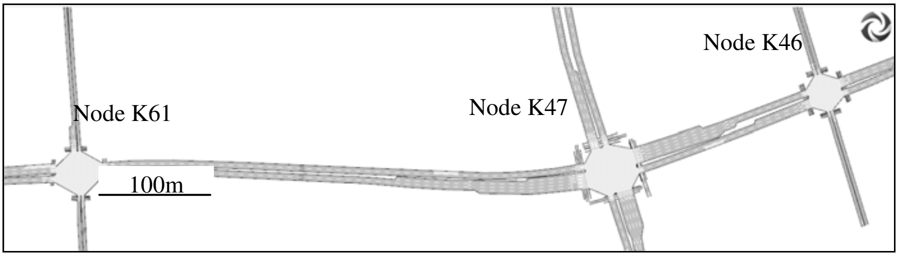


Fig. 1. Simulation model of the test area in Braunschweig, Germany

The simulation had to be specified concerning several aspects. The traffic demand has been deduced from real measurements on-site that have been taken on May 12, 2011 including detector counts over the whole day. Beside individual motorised traffic and public transport also cyclists and pedestrians are regarded within the simulation, which runs from 6 am to 10 pm. The pollutant emissions are calculated with factors taken of the German HBEFA emission model. For this purpose a total of seven vehicle types are defined to take account of the different pollutant emissions resulting from cars with gasoline or diesel engines, vans, lorries etc. The simulation model also maps the four essential system subcomponents, i.e., the signal program optimizer, the equipment rate of the V2I communication device, the automatic vehicle speed control for the approach strategy, and the tailback estimator. Table 3 gives an overview of how the parameters of these subcomponents are differentiated within the defined 12 scenarios. The vehicles penetration rate with V2I communication devices in scenarios Test 7 to Test 10 is chosen after the outcomes of scenarios Test 2 to Test 6 to give proper information about a suitable value to produce significant results.

The simulations' computed and semi-aggregated output data of 15-minutes-intervals comprises mean values and deviations for common traffic flow measures, as well as pollutant emissions. 25 replications with different random seed numbers have been run and the average of all obtained values has been calculated.

Table 3. Parameter differentiation within the Simulated Scenarios

Scenario	Applied Traffic Control	Vehicle Approach Strategy	Vehicle Penetration Rate	Tailback Estimation
Status Quo	Status Quo	No	0%	-
Reference	TRANSYT optimized	No	0%	-
Test 1	Signal Program Optimizer	No	0%	-
Test 2	Signal Program Optimizer	Yes	5%	detectors + vehicle data
Test 3	Signal Program Optimizer	Yes	10%	detectors + vehicle data
Test 4	Signal Program Optimizer	Yes	15%	detectors + vehicle data
Test 5	Signal Program Optimizer	Yes	20%	detectors + vehicle data
Test 6	Signal Program Optimizer	Yes	25%	detectors + vehicle data
Test 7	Signal Program Optimizer	Yes	5 -25%	detectors + vehicle data
Test 8	TRANSYT optimized	Yes	5 -25%	detectors + vehicle data
Test 9	Signal Program Optimizer	Yes	5 -25%	-
Test 10	Signal Program Optimizer	Yes	5 -25%	detectors

3.2 Evaluation Software

In a preparing step the voluminous AIMSUN generated MS Access Database of each scenario is filtered to leave only the average values of those 12 sections, which are accessing the three nodes under investigation. Since AIMSUN does not put out the number of stops and stop time for each lane as required by the German HBS, these values have to be estimated from the respective turn values and the ratio of turn tailback lengths. In the main step each one of the three evaluation procedures is applied to the database with JAVA software modules developed by the DLR-Institute of Transportation Systems. The results are stored in the respective scenario database.

4 Evaluation Results and Recommendations

The LOS and PI for bikes and pedestrians remain the same throughout the day at each of the three nodes since no changes to the signal plans are applied in the Status Quo scenario. Table 4 summarizes all relevant values for each node and in comparison PI evaluation outcomes for a project with multiple junctions in the German city of Münster [10], whereas Münster 0 stands for the baseline with conventional traffic light systems and Münster II for an optimized traffic adaptive regime. The LOS denotes for both East↔West and North↔South crossings. As stated in Sect. 2.3 the cyclists' PI is originally calculated without taking stops into account. It significantly rises when including this comfort aspect as the PI* values show.

Table 4. LOS and PI for pedestrians and cyclists

Node	Pedestrians			Cyclists			
	LOS E↔W	**LOS N↔S**	**PI**	**LOS E↔W**	**LOS N↔S**	**PI**	**PI***
K61 (West)	A	D	22	A	C	22	64
K47 (Middle)	D/E	E	36	C/D	D	37	92
K46 (East)	A	C	20	A	C	20	60
Test area	n.a.	n.a.	25	n.a.	n.a.	30	78
Münster 0	n.a.	n.a.	21	n.a.	n.a.	22	
Münster II	n.a.	n.a.	32	n.a.	n.a.	31	

* Performance Index for cyclists taking stops into account

The LOS values for the secondary crossing points in North↔South direction at the small junctions and for all crossing points at the big middle junction are only acceptable or even poor. Nonetheless the PI at both small junctions seems to be good in comparison to Münster. The PI for the big junction mirrors the bad LOS classification.

The PI of public buses and motor vehicles throughout the day is shown exemplary for the big junction K47 in Fig. 2. The morning peak time is depicted by the obviously augmented graph, heavily influenced by the bad performance of public buses. There seems to be no afternoon peak, but in the evening the PI shows the best values. A comparison with the morning and afternoon peak values from the Münster II optimized scenario give an idea which enhancements could be achieved within the KOLINE project.

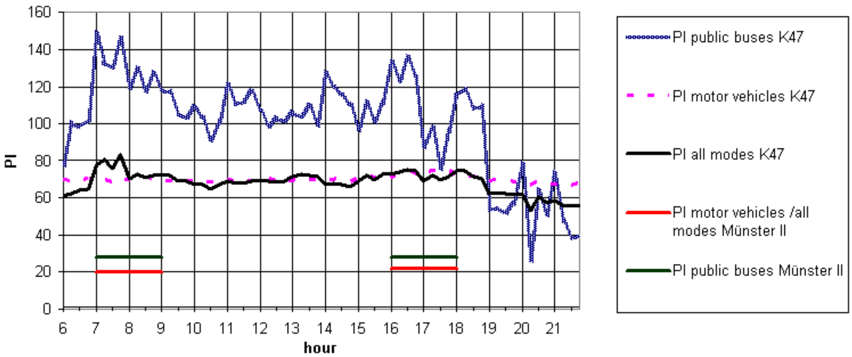


Fig. 2. PI for public buses and motor vehicles at node K47 and in Münster

The advantages of the PI compared to the LOS become clear. A PI allows to weigh the LOS of each crossing point and each transport mode by including the different amounts of affected people. It also enables the summarization of multiple junctions which lay in a row or within an area.

Nonetheless this Performance Index only enables a ranking of different scenarios and their quantifying rate against each other. It does not pay attention to necessary investments into technical equipment and engineering work and thus could prefere scenarios with excellent traffic improvements but a negative cost-benefit-ratio due to high costs. Therefore a CBA should be executed at all time for new ITS applications. Recommended monetizing values for all EU 25 countries are stated in [13].

Acknowledgements. The KOLINE project was granted by the German Federal Ministry of Economics and Technology (BMWi) according to a decision of the German Federal Parliament within the 3rd transport research framework "Mobility and Transport Technologies". The authors like to thank additionally our project partners Institut für Automation und Kommunikation e.V. Magdeburg (ifak), Institute of Control Engineering TU Braunschweig, Transver GmbH Munich, and Volkswagen AG Wolfsburg (VW).

References

1. Transport Canada (publisher): Development of a Project Evaluation Methodology Framework for Canadian Intelligent Transportation Systems. TP 14755E, Ottawa (2007)
2. Boltze, M., Jentsch, H., Friedrich, B., Bastian, M.: OptiV – Erschließung von Entscheidungs- und Optimierungsmethoden für die Anwendung im Verkehr. TU Darmstadt (2006)
3. FESTA project syndicate (publisher): Field opErational teSt supporT Action (FESTA) Handbook, Deliverable D6.4, [3a] Chapter 5, [3b] Chapter 10 (2008)
4. Bundesministerium für Verkehr, Bau- und Wohnungswesen (publisher): Bundesverkehrswegeplan 2003, Grundlagen für die Zukunft der Mobilität in Deutschland. Beschluss der Bundesregierung vom 2, Berlin (Juli 2003)
5. ORINOKO project syndicate (publisher): Anlage Ausführlicher Ergebnisbericht mit Konzept zum Weiterbetrieb (2008)
6. Dmotion project syndicate (publisher): Düsseldorf in Motion, AK 600 – Feldversuche und Evaluierung des Gesamtsystems. Project sheet (2006)
7. CityMobil project syndicate (publisher): Ex ante evaluation report. Deliverable D5.3.1a (2009), `http://www.citymobil-project.eu/downloadables/Deliverables`
8. FAIR project syndicate (publisher): Fully Automatic Integrated Road Control Work-Package 4.1: Cost Benefit Analysis. Deliverable D04, Berlin (2006)
9. FGSV (publisher): Traffic Light Actuated Intersections. In: Handbuch für die Bemessung von Straßenverkehrsanlagen (HBS), ch. 6. FGSV Verlag GmbH, Köln (2001) ISBN 3937356444
10. Wietholt, T.: Einsatzbereiche Grüner Wellen und verkehrsabhängiger Steuerungen. Dissertation, Schriftenreihe des Lehrstuhls für Verkehrswesen Ruhr-Universität Bochum, Heft 33, Bochum (2009) ISSN 1437-8299
11. FGSV (publisher): Empfehlungen für Wirtschaftlichkeitsuntersuchungen von Straßen (EWS). FGSV Verlag GmbH, Köln (1997)
12. Transport Simulation Systems S.L. (TSS): AIMSUN 6.1 User's Manual, Barcelona (2011)
13. HEATCO project syndicate (publisher): Developing Harmonised European Approaches for Transport Costing and Project Assessment, HEATCO Deliverable 5. Stuttgart (2006)

Electric Vehicle for the Students' Shell Eco-Marathon Competition. Design of the Car and Telemetry System

Kamil Sternal, Adam Cholewa, Wojciech Skarka, and Mirosław Targosz

Silesian University of Technology, Faculty of Mechanical Engineering
Konarskiego 18 A, 44-100 Gliwice, Poland
Sternal.Kamil@draexlmaier.de
{adam.cholewa,wojciech.skarka,miroslaw.targosz}@polsl.pl

Abstract. This paper describes the design of a prototype racing car designed for racing Shell Eco-marathon. In particular, the assumptions have been made to the vehicle simulation model used to predict the race strategy. These assumptions are the basis for an integrated methodology for designing racing cars, which provide a three-year development plan of the vehicle. Additionally, there are modular telemetry system, and computer applications for data visualization in the race.

Keywords: telemetry system, race car, Shell Eco-marathon, simulation model, energy consumption.

1 Introduction

The Shell Eco-marathon is one of the most prestigious students' competitions across the world. The main objective is to design, build, and, of course, drive vehicles devised and engineered for one purpose i.e. to be the most-fuel efficient.

Onset of the Shell Eco-marathon dates back to 1939, and now in a European edition itself more than 3000 students from 27 countries take part in it.

This year, which was the 28th edition of the race, to make it even more spectacular and demanding, the race took place at a street track in the area around the Ahoy Arena in Rotterdam.

There are 2 racing classes:

- Prototype – vehicles specifically designed to lower the drag coefficient.
- Urban Concept – the future of urban mobility. Vehicles resembling cars, however with high fuel efficiency.

Each class can be driven by different types of energy sources:

- Internal combustion engines
    - Shell FuelSave Unleaded 95 (the UE i Asia) / Shell Regular 87 (the USA) Petrol/Gasoline,
    - Shell FuelSave Diesel (the UE) / Shell Diesel (Asia and the USA),
    - Shell Gas To Liquid (100% GTL) (the UE and Asia),

J. Mikulski (Ed.): TST 2012, CCIS 329, pp. 26–33, 2012.

  - Biodiesel (Fatty Acid Methyl Ester (100% FAME)),
  - Ethanol E100 (100% Ethanol),
- Electric motors:
  - Hydrogen,
  - Solar Energy,
  - Electric Energy.

The decision to take part in Shell Eco-marathon originates from the experience in highly successful and world-renowned Silesian Greenpower [7] and the participation in Greenpower Corporate Challenge in the years 2010, 2011 and 2012. It was due to the fact that there are greater requirements and higher technical level of Shell Eco-marathon as well as greater possibilities resulting from the lack of regulation limits. It has been decided that the team will built a vehicle in the category Prototype driven by electric engine fed by batteries, category Battery Electric. The range of design development is foreseen for three years. The first year tasks included:

- construction of composite system of low mass and composite-metal suspension durable for three-year- consecutive starts
- construction of steering and suspention system made of ready-made components
- construction of our own engine steering system, which is capable of implementation of different algorithms and control strategies
- construction of module data acquisition system which would be able to further expand and communicate with host computer

As a sport aim of this year it has been decided to:

- take part and finish the race and in this way get data for further analysis
- get a satisfactory result as for a debut in so renowned competitions

The team of the Mechanical Engineering Faculty has chosen the name Smart Power and the car was called MuSHELLka [3]. Smart Power Project was established in October 2011. The race took part on a street track in Rotterdam in May 2012 and the aim was to drive a route of about 16 km in a time limit no longer than 39 minutes. After the ride the judges read the energy consumption, depending on the class it can be liquid fuel, hydrogen or electrical energy and make calculations in order to present the results in a form of number of kilometers per energy unit e.g. km/l or km/kWh. In a given class a vehicle with the longest distance covered wins.

2 The Concept of a Shell Eco-Marathon Race Car

The composite body of the race car was designed using Catia V5, Ergonomics Design Toolbox, for optimal dimensions to fit the drivers [9]. Parallely, designers worked on identifying optimal aerodynamics shape using ANSYS software. Different shapes of body were analyzed, exploring the aerodynamic resistance, stability and turning radius. After a series of analyses the shape of the front wheels was placed outside body. Such a structure is characterized by a greater Cx - aerodynamic drag coefficient, but less surface area and a smaller mass. Such a set of wheels greatly enhanced stability and improved turning radius.

After designing the final version of the body we moved to strength calculations, using FEM in ANSYS environment, where they have been used to optimize the type of fabric, number of layers and their orientation in composite body. Parallel to the work on the body, work on the powertrain system, battery system, control and telemetry proceeded.

Powertrain is composed of highly efficient, BLDC electric motor power of 530 W and a belt transmitting torque to the rear wheel of the vehicle. As a source of energy a set of lithium batteries is used.

Motor controller was designed in which all components were made to raise the efficiency and optimal control. The car was equipped with a data acquisition system that can work in offline state when all the data is collected on a micro SD card or online state and then using telemetry system based on GSM technology, data is transmitted to the central computer. Acquisition and telemetry system collect data such as motor speed, car speed, current, battery voltage, power and temperature. The unit was designed to allow modular extension with additional sensors, e.g. measuring wind speed, etc. On the display located on the steering wheel a driver sees the last lap time, race time, speed, time to the end of race. The software on the central computer can analyze data in real time and allows to calculate time for the next rounds, prediction of energy consumption, etc. The team developed a special application available on BlackBerry smartphones and tablets which show car parameter in real time.

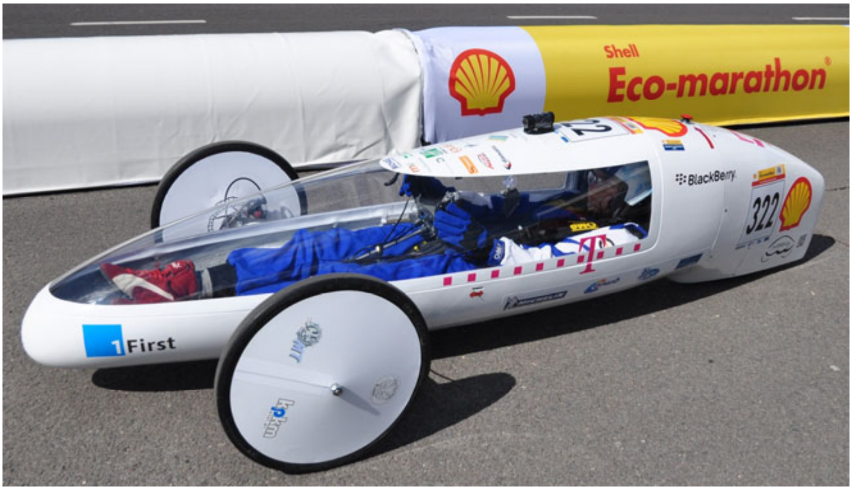


Fig. 1. Mushellka during Shell Eco-marathon Europe 2012

3 Race Car Development Plan

At various stages of race car development, which is aimed at achieving the best sports' result, the team has different level of experience. Experts usually form the basic source of knowledge however, in case of the lack of suitable experience knowledge is derived from literature or research is carried. Design development usually is made by construction of consecutive subassemblies and verification of tests results as well as participation in competitions. The gained experience allows identification of problems and evolution of new solutions. Based on these experiences new versions appear. It is not unusual that mature version which guarantees good results appears after a few years of experience and some iterations. To decrease the number of these iterations is the primary aim of the designing team and it significantly improves the economic result of the whole venture. It is especially important at the early stage of starts and creation of first versions where the lack of experience makes it difficult to even determine basic constructional features. The methodology [9] which aids the designing team of a new sports design has been elaborated. It allows determination of the influence of particular design features and external parameters on the sport result. The core of the methodology is a simulation model which takes into account the external factors. Due to different level of knowledge of the team, some main stages of development have been foreseen as well application of the simulation model:

- construction of module theoretical model including identification of phenomena which occur during the race and description of them, based on theoretical dependencies with the use of analytical forms of the model,
- the use of simulation model for determination of the concept of the construction by means of race simulation and the influence of particular design features on the result,
- gradual test stand verification of particular parameters of simulation model and verification of design features during construction of prototype of race car,
- elaboration of race strategy and its optimization,
- further verification of simulation model on the basis of research results during the race, improvement of the model with experimental forms of the model,
- simulation research aiming at identification of required changes in design which have the greatest influence on the result,
- introduction of design changes and further research focusing on verification of the simulation model,
- adaptation of simulation model for controling the object in real time and implementation of control on the object .

Currently, the team has the test results carried before the race and the results of the race vehicles. Additionally, two research stands are being built for particular subassemblies and based on the gained experience a research plan is being developed. The up-to-date development of the car as well as its further progress is made on the basis of data from the data acquisition system and their visualization.

4 Description of Data Module Acquisition System and Data Visualisation

One of the main criteria while elaborating the concept of on board systems for Mushellka race car was its modularity. It was designed to assure the greatest flexibility of the system with minimal interference to its subsystems. According to that each system element should form a separate unit, which performs precisely specified tasks. As an example engine drive can be given, whose function is to control engine performance depending on given parameters. CAN bus is the technology here which enables the fulfillment of the above mentioned criterion. Its principle of construction is based on two parallel lines, which form line CAN_H and CAN_L, where consecutive systems are connected to. Further information is transmitted on the bus and they reach each module which in turn checks if information identification is on its list of information. The advantage of the system is that the order of connections of particular modules to the CAN bus is not important.

The basic system for Mushellka car has the following systems:

- engine control,
- data acquisition,
- telemetry.

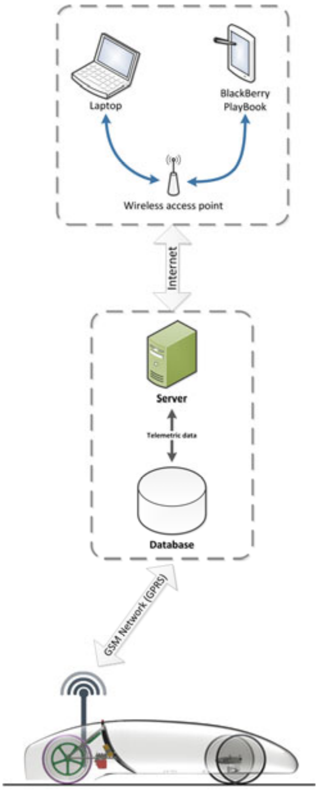


Fig. 2. Schema of the telemetry system

It is the indispensable piece of equipment which enables the car to compete with the best cars taking part in the Shell Eco-marathon. There is also a possibility to expand its design with next modules of a given function and connect them to the CAN bus [1].

Data transmitted on the CAN net in a car is read by telemetry module where it is stored on a memory card and then sent by GSM net to the server. The recording frequency on memory card is 4Hz and transmission by mobile net is 1Hz. This solution guarantees safe data collection of great time resolution. It would be more difficult to send that amount of data via GPRS, which is a drawback of the solution. Other technologies which were taken into consideration such as WiFi or Bluetooth would need many access points due to the large area of the race. The general outline of the whole system has been presented in Fig. 2 [5].

Application written in C# language is responsible for communication with the car on the server side. Its main tasks include two way communication with the car, data storing in the base and visualization of such parameters as speed, both instantaneous and mean, engine speed and instantaneous power. Fig. 3 shows the user’s interface which was designed to be possibly best readable and intuitive [6].

In order to facilitate preview of telemetric data for many people, an application for tablets has been written. It was used for previewing the most important parameters of a car but thanks to the advantages of BlackBerry PlayBook [2], [4] it is very convenient and mobile solution. Fig. 4 shows the application view for a mobile unit.

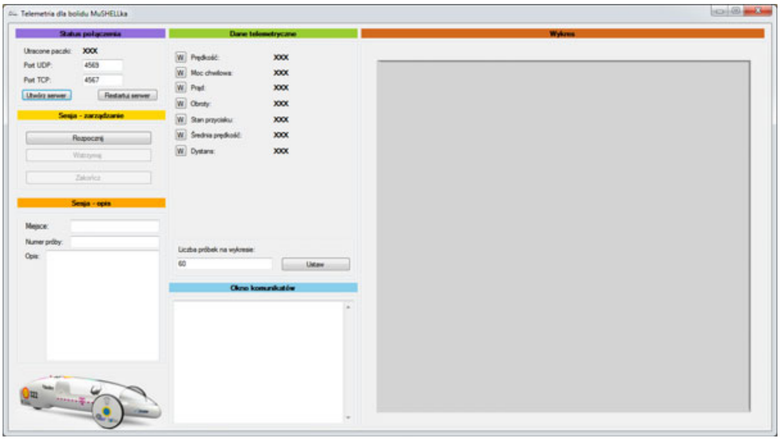


Fig. 3. Window of application for the analysis of telemetric data

5 Examples of System Application during the Race

The main task of the telemetry system is to collect and send telemetric data. Depending on the fact if the car is taking a test drive or in a race, the utility of data changes. During test drives, all data regarded as useful is data recorded on memory card, which allows very precise analysis of a car. Simulator has been elaborated especially for that reason which maps the most important parameters of a car and of the environment where it is tested. The first step was to define all physical phenomena which affect the car while driving and then to record the data in a form of equations. However, such a model is incomplete and requires real information on the performance during the race. Thanks to them it is possible to determine coefficients such as air resistance which result in 50% of energy loss. The tuned model helps to prepare the strategy of a car movement of the track during a race. This is when remote telemetry is used. It is possible to correct any deviations from the strategy due to the access to see data in real time. OLED screen mounted into the steering wheel is very useful here with a controller linked to CAN bus.

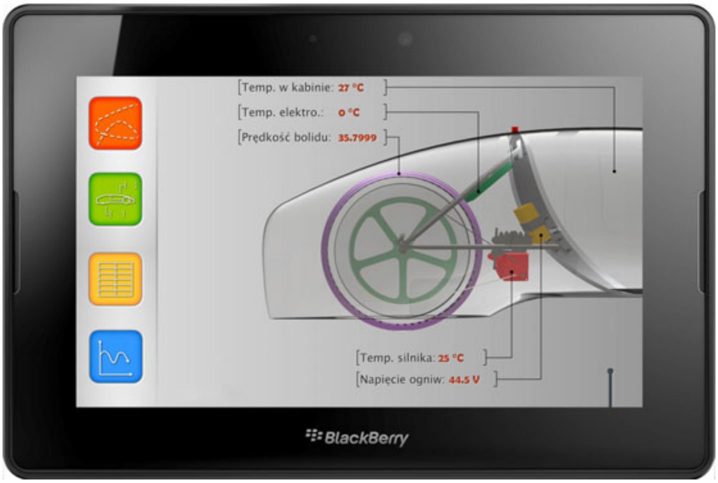


Fig. 4. Mobile application for the analysis of basic data from telemetry system

6 Plans to Develop the System

Current, first generation of board system of the car implies its further development so that its range will increase. Thanks to the use of CAN bus as a medium for data transmission of car systems it is possible to add such modules as a unit for autonomous data analysis without the need to process data. The whole process of information collection from CAN bus, its processing and showing messages for a driver is realized within the car. It facilitates the acceleration of the whole process, which with telemetry system would be delayed and that in turn can worsen its performance.

Decisions on how a driver should behave are shown on a driving wheel in a form of a message, which tells him/her if he/she follows the strategy or a correction is needed.

It is also planned to develop the safety system which is to be based on distance sensor placed in front and at the rear of a car. It should prevent collisions by informing a driver about the situation around the car.

7 Conclusion

Global Shell Eco-marathon competitions have become a meeting point for teams which build vehicles with minimal power consumption for many years. The best results in the class of electric cars this year was 732.3 km/kWh. The car of our debut team got the result 425.3 km/kWh, which is the best result among Polish teams and the best Polish record so far. Therefore, the aims for the first year have been fully realized. The accepted three-year-plan of design development includes wide range of research and optimization, which due to the lack of time were not feasible in the first year. The obtained results confirm stable and repetitive performance and the result itself places us in the lead of teams. Telemetry system aims at data transfer to the control center from data acquisition system. At present it allows current observation of working parameters of particular subassemblies and the state of car movement. According to the strategy when we have voice communication with a driver it is possible to react simultaneously and correct the vehicle movement. The elaborated applications on popular BlackBerry platform facilitate data tracking. The whole system is built based on industrial standards in a modular way and is ready to expand its functionality with additional sensor, safety and energy recovery systems. It is foreseen to have fully automated control system of a vehicle during the race.

References

1. Henry, R., Carden, F., Jedlicka, R.: Telemetry system engineering. Artech House, Inc., Boston (2002)
2. Gorton, B., Taylor, R., Yamada, J.: Adobe AIR Bible. Wiley, Delhi, India (2008) ISBN 978-0470284681
3. Koło naukowe Modelowanie Konstrukcji Maszyn, `http://mkm.polsl.pl` (accessed on July 01, 2012)
4. Miller, M.: BlackBerry PlayBook Companion. Wiley, Indianapolis (2011) ISBN 978-1118026489
5. Nawrocki, W.: Rozproszone systemy pomiarowe. Wydawnictwa Komunikacji i Łacznosci, Warszawa (2006)
6. Shneiderman, B., Plaisant, C., Cohen, M., Jacobs, S.: Designing the User Interface: Strategies for Effective Human-Computer Interaction, 5th edn. Addison Wesley, Electronic Edition (2010)
7. Silesian Greenpower, `http://www.sg.polsl.pl` (accessed on July 01, 2012)
8. Skarka, W.: Metodyka określania wpływu cech aerodynamicznych kadłuba kompozytowego na wyniki sportowe. In: IX Ogólnopolska konferencja Problemy Naukowo-Techniczne w Wyczynowym Sporcie Żeglarskim. STS Pogoria, pp. 107–108, Warszawa (2012)
9. Targosz, M., Skarka, W.: Innovative design of battery powered race car. In: Conference Proceedings of 16 International Dresden Lightweight Engineering Symposium, Dresden (2012)

Analysis of Chosen Aspects of a Two-Car Crash Simulation

Jarosław Zalewski[1] and Jerzy Kisilowski[2]

[1] Warsaw University of Technology, Faculty of Administration and Social Sciences
Noakowskiego 18/20, 00-668 Warsaw, Poland
j.zalewski@ans.pw.edu.pl
[2] Technical University of Radom, Faculty of Transport and Electrical Engineering
Malczewskiego 29, 26-600 Radom, Poland
jkisilow@kisilowscy.waw.pl

Abstract. The article presents a computer simulation of vehicle collision using the PC-Crash 8.0 software. The simulation results were compared to analytical calculations by the Routh method with the use of the coefficient of restitution for deformations or velocities. Calculations were prepared based on the crash theory. Both the stiffness based and the use mesh based model of a crash was considered. However, the use mesh based model was used, because the stiffness of car bodies in both models was not known.

The analysis concerning the chosen aspects of the consequences of car crash was made, mainly due to the change of both the mass – inertia parameters in the car body and its stiffness. The given simulation was conducted several times in order to specify the change in the after crash values.

Keywords: road accident analysis, simulation, car crash.

1 Introduction

The modelling of a car crash may be considered in different aspects [2]. This paper focuses on a front impact and a side impact crash, as the most frequent in Poland in the period of 15 years (Fig. 1). As a result of the side impact fatalities or several injuries of passengers may occur, as well as widespread damage to the car, following the change in its mass – inertia parameters.

Simulations of an oblique side impact and a frontal oblique impact was conducted using the PC-Crash 8.0. The vehicle model being hit is the Volkswagen Passat 2.0, while striking – the Opel Vectra 2.2 DTI, both from the years of production 2002 – 2005. The collision simulation was prepared at a high equal velocity on an intersection, and on a straight section of single carriageway road. This case reflects the event in which the Volkswagen was struck at the intersection while enforcing the right of way. In the event of a frontal collision the Opel left for the opposite lane.

J. Mikulski (Ed.): TST 2012, CCIS 329, pp. 34–44, 2012.

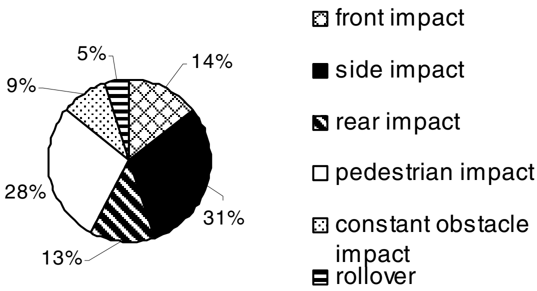


Fig. 1. Proportional participation of the types of impact in the overall number of collisions between 1995 and 2010. Source: own research.

2 Assumptions

The duration time of the simulation was 2 s, of which the time of the impact itself was 0.9 s. Basic assumptions:

- vehicle models are linear, the mass – inertia parameters were adopted according to the producers data, however some were modified for the simulation, the bodies are treated as collections of rectangular elements with constant stiffness;
- the motion takes place on a dry surface with a coefficient of adhesion equal to 0.8;
- the vehicles are moving at velocities: – 50 km/h, Vectra – 80 km/h;
- the initial Volkswagen Passat mass of 1370 kg has been increased by the weight of the driver and passengers equal to 272 kg in accordance with [1], while the initial

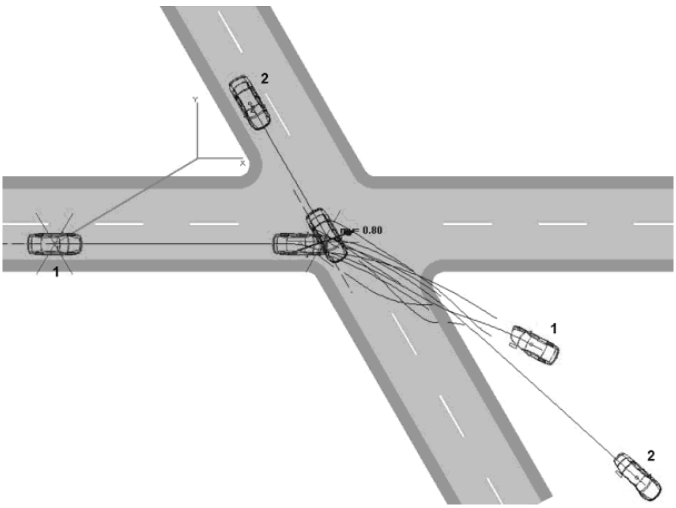


Fig. 2. Location of both cars before and after side impact collision. 1 – Opel, 2 – Volkswagen. Source: PC-Crash.

Opel Vectra mass of 1510 kg has been increased by the weight of the driver and three passengers equal to 272 kg, also in accordance with [1]. No baggage was included;

- according to [1] the height of the centre of mass for the laden vehicle was assumed at 0.56 m for both Opel and Volkswagen;

The assumption that the driver and passengers weighed 68 kg is consistent with the procedures for determining the allowed load of a car described in [1].

In Fig. 2 the location of both cars before and after the side impact collision is shown, whereas in Fig. 3 – before and after the front impact collision.

The main aim of the computer simulation was to conduct cars collision with the defined parameters as well as verification of the simulation results through analytical calculations. Another aim was to check the differences in the simulation results after running it several times. So called “use mesh based impact model” was used. It is a model of impact, where a vehicle body is divided into little polygons and the kinetic energy of impact into the deformation of each polygon.

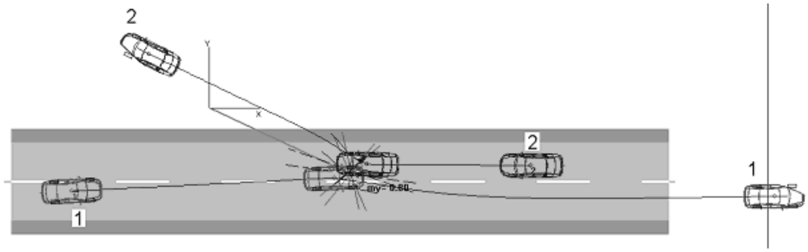


Fig. 3. Location of both cars before and after front impact collision. 1 – Opel, 2 – Volkswagen. Source: PC-Crash.

3 Description of the Simulation and the Selected Elements of PC-Crash Software

In the PC-Crash a vehicle body is treated as a single rigid block [3]. The vehicle body is represented by a rectangular prism of a certain mass, moments of inertia, structural* and torsional rigidity.

The aspect of structural rigidity is taken into account in two ways. In the mesh based impact model the stiffness of the elements, into which the body is divided, is the same for the whole body. The kinetic energy of the impact is transformed into the deformation energy, and the deformed elements are stiffer than non-deformed ones. In the stiffness based impact model the stiffness is described by a linear function. For the calculation of the impact elasticity a coefficient of restitution may be assumed based on the Newton hypothesis, i.e. the ratio of the impulse from the phase of deformations disappearance to the impulse from the phase of increasing deformations.

* Structural stiffness according to [3] is defined on the basis of the depth of deformation associated with the weight of car at rest.

In the PC-Crash the parameter defined as the rigidity of the body is described as a depth of deformations in relation to the weight of the vehicle model. The stiffness of wheel models is half the value specified for the body. In models of passenger cars a quarter of the stiffness of the lower vehicle body is assumed for the roof plate.

A change in the torsional rigidity of the body in PC-Crash 8.0 in terms of closing the gap in the normal thrust of wheels on the road surface was not taken into account, as the vehicle bodies are treated as rigid solids. In addition, the values of the vehicle wheel loads before and after the collision are presented in the crash protocol (Table 1 – 3). It is also assumed that during the simulation the vehicle models move along a flat surface with a specified coefficient of adhesion. It was assumed that in this case the body torsional stiffness is constant.

Table 1. Values of normal reactions of the road on wheels before collision

	Opel Vectra	Volkswagen Passat
Load of wheel 1 [N]	4337	3993
Load of wheel 2 [N]	4337	3993
Load of wheel 3 [N]	4403	4060
Load of wheel 4 [N]	4403	4060

Table 2. Values of normal reactions of the road on wheels after the side impact

	Opel Vectra	Volkswagen Passat
Load of wheel 1 [N]	4483	3991
Load of wheel 2 [N]	4195	3997
Load of wheel 3 [N]	4508	4057
Load of wheel 4 [N]	4220	4063

Table 3. Values of normal reactions of the road on wheels after the front impact

	Opel Vectra	Volkswagen Passat
Load of wheel 1 [N]	4329	3988
Load of wheel 2 [N]	4341	3995
Load of wheel 3 [N]	4396	4056
Load of wheel 4 [N]	4408	4063

The simulations of both collisions were repeated five times to test the repeatability. Table 4 shows the results for each repetition.

The first four simulations were carried out in succession without restoring the initial state, i.e. no return of vehicles to their original positions (to press "stop"). In the event of a side impact in repetitions 2, 3, and 4 there were secondary collisions that took place in the further movement of vehicles after the first collision. Nevertheless, the results of successive iterations differ from the first simulation by about 2 - 5%. All the velocities before and after the collision are different from the assumed (80 and 50 km/h) due to the collision detection option, through which the realism was preserved, i.e. braking just before the collision.

Table 4. The results of five repetitions of the same simulation

	side impact		front impact	
car	Opel Vectra	Volkswagen Passat	Opel Vectra	Volkswagen Passat
simulation 1				
before-crash velocity	76.49 km/h	48.22 km/h	75.23 km/h	45.13 km/h
after-crash velocity	52.93 km/h	63.07 km/h	64.39 km/h	35.59 km/h
simulation 2				
before-crash velocity	78.30 km/h	47.74 km/h	75.28 km/h	45.11 km/h
after-crash velocity	52.87 km/h	63.26 km/h	63.02 km/h	34.37 km/h
simulation 3				
before-crash velocity	79.50 km/h	48.83 km/h	76.96 km/h	46.90 km/h
after-crash velocity	52.74 km/h	63.44 km/h	64.47 km/h	35.80 km/h
simulation 4				
before-crash velocity	80.14 km/h	49.36 km/h	76.95 km/h	46.84 km/h
after-crash velocity	52.72 km/h	63.46 km/h	63.50 km/h	34.89 km/h
simulation 5				
before-crash velocity	76.49 km/h	48.22 km/h	75.23 km/h	45.13 km/h
after-crash velocity	52.93 km/h	63.07 km/h	64.39 km/h	35.59 km/h

The fifth simulation was carried out after resetting the previous, or back to the initial positions of vehicles. The results are equal to the ones of the first simulation. Obtaining consistent results thus involves the need to simulate with starting at the initial position of vehicles.

For further analysis the results of simulation no. 1 were used. Table 5 shows selected parameters of the protocol for the initial and final phase of the collision with the segmentation for the side and front impact. The depth of the body deformation is noticeable, which in the side impact velocity of about 80km/h is 0.34 m for the Volkswagen, bearing in mind the so-called mutual penetration of the body. The front of the striking car (Opel) was strongly deformed and the deformation depth is 0.40 m This shows that for such type of collision and for cars of similar weight and size the depth of deformation is larger for the front of the impacting vehicle (rigid passenger). The confirmation of this can be found in the section on the front impact collision. For a vehicle moving at higher velocities (Vectra) the depth of deformation is greater. It follows that the impact velocity affects the resulting strain, which is consistent with the basic equations of the theory of collisions [6, 7].

The values of the angular velocities around the vertical axes in the initial phase of collision were different from zero. Following the adoption of the centre of mass of vehicles greater than zero, phenomena of roll occur, and the moments of inertia for all

Table 5. The protocol of side and front impact crash. The moments of inertia after the load with passengers are shown in brackets.

car	**side impact** Opel Vectra	VW Passat	**front impact** Opel Vectra	VW Passat
VALUES AT THE BEGINNING OF THE CRASH				
before-crash velocity [km/h]	76.49	48.22	75.23	45.13
vehicle angle [deg.]	-1.27	-60.40	2.79	178.61
velocity direction [deg.]	-2.76	278.76	0.61	174.88
angular velocity around z axis [1/s]	-1.11	-0.24	-0.25	-0.90
velocity along z axis [km/h]	0.02	-0.02	-0.77	0.83
angle of lateral tilt [deg.]	-0.02	-0.01	-0.61	0.38
roll angle [deg.]	0.01	0.01	0.54	-0.75
angular velocity around x axis [1/s]	-0.01	-0.05	-0.17	0.18
angular velocity around y axis [1/s]	0.01	0.01	0.20	-0.32
moment of inertia around x axis [kgm^2]	714 (822.98)		661.9 (771.54)	
moment of inertia around y axis [kgm^2]	2379.9 (2743.16)		2206.2 (2571.66)	
moment of inertia around z axis [kgm^2]	2379.9 (2743.16)		2206.2 (2571.66)	
impulse of the impact force [Ns]	13096.46		7396.46	
VALUES AT THE END OF THE CRASH				
after-crash velocity [km/h]	52.93	63.07	64.39	35.59
the change of velocity dv [km/h]	26.46	28.71	14.94	16.22
vehicle angle [deg.]	-1.27	-60.40	2.79	178.61
velocity direction [deg.]	-13.97	332.15	-7.87	156.04
angular velocity around z axis [1/s]	-3.04	1.31	-1.17	-2.16
velocity along z axis [km/h]	0.11	-0.12	-0.47	0.51
angle of lateral tilt [deg.]	-0.02	-0.01	-0.61	0.38
roll angle [deg.]	0.01	0.01	0.54	-0.75
angular velocity around x axis [1/s]	-0.60	1.82	-0.66	-0.61
angular velocity around y axis [1/s]	0.47	-0.09	0.36	-0.00
depth of deformations [m]	0.40	0.34	0.63	0.55
coefficient of restitution k	0.10		0.10	

axes are included [3]. Doubts about certain simulation results are raised by the velocity change in the protocol (dv) that do not correspond to the difference in before- and after-crash velocities for both vehicles. Unfortunately it was not possible to determine how these values were calculated.

4 Verification on the Basis of Analytical Calculations

Calculations were prepared for comparison with the simulation results according to the method for the collision issues including the tangential velocity restitution. The data for the calculation was assumed according to Table 5, however the inputs of tangential and normal velocities were obtained by the transition from the Cartesian coordinate system (Fig. 4) to the natural local coordinate system (tangential and normal, Fig. 5, 6). The velocity vector of the striking vehicle (Opel) had to be projected orthogonally on the axes adopted in accordance with Fig. 5. It was assumed that the velocity vector of the impacted vehicle is parallel to the tangent axis (*t*), and the striking vehicle vector coincides with the normal to the collision (*n*).

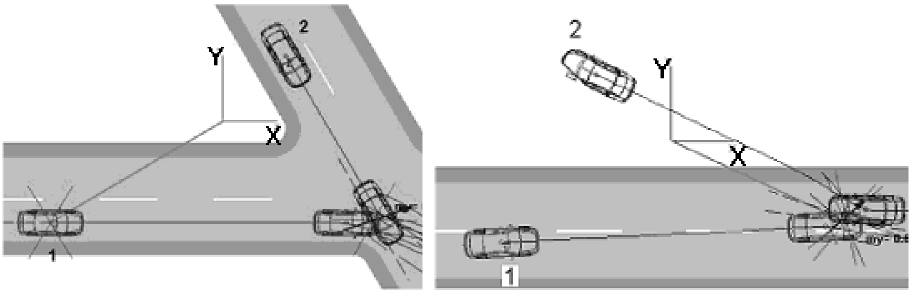


Fig. 4. Coordinate system x-y for side and front impacts. Source: PC-Crash.

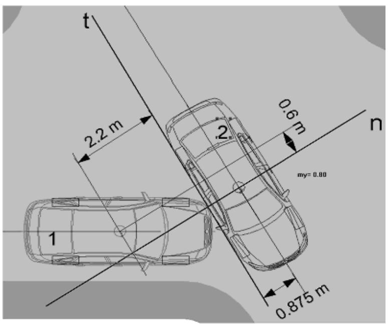


Fig. 5. Location of the local coordinate system for the side impact. Source: PC-Crash.

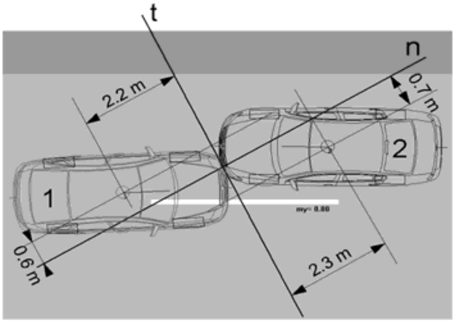


Fig. 6. Location of the local coordinate system for the front impact. Source: PC-Crash.

In the side impact, the angle of the vehicle taken from Table 2 was assumed as the impact angle, since it is measured relative to the x axis in the Cartesian coordinate system x-y. It is also easy to situate the tangential axis (along the edge) of the struck vehicle and the normal axis (perpendicular to the tangent). It is more difficult to realize this for a frontal collision, where the vehicles are positioned almost parallel to the x axis. Using the trigonometry, knowing the distance of the vehicles from the centre of collision (Fig. 6) the angles to the axes x and y can be calculated.

Based on paper [4] concerning the collision theory, a theoretical analysis of the collision issue was performed. The issue of restitution of tangential velocity was taken into account. It was assumed that during the collision, except for the so-called volumetric strains, also non-dilatational strains occur, associated with the stress in the tangential direction occurring on the surfaces of vehicles. By analogy with Newton's hypothesis a formula to determine the coefficient of restitution of tangential velocities [2] is presented.

$$\theta = \frac{w_t^{'}}{w_t} \tag{4.1}$$

The relative tangential velocity in the nonslip collision was described by formula (4.2).

$$w_t^{'} \equiv w_t - \alpha_{tt} S_t - \alpha_{nt} S_n = \theta w_t \tag{4.2}$$

The relative normal velocity

$$w_n - \alpha_{nt} S_t - \alpha_{nn} S_n = -R w_n \tag{4.3}$$

where [4]:

$$\alpha_{tt} = \frac{1}{m_1} + \frac{1}{m_2} + \frac{n_1^2}{I_1} + \frac{n_2^2}{I_2}, \qquad \alpha_{nn} = \frac{1}{m_1} + \frac{1}{m_2} + \frac{t_1^2}{I_1} + \frac{t_2^2}{I_2},$$
$$\alpha_{nt} = \alpha_{tn} \equiv \frac{n_1 t_1}{I_1} - \frac{n_2 t_2}{I_2} \tag{4.4}$$

Equations (4.2) and (4.3) were solved with respect to impulses, which in turn allowed determining the value of the tangential and normal impulses (4.5).

$$S_t = \frac{-(1+R)\alpha_{nt} w_n + (1-\theta)\alpha_{nn} w_t}{\alpha_{nn}\alpha_{tt} - \alpha_{nt}^2}$$
$$S_n = \frac{(1+R)\alpha_{tt} w_n - (1-\theta)\alpha_{nt} w_t}{\alpha_{nn}\alpha_{tt} - \alpha_{nt}^2} \tag{4.5}$$

The solution of this problem for a car crash on rough surfaces requires the knowledge of three factors: the dynamic coefficient of friction *f* (in the Routh system for the transient impulse $S_t = fS_n$), the coefficient of restitution for normal *R* and tangential θ velocities. The kinematic state after the collision is described by formulas (4.6).

$$v_{1t}' = v_{1t} - \frac{S_t}{m_1}, \qquad v_{2t}' = v_{2t} + \frac{S_t}{m_2},$$
$$v_{1n}' = v_{1n} + \frac{S_n}{m_1}, \qquad v_{2n}' = v_{2n} - \frac{S_n}{m_2},$$
$$\omega_1' = \omega_1 + \frac{S_t n_1}{m_1 i_1^2} + \frac{S_n t_1}{m_1 i_1^2}, \qquad \omega_2' = \omega_2 + \frac{S_t n_2}{m_2 i_2^2} + \frac{S_n t_2}{m_2 i_2^2}. \tag{4.6}$$

The results of analytical calculations are shown below. Distances n_1, n_2, t_1, t_2 of both centre of mass to the centre of collision were measured in the PC-Crash with both cars located at the point of a contact position. Using formulas (3.6) a system of six equations with six unknowns was obtained. Those unknowns were sought for the final moment of impact (4.6). In the shear and compression instantaneous velocity formulas [4], the instantaneous values were replaced by the values from the beginning of the collision. Then the coefficients *f*, *R*, and θ were so chosen, that the values of after-crash velocities were as close as possible to those obtained in the simulation.

The results of calculations for the side impact:

a) before crash:

$$v_{1n} = 10.62\, m/s, v_{1t} = 18.4\, m/s, v_{2n} = 0, v_{2t} = 13.39\, m/s, \omega_1 = -1.11\frac{1}{s}, \omega_2 = -0.24\frac{1}{s};$$

b) after crash:

$$v_{1n}^{'} = 5.09 m/s, v_{1t}^{'} = 15.18\, m/s, v_{2n}^{'} = 5.99\, m/s, v_{2t}^{'} = 16.88\, m/s, \omega_1^{'} = 1.33\frac{1}{s}, \omega_2^{'} = 1.71\frac{1}{s};$$

the whole impulse S=11392 Ns.

With coefficients f=0.58, R=0.01, θ=-0.8, the values of velocities in the final phase of collision were: $v_1^{'} = 57.66\, km/h, v_2^{'} = 64.5\, km/h.$

The results of calculations for the front impact:

a) before crash:

$$v_{1n} = 16.68 m/s, v_{1t} = 12.57 m/s, v_{2n} = 10.51 m/s, v_{2t} = 6.82 m/s, \omega_1 = -0.25\frac{1}{s}, \omega_2 = -0.9\frac{1}{s};$$

b) after crash:

$$v_{1n}^{'} = 14.31 m/s, v_{1t}^{'} = 10.95 m/s, v_{2n}^{'} = 13.08 m/s, v_{2t}^{'} = 8.59 m/s, \omega_1^{'} = 1.14\frac{1}{s}, \omega_2^{'} = 0.54\frac{1}{s};$$

the whole impulse S = 5121 Ns.

With coefficients f=0.71, R=0.01, θ=-0.8 the values of velocities in the final phase of collision were: $v_1^{'} = 64.89\, km/h, v_2^{'} = 56.35\, km/h.$

As it can be seen from the calculations, the coefficient of restitution of the normal velocities is greater than the default assumed in the PC-Crash (R=0.1). Moreover, the values of the angular velocities of both vehicles are much larger than those given in the protocol of computer simulation. Progressive values of impact velocities in the final phase differ slightly from the results of simulation in the PC-Crash. While the total value of the impulse is smaller by about 1300Ns for the side impact and by 2300 Ns for the front impact than the value obtained in the simulation.

5 Conclusion

The obtained results can be used for further research. The examination of the influence of disturbance in the centre of mass and moments of inertia for different cars and stability examination for so disturbed vehicle may be one aspect of the impact modelling [5].

Both the coefficient of restitution of normal and tangential velocity plays an important role in modelling the collision of vehicles in real conditions. The values of both coefficients depend on the type of collision and the angle of the resultant force of impact. The value of the impact force impulse was considered particularly in the case of the side impact, because in this case it is difficult to determine the point of force application. Besides, it depends on the impact velocity, the masses of colliding vehicles and moments of inertia of cars involved in the collision.

Such qualitative evaluation leads to arbitrary assumptions adopted in the quantitative sense, but with reference to reality.

References

1. Reimpell, J.: Podwozia samochodów. WKŁ, Warszawa (2004) (in Polish)
2. Wicher, J., Stawicki, R.: Modelowanie zderzenia samochodów. Zeszyty Naukowe Instytutu Pojazdów, PW, SIMR 1(31) (1999) (in Polish)
3. PC-Crash a simulation program for vehicle accidents, Operating Manual, Version 8.0, Dr Steffan Datentechnik, Linz, Austria (2006)
4. Gryboś, R.: Teoria uderzenia w dyskretnych układach mechanicznych. PWN, Warszawa (1969) (in Polish)
5. Zalewski, J.: Modelowanie wpływu zaburzeń geometrii nadwozia na stateczność ruchu pojazdu samochodowego. Doctoral Thesis, Warszawa (2011) (in Polish)
6. Huang, M.: Vehicle crash mechanics. Hardbound (2002)
7. Osiński, Z.: Mechanika ogólna. PWN, Warszawa (2000) (in Polish)

Theoretical Model and Activities to Change the Modal Split of Traffic

Grzegorz Sierpiński

Silesian University of Technology, Faculty of Transport
Krasiński Str. 8, 40-019 Katowice, Poland
grzegorz.sierpinski@polsl.pl

Abstract. The article deals the most popular modal split model of traffic. Actualization of this model was also proposed. Especially with information from White Papers. Some of these documents have clearly indicated the development directions of transportation after 2010.

Contemporary transport requires complex solutions. There are more possibilities to urban travels like carpooling and urban bicycles etc. Full integration aimed at achieving the goal of implementing the sustainable transportation requires undertaking the physical activities from the traffic engineering domain, implementing a right transportation policy and informing the society about the changes undertaken. In this meaning activities to change modal split of traffic was shown.

Keywords: modal split, urban travels, alternative modes of transportation sustainable development.

1 Introduction

Despite a number of actions undertaken in years aimed at reaching the state of sustainable development of the transportation, in many cases the results are not as significant as it had been assumed. The volume of the road transportation runs in private cars is still growing. The number of vehicles on the road grows, moreover, in the countries like Poland the share of vehicles older than 10 years participating in the road traffic grows too. Despite numerous attempts, the task of efficient shaping the modal split in such a way that the conditions of road traffic are improved has failed (in general). Also other detrimental phenomena, including the level of harmful emissions, have not been reduced. It need to remind the general definition of sustainable development was focus in the Chapter 2: Towards Sustainable Development of ‘Brundtland Report’: ‘Sustainable development is development that meets the needs of the present without compromising the ability of future generations to meet their own needs. It contains within it two key concepts:

- the concept of 'needs', in particular the essential needs of the world's poor, to which overriding priority should be given; and
- the idea of limitations imposed by the state of technology and social organization on the environment's ability to meet present and future needs [14].

J. Mikulski (Ed.): TST 2012, CCIS 329, pp. 45–51, 2012.

These two key concepts are important also for transportation development because transportation has a lot of influence on many aspects of our lives. It is important to have knowledge about real needs of travelling persons[1]. It will help to construct better transportation system because if a constructed system of sustainable urban transportation is adapted to the transportation-related needs of the inhabitants, then the system will be truly implementable and effective operationally. Changes are necessary. Innovations like carpooling and urban bicycles etc. can change image of alternative means of transportation. In that meaning the article deals modal split model of traffic and suggestions to change it. Activities to change modal split of traffic were also described.

2 The Modal Split Model of the Traffic

The choice of transportation mode is probably one of the most important classic models in transportation planning. [13, 21] One of the most often quoted methods of modelling of the modal split is a logit model. For individual travel motivations the split into individual modes of transportation is determined according to the following procedure:

1. A share of the non-pedestrian trips is defined using a logit example formula [8] (1):

$$U_{K\,ij} = a_I + a_{II} \cdot \frac{\exp\left(a_{III} \cdot \left(L_{ij} - a_{IV}\right)\right)}{a_V + \exp\left(a_{III} \cdot \left(L_{ij} - a_{IV}\right)\right)} \tag{1}$$

where:

U_{Kij} – the share of non-pedestrian trips in the traffic from transportation zone i to zone j;

L_{ij} – the distance of the gravity points of the transportation zones ij;

a_I, a_{II}, a_{III}, a_{IV}, a_V – the parameters of the model.

2. Next, a split of individual and public transportation trips (among all the trips considered) is defined using example formula (2):

$$U_{KI} = \frac{W_m}{1000 \cdot W_z} \cdot f_1 \cdot f_2 \cdot f_3 \tag{2}$$

where:

U_{KI} – the share (in the total of the non-pedestrian trips) of the trips made using individual transportation (in general – by a passenger car)

W_m – a motoring index;

W_z – an employment index;

f_1 – passenger car choice coefficient;

[1] This will be possible for example by using questionnaire polls or innovative proposition - the computer science tools using the GSM and GPS technologies [15].

f_2 – passenger car occupancy coefficient;

f_3 – the coefficient of the share of professionally active persons in the passenger car traffic.

The share of trips made by passenger cars in the total number of non-pedestrian trips must include several elements.

One of the main factors influencing the share of individual transportation is the passenger car choice coefficient. This coefficient should be determined individually for every trip motivation investigated.

Another element which is important from the point of view of the traffic split into individual and public transportation and allowing expressing the traffic volume in cars (and not trips) is the passenger car occupancy coefficient. In the subject matter literature the value of this coefficient is specified as 1.2-1.25.

A motoring index is a value in case of which the discrepancies between the statistical and empirical data reveal, amounting sometimes up to 30%. At present the value of this coefficient is growing but in the coming years its growth trends may change due to the fact that the elements of sustainable development of the transportation are being introduced in the cities.

The communication related behaviours and decisions of the inhabitants are influenced by the socio-economic transformations including the rapidly growing ones such as the mobility of the population as well as the motoring index. For this reason final formula of modal split model will be different on different places of world. Examples for Lyon Conurbation, Upper Silesian Conurbation or for China, Indonesia and Istanbul was described in [1, 2, 4, 8, 17, 19, 25].

In the description above only three typed of displacements were identified: a pedestrian trip, a trip using a passenger car (identified with the individual transportation) and a trip with the public transportation. It needs to be noted that as the sustainable development related activities are undertaken and the means of travelling alternative to a passenger car are being promoted, it seems reasonable that the first category is subdivided in more detail. The detailed model should therefore allow for the following options:

- pedestrian trip (travel on foot)
- non-pedestrian trip (travel other than on foot):
  - by individual transportation
    - by private transportation
      - on a bicycle
      - on a motorcycle/motorbike
      - by a passenger car
    - urban transportation
      - urban bike
      - urban passenger car
  - urban public transportation
    - by bus
    - by tram
    - by trolleybus
    - by underground
    - by fast urban railway.

Obviously, the availability of individual ways of travelling is dependent on the availability of specific means of transportation or of a dedicated infrastructure permitting a specific kind of travel. A subject of discussion may be the categorisation of the carpooling type of systems (a use of one vehicle by more than one person travelling together in one direction). The proposition above need more study covers real conurbation area., but it illustrates the scale of the problem.

3 Activities Influencing the Modal Split of the Traffic

Taking into account the guidelines provided, among others, in the White Papers [22, 23], it becomes necessary – in particular on the urbanised areas – to identify not only an actual modal split of the traffic but also of the real transportation related needs of the travelling persons. It was as early as in the Bruntland Report that the concept of needs had been pointed out as one of the conditions of the right approach to sustainable development. Another step towards a more effective use of the urban space is taking up activities influencing the behaviours and choices of the travellers, in particular in relation to the choice of a mode of transportation. It needs to be remembered, however, that in each case when some type of restrictions are introduced for the passenger cars, there is a necessity to provide an alternative means of transportation with the right availability. The examples of the activities targeted at achieving the changes in the model split are shown in Table 1.

Table 1. The examples of activities influencing the modal split of the traffic towards the sustainable development of the transportation

Means of transportation	Examples of solutions
Solutions supporting alternative forms of transportation	
Public urban transportation	Dedicated bus lanes
	Right forms of platforms and bays
	Right priorities at the crossings
	Other solutions from the domain of transportation telematics
	Optimisation of routes
	Clear fare policy and shared ticketing systems
A bicycle as means of urban transportation	Logical and safe bicycle routes
	Safe bicycle parks
	A concept of urban (public) bicycles
Car-pooling	Dedicated lanes
	Discounted parking charges
	Discounted zone entry fees/congestion fees
Limitations for passenger cars	
Passenger (private) car	Paid parking spaces
	Means of traffic calming
	Congestion fees and charges
	Traffic free zones

Source: own research and [6, 7, 11, 12, 24]

As far as the private cars are concerned, the paid parking zones are one of the softest forms of restriction. A next step would be a charge for entering a specific urban zone. The latter solution has been applied in a number of cities in the world, such as Singapore (1975), Rome (2001), London (2003), Stockholm (2006) and Milan (2008) [18].

The subject matter literature quotes a range of subjective opinions and statements of the passengers on different ways of travelling (cf. Table 2 and Table 3). These opinions are important from the point of view of the improvement of the perception of specific means of transportation.

Table 2. The advantages of travel by different means of transportation

On foot	By bike	By passenger car	By public urban transportation
• independent decision about the route • other activities performed during travel • communication/interactions with other people • healthy option (physical exercise) • environmentally friendly • cheap – no vehicle needed • opportunity to observe the environment	• independent decision about the route • communication/interactions with other people travelling in the same direction • healthy option (physical exercise) • environmentally friendly • cheap, compared with the price of a car • opportunity to observe the environment • Faster than travelling on foot	• independent decision about the route • practical in urgent cases • load carrying possibility (shopping, luggage, etc.) • independent from the weather • short travel time • convenience • possibility of giving a lift to other people • personal safety	• communication/interactions with other people travelling in the same direction • other activities performed during travel • passive – 'being carried' to the destination • cheap – no need to buy a car • opportunity to observe the environment

Source: own research and [9], [10], [26]

Table 3. The disadvantages of travel by different means of transportation

On foot	By bike	By passenger car	By public urban transportation
• dependent on the weather • inconvenient on longer distances • physical exercise causing tiredness • prone to accidents • low personal safety • long travel times	• dependent on the weather • inconvenient on longer distances • physical exercise causing tiredness • prone to road accidents • low personal safety • long travel times	• at high traffic – not environmentally friendly • the need to focus the attention on the road • in case of high congestion long waiting times make travel longer • if travelling only by car – insufficient physical exercise	• a need to adapt to the general timetable • high loading – many passengers • no possibility about deciding about the route • low personal safety • in case of high congestion long waiting times make travel longer • a need to wait at the stops • frequently no seating places • frequently incomprehensive or unclear information at the stops

Source: own research and [9], [10], [26] The technological solutions available today allow installation of new alternative systems, such as Autolib' set up in Paris in December 2011 [1, 3]. In that system travellers may use (at a charge) urban electric passenger cars (the system operates similarly as Velib' – a rental system of urban bikes [20]). The electric vehicles have been prepared according to the EU initiatives (such as the White Paper [22, 23] or Green eMotion programme [5]).

Companies should also support the activities aimed at the improvement of pro environmental transportation. [16]

4 Conclusion

The reduction of congestion in the cities is possible by changing the modal split of traffic. New modal split models must reflect innovative mean of transportation – alternatives to private cars – like carpooling systems, car-sharing systems, clean buses, trams and urban bicycles etc. High dynamic of changes enforced by the regulations but also stemming from the current situation makes the future shape of transportation systems in the cities hard to predict. New technologies being introduced as test schemes will then be modified as to their functionality; a fixed and permanent factor, however, remains to be the need of periodic investigation of the needs of persons travelling over a certain area.

References

1. Autolib' Official Site, `http://www.autolib.eu/`
2. Bonnel, P.: The Estimation of Aggregate Modal Split Models. In: European Transport Conference Strasbourg, France, October 8-10 (2003)
3. Brzeziński, A., Masłowski, K., Piasecka, A.: Samochody wykorzystywane inaczej. In: Konferencja Naukowo Techniczna "Miasto i Transport 2012. Innowacyjność Transportu – Oszczędzanie Energii", Politechnika Warszawska, 18 kwietnia (2012)
4. Celikoglua, H.B., Akada, M.: Aggregated Binary Logit Modal-Split Model Calibration: An Evaluation for Istanbul, `http://www.trafficforum.ethz.ch`
5. Green eMotion programme, `http://www.greenemotion-project.eu/about-us/index.php`
6. Janecki, R., Krawiec, S., Sierpiński, G.: Publiczny transport zbiorowy jako kluczowy element zrównoważonego systemu transportowego Górnośląsko-Zagłębiowskiej Metropolii Silesia. In: Pyka, R. (ed.) Sposób na Metropolię. Idee a społeczne oczekiwania wobec projektu utworzenia śląsko-zagłębiowskiej metropolii, UM Katowice, RSS MSNP UŚ, Katowice, pp. 105–132 (2010)
7. Janecki, R., Krawiec, S., Sierpiński, G.: The directions of development of the transportation system of the metropolitan area of Upper Silesia until 2030. In: Proceedings of International Scientific Conference Transbaltica, Vilnius, pp. 86–91 (2009)
8. Karoń, G., Janecki, R., Sobota, A., Celiński, I., Krawiec, S., Macioszek, E., Pawlicki, J., Sierpiński, G., Zientara, T., Żochowska, R.: Program inwestycyjny rozwoju trakcji szynowej na lata (2008-2011), Analiza ruchu. Praca naukowo-badawcza NB-67/RT5/2009
9. Khattak, A.J., Rodriguez, D.: Travel behavior in neo-traditional neighborhood developments: A case study in USA. Transportation Research Part A 39, 481–500 (2005)

10. Lu, X., Pas, E.I.: Socio-demographics, activity participation and travel behavior. Transportation Research Part A 33, 1–18 (1999)
11. Macioszek, E., Sierpiński, G., Czapkowski, L.: Methods of Modeling the Bicycle Traffic Flows on the Roundabouts. In: Mikulski, J. (ed.) TST 2010. CCIS, vol. 104, pp. 115–124. Springer, Heidelberg (2010)
12. Macioszek, E.: The passenger car equivalent factor for heavy vehicles on roundabouts. In: Janecki, R., Sierpiński, G. (eds.) Contemporary Transportation Systems. Selected Theoretical and Practical Problems. The Development of Transportation Systems, pp. 127–137. Silesian University of Technology, Gliwice (2010)
13. Mathew, T.V., Krishna Rao, K.V.: Introduction to Transportation Engineering. Civil Engineering – Transportation Engineering. IIT Bombay, NPTEL Online, http://www.cdeep.iitb.ac.in/nptel/Civil%20Engineering/
14. Our Common Future. Report of the World Commission on Environment and Development, Transmitted to the General Assembly as an Annex to document A/42/427 - Development and International Co-operation: Environment (1987), http://www.un-documents.net/wced-ocf.htm
15. Sierpiński, G., Celiński, I.: Use of GSM Technology as the Support to Manage the Modal Distribution in the Cities. In: Subic, A., Wellnitz, J., Leary, M., Koopmans, L.C. (eds.) Sustainable Automotive Technologies 2012, vol. 104, pp. 235–244. Springer, Heidelberg (2012)
16. Sierpiński, G.: Travel Behaviour and Alternative Modes of Transportation. In: Mikulski, J. (ed.) TST 2011. CCIS, vol. 239, pp. 86–93. Springer, Heidelberg (2011)
17. Silitonga, S.P., Sulistio, H., Djakfar, L., Wicaksono, A.: Modal Split Model for Public Transport Development in Indonesia. Journal of Applied Sciences Research 7(12), 2036–2041 (2011)
18. Suchorzewski, W.: Opłaty za wjazd do obszarów śródmiejskich – sukcesy i porażki. In: Kaczmarek, M., Krych, A. (red.) Skuteczne Zmniejszenie Zatłoczenia Miast, pp. 195–206. Poznań – Rosnówko (2009)
19. Sulistio, H., Silitonga, S.P.: Scenario of Modal Split and Public Transport Utility in Palangka Raya, Malang and Surabaya Cities of Indonesia. Journal of Mathematics & Technology (3), 101–104 (2010)
20. Velib' Official Site, http://www.velib.paris.fr/
21. Transit Estimation and Mode Split. NHI course on Travel Demand Forecasting (152054A), http://www.ctre.iastate.edu/educweb/ce451
22. White Paper: European transport policy for 2010: time to decide. COM (2001) 370 (2001)
23. White Paper: Roadmap to a Single European Transport Area – Towards a competitive and resource efficient transport system. COM (2011) 144 (2011)
24. Zamkowska, S.: Activities towards the implementation of sustainable mobility in the cities. In: Janecki, R., Sierpiński, G. (eds.) Contemporary Transportation Systems. Selected Theoretical and practical Problems. The Development of Transportation Systems, pp. 267–273. Silesian University of Technology, Gliwice (2010)
25. Zhang, S., Jiang, K., Liu, D.: Macro-level Transport Modal Split Model: Theory and Application in China. In: International Conference on Science and Engineering Management, ICMSE 2006, pp. 2033–2038 (2006)
26. Zwerts, E., Allaert, G., Janssens, D., Wets, G., Witlox, F.: How children view their travel behaviour: a case study from Flanders (Belgium). Journal of Transport Geography 18, 702–710 (2010)

Driver's Biocybernetic Monitoring and Early Warning System

Andrzej W. Mitas[1], Artur Ryguła[2], and Bartłomiej Pyciński[3]

[1] Transport Systems Institute, Warsaw, Poland
andrzej.mitas@ist.edu.pl
[2] APM Konior Piwowarczyk Konior, Bielsko Biała, Poland
artur.rygula@apm.pl
[3] Silesian University of Technology, Gliwice, Poland
bartlomiej.pycinski@polsl.pl

Abstract. The paper describes a method and a device for a current permanent observation of driver's response in the road traffic was described. The task of presented system is not only the "on-line" analysis of the human behavior but primarily the determination of the physiological changes and their fuzzy classification in order to determine the states considered sub-critical. The article is based on authors research and results of experiments, related to the international scientific state of art, in terms of excessive fatigue and aggression as essential biocybernetic factors increasing the road accident risk. A variant of the warning system for notification of driver's psychophysical condition changes was presented. The system is based on a revised analysis of the human metabolism under a monotonous long-term load. The results of simulations and real measurements are prerequisites for the development of the presented system structure.

Keywords: early warning system, driver's biocybernetic system.

1 Introduction

The level of traffic safety is conditioned by many factors ranging from the state of road infrastructure, weather, traffic, vehicle technical efficiency to the psycho physiological condition of the driver. As police statistics show, 55% of events occurred on straight parts of the road and most accidents happen during the summer, when the traffic conditions are close to optimal. These data clearly indicate how significant impact on the potential risk of accident has the psycho physiological state of the driver, including his capacity for a proper traffic risk prediction and for taking actions adequate to the assessed level of risk. According to published the information [1] fatigue was the main reason in about thirty per cent of serious accidents. There search carried out at the same time showed that the second of the factors - road aggression - had a direct impact on nearly 50% of road accidents [12].

J. Mikulski (Ed.): TST 2012, CCIS 329, pp. 52–59, 2012.

2 State of Art in Driver's Safety Systems

An increasing number of research on driver's safety system includes now the use of the electromyography, electroencephalography, electrocardiography and electrooculography [3, 4, 5]. In addition, in the process of driver's fatigue and aggression assessment, systems analysing facial expressions [6], head movements [7] and the speech signal [8] are used.

The analysis of changes in the driving dynamics is an equally important issue in the road safety systems. The subjects of studies are the detection of extreme mental and physical states such as fatigue, lethargy or aggression [9], as well as the modelling of a driver behaviour in order to adjust the technical parameters of the vehicle components [10], the user authentication [11] and the detection of subsequent road manoeuvres[12]. The basic measures of driving style include the brake and the accelerator pedal position and force [13], signals from the steering wheel [14] and its derivative in the form of tracking vehicle trajectory on the road lane [15]. In these cases, the signal frequency and amplitude changes are a reliable indicator of driver fatigue as well as of increased aggressiveness.

Because of the availability of statistical data, the vehicle speed is a particularly important indicator of the driving style as well as its derivatives as functions of time and distance. Conducted studies have demonstrated a low degree of scientific use of this parameter as an estimator of the current driver's psychophysical condition. Authors studies published in [16] are cited in the materials describing the currently on-going work in international scientific centres. One of the few studies were described in paper [17], in which the authors have compared the speed profile of four professional drivers, on the selected section of the road. The results showed that the variation in speed is a highly promising parameter for determining the individual driving characteristics. A low prevalence of this topic was also showed in [18], where authors emphasized that the results presented in paper [16] were the first results of the analysis of driver behaviour without the use of video cameras.

3 Monitoring and Early Warning System

The described system was established on the basis of doctoral dissertation [19]. The basic idea of an early warning system is real-time monitoring of changes in driver's behaviour, which could have a meaningful impact on defined driving style indicators (Fig.1).

Monitoring the variability of indicators allows the assessment of changes in the way of driving and thus of the increased risk of an accident. The Principle Component Analysis (PCA) was used to determine a set of vectors describing variances of the input data. Each time, after registering the defined road distance, elementary parameters of PCA methods (eigen values, eigenvectors and new observations' values) are computed. A set of first eigenvectors each time describes the variability of input data.

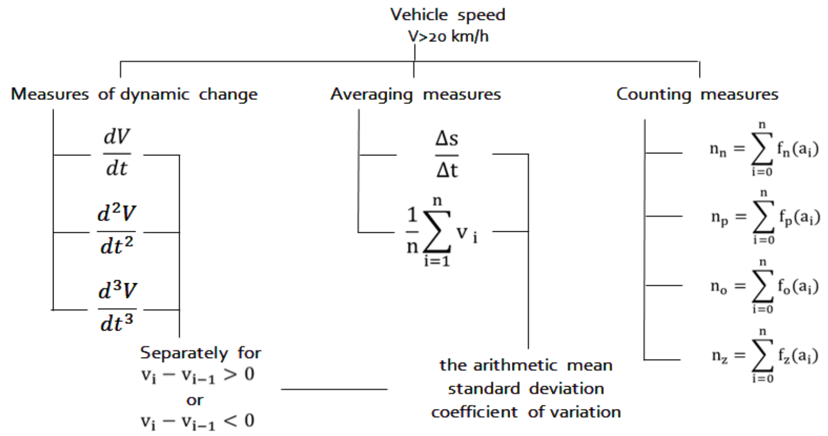


Fig. 1. Driving style indicators

The distribution of the registered value is irregular - the relative position of individual points is not constant but it depends on the information derived from subsequent brackets data. This change results from the assumptions of the PCA method, which each time determines the current vector that describes the highest variability of the input data. The data from the next driving bracket influence the direction and the value of particular eigenvectors by changing the participation of particular driving indicators. In view of these considerations, the percentage distribution of each indicator as a function of driving brackets was determined (1).

$$f(x) = 100 \cdot \left(\sqrt{\frac{(x_{k_i})^2}{\sum_{i=1}^{n} {x_{k_i}}^2}} - \sqrt{\frac{(x_{k-1_i})^2}{\sum_{i=1}^{n} {x_{k-1_i}}^2}} \right) \quad (1)$$

where: x_{k_i}– the value of the i-th feature in a particular principle component in the k-th distance bracket, n– feature number, n = 26.

Fig. 2 shows the distribution of normalization function (equation 1) for an example of driving data.

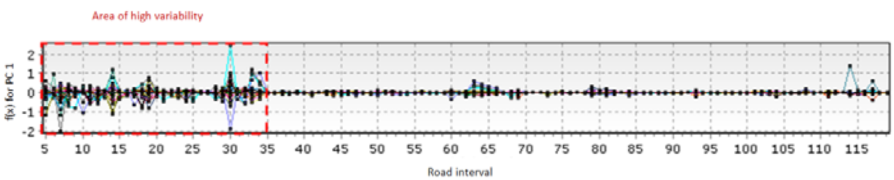


Fig. 2. Example chart of features participation in the selected (first) principal component

Characteristics of the curves indicate the area with a distinct variation of data and the area with a relative stability. The variability may basically result from: the adaptation process of user to driving conditions, the statistical normalization of the selected method or environmental conditions variation. In practice it is difficult to determine the contribution of individual elements to the normalization process, however empirically, for the data analysed in paper [19], it can be assumed that this phenomenon lasts for a section of 50 kilometres(for the speed of data recorded at a frequency of 1 Hz and accuracy of 1 km / h).

The criterion designation of the driving style changes was the next step in defining a nearly warning system. As deviation values of the current driving style characteristics occur, the selected method of assessing the position of the last recorded data points out the area of specified principal components. The area was limited to a set of the first three PC vectors, which according to paper[19], describe nearly 80% of the data variation. The rating process is done by verifying the following boundary conditions:

$$1.\ x_n > \max_{i=1,\dots,n-1} x_i \ \vee\ x_n < \min_{i=1,\dots,n-1} x_i$$

or

$$2.\ y_n > \max_{i=1,\dots,n-1} y_i \ \vee\ y_n < \min_{i=1,\dots,n-1} y_i \tag{2}$$

or

$$3.\ z_n > \max_{i=1,\dots,n-1} z_i \ \vee\ z_n < \min_{i=1,\dots,n-1} z_i$$

where: x_n, y_n, z_n– coordinates of actual registered observation, described by the first, second and third PC vector, x_i, y_i, z_i– coordinates of i-th registered observation, described by the first, second and third PC vector,n– the number of registered observations (distance brackets).

3.1 Method Algorithm

According to the assumptions of the early warning method in the vehicle-driver structure there is a feedback between the mental and physical state of the driver (biological process), and his behaviour (the object). The driver, on the basis of traffic conditions, his mental state and other external factors, makes decisions that affect the driving performance. The system, because of the earlier described method, monitors the driving style indicators and sends a warning in the case of detection of abnormal parameters. Driving style features take into account personally specific driving characteristics, defining a set of critical values, individually for each driver (Fig. 3).

The diagram shown in Fig.3 presents the early warning system, which consists of the following components:

- module of the driving style indicators - responsible for calculating the driving characteristics defined for a specific road section,
- module of verification if the last observation meets the criteria described in equation 2,

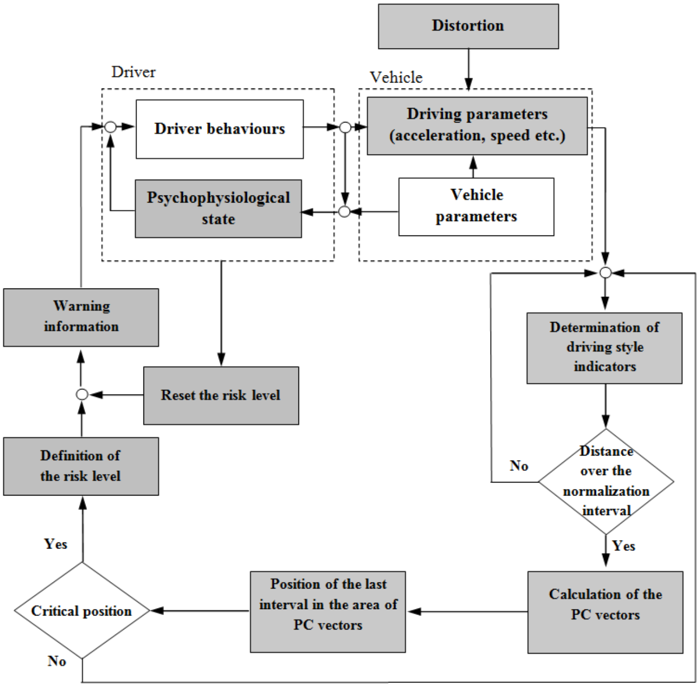


Fig. 3. Diagram of an early warning system

- module of PCA vectors calculation - each time after driving the defined road section, the segment is responsible for the PCA parameters determination in order to compute the degree of variability of the recorded driving style data,
- module of identification of the last section position in the PC space,
- module of the risk level determination –the level arising from the driving style changes, which is defined using five levels of risk (Fig. 4),
- module of warning information - responsible for sending a light and sound signal adjusted to the present level of risk,
- module of risk level reset- responsible for the process of manual reset when the driver accept the risk. The reset is done by pressing a button on the driver's system panel.

3.2 Results

The defined algorithm of the early warning system has been examined with twenty speed driving waveforms obtained from real accidents. In fourteen cases the level of high and increased risk (level 2 and 3) has been observed. In seven cases those levels occurred at a distance of 35 kilometres before the accident. For three data sets the changes have increased in the section of the last 50 kilometres. In a particular case, at the distance of 35 kilometres before the accident, the driver would receive the warning information three times.

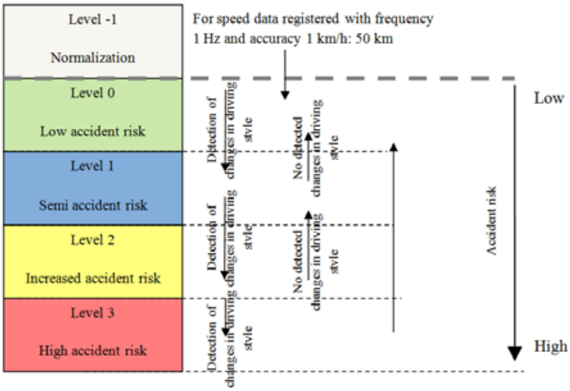


Fig. 4. Diagram of the level of risk determination

4 Biomedicine Reference System

The presented results of the early warning system show a promising performance of the proposed method. In a situation where there is no technical possibility to create databases of authoritative records of accidents, the system effectiveness can only be verified by using additional references. According to the described studies and authors knowledge the authors have decided to use a biomedical system recording the driver's psychophysiological parameters.

According to medical knowledge, the reduction of driver's mental, sensory and motor efficiency is strongly associated with the activity of the autonomic nervous system, especially with the sympathetic parts. As driver's subjectively perceived fatigue increases, the sympathetic activity is rising. A similar relationship exists if the driver is subjected to stressful factors, however the dynamics of changes in the system activity may vary. The methods for testing the activity of the sympathetic autonomic nervous system comprise both direct measurements of the concentration of epinephrine in the blood serum, as well as indirect methods such as measurements of exocrine glands activity, measurements of the heart rate and skin temperature of specific body parts or the study of skin-galvanic reaction. Indirect measurements are based on the assumption that the activity of the sympathetic nervous system and its influences on the organs through adrenergic receptors is the primary regulator of the specified parameter.

Preliminary research, described in[19,20], has shown relationships of the presented early warning method with selected biomedical features. The changes in the temperature, the ratio of HRV and eye movements were particularly important. Expanded research will be considered in authors' future studies.

5 Conclusion

The processing of information about the driving parameters recorded by the commonly available measuring devices is the main attribute of the presented early warning method. The proposed method can be implemented with the use of tachometric recorders as well as GPS receivers. The first group of devices is mandatory for professional drivers, while the second is a common component of mobile devices.

The method has been defined and verified on the basis of absolutely reliable data that have been registered for travel sended with a traffic accident. Despite the lack of a reliable description of the psychophysical condition before the accident, it was possible to identify a hypothetical state of increased risk.

The reference to pre-verified biomedical reference methods will be twofold: the verification (to demonstrate the effectiveness of the defined early warning method) and the analysis (a detailed analysis of the reasons of driving style changes as well as exploring the mechanisms and symptoms of mental and physical state changes of the driver).

References

1. Haworth, N.L., Triggs, T.J., Grey, E.M.: Driver Fatigue: Concepts, measurement and crash countermeasures, Human Factors Group, Department of Psychology Monash University (June 1988)
2. AAA Foundation for Traffic Safety, Aggressive Driving: Research Update, Washington, D.C. (April 2009)
3. Benoit, A., Bonnaud, L., Caplier, A., Ngo, P., Lawson, L., Trevisan, D.G., Levacic, V., Mancas, C., Chanel, G.: Multimodal Focus Attention Detection in an. Augmented Driver Simulator. Journal on Multimodal User Interfaces 1(1), 49–58 (2007)
4. Healey, J.A., Picard, R.W.: Detecting Stress During Real-World Driving Tasks Using Physiological Sensors. IEEE Transactions on Intelligent Transportation Systems 6(2), 156–166 (2005)
5. Qing, W., Bing Xi, S., Bin, X., Junjie, Z.: A PERCLOS-Based Driver Fatigue Recognition Application for Smart Vehicle Space. In: 2010 Third International Symposium Information Processing (ISIP), pp. 437–441 (2010), doi:10.1109/ISIP.2010.116
6. Mitas, A., Bugdol, M., Ryguła, A.: Simultaneous analysis of driver's physiological and behavioural parameters under the aspect of transport safety. Journal of Medical Informatics & Technologies 13, 241–247 (2009) ISSN-1642-6037
7. Shunji, K., Masayoshi, A.: Estimation of Rhythmic Movement of Driver's Head from Nostrils Detection. Papers of Technical Meeting on Intelligent Transport Systems. IEE Japan ITS-02(31-34), 13–18 (2002)
8. Krajewski, J., Trutschel, U., Golz, M., Sommer, D., Edwards, D.: Estimating Fatigue from Predetermined Speech Samples Transmitted by Operator. In: 5th International Driving Symposium on Human Factors in Driver Assessment, Training, and Vehicle Design. Driving Assessment, pp. 468–474 (2009)
9. Boyraz, P., Hansen, J.H.L.: Active Accident Avoidance Case Study: Integrating Drowsiness Monitoring System with Lateral Control and Speed Regulation in Passenger Vehicles. In: 2008 IEEE International Conference on Vehicular Electronics and Safety, Ohio, USA, September 22-24 (2008)
10. Miyajima, C., Nishiwaki, Y., Ozawa, K., Wakita, T., Itou, K., Takeda, K., Itakura, F.: Driver Modeling Based on Driving Behavior and Its Evaluation in Driver Identification. Proceedings of the IEEE 95(2), 427–437 (2007)

11. Ryguła, A.: Driving style as a behavioural biometric in the process of user identification and the current analysis of access to the vehicle. Biometry (2011); editor-in-chief: Mitas, A.W.
12. Kuge, N., Yamamura, T., Shimoyama, O., Liu, A.: A Driver Behavior Recognition Method Based on a Driver Model Framework. Structure 109(Idm), 469–476 (2000)
13. Malta, L., Miyajima, C., Takeda, K.: A Study of Driver Behavior Under Potential Threats in Vehicle Traffic. IEEE Transactions on Intelligent Transportation Systems 10(2), 201–210 (2009)
14. King, D.J., Mumford, D.K., Siegmund, G.P.: An algorithm for detecting heavy-truck driver fatigue from steering-based measures (98-S4-O-10). In: Proc. of 16th International Technical Conference on the Enhanced Safety of Vehicles, pp. 873–882. National Highway Traffic Safety Administration, Washington, DC (1998)
15. Friedrichs, F.: Drowsiness monitoring by steering and lane data based features under real driving conditions. Signal Processing, 209–213 (2010)
16. Ryguła, A.: Driving style identification method based on speed graph analysis. In: Proceedings of 4th International Conference on Image Analysis and Biometrics and International Conference on Kansei Engineering & Affective Systems, Cieszyn, June 25-28 (2009)
17. Cherrett, T., Pitfield, D.: Extracting Driving Characteristics from Heavy Goods Vehicle Tachograph Charts. Transportation Planning and Technology 24(4) (2001)
18. Islinger, T., Köhler, T., Wolff, C.: A Functional Driver Analyzing Concept. Advances in Human-Computer Interaction, Article ID 413964, 4 pages (2011), doi:10.1155/2011/413964
19. Rygula, A.: Early warning method in the aspect of tra_c safety. Doctoral Thesis. Faculty of Transport and Electrical Engineering, Technical University of Radom (2012)
20. Mitas, A.W., Ryguła, A., Pyciński, B., Bugdol, M.D., Konior, W.: Driver Biomedical Support System. In: Piętka, E., Kawa, J. (eds.) ITIB 2012. LNCS, vol. 7339, pp. 277–285. Springer, Heidelberg (2012)

Delivering Payment Services through Manual Toll Collection System

Zbigniew Kasprzyk

Warsaw University of Technology, Faculty of Transport
Koszykowa 75, 00-662 Warsaw, Poland
zka@it.pw.edu.pl

Abstract. A rapid development of the transport infrastructure leads to ever-stringent quantitative and qualitative requirements set to transport systems. In the qualitative context, they boil down to the quality of service delivered by transport facilities to users of the toll collection system. That quality is determined by meeting defined timeframes. Hence, it caters for user expectations that the toll collection service on highways meets certain standards. This paper discusses the concept of using evaluation methods and tools supporting the decision-making process for maintenance-related issues, in the context of manual toll collection systems.

Keywords: quality of customer service in toll collection system, methods and tools for maintenance decision support, toll collection system.

1 Introduction

As the Polish network of motorways and dual-carriageways expands, issues related to tolls and toll collection systems come to the fore. In line with the current standards, the toll collection process has to be efficient and to have the lowest possible impact on road users. The quality of customer service at toll collection plazas depends among other on meeting defined timeframes, especially today when the electronic toll collection (ETC) is ever-popular. A manual toll collection system could be perceived as a system distorting the transport process through extending the time to complete that process. The ETC system will initially, i.e. over the forthcoming few years, process lorries heavier than 3.5t of gross vehicle weight. Then the system is planned to be expanded to passenger cars. That perspective, however, is rather distant. Thus the toll collection system in Poland over the next few years will be a mixed system combining the manual toll collection system and the electronic toll collection [1, 2]. In practice, motorway management units will have to provide toll collection services using a manual toll collection system. In that case, the fact of making the payment via the manual toll collection system generates downtimes, which from road user's perspective are undesirable and should occur on a scarce basis. The majority of Polish toll roads are open toll system roads, where mainline barrier toll plazas are located along the motorway. Every driver pays fees (tolls) for using a given section, bridge or

J. Mikulski (Ed.): TST 2012, CCIS 329, pp. 60–68, 2012.

tunnel, based on the vehicle category. This toll collection method disrupts traffic flow as vehicles queue at mainline toll plazas causing congestion. Thereby the manual toll collection has to become ever-efficient on Polish roads.

2 Evaluation Methods and Tools for Maintenance Decision Support in the Context of Manual Toll Collection Systems

The above-mentioned factors call for using on Polish road over the forthcoming years manual toll collection systems. This in turn means for a toll road management the need to deliver toll payment services at a predetermined level. High enough quality toll services using manual toll collection systems should enable both the systems used on Polish roads to interoperate closely and reliably. The following concept of using evaluation methods and tools for the maintenance decision support, in the context of manual toll collection systems, enables attempting to aid manual toll collection systems in order to deliver a better quality customer service, satisfactory for road users. The discussed evaluation method for a manual toll collection system has been presented as a model of manual toll collection system in terms of capacity to deliver payment services of high enough quality. The tool for the maintenance decision support in manual toll collection systems has been presented as a proprietary application supporting the decision-making process in terms of evaluating customer service quality delivered to toll system users.

2.1 Model of a Manual Toll Collection System in Terms of Capacity to Deliver Payment Services of High Enough Quality

The toll road management is obliged to maintain a continuous and reliable toll collection process and to assure a high driving comfort on national roads. This consequently entails assuring high quality customer service for users of tool roads [3]. Therefore the analysis of manual toll collection system should - apart from investigating system maintenance and factoring in the reliability analysis of electromechanical components - answer the query whether the analysed system delivers the required service level defined based on categories determining service quality provided to users of the toll collection system. On the other hand, the toll road management concentrates of minimising capex outlays on building the toll collection system thus decreasing the number of toll gates, which in turn leads to a lower customer service level (possible traffic congestion at gates). Service levels imposed on the road management can be determined through the evaluation method for the model of toll collection system in terms of capacity to deliver payment services of high enough quality. Failing to meet nominal customer service levels levies on road management contractual penalties. It is therefore evident that the problem of assuring high enough service quality for motorway users is important and beyond doubt current, especially as far as the perspective of expanding road network in the following years is concerned.

When analysing the behaviour of a manual toll collection system in terms of capacity to deliver high enough quality toll services, a distinction can be made between operating modes of toll collection system:

- U1 - overload (state of reached operational capability) - critical time-per-vehicle values have been exceeded and there is no capacity to process any more vehicles in the predefined time - congestion,
- U2 - state of full operational capability (normal operation) - all vehicles go through mainline barrier toll gates within the predefined time,
- U3 - normal operational capability with a possible safety breach (state of partial operational capability) - queue has formed within assumed limits, however reaching operational capability could not be guaranteed. Some preventative measures could be deployed in that situation, which could prevent congestion from forming through e.g. opening another toll gate.

The considered system can operate in the U1 mode, i.e. overload (state of reached operational capability) - critical time-per-vehicle values have been exceeded and there is no capacity to process any more vehicles in the predefined time - congestion occurs. The system time after T_{13} can switch to the U3 mode with probability p_{13}. This is a safe mode, however, some capacity issues may arise (state of partial operational capability). Should that situation occur, it is fair to say the congestion has been alleviated to the extent, where the queue of vehicles going through the toll collection system is contained within admissible limits, however, entering the emergency mode could not be guaranteed. Some preventative measures could be deployed in that situation, which could prevent congestion from forming through e.g. opening another toll gate. A system operating in mode U3 can pass to mode U1 after time T_{31} with probability p_{31}. For that passage parameter K was introduced, which determines the probability of making an unsuccessful attempt to prevent congestion. Hence, there are no 100% effective congestion preventing measures. The K parameter depends on the tolerance of adopted timeframes either decreasing or increasing for different service levels. A system in mode U3 after time T_{32} can pass to mode U2 (state of full operational capability - normal operation) with probability p_{32}. This state means alleviating congestion at toll gates - all vehicles go through within the predefined time. A system operating in the U2 mode can pass to the U3 mode after time T_{23} with probability p_{23}. For that passage the K parameter has to be factored in, which determines the probability of making an unsuccessful attempt to prevent congestion. This transition means the system in the state of normal operating mode has generated a queue within assumed limits, however reaching operational capability could not be guaranteed. A system operating in mode U2 can pass to mode U1 after time T_{21} with probability p_{21}. For that pass the K parameter has to be factored in, which determines the probability of making an unsuccessful attempt to prevent congestion. This transition means a queue of vehicles has formed, where critical time-per-vehicle values have been exceeded and there is no capacity to process any more vehicles is the predefined time.

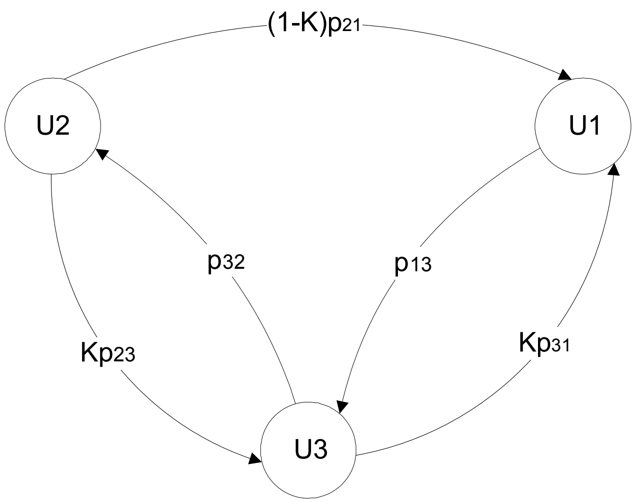


Fig. 1. State diagram of toll collection system in terms of assuring the required service quality
Source: own development

The diagram presented in Figure 1 could be represented by Figure 2. Notations are the following:

- P_{U2} – steady-state probability of system in the state of full operational capability (normal operation) - all vehicles go through mainline barrier toll gates within the predefined time;
- P_{U1} – steady-state probability of system in the state of overload (state of reached operational capability) - critical time-per-vehicle values have been exceeded and there is no capacity to process any more vehicles in the predefined time - congestion,
- P_{U3} – steady-state probability of system in the state of normal operational capability with a possible safety breach (state of partial operational capability) – a queue has formed within assumed limits, however reaching the operational capability could not be guaranteed.
- λ_{21} – transition rate from the state of full operational capability to the state of reached operational capability, where critical time-per-vehicle values have been exceeded and there is no capacity to process any more vehicles in the predefined time;
- λ_{23} – transition rate from the state of full operational capability to the state of partial operational capability, where a queue has formed within assumed limits, however reaching the operational capability could not be guaranteed;
- λ_{31} – transition rate from the state of partial operational capability to the state of reached operational capability, where critical time-per-vehicle values have been exceeded by the length of queue and there is no capacity to process any more vehicles in the predefined time. This transition means that

anti-congestion algorithms are not 100% efficient. This is reflected by the behaviour of the actual system.

- μ_{13} – transition rate from the state of reached operational capability to the state of partial operational capability. This transition means diminishing the queue of vehicles down to volumes within assumed limits, however reaching the operational capability could not be guaranteed;
- μ_{32} – transition rate from the state of partial operational capability to the normal operating mode. This transition means diminishing the queue of vehicles down to volumes which could be processed in the given timeframe whilst assuring the highest service level for users of toll collection system;
- K - probability of making an unsuccessful attempt to prevent congestion.

Transition rates between the system transfers in terms of assuring high enough service quality were generated by a proprietary application for the maintenance decision support described in the following chapter 2.2. It uses a video system to process, analyse, edit and store video footage recorded digitally. An additional source of data comes from continuous traffic monitoring facilities [4] obtained upon the consent of General Directorate for National Roads and Motorways (GDNRM) as the information used for scientific purposes. Paper [4] contains the data from the continuous traffic monitoring itself and its analysis dating back to 2010 - 2011. The continuous traffic monitoring is automated and uses the so called Continuous Traffic Monitoring Sites (CTMS). They are located along Polish dual-carriageways and motorways in 66 locations.

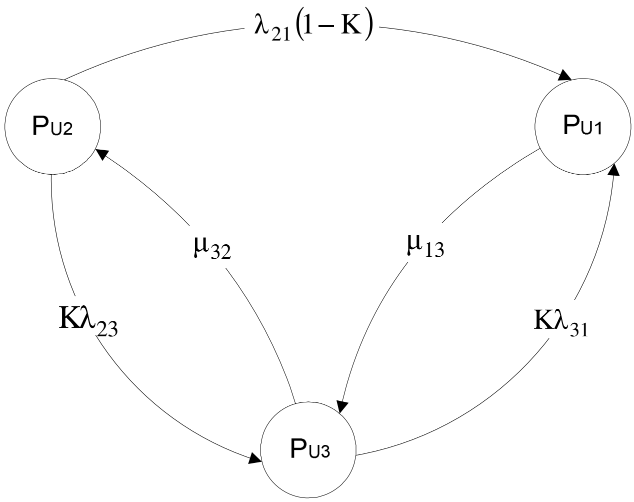


Fig. 2. State diagram of toll collection system in terms of assuring required service quality Source: own development

For states of toll collection system in terms of assuring high enough service quality defined above in Fig. 2, the probability changing during different states could be given by the following system of differential equations:

$$\frac{dP_{U1}(t)}{dt} = -\mu_{13}P_{U1}(t) + \lambda_{21}(1-K)P_{U2}(t) + K\lambda_{31}P_{U3}(t)$$

$$\frac{dP_{U2}(t)}{dt} = -\lambda_{21}(1-K)P_{U2}(t) - K\lambda_{23}P_{U2}(t) + \mu_{32}P_{U3}(t) \tag{1}$$

$$\frac{dP_{U3}(t)}{dt} = \mu_{13}P_{U1}(t) + K\lambda_{23}P_{U2}(t) - \mu_{32}P_{U3}(t) - K\lambda_{31}P_{U3}(t)$$

After numerous transformations we arrive at one the most important reliability metric, namely the Mean Time To Failure (MTTF). Because the moment of failure of manual toll collection system is defined as exceeding critical time-per-vehicle values and no capacity to process any more vehicles in the predefined time, in this case the MTTF defines the mean time to exceeding critical time-per-vehicle values and congestion.

In the discussed case the parameter is given by:

$$MTTF = \frac{2\mu_{32} + K\lambda_{31}}{(\lambda_{31}\cdot\lambda_{23} - \lambda_{21}\cdot\lambda_{31})\cdot K^2 + (\lambda_{21}\cdot\lambda_{31} - \lambda_{21}\cdot\mu_{32})\cdot K + \lambda_{21}\cdot\mu_{32}} \tag{2}$$

This result is an important parameter that determines the reliability of a toll collection system, expressed by parameters λ and μ of vehicle service and the K parameter.

The presented model of a toll collection system in terms of capacity to deliver high enough quality toll services enabled describing the behaviour of the entire toll collection system. The analysis of graphically represented process of delivering high enough quality toll services enables determining characteristics as steady-state likelihood functions for system transients and the reliability of toll collection system as the MTTF parameter - mean time to failure. In the case of toll collection system the moment of failure is defined as breaching customer service levels and exceeding capacity to assure continuous traffic flow.

2.2 Maintenance Decision Support Application for Manual Toll Collection Systems

In order to conduct a reliable evaluation of a manual toll collection system in terms of capacity to deliver high enough quality toll services, apart from the continuous traffic monitoring data obtained from the GDNRM [4], a proprietary maintenance decision support application was used for a manual toll collection system. The maintenance decision support application for the toll collection system was used to obtain the data to verify service levels defined in the specification [5]. The proposed solution turned out to be a very good tool indeed, for obtaining the data to verify customer service levels. The system also enables the verification to be carried out in any way from any place, both based on a live video feed and pre-recorded footage stored in a video surveillance system (Figure 3). To verify not only the toll collection process itself, but also the capacity of the system, both the toll collection plaza and the timeframe were selected, which already had been scrutinised by means of traditional methods. In both cases, the objective is to determine the waiting time for vehicles at the toll collection plaza. Hence, the following parameters are recorded at the toll gate (TG):

– time of admittance;
– starting time of vehicle processing at a given gate;
– finishing time of vehicle processing at the given gate;
– current queue at the given gate.

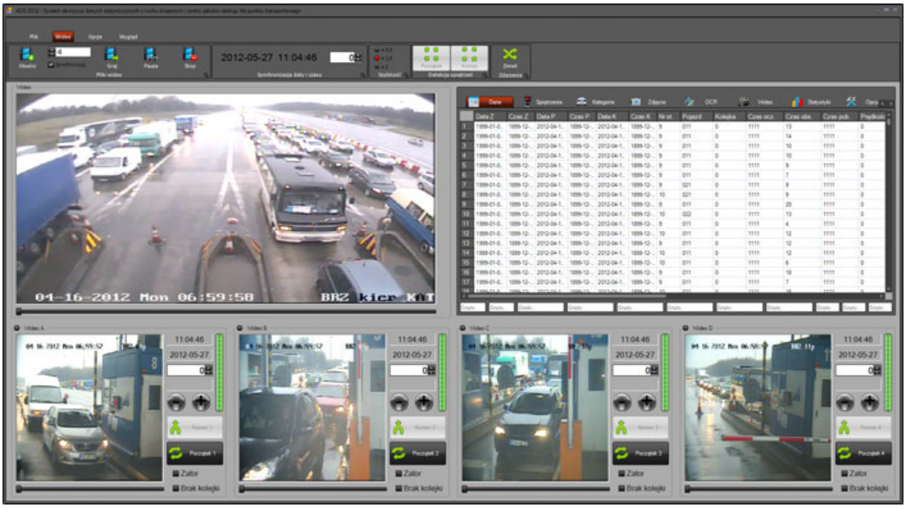


Fig. 3. Screenshot of software processing archival data. Source: own development.

Gathering that data enables determining parameters key for the service level at a toll collection plaza and input parameters for the model of manual toll collection system in terms of delivering high enough quality toll services.

The waiting time at the toll plaza (TP) was measured in two stages. Firstly, vehicle accumulation periods are identified, where there is the highest probability of missing service level standards for users of the toll collection system. Secondly, waiting times for users of the toll collection system are measured, based on stage one results. The software (Figure 3) allowed identifying at the first stage the traffic build-ups, where there is the highest probability of missing service level standards for users of the toll collection system. The software registered initial and end times for periods, where the queue has formed within assumed limits, however reaching the operational capability could not be guaranteed. That state means exceeding critical time-per-vehicle values and there is no capacity to process any more vehicles in the predefined time. At the second stage the periods selected based on measurements of traffic build-up periods taken at stage one were scrutinised. In order to obtain the most reliable and accurate measurements, the key stage involving using the maintenance decision support application for the toll collection system was split into two phases. In the first phase, processing times were measured for individual lanes and TG. Thus the processing starting times were measured, i.e. upon the vehicle approaching TG. Then processing end times were recorded, i.e. the moment when vehicle passed the barrier. Although the most obvious moment to determine the completion of processing is seemingly the moment of barrier lifting, this option was rejected. Observations proved that toll attendants in order to speed up the toll collection process lift the barriers once the driver makes the payment as opposed to when the processing is finished. Determining both those times (start and end of toll collection process) is synonymous with determining processing times at TG. At the second stage of measurement every vehicle whose start and end times have been determined at the first stage, are identified in the video footage using the Optical Character Recognition as the vehicle arrives at TP and joins the queue at TG. The arrival time was recorded as the entry to the toll collection system. The difference between processing finished (departure from TP) and processing started determines how long the vehicle remains in the system. This time is the total waiting time and has to comply with the imposed service level standards. In that manner the total processing time was determined for all vehicles passing through the analysed toll lanes. These results enabled a real-time monitoring of the toll collection at analysed TG. Consequently, correct maintenance decisions could be made regarding the manual toll collection system.

3 Conclusion

Issues discussed in this paper are only part of broader deliberations [6, 7, 8] concerning payment services using a manual toll collection system. The problem of assuring required service levels by manual toll collection systems is and will become in near-term a current issue as far as implementing electronic toll collection (ETC) on Polish roads is concerned. This problem is key in the context of using mixed toll collection systems, where the crucial parameter determining customer service levels is meeting timeframes for processing vehicles. Thereby it is highly justifiable to devise new evaluation methods and tools supporting the decision-making process for maintenance-related issues, in the context of manual toll collection systems. The project of a maintenance decision support application for manual toll collection systems caters for that demand. The system has shown a lot of promise already during trials by being flexible and reliable. In a substantial manner it could potentially help to evaluate reliably manual toll collection systems and support maintenance decision-making processes.

References

1. Act of 7 November 2008 amending Act on public roads and some other acts, consolidated text Journal of Laws 2008, No. 218 Item 1391
2. General Directorate for National Roads and Motorways: Terms of Reference. Closed tender for National Toll Collection System, including services related to Electronic Toll Collection, Warsaw (2010)

3. Act of 21 March 1985 on Public Roads, consolidated text, Journal of Laws 2007 No.19 Item 115 with subsequent amendments
4. General Directorate for National Roads and Motorways: Records and analysis of data from continuous traffic monitoring sites 2010 - 2011, Warsaw (2011)
5. General Directorate for National Roads and Motorways: Key Technical Requirements for Designing, Building and Maintaining Motorways: Operational requirements - reliability and availability of system, Warsaw (2007)
6. Kasprzyk, Z., Rychlicki, M.: Telematic system for data acquisition for purposes of evaluating quality of transport facility. Publishing House of Institute of Logistics and Warehousing in Poznan, Journal Logistics (4), 408–417 (2011)
7. Kasprzyk, Z., Sumila, M.: Evaluating operational reliability of toll collection system. Publishing House of Institute of Logistics and Warehousing in Poznan, Journal Logistics (4), 418–427 (2011)
8. Kasprzyk, Z., Rychlicki, M.: Evaluating customer service levels delivered by toll collection system. Publishing House of Institute of Logistics and Warehousing in Poznan, Journal Logistics (3) (May-June 2011)

Compass Made Good Correction with MTE

Włodzimierz Filipowicz

Gdynia Maritime University, Faculty of Navigation
Aleja Jana Pawła II 3, 81-225 Gdynia, Poland
wlofil@am.gdynia.pl

Abstract. Imprecise and uncertain data dominate in maritime transportation. The Mathematical Theory of Evidence (MTE) [12, 14] extended to a possibilistic platform [16] enable processing the uncertainty. The Mathematical Theory of Evidence enables upgrading new models and solving crucial problems in many disciplines. The evidence combining scheme as a mechanism enabling enrichment of the initial data information context is useful in many cases. In nautical applications it can be used to make a fix and to evaluate its accuracy. The MTE delivers a new unique opportunity once one engages fuzzy values. Approaches towards a theoretical evaluation of tasks including imprecise data are to be reconsidered. A compass made good correction[1] evaluation is one of such problems. To calculate the correction one has to know the direction towards a landmark or a celestial body. Taking two bearings to landmarks situated at opposite sides is also sufficient. Landmarks situated at counter bearings locations are not available very often.

Keywords: Mathematical Theory of Evidence, navigation, compass made good.

1 Introduction

Plenty of various qualities of data are available in the marine navigation. Indications meant as fixed position delivered by navigational aids, bearings, distances etc. are most frequently used. All of them include fixed and random errors, therefore they affect the result of reasoning in different ways. A bearing is the direction towards an object from a vessel. A bearing is ordinarily calculated in a clockwise direction starting from a reference direction of the true north and moving up to 360 degrees. Taking bearings to various objects is quite often in navigation. Obviously the bearing contains a fixed and random deflection from the true value. In order to get rid of the fixed deviation one can engage a differential approach once the position fixing is considered. Nonetheless the knowledge of a systematic compass error is required quite often. The error evaluation involves the reference value that can be a celestial body azimuth or the leading lights that are two beacons in a line. The beacons are separated in distance and elevation, so that when they are aligned, with one above the other, they provide a bearing. Opportunities are not always available. Thus yet another method of calculation is looked forward.

---

[1] From the Publisher: Author had in mind the specific correcion of the CMG.

J. Mikulski (Ed.): TST 2012, CCIS 329, pp. 69–78, 2012.

2 Navigational Observations

The taking of a bearing gives an imprecise value. The measurement contains a systematic and random deflection (see Figure 1). In the presented example random errors are governed by the bell function distribution.

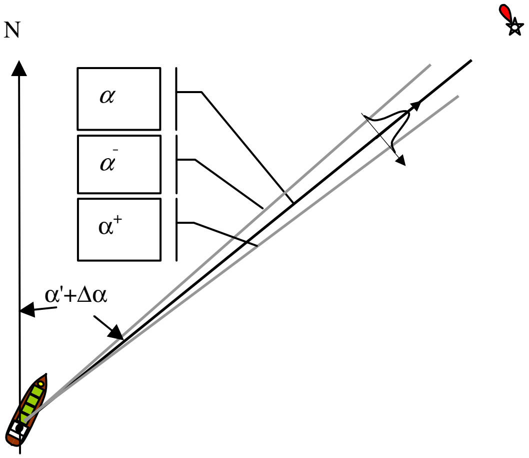


Fig. 1. Result of taking a bearing is imprecise

Crisp valued standard deviation of a measurement is inadequate. In recent navigation books the measurement mean error is described as an imprecise interval value usually as: [$\pm\sigma^{-}_{\alpha}$; $\pm\sigma^{+}_{\alpha}$]. Being interested in an isoline deflection one considers [$\pm m^{-}_{\alpha}$; $\pm m^{+}_{\alpha}$]. The following relation holds in the mentioned two cases.

$$m_{isoline} = \frac{\sigma_{measurement}}{|g|} \tag{1}$$

where - $|g|$ is an isoline gradient module.

Seafarers know that the mean error of a distance measured with a radar variable range marker is within the interval of [±1% ÷ ±1.5%]. In the same condition the mean error of a bearing taken with a gyrocompass is within the range [±0.5° ÷ ±1°], using a medium class radar the same error is within [±1° ÷ ±2°] [7]. With the fuzzy arithmetic notation the latest can be written as a quad (-2,-1, 1, 2). It means a fuzzy value with the core of [-1°, 1°] and the support of [-2°, 2°]. In Fig. 2 the dashed line shows that there is a membership function that returns the possibility of certain point inclusion into [$-m_1$; $+m_1$] range.

Observational random errors are assumed to follow the normal distribution. The mean error estimates the standard deviation (square root of a variance) of the distribution. Figure 2 shows values of confidence intervals related to two distribution functions. A measurement falls into the confidence range with certain probability. It is assumed that confidence intervals are symmetrically placed around the mean. The confidence equal to 0.683, for a Gauss probability density function, is related to the interval of [-m_α; +m_α]. Since confidence intervals have imprecise limits the last can be rewritten as: $\left[-[m_1^-, m_1^+], +[m_1^-, m_1^+]\right]$[2]

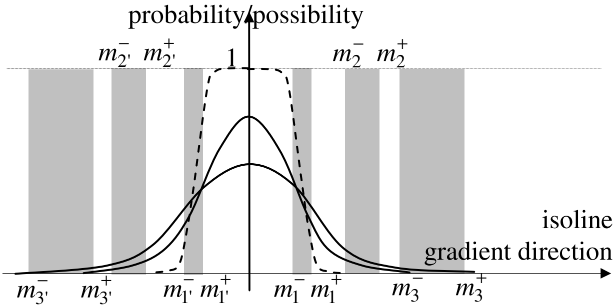


Fig. 2. Modelling imprecise measurements invokes probabilistic and possibilistic approaches

Two distribution functions related to the same observation, introduce imprecision that is usually expressed by a fuzzy value or a fuzzy set. Figure 2 shows a trapezoid-like membership function that locates adjacent bearings within the defined set. The function returns the possibility regarding given *x*, it attributes *x* degree of inclusion within the set. For example the abscissa $x = \alpha+0.5$[3] fully belongs to the given set, contrary to another point $x = \alpha+1$, its inclusion within the set is partial since the degree of its membership is equal to 0.5. Different membership functions intended for nautical application were discussed by the author in his previous papers [4, 5].

A new scheme enabling inclusion of knowledge into the compass fixing process is presented in the paper. The way of computation of belief and plausibility as well as of location vector grades can be found in another paper [3]. The location vectors constructed based on various distributions of measurement errors were presented in [6].

The paper is devoted to the new idea to be followed in the terrestrial navigation. A graphic presentation of the main idea of the considered problem is presented in Fig. 3. It is assumed that two bearings and two distances for two landmarks were taken. Distances and bearings are imprecise values. Bearings were established for objects located at opposite sides of the observer location. Thus their gradient directions are opposite but almost collinear. Distances are of secondary meaning; their precision is sufficient to solve the problem that is aimed at identification of bearing systematic error. The situation scheme is depicted in the insertion. Chord d_b with endpoints lying at the circumference of circle d_2 and marked by both bearings is explored. A point located at the line segment for which the credibility of representing the true position

---

[2] m_α - is modal for $[m_1^-, m_1^+]$.

[3] Example value expresses relative to mean error, the distance from the considered isoline.

of bearings intersection is sought. The point is expected to be the solution of equation defined by formula (2).

$$\frac{\Delta\alpha + \delta_{\alpha_1}}{57.3} \cdot d_1 + \frac{\Delta\alpha + \delta_{\alpha_2}}{57.3} \cdot d_2 = d_b \tag{2}$$

where:

- $\Delta\alpha$ - permanent, systematic error of bearings
- δ_{α_i} - random error of the *i*-th bearing
- d_i - distance to the *i*-th landmark

More precisely equation (2) is to be solved with respect to $\Delta\alpha$. Random components δ_{α_1} and δ_{α_2} should receive maxima of their belief and plausibility measures in the solution. Unfortunately the solution to such problem is trivial.

$$C_i \cdot m_1 + C_i \cdot m_2 \geq d_b \tag{3}$$

An alternative approach comes out of Formula (3). The formula specifies that the required range can be covered through sufficiently large random errors. Error values for extreme observations should cover the considered distance.

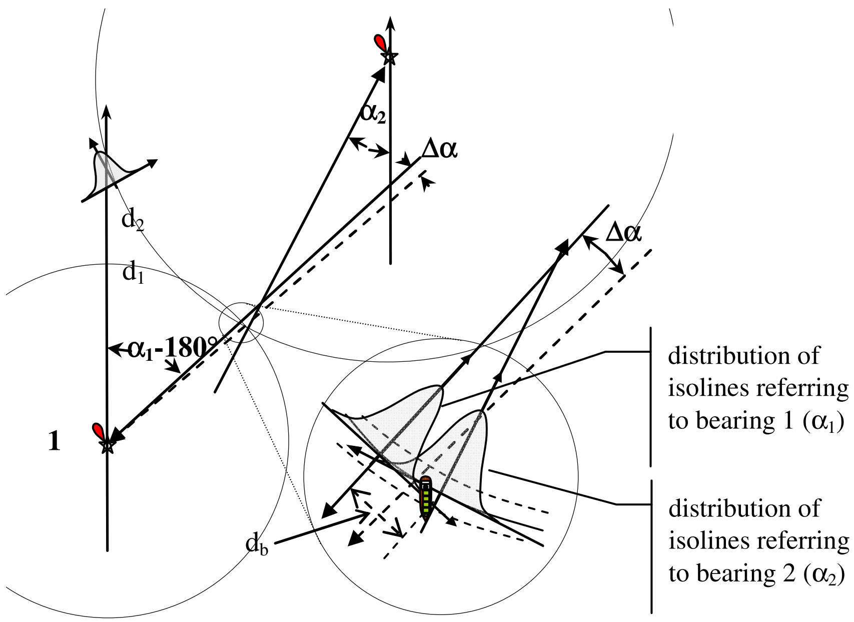


Fig. 3. Scheme of problem considered in the paper

The goal can be achieved with the Mathematical Theory of Evidence. An iterative combination algorithm is used starting from initial mean error values. During iterations the error values are increased, and the inconsistency, belief and plausibility measures recorded. Once the maximum of belief is obtained the process stops, provided that the level of inconsistency is satisfactory.

3 Imprecise Evidence Encoding

In navigation an isoline joins points of equal bearing, distance or horizontal angle. The isoline of a bearing is a line. Mean errors of taken bearings distribution are the same for observations taken with the same compass. For a given set of bearings taken to landmarks respective isolines distributions depend on distances from the landmarks. In the possibilistic approach an uncertain evidence is represented with sets and masses of confidence attributed to these sets. Sets embrace relations between hypothesis and evidence spaces. Relations can be binary or fuzzy ones. Fuzzy sets embrace grades expressing possibilities of consecutive hypothesis items belonging to the sets related to each piece of evidence. Therefore appropriate relations between considered spaces are encoded into the evidence representation, which takes the form specified by Formula (4) [6, 9, 10].

$$\mathrm{m}(e_i) = \{(\mu_{i1}(x_k),\ f(e_i \to \mu_{i1}(x_k)), \cdots, (\mu_{in}(x_k),\ f(e_i \to \mu_{in}(x_k))\} \tag{4}$$

In the presented application the evidence representation consists of pairs: fuzzy vectors representing locations of a set of points within sets related to each piece of evidence – degrees of confidence assigned to these vectors. Degrees of confidence reflect the probability of a true isoline being located within the given strip area. Appropriate imprecise values are obtained based upon statistical investigations of observations distributions.

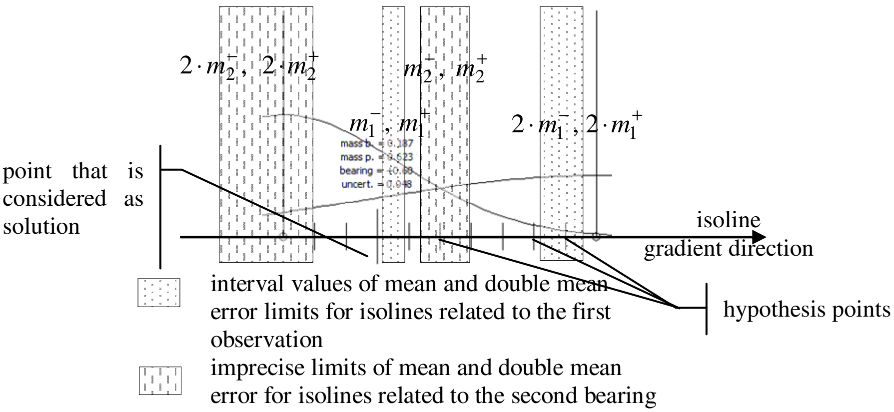


Fig. 4. Scheme of two isolines distributions with an example frame of discernment

Two example fragments of probability distributions, imprecise confidence intervals and hypothesis frame are presented in Fig. 4. Respective membership grades along with their credibility values are shown in Table 1. The table contains two examples of practical evidence representations.

The structures combination is carried out in the Mathematical Theory of Evidence [1, 2]. During the combination all pairs of location vectors are associated and the product of involved masses is assigned to the result set. The obtained assignment is supposed to increase the information context of the initial structures. The combination of structures embracing the measurements data is assumed to result in the determination of the systematic error. The goal can be achieved provided that the association of sets enables the selection of common points located within the intersection of introduced ranges. An adequate selection can be performed with T-form operations [13] used during the association. The simplest T-form results in smaller values being taken from consecutive pairs of elements in associating vectors. This operation is used in numerical examples presented further in the paper.

Table 1. Two example evidence representations

		1	2	3	4	5	6	7	8	9	$m(..)$
bearing 1	$\mu_{11}(x_i)$	{1.00	0.99	0.04	0.00	0.00	0.00	0.00	0.00	0.00}	0.615
	$\mu_{12}(x_i)$	{0.00	0.00	0.95	0.99	0.76	0.00	0.00	0.00	0.00}	0.245
	$\mu_{13}(x_i)$	{0.00	0.00	0.00	0.00	0.24	0.99	0.95	0.31	0.00}	0.036
	$\mu_{14}(x_i)$	{0.00	0.00	0.00	0.00	0.00	0.00	0.04	0.68	0.99}	0.004
	$\mu_{1u}(x_i)$	{1	1	1	1	1	1	1	1	1}	0.1
bearing 2	$\mu_{21}(x_i)$	{0.00	0.00	0.00	0.05	0.72	0.99	1.00	1.00	1.00}	0.546
	$\mu_{22}(x_i)$	{0.98	0.99	0.99	0.94	0.27	0.00	0.00	0.00	0.00}	0.218
	$\mu_{23}(x_i)$	{0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00}	0.036
	$\mu_{2u}(x_i)$	{1	1	1	1	1	1	1	1	1}	0.2

$\mu_{iu}(x_i)$ stands for uncertainty vector

Results of combination can be further examined to calculate probabilities of representing the true bearing by each of hypothetical points. A support plausibility measure may be calculated using formula (5). In the presented expression component $m(\mu_k(x_i))$ is a credibility mass attributed to the k-th intersection of ranges. Factor $\mu_k(x_i)$ reflects fuzzy locations of hypothetical points within ranges intersections.

$$pl(x_l) = \sum_{k=1}^{n} m(\mu_k(x_i)) \cdot \mu_k(x_l) \tag{5}$$

A support belief for certain hypothesis embraced in the given family of sets can be calculated with formula (6). The formula intended for singletons as reference sets is a reduced version of the general expression [2]. The reduction is made due to a special kind of the reference fuzzy set used in the considered applications. To calculate the belief value one has to find the minimum among complemented grades (¬) of each set, the grade of interest is to be omitted. It should be noted that a multiple point presence within the given intersection of ranges causes that the belief for each of points is zeroed. For this reason, in the presented form, the belief cannot be considered a primary factor for solving the position fixing problem. Nevertheless it can be useful while identifying uniqueness of the solution.

$$bel(x_l) = \sum_{k=1}^{n} m(\mu_k(x_i)) \min_{x_i \in \Omega;\, i \neq l} (\neg \mu_k(x_i)) \tag{6}$$

Results of calculations carried out with algorithm 1 are presented in Table 2. In the first row there are data obtained at the initial stage of calculations. In this case the sum of isolines mean errors is smaller than the considered range of two units. As a consequence the inconsistency is rather high; maxima of belief and plausibility refer to various points. The inconsistency occurs when a mass greater than zero is assigned to the empty set [15]. T-norm is a null producing operation.

```
                      Algorithm 1:
1. For one of given distances d_i calculate distance d_b
between extreme isolines as marked in figure 3
2. For given bearings errors distribution calculate mean
errors for each of available isolines, use gradient
modules for calculations
3. Create belief structures and combine them, record
inconsistency occurred during combination
4. Calculate belief and plausibility measures for all
hypothesis frame points
5. Select point with greatest belief/plausibility value
6. If stopping_condition then calculate compass
correction and quit
7. Increase isolines mean errors and go to point 3
```

Table 2. Results of structures combinations

$C_i \cdot m_1$, $C_i \cdot m_2$ isolines errors [cables]	maximum belief	maximum plausibility	bearing correction [deg]	uncertainty
0.5, 1.0	0.068[4]	0.427[2]	+0.40	0.350
0.67, 1.33	0.080	0.676	+0.60	0.045
0.70, 1.40	0.150	0.650	+0.60	0.045
0.74, 1.47	0.187	0.623	+0.60	0.048
0.78, 1.54	0.163	0.609	+0.60	0.048

[4] Values obtained for different hypothetical points.

Condition (3) is observed at points close to the solution. The uncertainty remains low but tends to increase in the next iterations. Belief and plausibility are rather high; they reach their maxima for the same point that is considered as the compass correction. Belief and plausibility measures for results presented in the shaded second last row of Table 2 are diagrammed in Fig. 4.

It should be also noted that the value of result's systematic error remains constant at the last stage of calculations.

Based on the above remarks the `stopping_condition` in the presented algorithm takes the form of the following implication: $C_i \cdot m_1 + C_i \cdot m_2 \geq d_b$ AND `bel_max`, where `bel_max` identifies reaching the maxima by belief and plausibility measures.

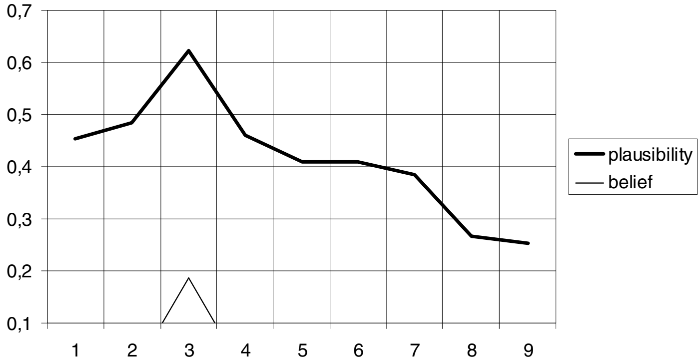


Fig. 5. Diagrams of belief and plausibility measures

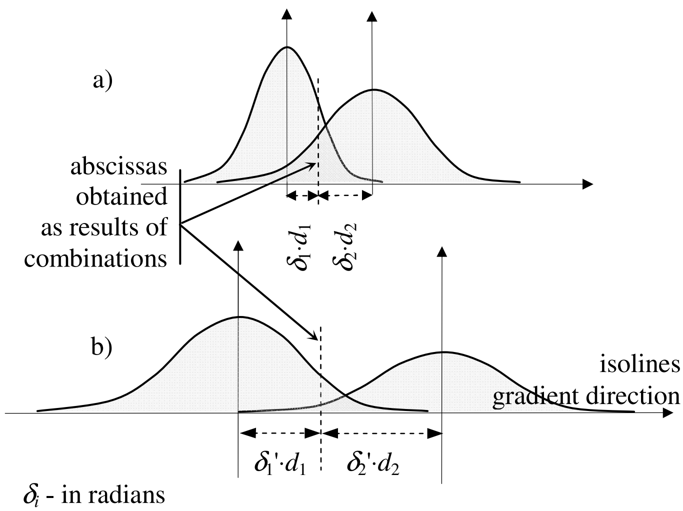
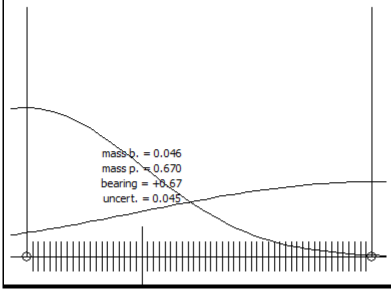


Fig. 6. Two cases of separated random variables distributions

Table 3. Results of belief structures combinations

diagram	$C_i \cdot m_1$, $C_i \cdot m_2$ isolines errors [cables]	maximum belief, plausibility	bearing correction [deg]	uncerta-inty
	0.64, 1.27	0.03, 0.59	+0.67	0.045
	0.67, 1.33	0.05, 0.67	+0.67	0.045
	0.70, 1.40	0.04, 0.63	+0.67	0.045

4 Conclusion

Results of the combination of evidence, related to two random variables that projected on a plane are separated by a certain distance, were presented in the paper. Considered variables referred to isolines related to the bearings taken for two objects. Results of observations include both systematic and random errors. The identification of a permanent measurement shift is an important practical nautical issue.

The identification can be achieved with an uncertain evidence combination scheme. It engages an iterative proportional increment of isolines mean errors. Fig. 5 shows two cases characterized by the same sets of initial data regarding observation conditions and their subjective assessments. Case a) does not embrace the systematic error; in situation b) the multiplied mean random errors cover a lengthy space due to the fix deflection, for both cases the condition $\delta_1/\delta_1' = \delta_2/\delta_2'$ holds; provided that $m_1' = C \cdot m_1$ and $m_2' = C \cdot m_2$ and engaging an imprecise evidence combination scheme.

With the Mathematical Theory of Evidence one should iterate over defining and combining belief structures and analysing the obtained solution. Iterations stop when the maximum belief measure is recorded and a sufficient level of inconsistency noticed.

The calculation of support belief multiple point presence within the given intersection of ranges causes that the belief for each of points is significantly reduced. Table 3 contains a diagram and three sets of results for the case with 60 points frame of discernment. The middle set presented in the shaded row embraces the final solution. The particular frame causes that within a smaller width of mean error there are about 20 search points. This is the reason that components of formula (6) deliver very small values. In this case each range refers to two angle minutes. Thus the relation between the width and mean error seems inadequate. It should be noted that the obtained results slightly differ from those achieved for a reduced frame. For both cases the accuracy of correction estimations is satisfactory.

References

1. Dempster, A.P.: A generalization of Bayesian inference. Journal of the Royal Statistical Society: Series B 30, 205–247 (1968)
2. Denoeux, T.: Modelling Vague Beliefs using Fuzzy Valued Belief Structures. Fuzzy Sets and Systems 116, 167–199 (2000)
3. Filipowicz, W.: Belief Structures in Position Fixing. In: Mikulski, J. (ed.) TST 2010. CCIS, vol. 104, pp. 434–446. Springer, Heidelberg (2010)
4. Filipowicz, W.: New Approach towards Position Fixing. Annual of Navigation 16, 41–54 (2010)
5. Filipowicz, W.: Belief Structures and their Application in Navigation. Methods of Applied Informatics 3, 53–82 (2009)
6. Filipowicz, W.: Evidence Representation and Reasoning in Selected Applications. In: Jędrzejowicz, P., Nguyen, N.T., Hoang, K. (eds.) ICCCI 2011, Part II. LNCS, vol. 6923, pp. 251–260. Springer, Heidelberg (2011)
7. Jurdzinski, M.: Principles of Marine Navigation. WAM, Gdynia (2008)
8. Lee, E.S., Zhu, Q.: Fuzzy and Evidence Reasoning. Physica-Verlag, Heidelberg (1995)
9. Liu, W., Hughes, J.G., McTear, M.F.: Representing Heuristic Knowledge and Propagating Beliefs in Dempster-Shafer Theory of Evidence. In: Federizzi, M., Kacprzyk, J., Yager, R.R. (eds.) Advances in the Dempster-Shafer Theory of Evidence, pp. 441–471. John Willey and Sons, Inc., New York (1992)
10. Liu, W., Hughes, J.G., McTear, M.F.: Representing Heuristic Knowledge in D-S Theory. In: UAI 1992 Proceedings of the Eighth Conference on Uncertainty in Artificial Intelligence. Morgan Kaufmann Publishers Inc., San Francisco (1992)
11. Klir, G.J., Parviz, B.: Probability – Possibility Transformations: a Comparison. International Journal of General Systems 21(1) (1992)
12. Kłopotek, M.A.: Identification of belief structure in Dempster – Shafer Theory. Foundations of Computing & Decision Sciences 21(1), 35–54 (1996)
13. Rutkowski, L.: Methods and Techniques of the Artificial Intelligence. PWN, Warsaw (2009)
14. Shafer, G.: A Mathematical Theory of Evidence. Princeton University Press, Princeton (1976)
15. Yager, R.: On the Normalization of Fuzzy Belief Structures. International Journal of Approximate Reasoning 14, 127–153 (1996)
16. Yen, J.: Generalizing the Dempster – Shafer theory to fuzzy sets. IEEE Transactions on Systems, Man and Cybernetics 20(3), 559–570 (1990)

Data Flows in an Integrated Urban Freight Transport Telematic System

Stanisław Iwan[1] and Krzysztof Małecki[2]

[1] Maritime University of Szczecin,
11 Pobożnego St., 70-515 Szczecin, Poland
s.iwan@am.szczecin.pl
[2] West Pomeranian University of Technology,
11 Żołnierska St., 70-515 Szczecin, Poland
kmalecki@wi.zut.edu.pl

Abstract. The fast growth of cities as well as the increasing complexity of processes taking place within them forces changes in planning and managing them. One of the key factors necessary for controlling the increasing disorder in the area of urban freight transport and conditioning its sustainable growth is ensuring effective data flows between individual parties engaged, to a smaller or greater extent, in its functioning. This paper is focused on assumptions and analysis of data sources for integrated UFT management systems supported by telematics systems.

Keywords: urban freight transport, city logistics, data flows, integration of management systems.

1 Introduction

The fast growth of cities as well as the increasing complexity of processes taking place within them forces changes in planning and managing them. The development of advanced societies, which translates into the increasing requirements regarding the provision of products and services necessary for well-being of the city users, businesses and visitors to the city result in more and more complex tasks for city administrators.

The research conducted in Germany has shown that 80% of the transport performed within cities is done by means of road transport [6]. Choosing this kind of transport for making deliveries in urban areas leads to many adverse effects connected mainly with environmental pollution, compromised safety on the roads and a negative image of a city along with diminishing its functionality. Another problem is the congestion effect which contributes to increasing costs of vehicle operation and infrastructure maintenance, time wasted by transport users and losses connected with the time needed for delivery making, etc. The majority of transport in urbanised areas is generated by industrial, commercial and service companies. The distribution function initiated by the businesses contributes to an increase in logistic flow streams within a limited area. Currently they are usually implemented in an uncoordinated and chaotic manner and it is extremely difficult to organise, manage and optimise them.

J. Mikulski (Ed.): TST 2012, CCIS 329, pp. 79–86, 2012.

A solution to stop all those negative processes is sustainable growth which deals mainly with analysing the reasons for excessive use and deterioration of the environment, and also with specifying a strategy for limiting the process in three main areas: ecological, economic, social and cultural. The most effective activities in a sustainable freight transport policy are those which concurrently take into account all the three aspects while minimising the costs connected with meeting the expected goals.

Sustainable growth is now one of the key challenges for more and more diversified and complicated economic systems. Developing appropriate assumptions for implementing the concept becomes of key importance for achieving a success and minimising the costs connected both with the implementation process itself and the subsequent use of the solution. One of the key factors necessary for controlling the increasing disorder in the area of urban freight transport and conditioning its sustainable growth is ensuring effective data flows between individual parties engaged, to a smaller or greater extent, in its functioning. Lack of knowledge regarding cargo streams and their direction, structure etc. makes it difficult to control and manage them so as to limit their negative impact on the city organism, particularly on the environment and inhabitants. Additionally, development of information society and continuous digitalisation of different areas of life make it necessary to get control over the data resources regarding urban freight transport, as it is becoming indispensable for its correct functioning.

2 Importance of Data Flows for UFT Efficiency

Utilization of telematics systems in the area of urban freight transport depends on [7]:

- reducing freight distribution costs
  - increasing productivity of local delivery vehicles;
  - increasing reliability of commercial vehicle operations;
  - increasing safety;
- increasing the capacity of urban freight systems (without providing additional traffic infrastructure).

The basis for efficient management of urban freight transport is an effective information flow and ability of knowledge extraction from the fast growing data resources. In view of the above assumptions, the following question remains unanswered: What is the actual demand for data and information voiced by the participants of the above mentioned transport processes within the city, their stakeholders and other beings (institutions, organisations, entities) connected with them and dependent on them? On the other hand, where should we acquire the specified data resources, what structure and scope they should have and what can be the costs connected with the data acquisition? And consequently, what can be the impact of determining the demand for data on urban freight transport effectiveness?

Specifying the information needs within the area of urban freight transport, and also the interdependencies between them and the other elements of the city's transport system will make it possible to take control over the increasing disorder in this area and enable its development in accordance with the principles of sustainable growth.

3 The UFT Solutions Based on Application of Telematics

It‘s possible to divide the present UFT solutions into many kinds. "BESTUFS Good Practice Guide" divides them into three major groups [1]:

- goods vehicles access and loading in urban areas,
- last mile solutions,
- urban consolidation centres.

This classification not clearly emphasises the needs at the area of IT and telematics systems application. All kinds of the above solutions include, to a greater or smaller extent, a necessity of information and communication technology implementation. More detailed classification is proposed in C-LIEGE project. The document "Definition of suitable set of actions/measures for an efficient and energy saving organization of goods transport and delivery in urban areas" indicates 17 categories of UFT soft measures in 7 activity areas [3]:

- administrative (access restrictions, access incentives, advance booking (un)loading areas, low Emission/Environmental Zone);
- financial (mobility credits schemes and congestion charging, vehicle financing models);
- organizational (freight traffic routing information, alternative delivery systems);
- technical (intelligent freight traffic routing, IT logistics tools);
- awareness (promotional and awareness campaigns, eco-driving, Freight Operators Recognition Scheme);
- governance (Local Freight Development Plan, distribution plan-scheme, Freight Quality Partnership);
- special urban planning conditions.

Table 1. Urban Freight Transport measures based on IT and telematics systems

Measure	Description
Freight traffic routing information	Channelling trucks that drive into cities of the urban agglomeration through designated truck routes, e.g. by setting up special road signs or providing special maps with designated routes and lorry-relevant road information.
Intelligent freight traffic routing	Integrating designated lorry routes and lorry-relevant information in navigation software. On this basis, data received from freight vehicles in traffic with regard to their current locations, loaded cargo and destination plans can be connected with real-time road traffic data.
Integrated logistics tools	Web-based logistics tools linking and coordinating producers, recipients and freight operators in order to optimise logistics flows.

It's possible to point out three kinds of measures which are based on IT implementation and telematics system application: freight traffic routing information, intelligent freight traffic routing and IT logistics tools (Table 1).

Also some of the other measures need the telematics support for proper functioning. The descriptions of the measures and types of support are presented in Table 2.

Table 2. Urban Freight Transport measures supported by IT and telematics systems

Measure	Description	Type of IT support
Access restrictions for (un)loading and transit	Access restrictions for loading/unloading operation as well as for moving/circulating related to (a) the type of transport means, and most commonly to vehicle emissions, weights and sizes; (b) access time within specified areas; (c) preferred truck routes and designated lanes; (d) loading and unloading zones; (e) based on licences.	Support for optimization, supervision and management
Advance booking of (un)loading slots	From an environmental, business and traffic flow/security point of view it is best for freight vehicles to avoid double lane stops and reduce waiting times for getting into a loading/unloading parking space. To this end a service can be offered that gives drivers the ability to book a delivery parking space before they reach their delivery point.	Support for slots' booking and management
Low Emission/ Environmental Zone	Institution of protected areas that include both vehicle access restrictions and incentives for environmental or historical/heritage reasons.	Support for supervision and pollution detection
Mobility credits schemes/ congestion charging	Limiting the access of freight vehicles to an urban area by making freight operators 'pay' for each access with mobility credits that were initially distributed by the public administration (or money payments for entries in excess of the assigned credits). Access control equipment in freight vehicles record every entry to the zone and permits the implementation of a mixed pricing / enforcement scheme for different users.	Support for supervision and management
Alternative delivery systems	Van-sharing, cycle-logistics, night-delivery service, packstations etc	Support for booking, supervision and management

4 The Data Sources for Integrated UFT Management System

IT and telematics systems application consists of the three following elements: data acquisition, data processing and information dissemination [7]. The data flow structure depends on interactions between the functioning system elements. Table 3 presents the major interactions in a typical transport system.

The most important factor in proper data acquisition for UFT management systems is the identification of appropriate data sources. Based on the analysis of urban freight transport functioning and structure of local road administration (the major area of analysis was West Pomeranian Voivodeship) as well as on literature research [4, 5, 7, 8, 9], the model of data flows in integrated UFT management system was developed. From the system usability point of view and its regional scope, it was possible to indicate 5 groups of entities:

Table 3. Interactions between major components of a transport system (source: own work based on [2])

Components	Description of information flows	Technology
Trackway – vehicle	Communication between the infrastructure and the vehicle with no direct intervention by the users.	GPS, automatic identification, electronic toll collection, advanced traffic control, automated highways.
Trackway – user	Communication between the road infrastructure and the user (static or dynamic).	Road signs, traffic lights, VMS, personalized information systems.
Vehicle – user	Information provided to the user from the vehicle	Sensors and on-board diagnostics systems.
Trackway – vehicle – user	Communication from the roadway to the vehicle and then to the user.	Real-time vehicle navigation systems, vehicle dispatch systems, driver and public transport information systems (in cars, bus and trams stops, computer kiosks, available through computer networks).
Other	By using a wider range of sensors and exploiting high-speed communications and computing, road authorities can access more information about current traffic conditions.	Detection systems, on-board computers.

- public administration units (including road authorities):
  - local, regional and national roads and motorways authorities,
  - local authorities,
  - road transport inspectorates ;
- safety and emergency units:
  - Police headquarters,
  - medical rescue services ,
  - fire service headquarters ;
  - engineering and technical services;
- traffic detection systems:
  - subsystem of transport environmental impact detection ,
  - subsystem of road and traffic conditions detection ;
- system users:
  - shippers, freight carriers, logistics companies and suppliers:
  - road users (drivers);
  - companies (purchasers);
- other institutions:
  - inspectorate for environmental protection,
  - insurance companies,
  - research and development units ,
  - community-based organizations .

Then, data flows were identified, taking into account their importance for the system efficiency as well as their relevance for the above mentioned entities (Table 4).

Table 4. Data Flows In the Proposed System

Terminator	Flow	Specification
Local, regional and national roads and motorways authorities	In1	Road infrastructure data.
	Out1	Data about: • present road situation, • congestion, • road incidents (collisions, accidents, etc.).
Local authorities	In4	Administration decisions in the area of transport system functioning. Road maintenance data. Regulations and rules in the area of transport system functioning.
	Out4	Data about present road situation. Statistical data.
Road transport inspectorates	In5	Regulations and rules in the area of transport system functioning.
	Out5	Data about delivering structure, traffic flows and present road conditions.
Police headquarters	In6	Regulations and rules in the area of transport system functioning.
	Out6	Data about: • the present road situation, • congestion, • road incidents (collisions, accidents, etc.).
Medical rescue services	In7	Emergency units location data.
	Out7	Data about road incidents (collisions, accidents, etc.).
Fire service	In8	Data about external factors (fires, floods, weather events, risk of pollution etc.). Fire units location data.
	Out8	Data about: • the present road situation, • road incidents (collisions, accidents, etc.).
Engineering and technical services	In9	Data about on-going road works. Engineering units location data.
	Out9	Data about: • the present road situation, • road incidents (collisions, accidents, etc.).
Insurance companies	Out10	Data about road incidents (collisions, accidents, etc.).
Subsystem of transport environmental impact detection	In10	Data about the current situation on the roads: • road traffic parameters: – vehicles detection, – vehicles classification, – speed measurement , – traffic intensity measurement, – dynamic vehicle weighing, • weather monitoring: – measurement of air temperature and RH, – measuring the temperature of road surface and the ground, – detecting icy road conditions, – measuring the wind velocity and direction, – atmospheric pressure measurement, – precipitation measurement, • detecting the road surface condition, • visibility measurement (fog detection).
Subsystem of road and traffic conditions detection	In11	Data on the environmental impact of the transport system: • measuring carbon oxide level • measuring nitrogen oxide level • measuring noise level Tracking vehicles carrying hazardous cargo .

Terminator	Flow	Specification
Shippers, freight carriers, logistics companies and suppliers	In12	Data about freight transport (directions, freight specifics, etc.).
	Out11	Data about present road situation. Data about regulations and rules in the area of transport system functioning, including tolls fees.
Road users (drivers)	In13	Vans and lorries location data.
	Out12	Data about the present road situation.
Companies (purchasers)	Out13	Data about the present road situation.
Community-based organizations	Out14	Information on the transport system impact on the social and economic situation in the Province (unemployment level, accident rate, etc.).
Research and development units	In14	Data from research works and results of scientific projects.
	Out15	Data and surveys necessary for research work and scientific projects implementation, including statistical data.
Inspectorate for environmental protection	Out16	Data about pollution, noise level, structure of hazardous cargo transport.

5 Conclusion

In view of the information society development, the more and more important aspect is getting to know the structures and directions of data flows found between objects being part of particular systems functioning in the economy, and their impact on its functioning both in micro scale as well as at the region or country level or even globally.

Efficient functioning of a transportation system is to a large extent dependent on a proper data and information flow as well as on acquisition of knowledge that enables effective management of the processes taking place in the transport system. On the other hand, the complexity of interdependencies occurring in the system and its considerable heterogeneity imply a need for research regarding the demand for specific data and data aggregates that are needed by individual beings engaged in the transport system functioning (transport system participants, self-governments, local administrative units, etc.). Organising information flows within the area of urban freight transport is a much more difficult task than any similar research carried out with regard to passenger transport. Attempts are made to formulate theoretical bases for determining data streams and their scopes, however, currently there is no universal model based on direct investigation of needs of particular objects of the system. The assumptions presented in this paper constitute an initial stage of work in this area for West Pomeranian Voivodeship.

Acknowledgements. This paper was prepared during the works connected with **C-LIEGE Project** implemented under **Intelligent Energy – Europe** programme, part of the **Competitiveness and Innovation Framework Programme**:

C-LIEGE
Clean Last mile transport and logistics management
Supported by INTELLIGENT ENERGY EUROPE

References

1. Allen, J., Thorne, G., Browne, M.: BESTUFS Good Practice Guide on Urban Freight Transport (2007)
2. Button, K., Stough, R., Bragg, M., Taylor, S.: Telecommunications, Transportation and Location. Edward Elgar Publishing Ltd. (2006)
3. Definition of suitable set of actions/measures for an efficient and energy saving organization of goods transport and delivery in urban areas. C-LIEGE Project Report (2012)
4. Ehmke, J.F.: Integration of Information and Optimization Models for Routing in City Logistics (International Series in Operations Research & Management Science). Springer (2012)
5. National ITS Architecture, `http://www.iteris.com/itsarch/` (date of access: June 19, 2012)
6. Figliozzi, M.A.: The impact of congestion on commercials vehicle tour characteristics and costs. Transportation Research Part E (2010)
7. Taniguchi, E., Russell, G.T., Yamada, T., van Duin, R.: City Logistics. Modelling and Intelligent Transport Systems. Pergamon (2001)
8. Taniguchi, E., Thompson, R.G.: Data collection for modeling and benchmarking city logistics schemes. In: Recent Advances in City Logistics: Proceedings of the 4th International Conference on City Logistics, pp. 1–14. Elsevier (2006)
9. Tseng, Y.-Y., Lin, H.-E., Taylor, M.A.P.: Provision of Real Time traffic information for VRPTW Model. In: Taniguchi, E., Thompson, R.G. (eds.) Innovations in City Logistics, pp. 171–184. Nova Science Publisher, Inc. (2008)

Telematic Systems in Emergency Medical Services, Applied in Treatment of Acute Coronary Syndrome of STEMI Type

Tomasz Ilczak[1,2] and Monika Mikulska[1]

[1]University of Bielsko-Biala, Faculty of Health Sciences
Willowa2, 43-309 Bielsko-Biala, Poland
tilczak@ath.bielsko.pl
[2] Bielsko-Biała Emergency Services Emilii Plater 14, 43-300 Bielsko-Biala, Poland

Abstract. Teletransmission systems play an ever-growing role in the National Emergency Medical Services System. In Poland, commencing from the year 2006, this technology has become the basis for a wide diagnosis and treatment of Acute Coronary Syndrome (ACS) of STEMI type. This paper describes the application of teletransmission systems based on the example of Bielsko-Biała Emergency Services. Various aspects ensuing from functioning of the system in the field of Emergency Medical Services were subjected to analysis. Operating time and healthcare facilities to which patients with acute coronary syndrome were subsequently transported were analysed. The effect of time factor on the prognosis for patients diagnosed with STEMI was assessed. The conclusions drawn confirm usefulness of teletransmission systems in Emergency Medical Services.

Keywords: Emergency Medical Services, Acute Coronary Syndrome, Myocardial Infarction, Telemedicine, Teletransmission.

1 Introduction

The history of application of telematic systems used in healthcare goes back to the year 1985 when in the United States the first medical data transmission in the form of CT scans sent to remote specialised healthcare centres [1] was used. In the Polish healthcare history, teletransmission was used for the first time in 2005 in pilot programmes. In 2006 there were implemented systems of early treatment of Acute Coronary Syndrome (ACS) of STEMI type, taking advantage of telematic methods of diagnosis.

1.1 System Changes in Emergency Medical Services with Reference to Telematic Systems Application

The system of National Emergency Medical Services in Poland has been undergoing dynamic changes and steady development. The system changes affect a number of

J. Mikulski (Ed.): TST 2012, CCIS 329, pp. 87–93, 2012.

elements, starting from medical personnel and ending with application of more and more technologically advanced medical procedures. One of the main changes in the process of reorganization of the National Emergency Medical Services involves introduction of a basic and a specialised unit. The basic unit, in pursuance of the Act on the National Emergency Medical Services of September 8, 2006, comprises at least two persons, meeting the system requirements, to the exclusion of a physician. A specialised unit must include a specialised physician of the specialty specified under the Act referred to above [2]. The range of activities pertaining to a basic unit is extremely broad, while at the same time their possibility to make decisions, such as giving a diagnoses or referring a patient to the specialised unit without a prior consultation is relatively restricted. Under such circumstances, teletransmission teletransmission systems used by EMS basic units take on a whole new dimension.

1.2 Components of Teletransmission System for Emergency Medical Services

While describing possibilities offered by teletransmission, the components without which the application of this method would not be possible should be mentioned. In order to transmit data, a transmitter station is necessary - in this case a defibrillator or a cardiac monitor located at the specialised unit [1]. Proper teletransmission requires appropriate application of all indispensable elements. It is necessary to carry out a heart activity monitoring procedure, set forth under the relevant standard, by analyzing a 12-lead ECG curve. Electrodes must be placed appropriately on the chest and ECG test must be properly performed. The above elements are vital for collecting and transmission of high quality ECG data records [3].

1.3 Teletransmission in Acute Coronary Syndrome

Teletransmission used by Emergency Medical Services finds a particular application in diagnosing the Acute Coronary Syndrome (ACS) in a patient.

The Acute Coronary Syndrome comprises unstable angina, non-ST-segment elevation myocardial infarction (NSTEMI) and ST-segment elevation myocardial infarction (STEMI) [4]. In Poland, approximately 140 thousand patients are hospitalized with ACS, out of which 50 thousand with the final diagnosis of ST-segment elevation myocardial infarction and about 30 thousand with non-ST-segment elevation myocardial infarction [5]. Heart diseases, with coronary heart disease in the lead, make up the shortlist of causes of hospitalization. According to GUS (Central Statistical Office) every ninth adult resident of Lublin, Łódź and Świętokrzyskie Provinces suffers from coronary disease [6]. While examining the condition of an ACS patient, a particular attention should be drawn to the likelihood of STEMI incidence. This type of ACS is characterized by the fact that a proper assessment of vital signs and an analysis of a 12-lead ECG are sufficient to make a full diagnosis.

The basis of treatment of myocardial infarction lies in reperfusion therapy, i.e. primary percutaneous coronary intervention (PCI). The time factor is of particular importance in proper treatment of a patient with ACS of STEMI type. According to the guidelines set forth by the European Society of Cardiology (ESC) of 2005, the maximum length of time from the onset of myocardial infarction to admission to the hemodynamics laboratory to have PCI performed should not exceed 120 minutes [7]. While analyzing the condition of a patient with STEMI, one cannot underestimate the importance of teletransmission in shortening the time from the onset of myocardial infarction to admission to the hemodynamics ward.

2 Materials and Methodology

In order to carry out an analysis of teletransmission usefulness for the purpose of Emergency Medical Services, statistical data provided by Bielsko-Biała Emergency Services (BBES) were used. A retrospective analysis covered medical documentation (EMS Dispatch Cards) between the period from November 2011 and January 2012. Operating time from the moment of accepting the dispatch request until arrival at a medical treatment destination was subjected to analysis. The number of teletransmissions performed by particular EMS teams in relation to the number of patients transported to the hemodynamics ward at the American Heart of Poland in Bielsko-Biała (AHP) was also analysed. This analysis encompasses ambulance dispatch requests within the area of Bielsko-Biała district, responded to by the Emergency Medical Services of Bielsko-Biała Emergency Services (BBES). The whole fleet of BBES ambulances comprises six specialised EMS units and five basic EMS units owned by BBES, located across the whole district.

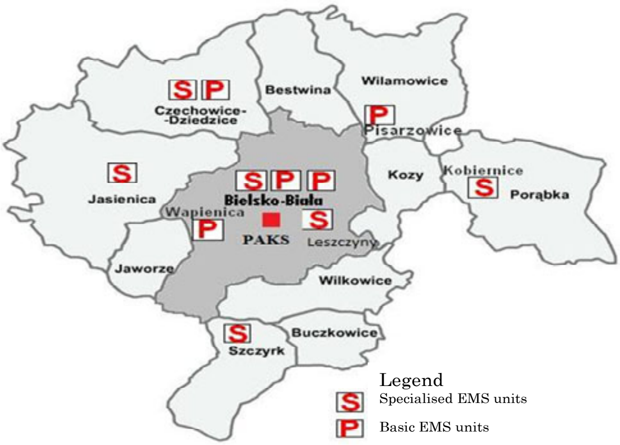


Fig. 1. Location of EMS units in Bielsko-Biała district in 2012

3 Results

On the basis of analysis of the results presented in table 1 it transpires that between November 2011 and January 2012 the number of STEMI ACS incidents accounted for the following number: November 2011- nineteen (27.53%), December 2011-ten

(14.49), January 2012-ten (14.49%), totalling thirty nine (56.52%) of all calls to cardiac conditions in the case of which a decision to perform a teletransmission was made. Other cases account for thirty calls (43.47%). It determines an approximate incidence rate of myocardial infarction, in the treatment of which the fastest possible transportation of a patient to an invasive cardiology ward is an absolute necessity.

Table 1. Calls to cardiac conditions, in the case of which decision to perform teletransmission was made

	Number of patients transported to AHP following teletransmission	Number of patients transported to another cardiology ward following teletransmission	Number of patients in the case of whom teletransmission was performed
November 2011	19 (27.53%)	12 (17.39%)	31(44.92%)
December 2011	10 (14.49%)	13 (18.84%)	23 (33.33%)
January 2012	10 (14.49%)	5 (7.24%)	15 (21.73%)
Total	39 (56.52%)	30 (43.47%)	69 (100%)

Table 2. The number of performed teletransmissions depending on the type of EMS unit

	Number of teletransmissions performed by specialised EMS units	Number of teletransmissions performed by basic EMS units	Total (%)
November 2011	14 (45.16%)	17 (54.83%)	31 (44.92%)
December 2011	10 (43.47%)	13 (56.52%)	23 (33.33%)
January 2012	5 (33.3%)	10 (66.6%)	15 (21.73%)
Total	29 (42.02%)	40 (59.97%)	69 (100%)

The performed analysis (table 2) indicates that basic EMS units support their activities with teletransmissions more often (59,97%) than specialised EMS units (42,02%). The decreasing number of teletransmissions performed between November 2011 and January 2012 signifies seasonal variation in the prevalence of ACS.

The table 3 shows the number of patients in the case of whom, following a teletransmission, a decision on transportation to the appropriate specialised medical facility was made. Out of 69 (100%) of performed teletransmissions, 39 (56.52%) ended up with transportation of a patient directly to the invasive cardiology ward, passing over the emergency cardiology ward. The remaining 30 patients (43.47%) were transported to another cardiology ward. Due to teletransmision, an invaluable time for patients with recent myocardial infarction of STEMI type was saved. None of the patients in the case of whom a decision on teletransmission was made was left at home. It shows that there was a suspicion of ACS of NSTEMI type or unstable angina without changes in ECG records.

Table 3. Transportation of patients to AHP or another emergency cardiology ward

	November 2011	December 2011	January 2012	Total
AHP	19 (27.53%)	10(14.49%)	10(14.49%)	39(56.52%)
Cardiology ward	12(17.39%)	13(18.84%)	5(7.24%)	30(43.47%)
Total	31 (44.92%)	23 (33.33%)	15 (21.73%)	69 (100%)

Table 4. Mean EMS response time to reach a patient and travel time required to transport a patient to the final treatment destination

	Mean EMS ambulance response time to reach a patient [min]	Mean travel time required to transport a patient to the final treatment destination [min]
Period of time from November 2011 until January 2012	6.7 minutes	46.3 minutes

Having analysed the EMS ambulance response time (table 4), it may be stated that reaching the emergency site took on average 6.7 minutes, whereas transporting a patient to the final treatment destination took 46.3 minutes. EMS teams need this time to carry out basic life support activities, such as physical examination, taking medical history, performing a 12-lead ECG, teletransmission and arriving at a specialised medical care facility. The key factor predetermining the time is the distance to be covered from the place of call to the place of treatment. According to the demographic data obtained for Bielsko-Biała district 70% of population live in rural area, far away from hemodynamics ward in hospital. The call received by a dispatcher, implying that a particular patient may be at risk of acute coronary syndrome (ACS), constitutes an indication that the ambulance response should be given the highest priority level.

The table 5 shows the average age of patients, in the case of whom teletransmission was performed. The average age of patients who were not diagnosed for ACS of STEMI type was 69 years, while the average age of patients with ACS other than STEMI, transported to the hemodynamics lab, was 57.6 years.

Table 5. Average age of patients in the case of whom teletransmission was performed related to the age of patients transported to AHP

	Average age in years
Teletransmission performed	69.1
Transportation to AHP	57.6

4 Conclusion

The results of the study illustrate the significance of teletransmission in the system of Polish Emergency Medical Services. The review of literature indicates a multidimensional application of teletransmission [1, 3, 5] with regard to the patient himself/herself and EMS team members. From the patient's perspective, teletransmission brings in a definite improvement in the quality of emergency services, it provides access to specialised consultation at home of the sick person and enables application of specialised treatment already on site. While analysing the incidents from the point of view of the EMS team, performance of teletransmission constitutes an extremely important element of a diagnostic process. The EMS team may avail themselves of specialist knowledge already on site. Analysing the situation further from the perspective of a two-member EMS team, it may be definitely concluded that teletransmission adds a new dimension as it constitutes an additional team member.

Papers [5, 9] indicated a particular importance of teletransmission by emphasizing that time saving is the main advantage of teletransmission, which, in the case of ACS of STEMI type, plays a crucial role. Thanks to the transmission of data there arises a possibility to transport a sick person directly to the facility where he/she will be administered treatment, without wasting time on additional diagnosing at the level of a municipal or district hospital. The data presented by the a.m. authors in their studies stress the fact that due to teletransmission the frequency rate of severe damage to the cardiac muscle diminishes and at the same time mortality rate induced by ACS of STEMI type [1, 7, 8] is also reduced. The present study confirms that the application of this method considerably influences the quality of life of persons who have suffered STEMI, it improves a sense of social safety by ensuring access to appropriate treatment methods in a very short period of time.

References

1. Lifenet system – system teletransmisji ratunkowej. Artykuł promocyjny. Na Ratunek 3(4), 56–57 (2009)
2. Journal of Laws 06.191.1410 Act of September 8, 2006 on National Emergency Medical Services, Art. 36 par. 1 and 2

3. Ilczak, T., Ćwiertnia, M., Mikulska, M., Odrzywołek, R.: Zastosowanie teletransmisji danych EKG w ratownictwie medycznym. Bezpieczeństwo zdrowia publicznego w zagrożeniach środowiskowych. Postępy metodologii badań / red. nauk. Jerzy Konieczny i in. Poznań: Garmond Oficyna Wydawnicza 720(5), s.: il. 23 cm (2012) ISBN 978-83-89250-67-4
4. Opolski, G., Filipiak, K.J., Poloński, L.: Acute Coronary Syndrome, vol. 280, pp. 2–3. Urban & Partner, Wrocław (2002)
5. Karcz, M.: Bezpośredni transport pacjentów z ostrym zawałem serca do pracowni hemodynamiki z pominięciem szpitala rejonowego po teletransmisji EKG z domu drogą telefonii komórkowej. Instytut Kardiologii im Prymasa Tysiąclecia Stefana Kardynała Wyszyńskiego, Warszawa, pp. 7–9 (2008)
6. Central Statistical Office, Department of Social Research, Statistical Publications Centre, Warsaw, Al. Niepodległości 208, p. 49, Warsaw (2011)
7. Guidelines for Percutaneous Coronary Interventions: The Task Force for Percutaneous Coronary Interventions of the European Society of Cardiology. Eur. Heart J. 26, 804–847 (2005)
8. Teletransmisja danych w relacji Zespół Ratownictwa Medycznego – Ośrodek Kardiologii Inwazyjnej. Artykuł promocyjny. Na Ratunek 4(2), 28–29 (2010)
9. Karcz, M., Rużyłło, W.: Bezpośredni transport pacjentów z zawałem serca do pracowni hemodynamicznej z ominięciem najbliższego szpitala po teletransmisji EKG - doświadczenia programu pilotażowego na Mazowszu, Post Kardiol Interw (2005)
10. The European Resuscitation Council, Guidelines for Resuscitation (2010)
11. Wipfler, E.J.: Medycyna Ratunkowa. In: Jakubaszko, J., Plantz, S.H. (red.) Seria Podręczników Medycznych do Egzaminów Testowych (NMS), Urban & Partner, Wrocław (2008)

Profitability and Effectiveness of Transport Investments

Beata Bujna[1] and Jerzy Mikulski[2]

[1] Jastrzębska Spółka Kolejowa Sp. z o.o.
Towarowa 1, 44-330 Jastrzębie-Zdrój, Poland
bbujna@jsk.pl
[2] Silesian University of Technology, Faculty of Transport
Krasinskiego 8, 40-019 Katowice, Poland
jerzy.mikulski@polsl.pl

Abstract. Paper is aimed at indicating a possible combination of economic and technical tools, which will allow an investor active in the transport sector to select the most effective project. Authors explains the terms of profitability and effectiveness and shows the methods for their calculation. The described profitability indices allow – to a various degree and at various designing stages – supporting the decision-making process taking into consideration the risk degree, the investment life or a variable time value of money. Instead, the effectiveness indicators present comprehensive methods for projects assessment taking into account not only financial issues, but also technical or social ones.

Keywords: transport investments, profitability, effecitiveness.

1 Introduction

For the transport investors the choice of the optimum investment decision is crucial, because it results in the company development and primarily in the generation of higher profits in the future. However, a series of analyses related to the investment profitability and effectiveness should be performed and the technical requirements for the specific project should be precisely specified to achieve the expected outcomes from the investment.

The process of selecting the proper remote traffic control system in the transportation, from the profitability and effectiveness point of view, is difficult and requires the investor to analyse numerous areas.

2 Term, Objective and Types of Investments

Businessmen associate today the term of investment with large financial expenditures and a high risk of losing the liquidity. However, the transport investments are the basic condition to ensure a stable and strong market position to the company. As the transport investments are a long-lasting process, the knowledge of issues related to their advisability and to the assessment of investment profitability and effectiveness is essential.

J. Mikulski (Ed.): TST 2012, CCIS 329, pp. 94–106, 2012.

Numerous definitions of ‘investments’ may be found in the literature, e.g. ‘these are expenditures on the manufacture or acquisition of fixed assets [1]. Hence the undertaking of an investment forces the investor to assess its effectiveness and profitability to achieve maximum benefits at an acceptable risk.

A company may invest to accomplish inter alia the following objectives [2]:

- reconstructing the fixed assets or replacing them to cut the production costs,
- work safety and environmental protection.

The assumed investment objective boils down to finding the proper relationships between the expected investment profitability and the risk taken.

A number of investment projects classification criteria can be encountered. Direct investments rank among the most popular used in the business practice [2]:

- reconstructional,
- modernisational (carried out parallel to the reconstruction investments),
- innovational,
- developmental,
- investments related to the working conditions improvement.

The investment decision may be divided into three groups [3]:

- used to reject or approve a specific project,
- related to the selection of specific investment option,
- related to the selection of the most favourable programme of the company development.

There are many evaluation methods for investment projects, supporting the decision-making process and allowing making the right and optimal choices. The most important of them include the assessment of investment projects profitability and effectiveness. In practice also other methods supporting the decision-making process are used, e.g.: the break-even point analysis, the sensitivity analysis or the SWOT analysis.

3 Investment Projects Profitability

The issue of investment projects profitability is one of most important factors conditioning the investment decisions making.

In a popular meaning the profitability stands for the gainfulness, yield, gains generating activity [4]. Therefore for a businessman it is important to determine, whether the capital expenditure is justified and whether with time it will generate income.

It is indispensable to calculate projects profitability in the case, where it is necessary to choose one out of a few projects. This will allow assessing various alternative solutions and selecting the most favourable project, i.e. the most profitable in the given conditions. However, we should consider the fact that the obtaining of the maximum benefits from the actions undertaken is optimal only, if appropriate risk level is taken into account in a predefined time horizon.

The risk is a probability of incurring a loss, related to making a specific business decision. The risk is a measure of uncertainty of results expected in the future as the effect of specific investment [5].

The occurrence of the following problems should be evaluated at the projects profitability analysis:

- identification of effects and expenditures and determination of their occurrence scale and place,
- valuation of effects and expenditures,
- including 'time' in the analysis (effects and expenditures in time),
- determination of intangible effects scale and scope,
- establishing, whether and what benefits will be brought by the investment.

The information about the following quantities is used in the profitability calculation:

- revenue on sales of production or services, treated as the investment effects,
- investment expenditures,
- operating costs,
- net benefits.

The procedure of project profitability assessment may be divided into three basic stages:

Stage I – determination of benefits (revenues) provided by the project implementation. Benefits may be obtained by increasing the sales or by reduction. The project implementation is frequently conditioned by the need to modernise the product or service.

Stage II – estimation of investment expenditures and operating costs after the project completion,

Stage III – calculation and comparison of indices using the profitability assessment techniques.

The implementation of all above stages will allow a proper assessment of the project profitability and the making of the right decision of choosing the most favourable investment option.

4 Techniques for the Investment Projects Profitability Assessment

Methods for the investment projects profitability assessment, from the time point of view, are divided into static and dynamic methods.

4.1 Static Methods

Static methods do not take into account a variable time value of money. This means that individual calculated values are not differentiated in the following years and the total of expected expenditures and effects or their average or values selected from a specific period are introduced to calculations. These methods are suggested to be used in the preliminary stages of the investment preparation process, for projects of a relatively short construction cycle and in the cases of small scale projects. The static methods most frequently used include [5]:

a) Payback Period (PP),
b) Accounting Rate of Return (ARR),
c) Return on Investment (ROI).

a) For the given project the payback period is calculated by determining the number of years, which will pass since the moment when the cumulative anticipated cash flows (the net profit and the annual depreciation) equal the initial expenditures. This is a simple and easy method in use and allows determining the period necessary to recover the expenditures incurred for the investment implementation. The estimation of cash flows before the return of the initial expenditures is very important in this criterion, while the cash flows after the return date are not examined.

If all financial surpluses generated by an investment project are equal, i.e. $CF_1=CF_2= \dots = CF_n$ then [7]:

$$PP = \frac{C_0}{CF_t} \tag{1}$$

where:

PP – the payback period,
C_0 – the investment expenditure,
CF_t – the financial surplus in year t (net profit + depreciation)

If the surpluses from the following periods are different, i.e. $CF_1 \neq CF_2 \neq \dots \neq CF_n$ then [7]:

$$PP = (T-1) + \frac{C_0 - \sum_{t=1}^{T-1} CF_t}{CF_T} \tag{2}$$

where:

T – the year, when the sum of subsequent financial surpluses covers the value of expenditures incurred,
t – 1, 2, … n – the consecutive years of the calculation period.

If the investor has determined the limit payback period, projects for which this limit is exceeded will be unacceptable.

b) The accounting rate of return is expressed in per cents and is determined by the ratio of average net profit (assumed for the whole project duration) and the amount of total expenditure (the capital employed). The accounting rate of return may be expressed using the relationship [7]:

$$ARR = \frac{\overline{Z}n}{N} \cdot 100 \tag{3}$$

where:

ARR – the accounting rate of return,
 - average annual generated profit,
N – the amount of invested owner's equity.

The higher the return on investment, the more favourable is the investment project.

c) Return on investment is a profitability index, which is used to examine the economic effectiveness of the project. The return on investment is a quotient of the operating profit and the wealth understood as the total assets. It may be presented by the formula [7]:

$$ROI = \frac{\dfrac{\sum_{t=1}^{n} CF_t - C_0}{n}}{C_0} \tag{4}$$

where:

ROI – the return on investment,
C_0 – the investment expenditure,
CF_t – the financial surplus in year t (net profit + depreciation)
n – the investment duration in years,
t – 1, 2, … n – the consecutive years of the calculation period.

The rate of return provides the information on the part of invested capital returned within one year as a profit. Knowing the rate of return it is possible to calculate the time needed for the predicted profits to return the invested amount.

The advantages of static methods use for the assessment of investment projects profitability comprise their simplicity, clarity of formulae, accessibility and the ease of the obtained results interpretation. Basic drawbacks of these method include the lack of variability of time value of money, the uncertainty of future income, as the expected value, and the disregarding of the inflation.

4.2 Dynamic Methods

Dynamic methods take into account irregular cash proceeds and expenditures during the investment duration and are considered the most precise tools for the project profitability assessment. The indices take into account not only the predicted cash flows but also the expenditures and effects distributed in time. Dynamic techniques allow bringing to comparability the expenditures and effects from the entire calculation period, expressing them in the present value, i.e. updated to the moment of assessment. The assessment applies to the whole period of project operation, that is the implementation period and the full period of investment operation [1].

The dynamic methods most frequently used include [5]:

a) Discounted Payback Period (DPP),
b) Net Present Value (NPV),
c) Internal Rate of Return (IRR),
d) Profitability Index (PI).

a) The Discounted Payback Period is calculated like the payback period, but before calculating the payback period the cash flows are discounted. This index provides the information about the investment duration so that the present value of net cash flows would equal zero. It is expressed by the relationship [7]:

$$DPP = (T-1) + \frac{C_0 - \sum_{t=1}^{T-1} \frac{CF_t}{(1+r)^t}}{\frac{CF_T}{(1+r)^T}} \tag{5}$$

where:

DPP – the Discounted Payback Period,

T – the year, when the sum of consecutive financial surpluses discounted to the moment of the expenditure incurring covers the value of expenditure incurred,

t – 1, 2, … n – the consecutive years of the calculation period,

C_0 – the investment expenditure,

CF_t – the financial surplus in year t (net profit + depreciation),

r – the discount rate.

It should be remembered that the principle of discounted payback period for the investment expenditures does not take into account the cash flows after the cut-off date and provides only the information about the expenditures liquidity.

b) The Net Present Value allows determining the present value of cash proceeds and expenditures related to the assessed project implementation and is based on the following findings [1]:

- a “zloty” today is more valuable than a “zloty” tomorrow,
- this value depends only on the cash flows predicted in the future and on the opportunity cost,
- the net present value of various projects is measured by the current (present) value.

Such projects are considered favourable, for which the net present value has values higher than zero. This means that the profitability rate of the examined project is higher than the limit rate, which is determined by the discount rate adopted for the calculations. The degree of predicted investment expenditure covering with expected net financial surpluses, after their discounting, may be determined as follows [1]:

$$NPV = \sum_{t=1}^{n} \frac{CF_t}{(1+r)^t} - C_0 \qquad (6)$$

where:

NPV – the Net Present Value,
C_0 – the investment expenditure,
CF_t – the financial surplus in year t (net profit + depreciation),
r – the discount rate.
t – 1, 2, … n – the consecutive years of the calculation period.

The discount rate in the formula expresses the minimum profitability rate of the capital employed in the investment – the opportunity cost. It reflects the time equivalent of money and also the risk related to the project implementation.
c) The Internal Rate of Return is a discount rate, at which the value of cash expenditures is equal to the cash receipts values, i.e. this is such a discount rate, at which the Net Present Value (NPV) is equal zero, so the following equation is satisfied [7]:

$$NPV = \sum_{t=1}^{n} \frac{CF_t}{(1+IRR)^t} - C_0 = 0 \qquad (7)$$

The internal rate of return is calculated by means of a linear interpolation method using the relationship [7]:

$$IRR = r_1 + \frac{PV \times (r_2 - r_1)}{PV + |NV|} \qquad (8)$$

where:

IRR – the Internal Rate of Return,
r_1 – the discount rate, at which the NPV is higher than zero,
r_2 – the discount rate, at which the NPV is lower than zero,
PV (positive value) – a positive NPV, calculated at r1,
NV (negative value) – a negative NPV, calculated at r2,

In reality, the search of internal rate of return consists in finding such a discount rate, for which the net present value is equal zero. However, this is a very arduous practice and requires substituting consecutive values of discount rates to the formula. The use of a spreadsheet may be considered another method for the internal rate of return determination.

d) The Profitability Index is a relationship between the project net present value and its expenditures. It is recommended to approve all projects, for which the profitability index is higher than one. The index is calculated according to the relationship [7]:

$$PI = \frac{NPV}{C_0} \tag{9}$$

where:

PI – the Profitability Index,
NPV – the net present value calculated for the given project,
C_0 – the investment expenditure,

The profitability index determines a relative earning capacity of the project and provides the information on the amount of profit brought by each monetary unit invested in the project.

The description of the discount methods presented above shows their numerous advantages helpful in making the project-choice decisions. However, it is necessary to remember that they also have drawbacks. The use of indices cannot be limited only and solely to mathematical calculations. A proper choice and a mitigation of the investment risk is possible by the use of a few criteria at the same time [1].

5 Investment Projects Effectiveness

In a general sense the effectiveness means that something is efficient and gives goods results [4]. It should be emphasised, however, that something effective is not always profitable. From the economic point of view, the assessment of investment project effectiveness means a result of undertaken actions, described by a relationship of obtained effects to the incurred expenditures [6]. Such projects are cost-effective, which guarantee a profit, while the most effective projects consist in the maximisation of the profit obtained from the capital employed [2].

The investment effectiveness assessment is carried out from various points of view:

- economic – a relation of the investment operation to the incurred expenditures is measured,
- technical – technical efficiency indicators of the new equipment are compared to the indicators of the equipment or technologies used so far,
- social – the relationship of all the cash and non-cash expenditures on the given investment to all the cash and non-cash effects obtained from the investment operation.

The effectiveness is measured using partial synthetic indices of productivity and of resources utilisation. The effectiveness may be identified on an ex-post (determination of specific actions results) and ex-ante (determination of predicted effects) basis [4].

From a material point of view, the effectiveness analysis for a specific project may be compared to a feasibility study performance. In practice this means an assessment of a few separate examination areas for the project (Table 1).

Table 1. Groups of criteria necessary to assess an investment project. Source: [own study].

Group of criteria	Detailed criteria
Financial	Profitability indices
Formal and legal	Strategic objectives accomplishment. Compliance with the land development plan
Social	Time saving Safety improvement Impact on the employment Influence on the economic standing
Environmental	Ecological impacts Acoustic impact areas Air pollution
Technical	Geological-engineering conditions Functionality and usefulness
Transportation	Bottlenecks removal Traffic safety improvement

It is noticeable, based on the table, that a feasibility study performance for an investment project allows assessing the project in a comprehensive way, both from the profitability and effectiveness point of view. So a precise specification of project assessment criteria is crucial. A decision, which is the most favourable in economic terms, may be not very favourable in social or political terms, and vice versa.

The choice of such product manufacture technology, which will ensure substantial profits and will feature a short payback period, but at the same time will result in a major environmental pollution, may be an example.

The calculation, which comprehensively takes into account financial, social and technical aspects of a specific project, captures the investment effectiveness in the fullest way. The investment effectiveness considered this way, combined with the performed profitability analysis, will allow the investor to make the right investment decisions [8].

6 Techniques for the Investment Projects Effectiveness Assessment

The performance of projects effectiveness assessment, like in the case of profitability assessment, is significant from the investor's point of view. The assessments are performed for comparative and decision-making objectives.

Multi-criteria evaluation methods allow analysing economic or material aspects of the project. These methods are carried out in two systems [3]:

- weighted,
- zero-one.

A multi-criteria method in a weighted system consists in assigning weights to criteria defined by the evaluator or by the investor. The assigned weights reflect the investors priorities. The total of all criteria weights should be equal 1.

Table 2 presents an example of a sheet for a multi-criteria evaluation in a weighted system, which is aimed at selecting the most effective project from a number of proposals.

The project, which will obtain the highest total weighted score, is considered the most effective by the evaluator. In decision making this method is frequently used as a complementary method, especially in the cases where the determined profitability indices for a few projects are similar. In the discussed method it is important that it is the investor, who specifies the project assessment criteria for the areas crucial from his point of view and that he assigns their importance in the form of weights.

A multi-criteria method in a zero-one system consists in the project approval or rejection in relation to the adopted criterion. This is a supporting system and it is used in the initial stage of projects assessment to eliminate the weakest projects.

In practice, simple methods are most often used for projects assessment. Deeper analyses, evaluating economic aspects of a project, are performed in the case of long-term projects related to large financial expenditures. The discount methods for the profitability assessment are then used, which evaluate the degree of risk at making the decision on a project implementation.

Table 2. A sheet for a multi-criteria evaluation in a weighted system for various projects (example weights). Source: [3].

Evaluation criteria	Weight	Project A		Project B		Project N	
		Evaluation	Weighted evaluation	Evaluation	Weighted evaluation	Evaluation	Weighted evaluation
Financial capacities	0,25						
Technical-production potential	0,20						
Company prestige	0,10						
Marketing activities	0,10						
Company location	0,10						
Management experience and managerial predispositions	0,10						
Staff qualifications	0,15						
Total:	**1,00**						

An example of feasibility study performance schedule for a given project, presented in Fig. 1, shows how many diverse issues are analysed and affect the decision on a project approval for implementation or rejection.

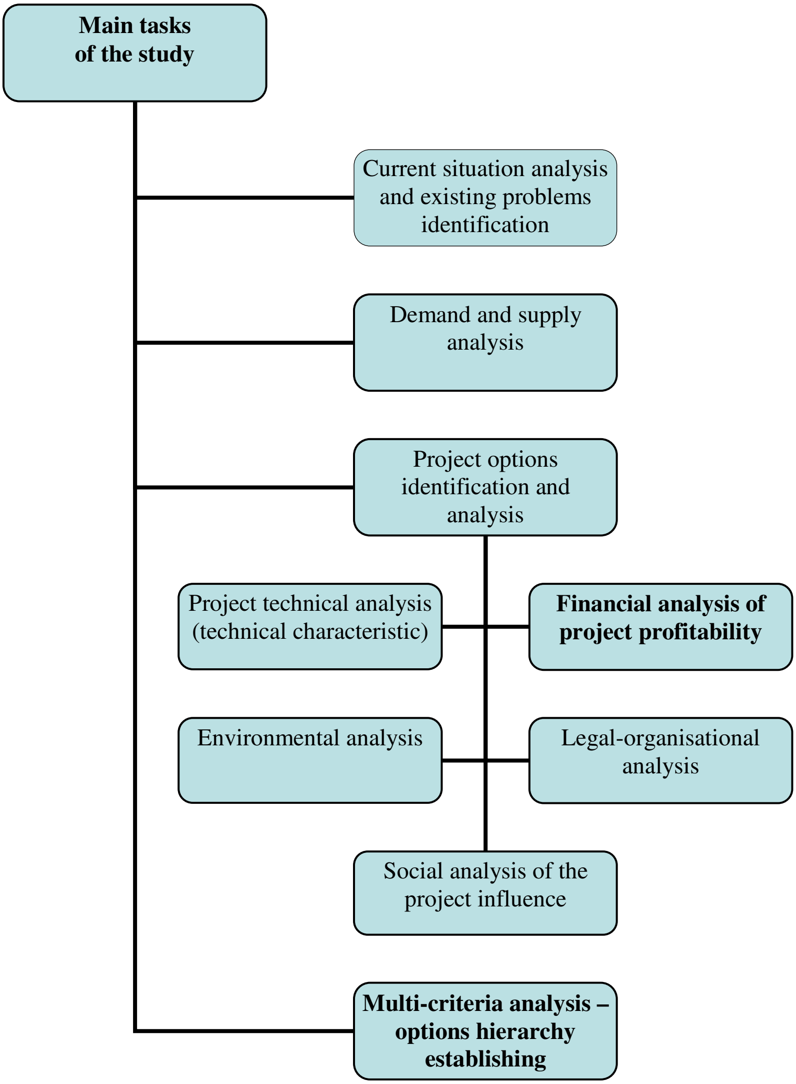


Fig. 1. Feasibility study layout. Source: [own study].

7 General Assumptions of Remote Control and the Requirements Accompanying the Installation of Railway Traffic Control Systems and Equipment

A set of equipment, which enables controlling the traffic from one signal box in the area of a few control sections may be referred to as a remote railway traffic control [9]. Remote control systems belong to the master layer of systems responsible for the railway traffic control. The movements inspector may be situated at a large distance, even a few dozen kilometres, from the operated equipment. The control orders and the messages about the equipment state and the traffic situation must be quickly and reliably transmitted between the staff, the control panel and the field equipment [9].

For poorly loaded lines the remote control allows a drastic cutting of the fixed costs of traffic maintaining on such lines [10]. The remote control is executed by special pieces of equipment (control offices) installed in stations and by the equipment situated in the control centre. The equipment ensures a full visualisation of the traffic situation in the area covered by the remote control and a possibility of controlling the equipment situated in this area from the centre.

Modern railway traffic control systems allow [11]:

- very simple itinerary control,
- remote control of traffic posts,
- control of track non-occupation with simultaneous train identification,
- automatic itinerary control,
- support of the traffic management, especially in emergency situations,
- management of a large number of executing devices from one signal box.

However, it is necessary to bear in mind that the remote control systems are most frequently installed in new stations or in such that have been comprehensively reconstructed. Because of that the majority of systems have been adapted to work with widely used relay and electromechanical systems. The outdoor devices like the point mechanisms, track circuits, and signals remained practically unchanged.

Because of the way of executing devices control the railway traffic control systems may be divided into [11]:

- Fully computerised systems,
- Hybrid systems using the relay-computer technique.

When considering a general characteristic of railway traffic remote control systems it is necessary to remember about requirements and conditions of the discussed systems installation.

The basic rules for the remote control equipment installation include [9]:

- the area of remote control of trains traffic should comprise the area of small and medium-size stations, groups of large stations tracks and junction signal boxes,
- the traffic will be managed in the area by one section movements inspector working in the remote control signal box,
- the remote area size should take into account perceptive capabilities of the movements inspector, the analysis of the duration and of the scope of activities,
- traffic posts in the part assigned to the train traffic must be equipped with electric, relay or electronic interlocking,
- tracks within the area and the boundary tracks must be equipped with non-occupation control devices or feature an ensured rear of a train control,
- a remote-control area should be provided with communication: with train drivers, with permanent-way gangs,

- a remote control signal box may be situated at any place,
- the maintenance of the track system and of the interlocking must guarantee the fulfilment of reliability criteria.

The installation of remote railway traffic control systems is a favourable project and having many advantages from the railway traffic management point of view, but it is also a costly project from the railway infrastructure manager point of view. It is assumed that the remote control systems should cause a reduction of train travel time, so their installation is important for an infrastructure user (carrier). However, it is necessary to bear in mind that the cost of infrastructure maintenance, costs of major repairs and of replacements translate directly into the rate of the railway infrastructure access to carriers. In practice, the remote control should also increase the capacity of a railway line, resulting in obtaining by the infrastructure managers of higher revenues on the railway line access. Taking into account the economic aspect, also the fact of downsized personnel of traffic posts (in some cases significant) supports the installation of remote control systems, which translates into savings in the payroll costs. However, the investment projects related to the installation of the discussed systems, due to their cost-intensive nature, should be carefully analysed in terms of project profitability and effectiveness.

References

1. Wiśniewski, T.: Analiza finansowa w przedsiębiorstwie (Financial Analysis in an Enterprise). Fundacja Rozwoju Rachunkowości w Polsce, Warszawa (1997)
2. Sierpińska, M., Jachna, T.: Ocena przedsiębiorstwa według standardów światowych (Company Assessment According to the World Standards). PWN, Warszawa (1996)
3. Czerska, J.: Metody oceny efektywności projektów inwestycyjnych (2002), http://www.zie.pg.gda.pl/~jcz/anal_efekt_inwest.pdf (June 15, 2012)
4. Arnold, A.: Słownik języka polskiego (Dictionary of Polish). Wydawnictwo Parki, Bielsko-Biała (2007)
5. Towarnicka, H., Broszkiewicz, A.: Przygotowanie i ocena projektów inwestycji rzeczowych (Preparation and Assessment of Tangible Investment Projects). The O. Lange Academy of Economics, Wrocław (1994)
6. Encyklopedia Zarządzania (Encyclopedia of Management) (June 09, 2012), http://mfiles.pl/pl/index.php/Strona_g%C5%82%C3%B3wna
7. http://www.kms.ae.katowice.pl/PL/dydaktyka/pchrzan/pliki/Metody-dys-cz2.pdf (June 29, 2012)
8. http://suchozebrski.pl.pl/attachments/004_7.%20Rachunek%20efektywnosci%20inwestycji.pdf (June 29, 2012)
9. Dąbrowa-Bajon, M.: Podstawy sterowania ruchem kolejowym (Basics of Railway Traffic Control). Publishing House of the Warsaw University of Technology, Warszawa (2007)
10. Rydziński, P.: Raport "Sterowanie ruchem kolejowym" (Railway Traffic Control). Rynek Kolejowy (3), 10–18 (2004)
11. Dyduch, J., Kornaszewski, M.: Systemy sterowania ruchem kolejowym (Railway Traffic Control Systems). Radom University of Technology, Radom (2003)

Installation of Rail Traffic Remote Control Systems in Terms of Profitability and Investment Effectiveness

Beata Bujna[1] and Jerzy Mikulski[2]

[1] Jastrzębska Spółka Kolejowa Sp. z o.o.
Towarowa 1, 44-330 Jastrzębie-Zdrój, Poland
bbujna@jsk.pl
[2] Silesian University of Technology, Faculty of Transport
Krasińskiego 8, 40-019 Katowice, Poland
jerzy.mikulski@polsl.pl

Abstract. Paper is aimed at analyzing the choice of rail traffic remote control system from the profitability and effectiveness point of view. The analysis was performed for the area of an example railway operator.The paper defines the basic groups of technical criteria for the choice of remote control system and also discusses systems of various manufacturers available on the Polish markets. Investors choose the most effective systems from those offered, according to their needs, technical conditions and financial capabilities. The paper assumes that a project, which is profitable and meets the technical criteria, is considered effective. Article is completed with an example of a practical application of profitability and effectiveness study techniques combined with technical criteria for the installation of rail traffic remote control system within a defined railway area.

Keywords: transport investments, profitability, effecitiveness.

1 Introduction

The choice of optimum investment decision is crucial for the investors, because it leads to a company development. However, precise technical requirements for the given project should be specified to achieve the expected results from the investment.

The process of selecting the proper rail traffic remote control system, from the profitability and effectiveness point of view, is difficult and requires analysing numerous areas: economic-financial, technical and frequently environmental or social areas. The study methods and techniques presented in the paper to a large extent facilitate this process and allow making the most effective choices.

2 Technical Criteria of Selecting the Type of Rail Traffic Remote Control System

The choice of appropriate remote control system for a given railway area depends on many factors. Four groups of criteria may be distinguished, providing the basis for a proper decision on the system selection:

J. Mikulski (Ed.): TST 2012, CCIS 329, pp. 107–116, 2012.

Group I – Traffic parameters criteria,
Group II – Technical criteria,
Group III – Functional criteria,
Group IV – Quality criteria.

Group I – the basic criterion differentiating the system scope and type consists of planned basic traffic parameters and the fact, whether the railway area appears as a newly constructed or modernised. The basic traffic parameter affecting the system type and hence equipping with the interlocking systems is the running speed limit on the lines (section) under construction or modernisation. Three ranges of speed are distinguished, differing in the required equipment with the interlocking systems [3]:

- up to 160 km/h – effective for the existing or modernised lines, which for various reasons cannot be adapted to higher speeds,
- to 200 or 250 km/h,
- to 300 or 350 km/h – for majority of newly constructed or reconstructed lines, for which the equipment must meet the requirements for the high speed railways system.

Group II – technical criteria, comprising relationships, which should be fulfilled for efficient and safe control of train traffic and of manoeuvres.

Table 1 presents the interlocking systems and groups, with which the work and communication is necessary [3].

Table 1. Technical criteria for interlocking systems and equipment dependence for the work with a rail traffic remote control system. Source: Own study.

System or device specification	Dependence description
Station interlocking (dependence systems)	Computer hardware or a relay device equipped with control panels or interfaces for the work with a master system. Devices should meet appropriate detailed technical requirements, adapted to a specific facility and planned parameters.
Block systems (line interlocking)	Possibility of both-directions working. Up to 160 km/h – automatic block system adapted to work with traffic management systems chosen for the line and with a remote diagnostics system and equipped with recorders of equipment states and traffic incidents. Above 160 km/h and on newly constructed sections – the ERTMS/ETCM system required (level 2); at the traffic of vehicles not equipped with ERTMS/ETCS equipment additionally and automatic block systems, like for speeds up to 160 km/h.
Track non-occupation control systems	The tracks non-occupation control should be executed by jointless track circuits or axle counters, observing the electromagnetic compatibility requirements, in particular the resistance to disturbances generated by the high-speed rolling stock.

Table 1. (*continued*)

Point mechanisms and switching systems	In the main basic tracks – non-trailable point mechanisms. In the main additional tracks – trailable point mechanisms of the holding force required by the regulations, depending on the point switch design, type of interlocking and switching systems used.
Traffic protection equipment on level crossings	The way of roads and pedestrian level crossings with the railway tracks protection should comply with relevant regulations. For speeds above 160 km/h and for newly constructed lines the crossings should be two-level ones. For trains running up to 160 km/h category A and B level crossings are allowed.
Signalling systems	Lines should be equipped with ATP class systems (signal repetition on the locomotive) of ERTM/ETCS train control. Depending on the planned speed limit it is justified to provide with the track-side signals (light signals). The light signals must protect each traffic posts and also stations. The driving acc. to track-side signals may occur at a maximum speed up to 160 km/h, above this speed the driving should be carried out only according to the guidance of on-board instruments for the train control.
Diagnostic systems	The diagnostic system should be adapted to work via interfaces with diagnostic subsystems of individual interlocking systems and devices. It must feature a modular design with a possibility of configuration change depending on the number of monitored objects.
Fire and burglar signalling system	The system should be adapted to a possibility of transferring the information about the existing fire or burglar hazard in objects and equipment working in the area covered by the remote control.
Control station	The computer hardware must satisfy requirements related to the traffic situation and the equipment state visualisation. It should feature a high reliability, a possibility of connecting in networks, a modular design, a possibility to connect with specialised equipment. Appropriate number of interlocking computers.

Group III – functional criteria related to the installed system operation and its maintenance. In this criterion the important factors are related to:

- guarantee servicing by the system manufacturer (the guarantee duration),
- repair and replacement of system subassemblies (subassemblies availability, anticipated subassemblies production time),
- time of response to a defect notification (service location),
- possibility of training for the staff responsible for the system maintenance.

Group IV – quality criteria are strictly related to the obtaining for a specific system the certificate of permanent operation permit, which shows the system users its reliability and observation of appropriate level of traffic safety.

The description of individual groups of criteria shows the complexity of the decision-making process for the remote control system choice both for the newly constructed and modernised lines.

3 Types of Rail Traffic Remote Control Systems Available on the Polish Market

In contemporary conditions an efficient, fast and safe railway requires such rail traffic control systems, which on the one hand – using appropriate documents or signals – provide the vehicle driver with the information about the driving speed and on the other hand enable preparing the routes in a coordinated and collision-free way [1]. The railway managers, in accordance with the needs, technical conditions and financial capabilities, are pursuing modernisation of obsolete rail traffic control systems using the offers of companies present on the Polish market, but also offers of foreign companies, which try to enter the Polish market.

The Bombardier Transportation (ZWUS) Katowice is the best known and experienced company in the rail traffic control sector on the Polish and European market. This company offers a wide range of products starting from modern remote control systems, station computer hardware through block systems and up to automatic level crossing signaling.

The Kombud S.A. Zakład Automatyki from Radom is another manufacturer of interlocking systems known on the Polish market.

Systems of other (foreign) manufacturers are also used in Poland, e.g. the Siemens Sp. z o.o. system or the ALCATEL SEL AG system.

The first of them is a secure microcomputer system, which can be used for small and large areas, for local and main-line services. The latter is a microprocessor switching system.

Because of a great technical progress newer and newer solutions appear in the interlocking systems sector. This applies both to Polish and also to foreign companies. The system of a Czech company PrvniSignalni is an example, which in its development assumptions will be using a wireless (radio) communication to execute the remote control function. Such a solution will allow obtaining substantial savings in the field of cable lines construction, modernisation or maintenance costs. Moreover, it is very important in the case, when the railway infrastructure is situated in the area of active mining subsidence influence. The very cable lines are the most frequently exposed to continual breaking, resulting in systems failures. Cable lines are also very often exposed to theft.

4 Matrix of Technical – Economic Aspects Relationship for the Installation of Rail Traffic Remote Control Systems – Options Hierarchy Setting

The specification by the investor of the investment objective and the investment type is the necessary basis to establish relationships between technical and economic aspects and to hierarchise the options of remote interlocking systems installation. This will allow establishing at the beginning the basic interlocking parameters of technical or economic criteria importance. The assumption is that the investor selects an effective project, i.e. such which is profitable and meets the technical criteria. The assumptions for individual technical-economic relationships are determined by the investor and broken down to those that must be fulfilled unconditionally and those, which execution is optional.

To assess and choose the most profitable project the profitability examination methods are selected and basic parameters are established, which must be fulfilled unconditionally and then optionally [2].

Table 2. Technical criteria evaluation sheet for four example Projects of rail traffic remote control system installation. Source: Own study.

Criteria	Project I	Project II	Project III	Project IV
UNCONDITIONAL CRITERIA **Group I – traffic parameters**				
Running speed limit - up to 160km/h	1	1	1	1
Group II – technical criteria:				
Computer hardware with interfaces to work with the master system	1	1	1	1
Track non-occupation control executed using axle counters, observing the electromagnetic compatibility requirements	1	1	1	1
Electronic system works and enables using the existing external devices	1	1	1	1
Work with the signalling system	1	1	1	1
Work with the traffic protection equipment on level crossings	1	1	0	1
System diagnostics	1	1	1	1
System operating in the “2 of 3” dependence system	1	1	0	1
Possibility of radio communication	1	0	0	0
OPTIONAL CRITERIA **Group III – functional criteria**				
1 year guarantee period	0	1	0	1
Easy access to components and anticipated period of their manufacture	1	0	1	1
Service location	1	1	0	0
Personnel training	1	1	1	1
Group IV – quality criteria				
System holds a certificate of permanent operation permit	0	1	1	1
It is allowed to carry out field tests to obtain a certificate of permanent operation permit	1	0	0	0

1 means that the specific parameter is fulfilled by the project
0 means that the specific parameter is not fulfilled by the project

There is a similar procedure for the technical analysis, specifying technical parameters, which must be absolutely fulfilled and such, which execution may be optional.

To illustrate the procedure, four different Projects[1] were subject to an example assessment.

Table 2 presents an example of an evaluation sheet in terms of technical criteria for four projects, which will allow eliminating projects not meeting the unconditional criteria and further on which will allow determining the optional criteria fulfilment.

Groups of formerly described criteria were used in Table 2, which will differ depending on the investor needs, the investment objective and the investment type.

Table 2 shows that example Projects I, II, and IV will be subject to further profitability analysis.

Table 3 presents an example of projects evaluation sheet in terms of investment profitability parameters, determined and indicated as unconditional and those, which may be optional. Profitability assessment techniques formerly described were used in the table [2], which will differ depending on the investor needs, the investment objective and the investment type.

Table 3 shows that two projects fulfil the unconditional technical and economic criteria. The selection of the most effective project may be performed using a multi-criteria method in a weighted system, in which optional technical-economic criteria are analysed for projects satisfying the unconditional relationships, assigning them importance in the form of weights. Table 4 presents an example sheet for multi-criteria evaluation in a weighted system. The weights assigned for individual optional parameters reflect priorities of the given investor. The total of all criteria weights should be equal 1. The most effective project is that, which will obtain the highest total score of weighted evaluation. Projects I and II will be evaluated in Table 4.

Table 3. Profitability parameters evaluation sheet for three example Projects fulfilling unconditional technical criteria for rail traffic remote control system installation. Source: Own study.

Criteria	Project I	Project II	Project IV
UNCONDITIONAL CRITERIA: Static methods:			
ARR for the investment is higher or equal to a limit rate, determined based on the market interest rate	1	1	1
Dynamic methods:			
NPV should be higher than zero	1	1	0
PI value should be higher than one	1	1	0
OPTIONAL CRITERIA: Static methods:			
Investment payback period is shorter than 5 years	1	0	0
Investment payback period is shorter than 7 years	0	1	0

1 means that the specific parameter is fulfilled by the project
0 means that the specific parameter is not fulfilled by the project

[1] Systems names are not important here, although the names and their manufacturers are known to the paper authors.

Table 4 shows that, for the presented example, Project I is the most effective, which implies that Project I is also profitable and meets technical criteria assumed in the example.

5 Analysis of Options of the Rail Traffic Remote Control System Selection for the Identified Railway Area in Terms of Project Profitability and Effectiveness

Possibilities to finance an investment consisting in the modernisation of traffic control equipment via the installation of a rail traffic remote control system were determined based on the current and predicted economic-financial results as well as the investment profitability criteria were determined using the static and dynamic methods. These assumptions and the other parameters necessary do determine the profitability indices are presented in Table 5.

Table 4. Multi-criteria in a weighted system evaluation sheet for two Projects fulfilling unconditional technical-economic criteria for a rail traffic remote control system installation. Source: Own study.

Evaluation criteria	Weight	Project I		Project II	
Optional technical criteria:		Evaluation	Weighted evaluation	Evaluation	Weighted evaluation
1 year guarantee period	0,07	0	0,00	1	0,07
Easy access to components and anticipated period of their manufacture	0,10	1	0,10	0	0,00
Service location	0,05	1	0,05	1	0,05
Personnel training	0,08	1	0,08	1	0,08
System holds a certificate of permanent operation permit	0,10	0	0,00	1	0,10
It is allowed to carry out field tests to obtain a certificate of permanent operation permit	0,10	1	0,10	0	0,00
Optional profitability criteria:					
Investment payback period is shorter than 5 years	0,30	1	0,30	0	0,00
Investment payback period is shorter than 7 years	0,20	0	0,00	1	0,20
TOTAL:	1,00	x	0,63	x	0,50

1 means that the specific parameter is fulfilled by the project
0 means that the specific parameter is not fulfilled by the project

Profitability indices were calculated for parameters assumed in Table 5, using dynamic and static methods.

Evaluation using dynamic methods:

1. Net Present Value (NPV) [calculated acc. to formula (**6**)] at the assumed discount rate of 9.26% is approx. PLN 54,000,
2. Internal Rate of Return (IRR) [calculated acc. to formula (**9**)] is 9.4%.

Formulae for the calculation are provided in [2].

Table 5. Basic data and assumptions necessary to determine the value of investment consisting in the modernisation of the interlocking and in the installation of a rail traffic remote control system. Source: Own study.

Specification	Value, PLN million
Defined maximum value of the investment	15
Maximum own contribution	10
Maximum value of a long-term loan	5
Forecast of annual revenue on sales is based on the historic data and on the analysis of demand for services in further years	Around 26
Cost forecast is based on the historic data	Around 97% of total revenue on sales
Investment depreciation calculated from January 2017 (after the investment completion). The other replacement costs of the investment in the railway infrastructure determined based on the forecast	Approx. 3 million a year
Assumed period of project forecast	5 years

Figures in the table are examples

Evaluation using static methods:

1. Payback Period (PP) [calculated acc. to formula (**1**)] is 3.5 years.
2. Accounting Rate of Return (ARR) [calculated acc. to formula (**3**)] is 33%,
3. Return on Investment (ROI) [calculated acc. to formula (**4**)] is 14.6%.

Formulae for the calculation are provided in [2].

Assuming the achievement of all parameters adopted in Table 5, for the investment of around PLN 15 million (unconditional criterion), the project will be profitable and cost-effective and the company, for the profitability indices assumed in this way, in the future period will feature liquidity, stability and solvency.

Therefore, to consider the project for the determined area effective, its value should not exceed PLN 15 million and it should fulfil the technical criteria specified in the evaluation sheet in Table 2. The value of each of both offers costs does not exceed PLN 15 million.

6 Summary

It results from the information contained in Tables 2 and 4 that only one offer – project I – satisfies all unconditional technical criteria and a majority of optional criteria. The system does not hold a certificate of permanent operation permit. For the presented considerations a hypothetic assumption was made that the manufacturer carries out tests in the field of the radio communication utilisation in the interlocking control. Such a solution seems to be the optimal. A system based on the radio communication would allow achieving substantial savings in the field of cable pathways construction and maintenance. However, the selection of this system would result in the investment elongation by the time of field tests duration.

Despite the fact, that not all the unconditional criteria were met, it is also possible to consider the acceptance of Offer II, i.e. a system, which holds a certificate of permanent operation permit. But this system is not working with the existing external devices. However, had the company continued considering both projects, it should carry out again a multi-criteria evaluation for them, in which – for specific technical parameters – the weights will be assigned in accordance with the company priorities. A suggestion of assigning weights, acc. to the company assessment, to individual criteria is presented in Table 6.

Table 6. A sheet for multi-criteria evaluation in a weighted system for both offers for the installation of a rail traffic remote control system in the railway area of the company. Source: Own study.

Evaluation criteria	Weight	X		Y	
Functional criteria		Evaluation	Weighted evaluation	Evaluation	Weighted evaluation
Possibility of modular installation	0,10	1	0,10	1	0,10
1 year guarantee period	0,07	1	0,07	1	0,07
Determinable predicted period of components manufacture	0,05	1	0,05	1	0,05
Easy access to various manufacturers components	0,08	1	0,08	0	0,00
Close location of the service	0,07	1	0,07	0	0,00
Personnel training	0,08	1	0,08	1	0,08
Qualitative criteria					
System holds a certificate of permanent operation permit	0,30	0	0,00	1	0,30
It is allowed to carry out field tests to obtain a certificate of permanent operation permit subject to indication premises of system innovation	0,25	1	0,25	0	0,00
TOTAL:	1,00	x	0,70	x	0,60

1 means that the specific parameter is fulfilled by the project

0 means that the specific parameter is not fulfilled by the project

The table shows that the company should accept Offer I, which according to specified technical-economic criteria will be most effective. As the offer predicts field tests, the installation and commissioning of the equipment must be carried out based on relevant agreements and contracts between the interested parties.

7 Conclusion

At present a fast and primarily safe railway is expected. Such parameters may be achieved inter alia by the installation of rail traffic remote control systems. As these are very expensive investments, it is necessary to analyse carefully the area, in which the system is to operate, to specify expectations related to technical parameters and primarily to determine the financial capabilities of the entity, which will become the investor.

Investments related to the installation of rail traffic remote control systems are long-term investments. So for the investor it is important to determine the investment profitability and effectiveness under changing economic-financial or organizational-legal conditions. The undertaken investment risk should adequately balance the technical needs, therefore it is important to carry out profitability and effectiveness analyses for the undertaken projects.

The selection of the most effective project of rail traffic remote control system installation, fulfilling all the profitability parameters and ensuring the implementation of technical conditions, is a difficult but necessary task for the investor.

References

1. Rydziński, P.: Report.: Sterowanie ruchem kolejowym (Railway Traffic Control). Rynek Kolejowy (3), 10–18 (2004)
2. Bujna, B., Mikulski, J.: Profitability and Effectiveness of Transport Investments. In: Mikulski, J. (ed.) TST 2012. CCIS, vol. 329, pp. 94–106. Springer, Heidelberg (2012)
3. http://www.siskom.waw.pl/komunikacja/kolej/KDP/synteza.pd (May 11, 2011)

Integrated Vessel Traffic Management Systems in the Function of Inland Waterway Traffic Optimization

Mihaela Bukljas Skocibusic and Mato Brnardic

University of Zagreb, Faculty of Traffic and Transport Sciences
Vukeliceva 4, 10000 Zagreb, Croatia
mihaela.bukljas@fpz.hr

Abstract. Awareness of an efficient and sustainable vessel traffic management is a result of raised traffic congestion as well as of increased waiting times at locks and inland ports. The established vessel traffic management systems will have to be modified by 2025, to challenge the expected increase in the transport demand on navigable European inland waterways. Substantial benefits to traffic flow optimization are anticipated by the dissemination of an accurate information regarding the waterway and the management of the locks in the waterway network. Positive effects resulting from a harmonised implementation of traffic management systems optimize also the use of waterway infrastructure by means of reducing dredging activities and lock utilization. The outcomes of this paper will give a concise assessment of the benefits of an integrated traffic management implementation to private and public stakeholders of inland navigation.

Keywords: optimization, traffic management, inland waterways, assessment.

1 Introduction

The implementation and overall efficiency of an intermodal freight transport system with inland navigation as a core mode is an important and challenging issue in the EU Transport Policy of the 21st century. On-going research studies that are being carried out on a European level determine in what way the European inland waterways transport could be used more efficiently and consequently integrated into the future European intermodal transport system. The vessel traffic management is aimed at making the inland shipping a competitive and safe means of transport, that makes the optimal use of the available capacity. This implies smooth shipping which strengthens the position of inland navigation as a cost-efficient link in the transport network and contributes to the efficient use of the available transport infrastructure and any future extensions to it [2]. The objective of this paper is to assess vessel traffic management systems with respect to the costs and benefits for the different target group actors. This involves a comprehensive cost/benefit analysis for private and public stakeholders of inland navigation, in general a shipper, freight forwarder, logistic chain manager, ship owner, skipper, barge operator, terminal operator, and the competent authorities for which the costs are supported by the public authority.

J. Mikulski (Ed.): TST 2012, CCIS 329, pp. 117–123, 2012.

2 Interaction between Private and Public Stakeholders of Inland Navigation

In order to understand the interaction between private and public stakeholders it is essential to analyse a decision process. The navigation as well as the support from traffic management services influence the decisions made by a navigator. The navigator makes decisions on courses to sail as well as on speeds, whilst a traffic management operator attempts to affect the decisions of a navigator by providing a relevant information that enables him to make correct navigation decisions.

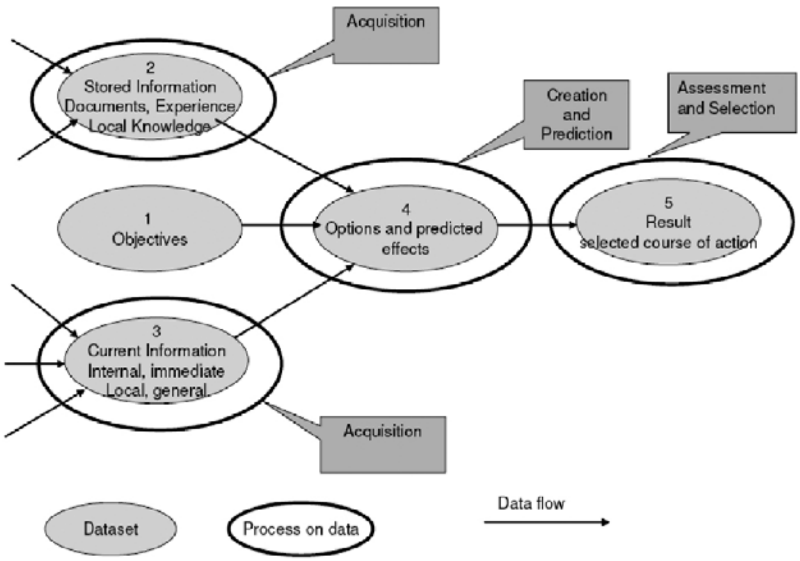


Fig. 1. Schematic review of an inland navigation decision process

The Fig. 1 contains ellipses which show various elements of the decision process in the inland navigation. The ellipse marked 1 represents the objectives for the decisions to be made. The ellipse marked 2 is the long-term information, such as documents, expert judgment, local knowledge and experience. The ellipse marked 3 represents the internal current information. It might also have a local and general information. In many cases the information stored as well as current should be acquired. This is indicated by the squares. A decision maker is affecting some processes on the datasets that are represented by ellipses 2 and 3. The results of the processes are options with their predicted effects. The process that is drawn around the 5th ellipse is the assessment of the different options and their predicted effects. The assessment as a final result, the selection of an option takes place taking the objectives into account. Figure 1 depicts the key to understanding the process of navigation and provided Vessel Traffic Management services [5].

3 Problem Solving Approach

The following fundamental relationships have been derived from the definition of vessel traffic management: the first basic relationship to be highlighted is between the management and the *safety and capacity* of the waterway traffic. In other words, the relationships between the level of information and knowledge provided to the waterway users and the reduction of the risk and of the number of accidents as well as the reduction of the time delays and the use of optimum drafts. The second basic relationship is the saving of *costs in infrastructure* investment as a result of traffic management implementation. Many locks have large deceleration works that are necessary for mooring the inland vessels that have to wait for a lock cycle. By precise planning of the RTA's, based on ETA's send by skippers the number of waiting vessels can be reduced and hence the length of the deceleration works can be reduced. On-going development of planning messages that are a part of VTM might help this process [4].

Moreover, by positively affecting the traffic flows throughout locks it is possible to make the traffic supply through a lock more uniform over the day, implying a better usage of the locks. An integrated VTM application on waterways would imply that the lock might be operated for a longer period without renewal or replacement of the lock. The investment costs may also be reduced. Benefits due to the reduction of fuel consumption and waiting times at locks in a RIS covered area, the reduction of dredging work, replacing VTS-centres by new established RIS centres, smaller investments in deceleration works and the delay of investment of locks could derive benefits.

3.1 Basic Indicators

Quantitative indicators for the traffic management implementation offer suitable basis for a set-up of cost benefit analysis framework assessment. Basic indicators consist of monetary resources that go into the financial aspect of VTM implementation process, inputs (technical components), outputs (traffic management services) as well as outcomes which occur as a result of integral implementation.

Table 1. Basic indicators

Indicators			
Costs	VTM inputs	VTM outputs	VTM outcomes
Monetary resources	Technical components - VHF - Radar - ENC - AIS - ICT infrastructure	- Traffic Planning - VTS - Lock management	- Safety - Reliability - Cost efficiency - Competiveness - Effectiveness

3.2 Calculation Methodology

A clear and structured methodology is essential for summing up efforts into general results. Approaches are based on practical considerations in relationship with the available data and information. Therefore, the methodologies used is this paper define and regroup benefits and make assumptions of the expected changes based on the input data, modelling and on the expert judgment.

The above-mentioned calculation is based on the following three-steps methodology:

1. Evaluation of the costs

Review of the "clusters" as regards the targeted actor group, the necessary related on board and shore equipment (hardware, software, communication means, etc…), as well as the navigation mode of operation (on-demand, continuous, etc.). For each of the "clusters" the determination of the capital cost, of the investment lifetime, of financial rates, of yearly operating cost according to the level/intensity of use, and of assumptions related to the level of the fleet equipped with.

2. Evaluation of the benefits

As an introduction to the calculation of the cost/benefit analysis it is important to highlight the most important benefits that VTM can provide to the private and public stakeholders: optimization of traffic delays, substitution of VTS and introduction of RIS centres, optimization of dredging costs, optimization of deceleration works and capital costs of locks.

3. Extrapolation of results

Results will comprise a calculation of the costs that an actor (public or private) could support when working without VTM services or with VTM. By comparing the results the authors will obtain quantified benefits in the final analysis. Authors developed a model for the period between 2005[1]-2020 on the river Rhine and its tributaries. It is assumed that a full integration of vessel traffic management services to users will occur in 2015, and that the efficiency will increase to 100 % by 2020 [6]. The following table provides modelling of the level of vessel traffic management implementation and contribution to the optimization of costs [3].

Table 2. Percentage of cost optimization to private and public stakeholders as a result of integrated VTM implementation

Level of VTM implementation and contribution to the optimization of costs				
Year	2005	2010	2015	2020
VTM efficiency	0%	50%	50%	100%
Optimization of traffic delays	0%	5%	10%	10%
Introduction of RIS centers	0%	50%	100%	100%
Optimization of dredging costs	0%	12.5%	25%	25%
Optimization of deceleration works	0%	12.5%	25%	25%

[1] RIS Directive (2005/44/EC) established framework for harmonised deployment of vessel traffic management systems across Europe within the scope of the author's model. As a consequence, year 2005 was taken as a reference in the model.

The cost optimization is calculated by multiplying the optimization percentage of table 2 with the statistical data gathered by the study "Assessment of INDRIS on the basis of the Rhine, Seine and Danube demonstrator". The results are discounted using the European Commission discount rate of 4 %. Also the Net Present Value is calculated for the above-mentioned cost optimization.

4 Cost/Benefit Analysis of Integral Vessel Traffic Management Implementation

A cost-benefit analysis links the cost of an action to its consequences expressed in monetary units. This analysis is based on all the positive and negative impacts of a project for all the actors involved in the project and on the society as a whole [1]. As stated, the private stakeholders benefit from direct short-term effects of VTM.

$$NPV = \sum_{t=1}^{n} \frac{B_t - C_t}{(1+i)^t}$$

Authors have calculated average values of the costs that the private parties would support in three particular situations and for two cases: without-VTM and with-VTM. If we compare the results of the situations we will obtain the following benefits due to the RIS implementation for the private stakeholders. For an investment to be justified from the economic evaluation theory point of view, the value of the Benefits/Costs (or Advantages/Costs) ratio should be equal or greater than 1 (B/C $\geq$ 1) [1]. The following table provides the B/C ratio as calculated for the private stakeholders. Benefits consist of the reduction of delays caused by the traffic congestion and locks utilization.

Table 3. Benefit/Cost ratio for private stakeholders

B/C private stakeholders	
	Net present value
Optimization of traffic delays	9.727.127
Total benefits, private stakeholders	9.727.127
On board equipment Rhine corridor	4.310.996
Total costs, private stakeholders	4.310.996
Benefit/Costs Ratio	***2.56***

The following table provides the B/C ratio calculated for the public stakeholders, the competent authority. The most relevant costs are the costs of RIS centres introduction as well as the extra costs to equip the locks with adequate RIS technology.

Table 4. Benefit/Cost ratio for public stakeholders

B/C public stakeholders	
	Net present value
Optimization of VTS costs	36.299.147
Optimization of dredging costs	6.721.217
Optimization of deceleration costs	15.373.510
Delay of investments in locks	14.513.822
Total benefits, public stakeholders	72.907.696
Introduction of RIS centers	65.753.883
RIS lock centers	4.785.933
Total costs, public stakeholders	70.569.816
Benefit/Costs Ratio	***1.03***

5 Conclusion

The Vessel Traffic Management system as a part of River Information Services is aimed at optimizing the use of waterway infrastructure as well as at facilitating the navigation safety. The measurement of benefits and their relationship to costs is a commonly used practice while evaluating infrastructural projects. The VTM system provides the parties involved with positive B/C ratios. The private stakeholders on the stretch of the river Rhine will have a large return (2.66 B/C ratio) for any investment that they are prepared to make. The public authorities may also have a positive B/C ratio (1.03) but the application of VTM will not bring them many benefits. However, as managers of the waterway infrastructure they may be expected to co-operate with a change that is clearly in the interest of the private stakeholders.

Authors conclude that the implementation costs per stakeholder are relatively easy to evaluate but the benefits can only be estimated based on the assumptions, modelling and on the expert judgment. It can be concluded that performed studies within this article come to a very high C/B ratio, making the Vessel Traffic Management system as part of RIS a prospective infrastructural project for the inland navigation.

References

1. Gerard, M.: Evaluation socio-economique de la mise en oeuvre de systèmes d'information pour la navigation fluviale européenne. ANAST, Université de Liège, Liège (2000)
2. Sys, C., Vanelslander, T.: Future Challenges for Inland Navigation. A Scientific Appraisal of the Consequences of Possible Strategic & Economic Developments up to 2030. Asp / Vubpress / Upa., Antwerpen (2011)
3. INDRIS (EU DG-VII, 4th FP Task 2.17) Methodological Framework for INDRIS Assessment, SWP. 13.1. Technical Report, Liège (1999)

4. INDRIS (EU DG-VII, 4th FP Task 2.17) Assessment of INDRIS on the basis of the Rhine, Seine and Danube demonstrators SWP13.4-5: C/B Analysis and Global Assessment. Final report, Liège (1999)
5. COMPRIS Consortium Operational Management Platform River Information Services, Deliverable 12.1: Report on Socio-economic assessment of RIS, Liège (2006)
6. Directive 2005/44/EC on harmonised river information services (RIS) on inland waterways, Bruxelles (2005)

The Analysis of Transmission Parameters in Railway Cross Level Protection Systems with Additional Warning of Car Drivers

Andrzej Lewiński and Lucyna Bester

Kazimierz Pułaski Technical University of Radom
Faculty of Transport and Electrical Engineering
26-600 Radom, Malczewskiego 29 St, Poland
{a.lewinski,l.bester}@pr.radom.pl

Abstract. The paper deals with problems of time parameters (delays, queues) in Additional Warning System of unguarded cross level systems., The open radio transmission standards (WiFi, WiMax, WSN.) applied in AWS system is analyzed with respect to queues and delays based on mathematical methods (Markov processes). These systems apply the wireless networks transmitting the signal between different dissipated objects and safety transmission standards require such t analysis corresponding to requirements defined in obligatory UE standards.

Keywords: open transmission standards, safety transmission, protection of cross level intersections, markov process and time analysis.

1 Introduction

Modern railway management and control systems are based on wireless solutions of transmission. In these new systems the problem of time delay in transmission is very important with respect to functionality and safety. The delay of proper processing of information connected with typical transmission parameters (etc. interferences, interrupts or integrity code faults) is related to number of transmitting devices and time of service. The open transmission standards connected with public computer networks include the measures and procedures identical for all users. The application of such solutions, recommended by UE railway organizations (UIC, ORE) requires the application of PN-EN 50159-2010 standard [9] (*Railway applications – Communication, signaling and processing systems – Safety-related communication in transmission systems*) including the earlier recommendations of design of safety computer system in rail control an d management PN129-2007 [8]. The new generation of railway control and management systems use the radio and wireless radio transmission solutions like WiFi/WiMax, and WSN [10] . These systems may co-operate with existing railway control systems but give more functionality connected with efficient computer monitoring and management. In the case of additional warning systems for unguarded cross level intersections is possible to

J. Mikulski (Ed.): TST 2012, CCIS 329, pp. 124–131, 2012.

significant increasing the level of safety (the probability of catastrophic failure may be 10-100 less). The scheme of implementation of such system for cross levels cat D (completely unguarded) or C (without barriers only equipped with signaling lamps for car drivers).is shown on Fig. 1.

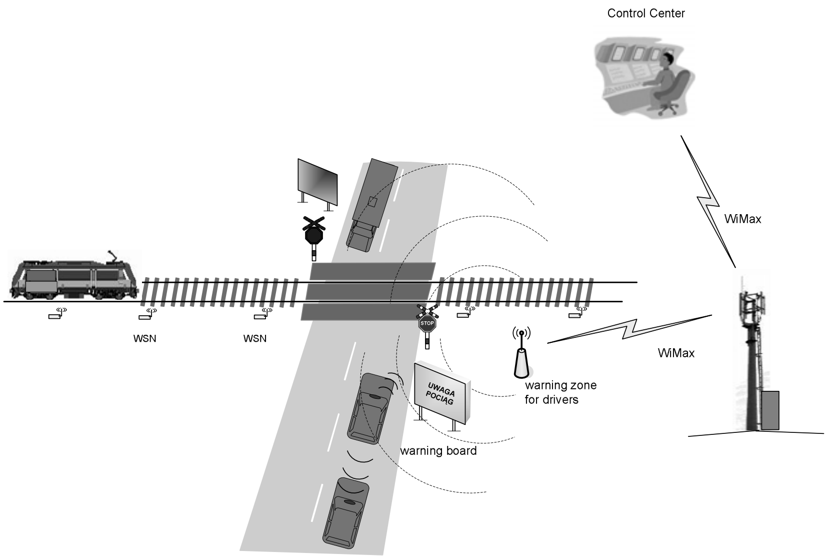


Fig. 1. Additional Warning System for cross level intersection cat. D and C.

2 Open Transmission in Additional Warning System (AWS) for Car Drivers

2.1 General Characteristics of Open Transmission Systems

The technical means which modernize the previous methods of train control are free transmission standard such as WiMax (*World Interoperability for Microwave Access*) or Wi-Fi (Wireless Fidelity) standard. These standards cooperate with each other and are based at DSSS (Direct Sequence Spread Spectrum) and OFDM (*Orthogonal Frequency Division Multiplexing*) protocol. These standards can be used for communication between locomotive driver and dispatcher center, they also improve the communication between vehicles and railway infrastructure and give the access to the nets in distant areas. It gives the warranty of QoS quality and small delay of transmission to 200km/h speed [12]. However to diagnose and monitor the brake temperature, wheel detection, the middle speed measurement to take advantage of GPS data, there can be used the WSN (*Wireless Sensor Network*) [5]. The WSN, thanks to fulfilling IEEE 802.15.4 standard, is prepared to cooperate with WiMax and Wi-Fi standard.

The WiMAX standard offers very low delay between the mobile device and base station (<10ms), while the delay between base stations is less than 100ms. The delay introduced from the radio link, which introduces a delay in data transmission, is of the order of several milliseconds and another millisecond delay introduces a wired technologies used in operation of the system, which is very often used as a link for access to wireless networks. In the case of sensor network, a lot of operating parameters is dependent on network topology and frequency band. For the higher frequency range, is much greater throughput and lower delays (from 15 ms) compared to devices operating in the range below 1 GHz. Taking into account, that some applications require a lower and constant delays, we can through the use of standard ZigBee (IEEE 802.15.4) adjust this parameter from 15 ms to 245 s. The packet delay T, in both the sensor network as well as WiMax standard (IEEE 802.16) and WiFi (IEEE 802.11) is the sum of the individual elementary delays, such as different times between frames, the time of retransmission and the time of transmission control frames [1, 2]:

$$T_{op} = T_{rt} + T_{ra} + T_{oc} + T_{ACK} + T_{ISF} \tag{1}$$

T_{rt} - time of packet retransmission in sec. (320μsec.)
T_{ra} - time of transmission of the packet, size of x bytes of data (100msec.)
T_{oc} - time for waiting to response (192μsec.)
T_{ACK} - time of transmission acknowledgement ACK (100msec.)
T_{ISF} - time between frames (192μsec.)

For the analysis of delays the total average time 900 ms is assumed.

2.2 Delays in the AWS System

For Sensor Network the average queue length N and waiting time Ts for service of telegrams sent from n number of simultaneously transmitting sensors can be calculate, using the equations defined in [7]

$$Ts = N \frac{1}{\lambda} = \frac{1}{\mu - \lambda} = \frac{\frac{1}{\mu}}{1 - \frac{\lambda}{\mu}} \tag{2}$$

where: λ – intensity of packets sent per unit time (1/λ – mean time between appearance of the packets), μ – the intensity of packet service (1/μ – average service time of packets).

For estimation two cases of service are assumed; one (normal operation) with 0.1s time of service per sensor, second (in faulty transmission processing) with 1s time of service.

Table 1. Queues and delays in the Sensor Network

The number of transmitting sensors in time 100 i 10 s.	**λ 1 (100s)**	**λ2 (10s)**	**$\mu=100^{-1}$ ms^{-1} T_N [s]**	**$\mu=10^{-1}$ ms^{-1} T_N [s]**
5	0,05	0,5	1,00E-02	1,01E-02
10	0,1	1	1,00E-02	1,01E-02
20	0,2	2	1,00E-02	1,02E-02
40	0,4	4	1,00E-02	1,04E-02
60	0,6	6	1,01E-02	1,06E-02

For medium service time of telegrams related to the number of sensors (n) intensity of incoming calls λ = 100 sec. i λ = 10 s., time of service μ = 100 ms and μ = 10ms the time parameters are presented in Table 1.

The Fig. 2 shows the difference of medium time of waiting telegrams corresponding to number of sensors and time of telegram service.

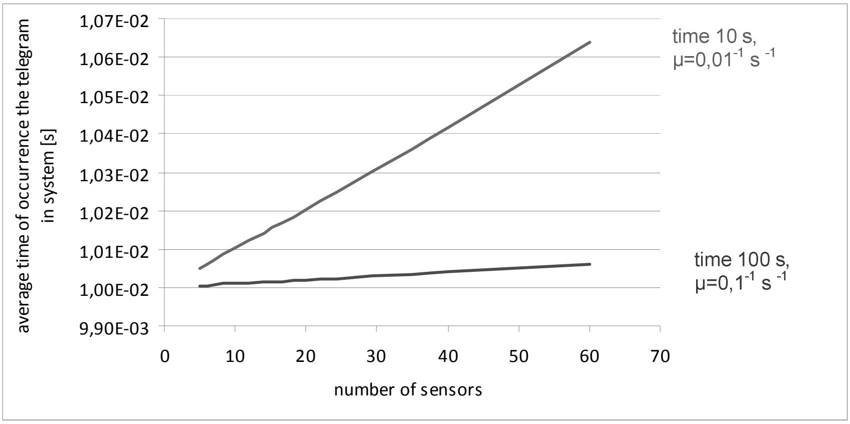


Fig. 2. The medium time of waiting for service together with service) for „n" sensors

The influence of delays and processing time for correct operation of AWS system the important problem is connected with time of transmitting the message to the car diver. The total delay time T_S is related to transmission time of the package (T_p), reaction time of car driver for message (T_c = 0,8), transmission delay connected with processing the message (T_{ot} = 3 ms) and global delay connected with transmission the messages from given number (n) of sensors (T_n) corresponding to Table 1. It is obvious that the reaction time of car driver for message T_c has a main part in the total delay.

$$T_S = T_p + T_C + T_{ot} + T_n$$
$$T_S = 10\,\mu s + 0{,}8\,s + 3\,ms + 12\,ms = 0{,}815\ s \tag{3}$$

The next step of AWS analysis is an estimation the probability of transmission errors BER (*Bit Error Rate*) corresponding to following formula (4):

$$P_{pop} = (1 - BER)^N \tag{4}$$

where: N – the length of telegram (in bits), BER – Bit Error Rate, P_{pop} – probability of telegram correct transmission

Table 2. Probability of faulty transmission with respect to length and BER

Probability of correct transmission	BER for package 128 bits (16 Bytes)
1	0
0,9	$8{,}22*10^{-4}$
0,8	$1{,}74*10^{-3}$
0,7	$2{,}78*10^{-3}$
0,6	$3{,}98*10^{-3}$
0,5	$5{,}40*10^{-3}$
0,4	$7{,}13*10^{-3}$
0,3	$9{,}39*10^{-3}$
0,2	$1{,}24*10^{-2}$
0,1	$1{,}78*10^{-2}$

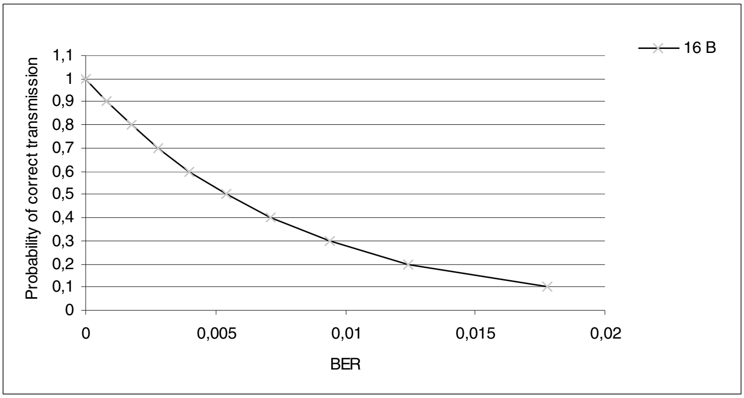


Fig. 3. The probability of faulty transmission with respect to BER

We can observe that for BER= $18*10^{-3}$ the probability of correct transmission is above 10^{-2}.

The CRC (*Cyclic Radundancy Code*) is always used in local network for detection of faulty packages. Corresponding to length of protected data the three CRC codes (CRC-8, CRC-16, CRC-32) are analyzed. The estimation of faulty transmission may be done as follows:

$$P_{prz} = \sum_{i=D}^{N} \frac{N!}{i!*(N-i)!} * p^{i} * (1-p)^{N-i} \qquad (5)$$

where: D – Hamming distance, N – number of protected bits p – probability of single bit fault (for radio transmission the 10^{-4} is assumed)

The Table 3 contains the typical estimated values of faulty transmission and general relations are presented as a form of chart in the Fig.4

Table 3. Probability of faulty transmission with respect to CRC code and number of bits

encoding	8 Byte	14 Byte	16 Byte	18 Byte	22 Byte	24 Byte
CRC 8 Hamming distance 4	$3,85*10^{-9}$	$6,69*10^{-8}$	$1,31*10^{-7}$	$2,38*10^{-7}$	$6,56*10^{-7}$	$1,01*10^{-6}$
CRC 16 Hamming distance 6	$4,39*10^{-15}$	$2,53*10^{-13}$	$6,59*10^{-13}$	$1,52*10^{-12}$	$6,36*10^{-12}$	$1,18*10^{-11}$
CRC 32 Hamming distance 7	$3,58*10^{-18}$	$3,80*10^{-16}$	$1,13*10^{-15}$	$2,99*10^{-15}$	$1,53*10^{-14}$	$3,11*10^{-14}$

We can see that for 192 bits in transmitted package the probability of faulty transmission for CRC-8, CRC-16 and CRC-32 is equal to 10^{-6}, 10^{-11} and 10^{-14}..

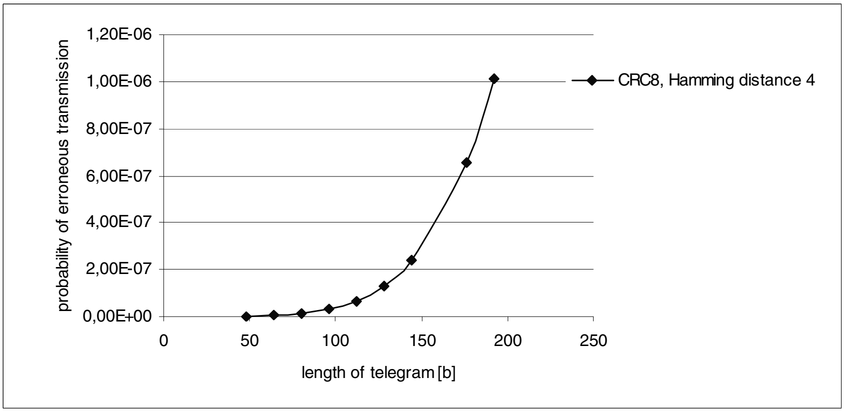
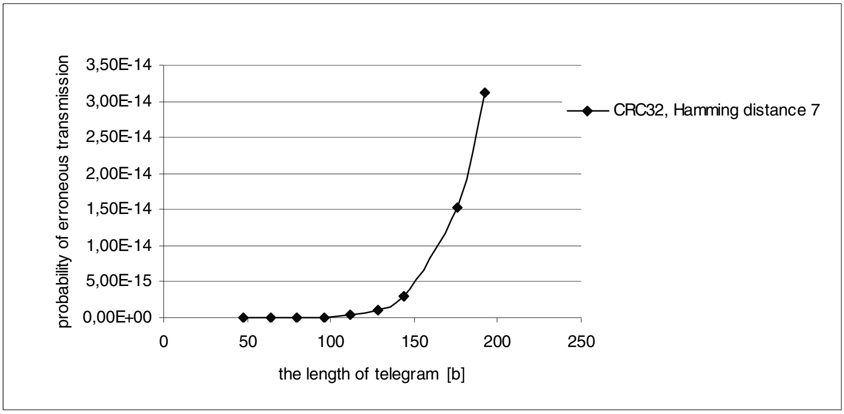


Fig. 4. The probability of faulty transmission as a function of number of bits (length of telegram)



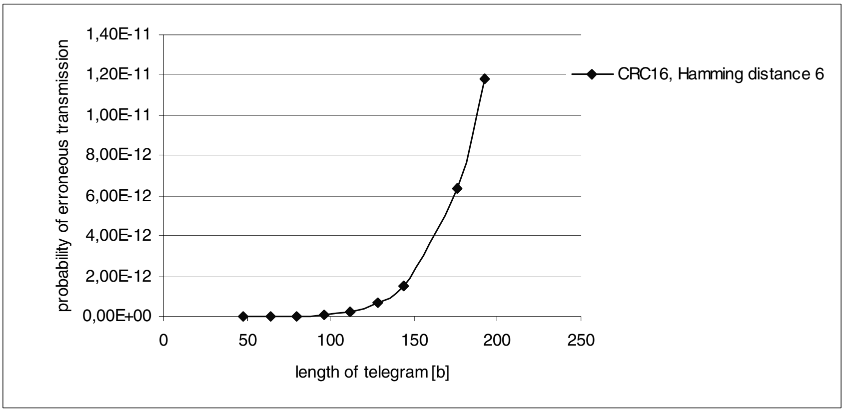


Fig. 4. (*continued*)

We can see that for 192 bits in transmitted package the probability of faulty transmission for CRC-8, CRC-16 and CRC-32 is equal to 10^{-6}, 10^{-11} and 10^{-14}.

3 Conclusion

The analysis of AWS system with open radio transition shows, that assuming typical delays related to transmission failures must regard the number of transmitters (sensors) and time of service in the concentrator (the number of cars before cross level) is has no significant influence to the proper work of the concentrator. For AWS system the optimal number of sensors may be reduce to several (3-6) for each rail in both directions because increasing queue of transmitted and not service messages for greater number may cause inoperability in processing the faulty messages in 1s time of service. The analysis of transmission errors in AWS shows very low level of faulty transmission probability. Such level of faulty transmission has no significant influence for safety of unguarded intersections protected by AWS because other factors (environment, weather, reliability of applied technical equipment) may more affect for critical failure (accident). The completely damage of AWS system recover the typical situation in cross level area. The presented AWS system may be treated as an additional radio transmission overlay into existing infrastructure of uncontrolled and unguarded cross level intersections.

References

1. Latré, B., De Mil, P., Moerman, I., Van Dierdonck, N., Dhoedt, B., Demeester, P.: Maximum Throughput and Minimum Delay in IEEE 802.15.4. In: Jia, X., Wu, J., He, Y. (eds.) MSN 2005. LNCS, vol. 3794, pp. 866–876. Springer, Heidelberg (2005)
2. IEEE Std 802.11-1999 Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) specifications
3. Jackowski, S., Chrzan, M.: Contemporary telecommmunications systems. Technical University of Radom (2008) (in Polish)
4. Lewiński, A., Perzyński, T.: The delay analysis in dissipated railway management and control systems. In: Telematic Transport Systems Conference (2005)
5. Standard PN-EN 50129:2007 Railway applications – Communication, signalling and processing systems. Safety related electronic systems
6. PN-EN 50159 – 2010. Railway applications – Communication, signaling and processing systems – Safety-related communication in transmission systems
7. Tanenbaum, A.S.: Computer Networks. Prentice Hall PTR, New Jersey (1996)
8. UIC No. O-2475: ERTMS/GSM-R
9. Quality of Service Test Specification form 11.09.2003 and parameters: GSM-R Interfaces Class 1 Requirements. SUBSET-093 z 11.10.2005
10. Lewiński, A., Bester, L.: Zastosowanie nowych standardów transmisji bezprzewodowej w systemach zarządzania i sterowania ruchem kolejowym. Prace Naukowe Politechniki Warszawskiej, Transport z.62, Stare Jabłonki (2007)

A Prototype of Economical and Universal On-board Events Recorder – A Black Box for Vehicles

Gabriel Nowacki and Anna Niedzicka

Motor Transport Institute, Transport Management & Telematics Centre
Jagiellońska 80, 03-301 Warsaw, Poland
{gabriel.nowacki,anna.niedzicka}@its.waw.pl

Abstract. The paper refers to the project to develop two economical prototypes, of the universal on-board devices: Event Data Recorder – black box for all types of the motor vehicles, Event Data Recorder - black box for all types of the motor vehicles, registering eCall notifications. The device will record much data concerning the vehicles technical condition, the way it was driven and RTS. The recorder may be used in the passenger, company and cars with the right of way, taxies, buses and trucks. The recorder may serve the purpose of the neutral witness for the police, courts and insurance firms, for which it will facilitate making the reconstruction of the road accidents events and will provide proofs for those who caused them. The device will bring about more law-abiding and efficient driving, which will significantly contribute to decreasing the number of the road accidents and limiting the environmental pollution.

Keywords: on-board device, event data recorder (EDR), black box.

1 Introduction

According to the World Health Organisation in 20 years time road accidents will take fifth place on the list of the most frequent causes of death in the world. It has been estimated that annual costs of accidents in Europe amount to 130 billion Euros, while in Poland itself - 5 billion Euros. Despite the measures taken to reduce this amount, the number of fatal road accidents' victims is in the European Union exceptionally high - 34 800 in 2009 [1].

The reason for such situation is a significant growth of the road transport, associated with the development of the European economy and citizens demands as far as mobility is concerned, which results in ever growing congestion of the road infrastructure, energy consumption increase as well as ecological and social problems.

European action program for the road traffic safety defines main areas of the activities: promoting more responsible attitudes of the drivers (obeying road traffic regulations and punishing dangerous behaviour on the roads), improving safety of the vehicles by introducing new technical solutions, improving road infrastructure, thanks to the information and communication technology.

J. Mikulski (Ed.): TST 2012, CCIS 329, pp. 132–140, 2012.

1.1 The Road Accidents Statistics

The statistical data suggests that the year 2010 was the safest in the last 20 years (Fig. 1) on the Polish roads (lowest number of the people killed). Despite this fact, in the statistics referring to the percentage decrease of the number of accidents in the last decade (2001-2010), we still hold one of the last places. For this period, an average percentage decrease of the number of accidents in the EU was 43 %, while in Poland - 29 %, in Lithuania and Estonia - 61 % and in Latvia – 58 %. European Commission regularly publishes reports on the safety on the roads of all countries belonging to the EU.

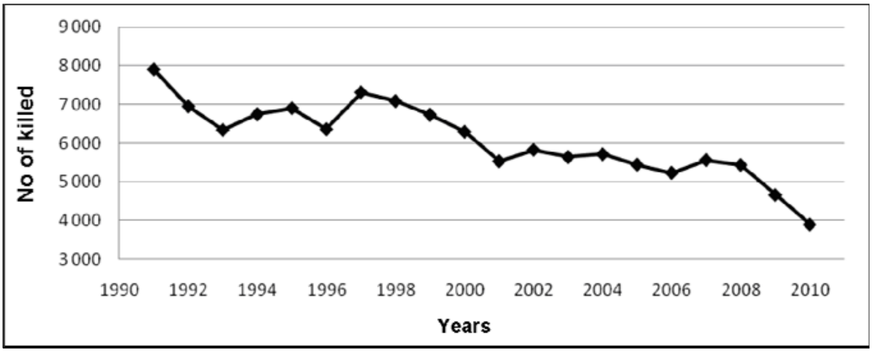


Fig. 1. Number of people killed on the roads during the years 1991-2010

The statistics are invariably devastating for Poland, which with respect to the number of victims and mortality in the road accidents is always at the end of the list**.** The threat of death in the road accident in Poland is three times higher than in the rest of the EU countries, and the statistics of the fatalities per 100 accidents we are worse even than in Lithuania. For Poland this number is 10,3, in Lithuania 9,9, Greece - 8,8, while the average in the EU is 3. In this context, a shockingly low mortality coefficient is noted in such countries as Germany and Great Britain (1,4), as well as in Austria (1,7) or Italy and Sweden (1,9). The probability of death of the accident participant is in Poland, on average, four times higher.

Moreover, in the statistics concerning the number of people killed per 1 million inhabitants we also hold one of the last places, since the EU average in 2010 was 61, while in Poland - 102, and worse are only Romania – 111 and Greece - 112.

Data of the National Police Headquarters [2] shows, that in 2011 there were 40 065 accidents (38 776 - 2010) on the Polish roads, as the result of which 4189 (3907 - 2010) people died and 49 501 (48 872 - 2010) people were injured. In comparison to 2010 there was an increase of the number of accidents by 1 233, i.e. by + 3,2%, increase of the number of fatalities by 282 people, i.e. by + 7,2%, and increase of the number of people injured by 549, i.e. by + 1,1%.

In Poland, accidents are regarded at present as one of the largest public health problems. More so, their consequence is not only a trauma of the injured ones and their families but also a significant burden to the public finances.

1.2 The EC Actions

The EC has decided (recommendation of 8 September 2011) to equip all cars with an on-board system and ensure the implementation, by the mobile networks operators, of the mechanism serving the eCall reporting indicator in their networks by the 31 December 2014 [3].

The designation of emergency reporting category means 8-bit value applied to the alerts transmitted through mobile communication networks in order to determine the specific nature of these reports (1 - police, 2 - ambulance, three - fire department, 4 - Marine guards, 5 - mountain rescue services, 6 - eCall activated manually, 7 - eCall activated automatically, 8 - reserve), as shown in Table 10.5.135d in ETSI TS 124 008. This implementation is to help the police identify the party responsible for the accident.

The ECall indicator means designation of the emergency reporting assigned to the eCall reporting according to ETSI TS 124 008 (e.g.,"6 - manually activated eCall" and "7 - automatically activated eCall"), which allows to distinguish between the reports made to 112 number from mobile terminals and the eCall to the 112 number from the on-board terminals, as well as manually activated eCall from the calls activated automatically.

Minimum set of data means information that must be sent to the entry point for reports of accidents in accordance with EN 15 722 standard.

Successful implementation in the entire EU of a harmonized interoperable eCall service requires the automatic transfer of voice and audio connection and the minimum data set of the accident, generated by the on-board system, to the appropriate public accidents reporting exchange.

Member States should commit its national authorities to notify the Commission, by the end of March 2012, about the measures that have been applied in response to this recommendation.

Furthermore, according to European Parliament resolution of 27 September 2011 on road safety in Europe for the years 2011 to 2020, there should be following legislative documents developed [4]:

- by the end of 2012, proposals for legislative changes to implement in the corporate and private vehicles, the integrated system for recording data immediately before, during and after the accident,
- by 2013, proposals for legislative changes, assuming that each new vehicle must be equipped with a system reminding about fastening the seat belts in the front and rear seats, operating based on acoustic and visual signal.

2 Characterization of Black Box

Event Data Recorder (EDR) “black box” is a device in a vehicle that stores event-specific data (e.g. vehicle speed and driver inputs when a crash, rollover, or other mishap occurs) [5].

The first EDRs or *black boxes* were used in the aviation industry in the late 1950s. In 1958, the Federal Aviation Act and corresponding regulations issued by the Civil Aeronautics Administration (the predecessor of the Federal Aviation Administration) made mandatory the use of black boxes or *flight data recorders* for commercial aircraft. In 1976, the National Transportation Safety Board (NTSB) issued regulations requiring the use of EDRs in commercial marine vehicles. In May 1995, the Federal Railroad Administration issued regulations requiring EDRs on heavy rail transportation. While the use of EDRs in automobiles and light trucks is currently voluntary, vehicle manufacturers such as General Motors and Ford have installed EDRs on many of their newer models [6].

The recorder proposed by the Motor Transport Institute, to a large extent, may help to reduce the number of accidents, a significantly shortening of travel time and energy consumption, thereby improving the quality of the environment, and will be useful as evidence in the disputable matters.

The proposed recorder - car black box, can be used to record data concerning the technical condition of the vehicle, the driving technique, and drivers compliance with the traffic rules and road traffic safety in the following cars:

- passenger, service and privileged cars – will allow the registration of the earlier indicated data and will provide evidence in case of accident,
- buses and taxis – apart from the recorded data, it will help to ensure the safety of the driver and passengers, will enable the location of vehicles in the event of theft,
- trucks - will ensure registration of the data on technical condition will enable the localization, will contribute to reduce the number of accidents and ensuring safety of the driver.

A significant disadvantage is the fact that currently there is no standardization of such devices. Therefore, one of the objectives of the project will also be proposal of creating the set of data recording (in the future it may be adopted as a standard data recording) so that their reading was possible without a specialist equipment or software.

2.1 Functional Requirements of the Recorder

The device will receive data from selected circuits of the vehicle via digital and analogue input ports. Sensors may be possible to be connected e.g. to the doors, lights, turn indicators or brakes.

The electronic system will record and remember the course of acceleration during the collision and remember the result of changing the vehicle speed. In addition, it will remember certain information prior to the accident and immediately after the accident.

The device will interpret digital information transmitted on the CAN bus and the FMS-CAN to record the following parameters (range, depending on the type of vehicle and equipment, such as truck equipped with a tachograph): speed, engine rpm, the position of the brake pedal, clutch pedal position, accelerator pedal position, the state of the cruise control, fuel level, mileage, total fuel consumption, tachograph - operation mode, tachograph - speed, tachograph functioning, dealing with the event status, the mileage remaining until the next tests, engine hours, coolant temperature, the axis load.

Additional sensors will be able to monitor: the time of release (reaction), a longitudinal, lateral acceleration, vehicle speed (the counter), the engine throttle (gas pedal), brake status (enabled or disabled), supply voltage, the position of the ignition, cushions signaling, the number of events, the time between events, horn, light switches, traffic lights, parking lights, turn indicators, the change in the car deflection (car rotational speed with respect to the vertical axis), the driver's seat belt status, events registration time. The actual sensors in car were presented on Fig. 2.

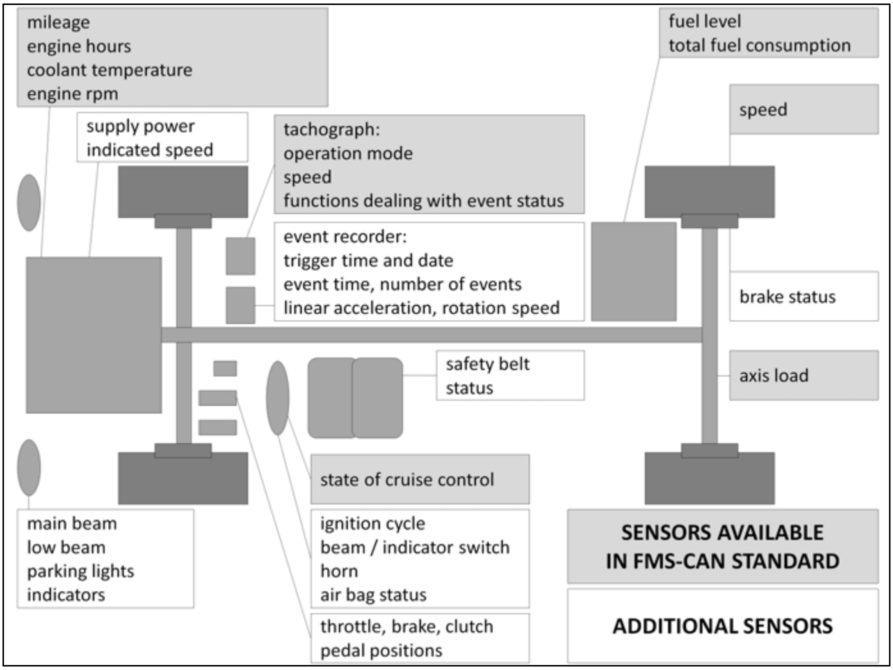


Fig. 2. Monitoring sensors in vehicle

2.2 Technical Requirements of the Recorder

One element of accident reconstruction is to recreate the time-space relationships of the event participants. Motion reconstruction process is based on the analysis of records of the parameters characterizing the motion of the car body. The forward motion is recorded as standard in the form of linear acceleration components (components: longitudinal, lateral, and vertical).

The device is designed for installation in all types of vehicles (passenger cars, trucks, buses) to record the driving parameters such as speed, acceleration, braking, use of direction indicators, etc. Such information can be extremely helpful in identifying those responsible for road accidents and will allow reconstructing the accident. They are also to replace the witnesses who are not always reliable.

The recorder will have a small size, will be much smaller than the car radio, and made of durable materials, and the place specially protected in it should be the SD card casing, on which the data will be stored.

The device will be equipped with the following elements (Fig. 3):

- Microprocessor module – the element controlling the operation of all other components.
- GPS - to receive the geographical coordinates.
- Module "black box" - memory chips, including SD card, capable of recording the driving and operating parameters.
- Digital inputs module - this module allows connection of digital signals.
- Analogue inputs Module - this module allows connection of analogue sensors.
- CAN inputs module - to connect CAN bus in the FMS standard.
- Battery backup module.
- Communication module GSM / GPRS version 2, for data transfer and entry to install the camera and the ability to record audio and video.
- Housing.

Intelligent algorithm for data collection will allow, at the moment of detecting an excessive acceleration by the built-in sensors, to switch to a continuous recording of all available parameters. In version 2, in the event of an accident, the automatic emergency call function will be activated in order to provide the assistance to the victims as soon as possible, and the GPS receiver will allow pinpointing where exactly the car is and sent, via the GSM network, the location data to an external server.

The device will save the measurement data on the SD card, and in the version 2, if the camera is installed, the unit will record audio and video, for a few seconds before the accident and a few seconds after the accident. The card can later be inserted into your computer to recall the data.

The device will be designed for commercial vehicles powered by 12V DC or 24V. The possibility of using backup battery is envisaged, to ensure continuity of work, e.g. in the case of basic voltage cut-off, for example, after a traffic collision.

Equipment must be installed in a protected, dry place. These places include the vehicle cabin. The optimal place is the area behind the dashboard.

Selecting the installation place must also be dictated by the visibility of satellites by the GPS antenna, in a place not covered by bodywork.

Devices will be connected to the CAN bus and to the selected electrical circuits of the car.

The design may not allow for manipulation or external sensors, such as their exclusion. It will be deprived of the possibility of changing the stored data, connected directly to the accident.

The device must be designed so as to minimize the reconstruction error of the motion parameters, in particular trajectory of the movement. This error should to the smallest degree be dependent on the vehicle load.

The unit will have an airtight case ensuring resistance to a short-term immersion in water and service liquids (fuel, oil, hydraulic fluid).

It will withstand exposure to direct flame and high overload for a period of several milliseconds.

Equipment should be protected against shock and vibrations in accordance with the standard defined in the EN normative documents (environmental conditions, electromagnetic compatibility). All mobile devices should withstand the following exposure: a one-time shock and falls from a height of 1 m. The devices should be able to withstand vibration, both sinusoidal and random.

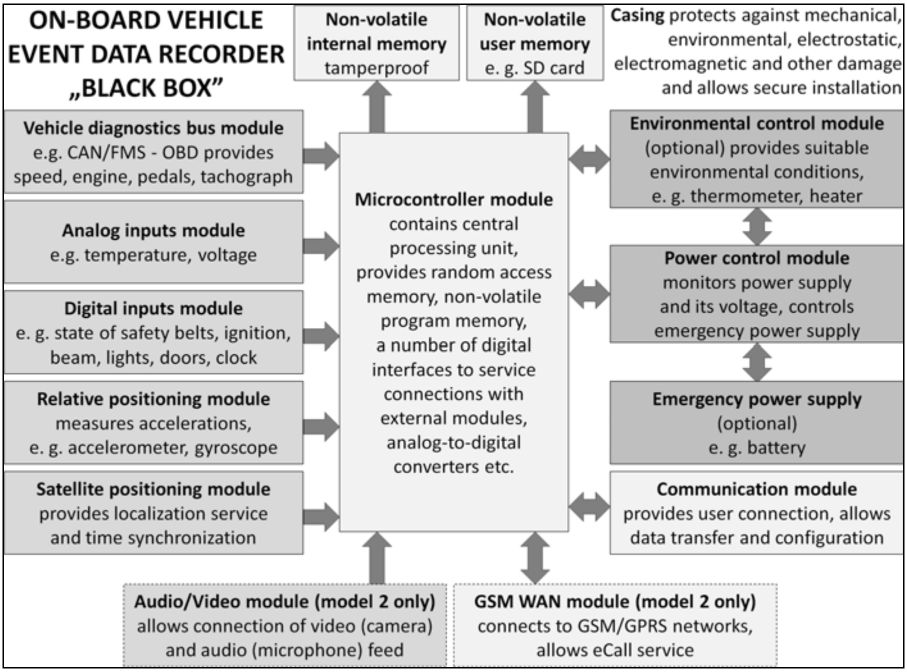


Fig. 3. On-board vehicle, event data recorder functional structure

For safe and reliable operation of all electrical and electronic equipment in the car it is necessary to ensure the electromagnetic compatibility of the recorder - the black box.

The device should be compatible with all environmental specifications, physical and compatibilities defined in the CEN, ISO and ETSI standards. It should meet all requirements, relating to this group of products, of the EU Council Directives, European standards and national legal regulations.

3 Conclusion

The end result of the project will be turning out two devices:

– prototype of the economic, universal simple event data recorder - black box,
– prototype of the economic, universal event data recorder - black box for all types of vehicles, taking into account the eCall reporting.

Recorder - car black box can be used to record data concerning the technical condition of the vehicle, the driving technique, and the driver's compliance with the traffic regulations and maintaining the road traffic safety in all motor vehicles.

The European Commission is currently considering the implementation of legislation in this area, prescribing the mandatory installation of black boxes in all vehicles.

Psychological impact of the black box will revolutionize road safety. Drivers will be more cautious, knowing that their every maneuver may be recorded, so in the event of an accident they will not be able to make false statements.

The device, connected to the vehicle monitoring sensors, will be installed behind the dashboard or under the driver's seat. Each sudden change of speed or opening of the airbag will activate it so that way also the collisions involving pedestrians will be recorded. In order not to violate privacy, car black boxes will store the data recorded 30 seconds before the accident and 15 seconds after it. The machine will automatically alert the emergency road services about the accident. In Britain, black boxes are standard equipment in many privileged vehicles. When in 1999 the London police installed them in a 3.5 thousand of company cars, within 18 months, the costs of road accidents fell by 2 million pounds. The devices are also placed in some newer car models.

In the U.S., black boxes are quite commonly used, and right now they belong to a standard equipment of over two thirds of new cars. U.S. Senate approved the bill, under which from the 2015 all new vehicles must be equipped with digital driving parameters recorders, known as black boxes. The failure install such equipment will result in punishment.

The studies conducted in the U.S. and the UK have shown that drivers who drive with black boxes, were 20% less likely to have participated in the fatal cases, the failure rate and repair bills for their cars fell by 25 percent. In Poland, the annual cost of road accidents alone are 5 billion, so if that gets reduced by about 20%, one will get the savings for the state – amounting to 1 billion annually.

References

1. Report on accidents, Care & national data. European Commission, Brussels (2012)
2. Symon, E.: Prophylactic and Analyze Group of BRD KGP, Warsaw (2012)
3. Commission Recommendation of 8 September 2011 on support for an EU-wide eCall service in electronic communication networks for the transmission of in-vehicle emergency calls based on 112 ('eCalls') 2011/750/EU) L 303/46, Brussels (2011)

4. European Parliament resolution of 27 September 2011 on European road safety 2011-2020. P7_TA-PROV(2011)0408, Brussels (2011)
5. Messerschmidt, W., Smith, B., Messerschmidt, K.: Minimizing the Risk of Losing Valuable Forensic Data When Downloading the Electronic Control Modules (ECMs) of Heavy Commercial VehiclesTechnical White Paper by Messerschmidt Safety Consultants Birmingham (2011)
6. Westbrook, R.: Evaluation of electronic data recorders for incident investigation, driver performance, and vehicle maintenance project Bd549-50 final report. University of South Florida (September 2009)

Selected Aspects of Message Transmission Management in ITS Systems

Marek Sumiła

Warsaw University of Technology, Faculty of Transport
Koszykowa 75, 00-662 Warsaw, Poland
sumila@it.pw.edu.pl

Abstract. This paper discusses selected aspects of data exchange in ITS systems. Characteristics of data sources and recipients in ITS systems have been determined. An example showing the scale of data exchange within a given telematic subsystem has been presented. Factors causing data feed latency have been identified. Key aspects of information flow management have been recognised.

Keywords: telematics-enabled devices, ITS systems, data transmission.

1 Introduction

We are now right in a historically important pivotal point for the development of transport systems. In this tipping point the telecommunication and IT technologies meet to take over the helm and decide about the future shape of transport services. The modern transport services market has become flexible and independent of rigid delivery schedules or timetables. This applies to what, how and in what quantities to transport. All those measures are deployed to minimise outlays incurred through the transportation process. At the same time teleinformatic resources were recognised as paramount to support transport. Consequently they also improve the traffic safety, the traffic management coordination and the information about travellers. Hence, ITS systems are to the benefit of all parties involved.

The common denominator for all those measures are research and implementation programmes concerning ITS. Those initiatives, first started in the seventies [9] in Japan, USA, Australia and later on in Europe, paved the way for developing many telematics devices, today commonplace in vehicles (e.g.: GPS, proximity sensors), in the streets (e.g. variable-message sign, weather stations), on public transport stops (e.g.: notice displays, request stop buttons), in parking lots (car park spaces display, parking guidance system, number plate recognition system) and in many other places [8, 10, 11, 13].

In terms of ITS development Poland is in a very good place at the moment. For several years, current thoroughfares have undergone modernisation and new ones have been built. What is more, Poland is not burdened with obsolete, non-depreciated telematic equipment. Because of negligence many former, imperfect systems had not

J. Mikulski (Ed.): TST 2012, CCIS 329, pp. 141–147, 2012.

been implemented in the first place. Hence, while building new roads or modernising the existing ones, state-of-the-art telematic solutions will find their place in the infrastructure consequently putting Poland firmly on the map in terms of countries using ITS. Substantial implementation costs call for a correct planning underpinning those projects. It is a common knowledge that ITS-related investment outlays incurred whilst building or modernising roads are a fraction of possible retrofit costs.

A considerable responsibility for developing new telematic systems rests on shoulders of transport research institutions. Their key objectives include:

- Projecting the demand and developing new telematic services,
- Case studies and devising assumptions for system,
- Developing new technical solutions and employing the existing ones for purposes of new telematic systems,
- Assuring ergonomic solutions,
- Impact analysis of telematic systems in terms of their influence on safety,
- Assessing the impact of implementing new systems.

Among some of issues currently under research, the problem of interoperability of telematic systems and their components came to the fore. This is particularly important in the context of peer data exchange between systems delivered by different developers. According to ITS America [3] the following are key:

- safety of vehicle to vehicle communications (V2V),
- safety of vehicle to infrastructure communications (V2I),
- real-time data acquisition,
- developing applications for a dynamic mobility,
- better access to weather and environmental information.

At present, Poland has to focus on implementing new ITS-enabled systems and on integrating them to assure a seamless data exchange.

In the remainder of this paper, some aspects of data exchange will be listed and elaborated on, in the context of exchanging telematics-based information within ITS systems.

2 State of the Art in the Data Exchange within the ITS

According to [3], the ITS aims to improve the transport safety and mobility and to optimise its efficiency through the integration of advanced telecommunication technologies for purposes of transport infrastructure and vehicle solutions. Hence, the ITS comprises a wide range of wireless and wire resources based on IT and electronics, as well as measures aiming to manage – based on the acquired and projected information – ITS-enabled facilities/sites.

Absolutely key for those systems is the end-to-end full and reliable information. It can be generated by both vehicles and transport infrastructure, and it is the first link in the dataflow chain. Majority of that information is concise and over-abundant in volume. Examples are satellite navigation, road surface sensor (roadway weather information), inductive loops embedded in road lanes, etc. [11].

The information generated by the infrastructure and vehicles has to be correctly disseminated. Converged wireless networks are the centrepiece in this process. From the technical point of view there are wide-area and local-area networks. Currently local-area wireless networks are in intensive development as far as V2V and V2I communications are concerned.

Once the information reaches its recipient it is processed. The information processing is one of the highest priority and urgent tasks in hand of standardisation offices. As the technical solutions are not integrated, there are no developed data exchange protocols and solutions delivered by different software houses are numerous. This state of affairs constrains the interoperation between telematic systems located within a given area. This violates the very idea of ITS and diminishes its impact on transport processes [12].

Should the aforementioned issues of information processing be ignored and it was perceived as unambiguous, it then comes as "mine of knowledge" for the traffic decision support & demand management systems. In the author's opinion, only then ITS objectives could be met.

The information flow is represented holistically, i.e. including available technologies and utilitarian factors in Fig. 1.

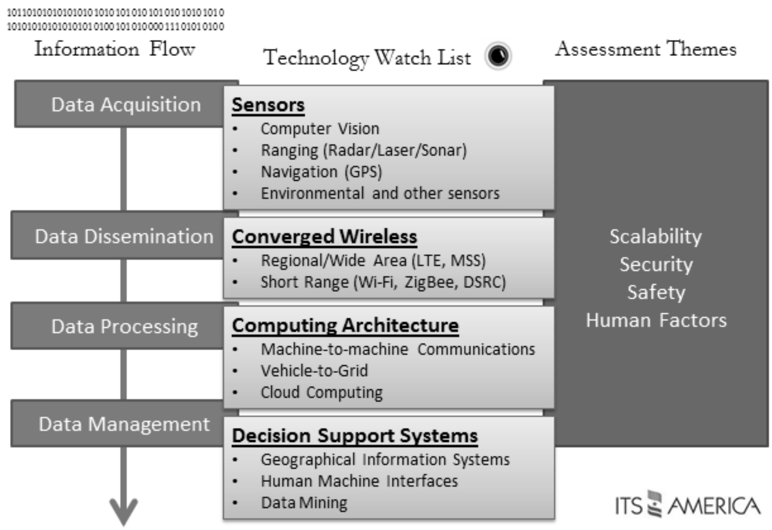


Fig. 1. Information flow in ITS systems. Source: [3].

2.1 Sources of Information in ITS systems

Vehicles are the greatest group of information sources and recipients in ITS systems. Telematics-enabled, independently of the driver, they can feed the location coordinates, the vehicle maintenance data and the transported cargo information to a logistic base. The information acquired by vehicle sensors can also serve purposefully to other drivers within the closest proximity to the vehicle. V2V communications could find its use here, to feed the data about the current speed or to make an e-call.

Another group of information sources comprises road infrastructure sensors. The data from those sensors can be transmitted in a twofold manner. On the one hand, normally over wires, the information is fed to management centres, on the other hand, over wireless and via access points it is fed to vehicles driving past the sensor. Infrastructural sensors can be used in a multifaceted manner. The most popular applications are the assessment of lane occupancy and of traffic volume. Apart from these two, important are roadway weather information systems and air quality sensors feeding the information about hazards created by e.g. road surface conditions. They are normally complex systems transmitting a processed information about road surface conditions based on many sensors and processing algorithms. Another popular use of ITS includes electronic toll collection on toll roads. An interesting, yet not the only, application for infrastructural sensors is predicting the congestion and rapidly changing road hazards. Intelligent IP cameras are best-suited for this role, whose built-in video filters can generate warning and emergency signals.

Another crucial source of information are satellite positioning systems. At this point in time, there are four major systems. American GPS, Russian GLONASS, European GALILEO and Chinese COMPASS [5, 7]. The American GPS is the most popular solution. It uses the NMEA-0183 protocol (*National Marine Electronics Association*) capable of sending data at 4800 baud.

2.2 Data Feed to End-Consumers

The acquired and processed telematics-based information is transmitted via various communication interfaces [1, 6]. That information is displayed in vehicles, by variable-message signs along roads, on public transport stops and at travel information centres [10, 13].

The information displayed to drivers should influence their decisions concerning driving itself and choosing optimum routes. Generally speaking, that information should enhance the driving comfort as well as improve safety. The data is fed to the driver over Man to Machine (M2M) interfaces, which might include video displays of built-in touch screens, head-up displays, IVR voice communication and voice recognition systems, and other more conventional features. Nowadays the development of those systems is dictated by stringent requirements for the driver safety. Telematic driver assistance systems can automatically reduce the speed upon joining the vehicle in front or even brake before a barrier. On the other hand, studies carried out for purposes of aviation in the seventies, and later for ITS, proved that an incorrect data fed to the driver causes distraction and absent-mindedness, which in turn could potentially cause an accident.

Variable-message signs are a bright indication of ITS at work. They are a replacement for traditional notice boards. Drivers obtain numerous pieces of information helpful in driving a vehicle. What is an important feature, those solutions operate in real-time, i.e. change the displayed information depending on the current weather conditions and traffic situation. Variable-message signs also influence the traffic in areas where so far they have not been used. Examples include car parks and urban agglomeration welcome signs.

3 Information Analysis in Telematic Systems

The task of analysing comprehensively the information volumes involved in ITS systems is a complex one, depending on numerous technical factors, but above all on time-constraints imposed by the dynamics of those systems. Hence the data transmission could be classified as continuous and discrete (discontinuous). A continuous transmission is normally associated with media streaming, whilst a discrete transmission is more popular and associated with the current status reporting and controlling equipment.

3.1 Data Volume

The issue will be illustrated with an example of an abstract ITS system supporting a tram operator. As far as the issue of positioning vehicles is solely concerned, assuming the ITS-enabled area to be the agglomeration of Warsaw whose current rail infrastructure is 276.5 km long, current (February 2012) number of carriages is 837 and each carriage will feed its position in 3 second intervals, the number of messages either accumulated on-board or fed to the surveillance centre could reach 1200 per hour. In the entire enterprise (the entire rolling stock) that number soars to 1004400 messages. Thus, over 24 hours this translates to tremendous 24 million messages, which have to be transmitted, accumulated and processed let alone for documenting purposes, to prove the meeting of contractual obligations to deliver timely the transportation services.

There are two things to note here. Firstly the update time for carriage positioning. The three second interval seems small, however, since vehicles travel at 40 km/h on average, between consecutive readings the carriage covers 33 metres. This is not a satisfactory enough metric from the standpoint of vehicle position over time. Secondly, and more importantly, this is one of several sources of information. Apart from positioning information, telematic devices can generate information about vehicle parameters, number of seats taken, security etc. This means, the amount of information would have been substantially greater than the above-mentioned.

3.2 Data Latency

The phenomenon of latency is commonly associated with a network delay. This is an inaccurate statement narrowing down the issue to information transmission only. Nevertheless, message latencies depend on multiple factors which include:

- delays generated by the equipment,
- packetisation time depending on the frame size used by the protocol,

- transmission delay of individual connections,
- node waiting time,
- node buffering time to compensate for jitter,
- target node buffering time,
- target node processing time (information management centre),
- response time.

All the above-mentioned factors are characteristic of package networks. In the case of using circuit-switching networks, a substantial part of delay-generating activities would have been eliminated. Unfortunately the circuit-switching networks are the least efficient and therefore it has to be assumed that packet-switching networks would have been used. One of arguments in favour of this solution is the ubiquity of internet which is based on IP packet-switching protocol.

3.3 Data Management in the System

The information management can be carried out in either central or distributed manner. The solution selection depends on the size of and on constraints on the message processing time. The literature gives comprehensive information about advantages and disadvantages of every data management method. The examples illustrated in paragraph 3.1 reveal the scale of the problem – the provided central model is adopted. The decentralisation augurs well for simplifying the management process, however it fails to give a broad, global overview of the system.

Regardless of the method selected for the ITS system, the transmitted information has to be methodised in terms of:

- weight/priority of message,
- access privileges/publicity of message,
- message demand,
- message-based inference,
- predicting subsequent messages based on the previously received ones (in case of message loss),
- others.

4 Conclusion

Issues discussed in this paper are only part of broader deliberations over data exchange within ITS systems. Should organisations such ISO or ITU-T introduce unified standards, it would give more consistency. The lack of such standards, on the other hand, leads to the discretion in developing telematics-enabled equipment. Consequently the rolled-out solutions are incompatible and time constraints are imposed on the data exchange.

Research institutes have to answer the following questions:

- How to build the data exchange and management systems?
- To what extent should the ITS communication systems be open?
- How to assure security of the ITS telecommunication systems?
- Is the information fed by the system reliable?
- How to define and assess reliability of communication system in ITS?
- To what extent is the ITS-fed information distracting for the driver?

Furthermore, both developers and their clients should consider when designing and implementing ITS systems:

- How many messages are received?
- What size are those messages?
- What resources are used to transmit those messages?
- In what order should they be processed and transmitted?
- What data-centric systems will process gathered information?
- What criteria should be adopted when deciding about the distribution of processed information?

It is fair to say, that once those questions will have been answered over the forthcoming years, the integration and management of ITS will have become easier.

References

1. Dabrowski, A., Dymarski, P.: Fundamentals of digital transmission, script of Warsaw University of Technology, Warsaw (2004)
2. `http://www.nmea.org/`
3. `http://www.itsa.org/`
4. Haykin, S.: Telecommunications systems. WKiL, Warsaw (2004)
5. Januszewski, J.: GPS, Galileo and other satellite systems. PWN, Warsaw (2009)
6. Mielczarek, W.: Serial digital interfaces. Helion (1993)
7. Narkiewicz, J.: GPS and other satellite systems. WKiL, Warsaw (2007)
8. Nowacki, G.: Origin and meaning of Transport Telematics, vol. 1, 2. ITS Digest no. 14, 15, Warsaw (2009)
9. Nowacki, G.: Road transport telematics Collective work. Publishing House of Road Transport Institute, Warsaw (2008)
10. Signalco: Information and advertising leaflets, Krakow (2005), `http://www.signalco.pl`
11. Sumila, M., Kasprzyk, Z.: Analysis of elements of telematic systems interaction and rules of their mutual communications. Archives of Transport System Telematics 4(1), 31–35 (2011)
12. Sumila, M., Rychlicki, M.: Structural analysis of data exchange in ITS systems Logistics 4/2011, pp. 908–915, Poznan (July-August 2011)
13. TRAX Elektronik: Information and advertising leaflets, Krakow (2005), `http://www.traxelektronik.pl`

Computational Algorithms Implemented in Marine Navigation Electronic Systems

Adam Weintrit and Piotr Kopacz

Gdynia Maritime University, Faculty of Navigation
Al. Jana Pawla II 3, 81-345 Gdynia, Poland
weintrit@am.gdynia.pl, kopi@nets.pl

Abstract. The authors aim to present the problems related to computational algorithms implemented in the software used in marine navigation electronic devices and systems (GPS/GNSS, AIS, ENC/ECDIS). From the early days of the development of basic navigational software built into satellite navigational receivers it has been noted that for the sake of simplicity and a number of other reasons, this navigational software is often based on simple methods of limited accuracy. Even nowadays the navigational software is sometimes used in a loose manner, adopting oversimplified assumptions and errors such as a wrong combination of spherical and ellipsoidal calculations in different steps of the solution of a particular sailing problem. The lack of official standardization on both the “accuracy required” and the equivalent “methods employed”, in conjunction to the “black box solutions” provided by GNSS navigational receivers and navigational systems (ECDIS and ECS) suggest the necessity of a thorough examination of the issue of sailing calculations for navigational systems and GNSS receivers.

Keywords: rhumb line, great circle, ECS/ECDIS, GIS, marine navigational software, telematics, marine navigation electronic devices, WGS-84.

1 Introduction

Under the provisions of the International Maritime Organization (IMO) Electronic Chart Display and Information System (ECDIS), which provides marine Geographical Information System (GIS) application which is currently the equivalent of legitimate traditional navigational paper charts used for centuries in international shipping, in particular on merchant ships - cargo and passenger (seethe International Convention for the Safety of Life at Sea, SOLAS, Chapter V: Safety of Navigation). In view of the fact that electronic charts, ECDIS and ECS (Electronic Chart System), as well as simulation of these systems used for training of crews - navigators have widely been installed on board vessels since 2012, there are a lot of commercial products on the market today offering competitive solutions in the field of marine GIS. A comparative analysis of calculation results obtained from navigational devices and the implemented software indicates different calculation procedures used by the producers, and thus obtaining at the output different results for the same input data.

J. Mikulski (Ed.): TST 2012, CCIS 329, pp. 148–158, 2012.

What is more, the instruction manuals, ‘as a rule’, lack of official information on the algorithms used in the implemented software. Noticeable serious mistake is the simultaneous use of different mathematical formulae originating from different geometric models, e.g. spheroidal and spherical in case of a global modeling the Earth's surface in computational algorithms used in the navigation software, and even in the professional literature. This leads to incorrect calculation of navigational parameters.

The analysis of the navigational calculations indicates the use of different calculating algorithms based on the same geometric model as well as various geometric models of the Earth (spheroid, sphere, Euclidean plane, triaxial ellipsoid, etc.) in the algorithm used without defining the boundary conditions (criteria) for their use and change today in modeling locally and globally. Numerical tests demonstrate that the difference in calculations of the distance based on spherical and spheroidal model with geodesics used in marine GIS at sea today is approximately 0.5%, which actually exceeds 15 Nm at distances ranging from about 6000 Nm to the maximum geodesic distance, i.e., between positions of antipodal points. Despite getting in applications satisfactory accuracy of calculations, the existence of approximating, differing in the structure of implemented calculation procedures leads to significantly different results of distance and angle calculation, which plays an important role in theoretical solutions and applications requiring high accuracy and precision.

Original contribution affects and verifies established views based on approximated computational procedures (mainly power series) used in the software of marine navigational systems and devices. It also lets to assess the quality of contemporary algorithms used in practical marine applications. Currently, on the pages of a reputable scientific magazine *Journal of Navigation*, there is a discussion on the problem of calculation procedures for marine navigation (great ellipse sailing (GES), a rhumb line sailing (RLS), great circle (GC)), as evidenced by works published by researchers from different countries and institutions. The authors also responded in their work to a new algorithm recently proposed in the literature (see: [7] and [8]), presenting the results of own research. It should be noted that an important step in the solution included in the above-cited work is simplification by the omission of the expansion part into power series of mathematical solutions, previously known from the literature, i.e., [9, 11], and reliance in the explanatory memorandum of application, in particular, on the amount of the available processing power of modern calculating machine (processor). In the authors’ opinion this criterion is relevant from a practical point of view, but temporary, given the growth and availability of computing power, including GIS. Calculating speed measured by the parameter CPU (Central Processing Unit called Speed Measurement) should not restrict the available accuracy, and this takes place in the proposed algorithm [7].

Vincenty’s algorithm allowing calculation of geodesic distances along lines of rotational ellipsoid is taken as a reference point for verification and determining the accuracy of the calculations. Numerical analysis associated with the power series development involves giving the solution to elliptic integrals of the second kind $\int \sqrt{1 - k^2 \sin^2 t}\, dt$, which occurs when calculating the length of the arc of a great ellipse of a spheroid in sailing called great ellipse sailing. The second of the arguments supporting the use of the above algorithm says about the threshold accuracy established in practice of 1 Nm in the global modeling and hence no need, according to the authors of the cited paper, to obtain greater accuracy than the proposed computational algorithm. The authors claim that this argument is, in principle, questionable but acceptable for many applications. The authors in the journal *Coordinates*[13, 14] and a research paper presented at the International Conference *TransNav* in 2011 [12] drew attention to the merits of navigational calculations based on geodesics, both locally and globally, stressing the importance of the function of curvature of the modeling surface. The validity of the discussed problem is also related to the fact that at present navigation algorithms based on the lines being arcs of great ellipse or a geodetic instead of a rhumb line used in practice for hundreds of years and cartographically associated primarily with conformal Mercator's projection is used more and more frequently. In addition, it should be noted that there are no standards on the use map projections in created present marine Geographic Information System. Today, the same system ECDIS / ECS can potentially use several different projections in which graphical presentation of the above mentioned lines used in navigation is differ significantly. This affects the correct interpretation of the information generated in GIS, particularly in navigation, by the more and more mass user of the system.

Scientific workshop employed to solve the problem makes use of various tools, i.e. of differential geometry, marine geodesy (marine navigation), analysis of measurement error, approximation theory, modeling and computational complexity, mathematical and descriptive statistics, mathematical cartography. Geometrical problems are important aspect of the tested models which are used as the basis of calculations and solutions implemented in contemporary navigational devices and modern electronic chart systems.

2 New Meridian Arc Formulas for Sailing Calculations in ECDIS -Navigational GIS

This chapter presents simple compact formulas for the computation of the length of the meridian arc. The proposed alternative formulas are to be primarily used for accurate sailing calculations on the ellipsoid in a GIS environment as in ECDIS and other ECS. Their validity and effectiveness in terms of the accuracy achieved and the CPU time required are assessed and compared to standard geodetic methods. The results of this study show that the proposed formulas are simpler, shorter and more than twice as fast as other geodetic methods of the same accuracy, used for sailing calculations on the ellipsoid, such as “rhumb-line sailing” and “great elliptic sailing”.

2.1 Preliminaries

The calculation of the length of the arc of the meridian is a fundamental element for many geodetic and navigational computations for precise positioning and reliable route planning. The precise calculation of the arc of the meridian has been a topic of on going research that started in the 18th century as a scientific debate between

British and French scientists on the proper ellipsoidal earth models proposed by Isaac Newton (1643-1727), Christian Huygens (1629-1695) and J.D. Cassini (1624-1712). The scientific interest and research on the calculation of the length of the arc of the meridian is still vivid, Bowring (1983) [7].

In traditional navigation, the computations are usually simplified by the use of a spherical Earth model and the assumption that the length of one minute of arc on the meridian is equal to the international nautical mile (1852 meters).

The discrepancies between the results on the spherical and the ellipsoidal model of the earth are in the order of 0.27% according to Tobler (1964), and in the order of 0.5% according to Earle (2006) [2]. In reality these discrepancies can exceed 13 nautical miles (about 24 km) for a number of common navigational routes. An example of such a discrepancy is shown through the calculation of the shortest navigational distance from a departure location in the west coast of USA such as the entrance of San Francisco bay ($\varphi = 37^\circ\ 45.047'N$, $\lambda = 122^\circ 42.023'W$) to a destination point in Yokohama Harbour, Japan ($\varphi = 34^\circ\ 26.178'N$, $\lambda = 139^\circ\ 51.139'E$). This calculation on the spherical earth model using spherical trigonometry and the classical assumption that 1 minute of a great circle arc is equal to the international nautical mile (1852 meters) yields a distance of 4489.9 nautical miles. The calculation of this distance on the WGS-84 ellipsoid, using very accurate methods for the calculation of long geodesics, as the method of Vincenty (1975), yields 4502.9 nautical miles.For this example the difference in calculated distances on the spherical model from those on the ellipsoid is 13 nautical miles (~24 km) [7].

In traditional navigation the calculations of shortest navigational distances are carried out on the “navigational sphere” which has the property that one minute of a great circle arc is equal to one nautical mile. Theoretically slightly better accuracies could be achieved with the use of the auxiliary geodetic sphere with radius equal to the semi-major axis of the WGS-84 ellipsoid, and the calculated results be transformed from meters to international nautical miles. Nevertheless, in practice there is not significant discrepancy between the calculations on this auxiliary geodetic sphere and the navigational sphere. Calculations of shortest navigationaldistances should not be carried out on the auxiliary geodetic sphere of Gauss, which has a radius equal to the geometric mean of the radii of curvature of the meridian and the prime vertical at the mean latitude of the two points. The sphere of Gauss is used only for very short distances and consequently, when applied to navigational distances, that normally are not very short, result in big errors.

Despite these discrepancies the use of the spherical model in traditional navigation for most practical purposes is considered satisfactory. Nevertheless for the case of sailing computations in GIS navigational systems such as ECDIS the computations must be conducted on the ellipsoid in order to eliminate these errors but without seeking the sub meter accuracies pursued in other geodetic applications. Seeking extremely high accuracy for marine navigation purposes does not offer any real benefit and requires more computing power and processing time. For these reasons and before proceeding with the adoption of any geodetic computational method on the ellipsoid for sailing calculations it is required to adopt realistic accuracy standards in order not only to eliminate the significant errors of the spherical model but also to avoid the exaggerated and unrealistic requirements of sub meter accuracy.

2.2 The Length of the Meridian Arc in Sailing Calculations

The calculation of the length of the arc of the meridian is a basic prerequisite for many accurate sailing calculation methods on the ellipsoid concerning both Rhumbline Sailing (RLS) and shortest sailings on the ellipsoid such as Great Elliptic Sailing (GES). A lot of specific papers present in detail the advantages and benefits of these methods [1, 2, 7] and [8].

It is noted though that in certain sailing calculation methods it is not necessary to calculate the length of the meridian arc. Typical examples of these methods concern:

i) RLS calculations by the employment of the general formulas of the Mercator projection and isometric latitude;
ii) calculation of shortest sailings paths on the ellipsoid by a geodetic inverse method such as the Andoyer-Lambert method from the Admiralty Manual of Navigation.

RLS calculations employing direct formulas on the ellipsoid, which require the calculation of the length of the arc of the meridian [1], are simpler than those employing the Mercator projection formulas and isometric latitude. In addition the formulas on the ellipsoid provide more flexibility for the solution of the direct problem for the calculation of the geodetic coordinates of an unlimited number of intermediate points for the purpose of the display of RLS routes on the electronic chart of the ECDIS and ECS systems.

Various numerical tests and comparisons show that discrepancies in the computed distances between the “geodesic” and the “great elliptic arc” are practically negligible for marine navigation [2, 7, 15]. Moreover GES calculations are much simpler and straightforward and can be easily implemented in navigational software. They provide the same and in some cases, higher accuracy than other methods and formulas for sailing calculations on the ellipsoid. An example is that GES calculations provide more accurate results than the Geodesic inverse solutions with the Lambert method.

GES calculations can be also used for the precise calculation of the geodetic coordinates of anunlimited number of intermediate points along the great elliptic arc, and thus be implemented in GIS navigational systems (ECDIS and ECS) for the display of navigational paths on the electronic chart.

2.3 Formulas for the Meridian Arc Length

The methods and formulas used to calculate the length of the arc of the meridian for precise sailing calculations on the ellipsoid are simplified forms of general geodetic formulas used in geodetic applications. In this section an overview of the most important geodetic formulas along with general comments and remarks on their use is carried out. For consistency purposes and in order to avoid confusion in certain formulas the symbolization has been changed from that of the original sources.

The fundamental equation for the calculation of the length of the arc of the meridian on the ellipsoid M_0^{φ} (Figure 1), is:

$$M_0^\varphi = \int_0^\varphi R_{Md\varphi} \tag{1}$$

In (1), R_M is the radius of curvature of the meridian given by (2).

$$R_M = \frac{a(1-e^2)}{(1-e^2\sin\varphi)^{\frac{3}{2}}} \tag{2}$$

In (2), *a* and *e* are the semi-major axis and the eccentricity of the ellipsoid.

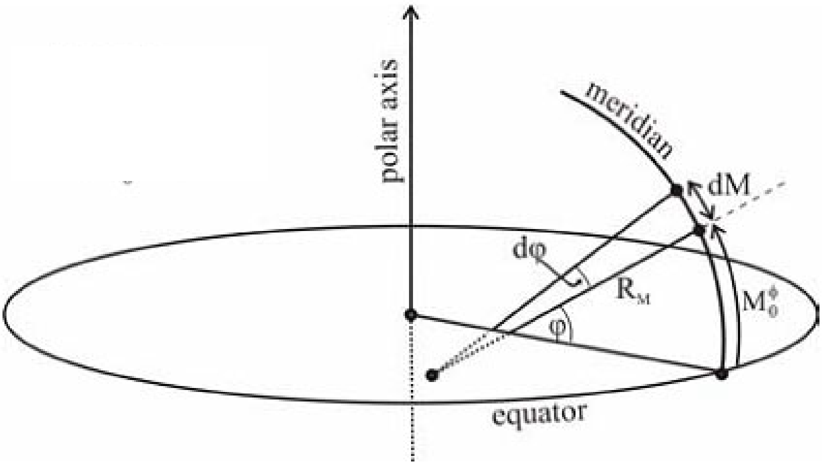


Fig. 1. The length of the arc of the meridian [7]

Replacing the value of R_M from (2) in (1), we obtain:

$$M_0^\varphi = \int_0^\varphi \frac{a(1-e^2)}{(1-e^2\sin\varphi)^{\frac{3}{2}}} d\varphi \tag{3}$$

Equation (3) can be transformed to an elliptic integral of the second type, which cannot be evaluated in a "closed" form. The calculation can be performed either by numerical integration methods, such as Simpson's rule, or by the binomial expansion of the denominator to rapidly converging series, retention of a few terms of these series and further integration by parts. According to Snyder (1987) and Torge (2001), Simpson's numerical integration does not provide satisfactory results and consequently the standard computation methods are based on the use of series expansion formulas [7]. Expanding the denominator of (3) by the binomial theorem yields:

$$M_0^{\varphi} = a\cdot(1-e^2)\int_0^{\varphi}\left(1+\frac{3}{2}e^2\sin^2\varphi+\frac{15}{8}e^4\sin^4\varphi+\frac{35}{16}e^6\sin^6\varphi\right)dx \quad (4)$$

Since the values of powers of e are very small, equation (4) is a rapidly converging series. Integrating (4) by parts we obtain:

$$M_0^{\varphi} = a(1-e^2)\left(\left(1+\frac{3}{4}e^2+\cdots\right)\varphi-\left(\frac{3}{8}e^2+\frac{15}{32}e^4\ldots\right)\sin 2\varphi + \right.$$
$$\left. + \left(\frac{15}{256}e^4+\frac{105}{1024}e^6+\cdots\right)\sin 4\varphi+\cdots\right) \quad (5)$$

Equation (5) is the standard geodetic formula for he accurate calculation of the meridian arc length, which is proposed in a number of textbooks such as in Torge's*Geodesy* using up to sin(2φ) terms, Torge (2001) [9] and in Veis' *Higher Geodesy* using up to sin(8φ) terms, Veis (1992) [11]. A rigorousderivation of (5) for terms up to sin(6φ).Equation (5) can be written in the form of equation (6) provided by Veis (1992) [11]

$$M_0^{\varphi} = a(1-e^2)(M_0\varphi - M_2 2\varphi + M_4 4\varphi - M_6 6\varphi + M_8 8\varphi + \cdots) \quad (6)$$

where

$$M_0 = 1+\frac{3}{4}e^2+\frac{45}{64}e^4+\frac{175}{256}e^6+\frac{11025}{16384}e^8+\cdots$$

$$M_2 = \frac{3}{8}e^2+\frac{15}{32}e^4+\frac{525}{1024}e^6+\frac{2205}{4096}e^8$$

$$M_4 = \frac{15}{256}e^4+\frac{105}{1024}e^6+\frac{2205}{8820}e^8+\cdots$$

$$M_6 = \frac{35}{3072}e^6+\frac{315}{12288}e^8+\cdots$$

$$M_8 = \frac{315}{130784}e^8+\cdots$$

Equation (7) is derived directly from equation (6) for the direct calculation of the length of the meridian arc between two points (*A* and *B*) with latitudes φ_A and φ_B. In the numerical tests for the assessment of the relevant errors of selected alternative formulas, we will refer to equations (6) and (7) as the Veis - Torge formulas [9], [11].

$$M_{\varphi_A}^{\varphi_B} = a(1-e^2)M_0[(\varphi_A - \varphi_B) - M_2(\sin 2\varphi_B - \sin 2\varphi_A) + M_4(\sin 4\varphi_B - \sin 4\varphi_A)\ \varphi_B - -M_6(\sin 6\varphi_B - \sin 6\varphi_A) + M_8(\sin 8\varphi_B - \sin 8\varphi_A)] \quad (7)$$

Equations (6), (7) are the basic series expansion formulas used for the calculation of the meridian arc. They are rapidly converging since the value of the powers of e is very small. In most applications, very accurate results are obtained by formula 6 and the retention of terms up to *sin(6φ)* or *sin(4φ)* and 8th or 10th powers of e.

For sailing calculations on the ellipsoid it is adequate to retain only up to *sin (2φ)* terms, whereas for other geodetic applications it is adequate to retain up to *sin (4φ)* or *sin (6φ)* terms. The basic formulas (6) and (7) can be further manipulated and transformed to other forms. The most common of these forms is formula (8). Simplified versions *Admiralty Manual of Navigation* of (8) (retaining up to $A_6$and e^6 terms only) are proposed in textbooks such as the.

$$M_0^{\varphi} = a(A_0\varphi - A_2 2\varphi + A_4 4\varphi - A_6 6\varphi + A_8 8\varphi \ldots) \quad (8)$$

where

$$A_0 = 1 - \frac{1}{4}e^2 - \frac{3}{64}e^4 - \frac{5}{256}e^6 - \frac{175}{16384}e^8 \ldots$$

$$A_2 = \frac{3}{8}\left(e^2 + \frac{1}{4}e^4 + \frac{1}{15}e^6 + \frac{35}{512}e^8 \ldots\right)$$

$$A_4 = \frac{15}{256}\left(e^4 + \frac{3}{4}e^6 + \frac{35}{64}e^8 \ldots\right)$$

$$A_6 = \frac{35}{3072}e^6 + \frac{175}{12228}e^8 \ldots$$

$$A_8 = \frac{315}{131072}e^8 \ldots$$

2.4 The Proposed New Formulas

The proposed new formulas [7, 8] for the calculation of the length of the meridian in sailing calculations on the WGS-84 ellipsoid in meters and international nautical miles are (9) and (10), respectively.

$$M_{\varphi_A}^{\varphi_B} = 111132.95251 \cdot \Delta\varphi - 16038.50861 \cdot \left(\sin\left(\frac{\varphi_B \cdot \pi}{90}\right) - \sin\left(\frac{\varphi_A \cdot \pi}{90}\right)\right) \quad (9)$$

$$M_{\varphi_A}^{\varphi_B} = 60.006994 - 8.660102 \cdot \left(\sin\left(\frac{\varphi_B \cdot \pi}{90}\right) - \sin\left(\frac{\varphi_A \cdot \pi}{90}\right)\right) \quad (10)$$

In both formulas (9) and (10) the values of geodetic latitudes φ_A and φ_B are in degrees and the calculated meridian arc length in meters and international nautical miles respectively. Formulas (9) and (10) have been derived from (7) for the WGS-84, since the geodetic datum employed in ECDIS systems is WGS-84. The derivation of the proposed formulas is based on the calculation of the M_0 and M_2 terms of (7) using up to the 8th power of *e*. This is equivalent to the accuracy provided by (8) using A_0 and A_2terms with subsequent *e* terms extended up to the 10th power since in formula (7) the terms $M_0, M_2, M_4\ldots$ are multiplied by $(1-e^2)$.

According to the numerical tests carried out, which are presented in the next section, the proposed formulas have the following advantages:

- They are much simpler than and more than twice as fast as traditional geodetic methods of the same accuracy.
- They provide extremely high accuracies for the requirements of sailing calculations on the ellipsoid.

3 Conclusion

The proposed new formulas [7, 8] for the calculation of the meridian arc are sufficiently precise for sailing calculations on the ellipsoid, because the maximum error for the calculation of the length of themeridian arc for very long distances is less than 17 meters. It is pointed out that they are about 235% faster than the alternative geodetic methods and formulas of the same accuracy. Higher sub meter accuracies can be obtained by the use of more complete equations with additional higher order terms. Seeking this higher accuracy for sailing calculations does not have any practical value for marine navigation and simply adds more complexity to the calculations. In other than navigation applications, where higher sub meter accuracy is required, the Bowring formulas showed to be approximately two times faster than alternative geodetic formulas of similar accuracy.

Despite the fact that contemporary computers are fast enough to handle more complete geodetic formulas of sub meter accuracy, a basic principle for the design of navigational systems is the avoidance of unnecessary consumption of computing power. Saving and reserving computer resources is always beneficial for the improvement of the systems effectiveness on the evolving new navigational functions and applications such as the handling of greater amounts of cartographic and navigational information, the capability for 3-D presentation.

4 Final Remarks

There are the following final remarks and conclusions:

- Setting the geometrical base of modeling structure implies the final analytic formulae essentially regarding the direct solutions or approximation;
- The shortest distance implies geodesic but geodesic does not imply the shortest distance (recall);
- CPU time requirements shows the green light for geodesic based algorithm instead of the RLS;

- Improper interpretation of cartographic projection presented in ECDIS/ECS;
- Need of applied computation procedure's unification and proper adoption of algorithms (GES);
- Transparency in used parameters (geometric and analytic) of mathematical models in local and global Earth modeling;
- The discrepancies between the results on the spherical and ellipsoidal model achieve ca. 0,3% in seagoing navigation (numerical test);
- Improper combining the spherical and spheroidal computational procedures;
- Need of defining the boundary conditions when changing the mathematical models in implemented navigational software (plane, sphere, spheroid, local surfaces of differing curvature);
- No defined specific map projections for use in navigational systems (no IMO standards) may gain improper visual perception;
- Expectations: much to yield in both practical and theoretical aspects of navigational software continuously following the nowadays requirements.

Acknowledgement. The research is supported by the Polish Ministry of Science and Higher Education research grant N N526 205540.

References

1. Bennett, G.G.: Practical rhumb line calculations on the spheroid. The Journal of Navigation 49(1), 112–119 (1996)
2. Earle, M.A.: Sphere to Spheroid Comparisons. The Journal of Navigation 59(3), 491–496 (2006)
3. Freire, R.R., Vasconcellos, J.C.: Geodetic or Rhumb Line Polygon Area Calculation over the WGS-84 Datum Ellipsoid. In: FIG Congress 2010, Facing the Challenges - Building the Capacity, Sydney, Australia, April 11-16 (2010)
4. Kos, S., Brcic, D.: Approximation Models of Orthodromic Navigation. TransNav, the International Journal on Marine Navigation and Safety of Sea Transportation 5(4), 519–526 (2011)
5. Lenart, A.S.: Solutions of Direct Geodetic Problem in Navigational Applications. TransNav, the International Journal on Marine Navigation and Safety of Sea Transportation 5(4), 527–532 (2011)
6. Nastro, V., Tancredi, U.: Great Circle Navigation with Vectorial Methods. Journal of Navigation 63(3), 557–563 (2010)
7. Pallikaris, A., Tsoulos, L., Paradissis, D.: New Meridian Arc Sailing Formulas for Calculations in GIS. International Hydrographic Review, 24–34 (May 2009)
8. Pallikaris, A., Tsoulos, L., Paradissis, D.: New calculations algorithms for GIS navigational systmes and receivers. In: Proceedings of the European Navigational Conference ENC - GNSS, Naples, Italy, May 3-6 (2009)
9. Torge, W.: Geodesy, 3rd edn. Walter de Gruyter, Berlin (2001)
10. Tseng, W.K., Earle, M.A., Guo, J.L.: Direct and Inverse Solutions with Geodetic Latitude in Terms of Longitude for Rhumb Line Sailing. The Journal of Navigation 65(3), 549–559 (2012)
11. Veis, G., et al.: Advanced Geodesy, ch. 3, pp. 13–15. National Technical University of Athens (1992)

12. Weintrit, A., Kopacz, P.: A Novel Approach to Loxodrome (Rhumb Line), Orthodrome (Great Circle) and Geodesic Line in ECDIS and Navigation in General. TransNav, the International Journal on Marine Navigation and Safety of Sea Transportation 5(4), 507–517 (2011)
13. Weintrit, A., Kopacz, P.: On geodesic based trajectories in navigation. Coordinates, a Monthly Magazine on Positioning, Navigation and Beyond VII(5), 27–31 (2011)
14. Weintrit, A., Kopacz, P.: Geodesic based trajectories in navigation. Coordinates, a Monthly Magazine on Positioning, Navigation and Beyond VII(6), 32–35 (2011)
15. Williams, R.: The Great Ellipse on the Surface of the Spheroid. The Journal of Navigation 49(2), 229–234 (1996)

Application of Closed Circuit Television for Highway Telematics

Mirosław Siergiejczyk[1], Jacek Paś[2], and Adam Rosiński[1]

[1] Warsaw University of Technology, Faculty of Transport
Koszykowa 75, 00-662 Warsaw, Poland
{msi,adro}@it.pw.edu.pl
[2] Warsaw Military University of Technology, Faculty of Electronic
gen. S. Kaliskiego 2, 00-908 Warsaw, Poland
JPas@wat.edu.pl

Abstract. The article presents questions connected with the transport telematic systems. Both analogue (CCTV) and TCP/IP controlled (CCTV IP) visual monitoring system are presented. New concepts of technical and functional solutions are also discussed, aimed to improve the operational effectiveness of integrated control and surveillance centres managing the transport telematic systems.

Keywords: visual monitoring system, transport telematics, security system.

1 Introduction

The transport telematics is a field of knowledge integrating IT and telecommunications, intended for purposes of organising, managing, routing and controlling traffic flows, which stimulates technical and organisational activities enabling higher efficiency and safety of those systems [18, 19]. Transport telematic solutions are used for a broadly defined transport safety at multiple systems and transport facilities [16].

This issue becomes ever-important in the transport perceived as a wide-area system. This applies not only to facilities with high number of staff using different transport solutions, but also to monitoring of passenger compartments (passenger safety during transport), load safety, monitoring of routes and key transportation structures (bridges, tunnels, overpasses etc.) since they are easy targets for terrorist attacks, which if successful could disrupt transport networks, hence the entire economy over substantial area.

Telematics-based security systems could be used as part of transport telematic systems. In that arrangement they assure safety during travel, which among other is the service provided by transport telematic systems. This functionality is delivered by systems installed at permanent structures of airports, railway stations [11], logistic bases, handling terminals as well as by the systems installed in moving objects (e.g. vehicles). Consequently, the security level of both the travellers and cargo increases

J. Mikulski (Ed.): TST 2012, CCIS 329, pp. 159–165, 2012.

[4, 15]. In terms of reliability analysis of those systems there are already numerous publications (concerning both the entire system [2, 8, 9, 10, 12], as well as its components e.g. power supply [14] and transmission media [13]), thus they will not be discussed.

The highway telematics entails using various IT systems along highways in order to considerably increase safety during travel and commercial transport. A range of other positive effects are generated, i.e. a lower environmental impact, a higher efficiency of transport processes through traffic solutions, a better use of road infrastructure, a stronger economic validity of highway operator business [20, 21].

Highway telematics' elements include centres controlling transport, passengers, vehicles (cars, coaches etc.), drivers and cargos. Intelligent Transport Systems, subsystems managing roads, vehicles, drivers and transport services based on real-time telecommunications create a logical sequence capable of managing moving people, vehicles and cargos under changing environmental conditions. The visual monitoring is one of the subsystems. Among other is enables to assess highway conditions by locating the area where an accident has taken place and detecting its type. It can also determine the length of any consequent traffic jams (Fig. 1).

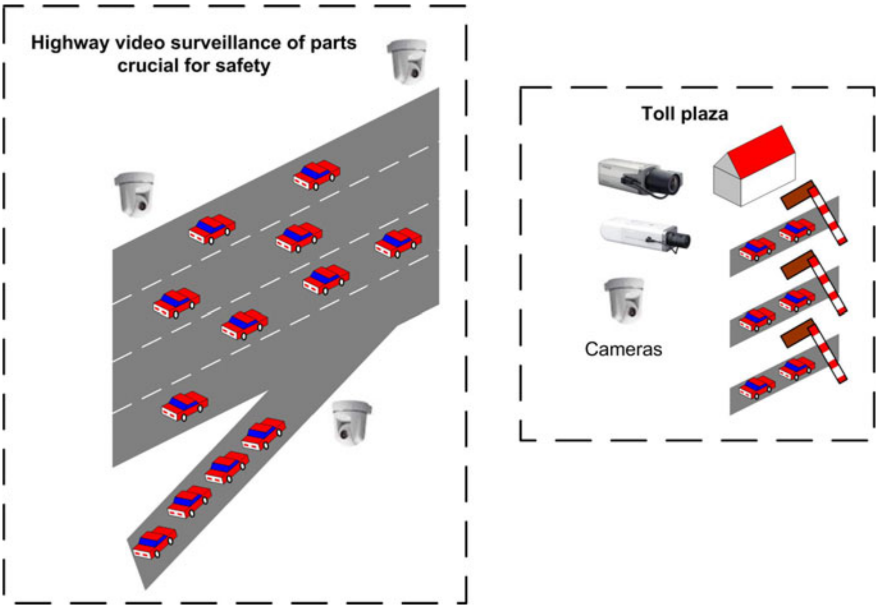


Fig. 1. Example of visual monitoring system in highway telematics

Currently there are many key solutions being implemented at visual monitoring systems concerning constructional design and organisation. Hence, the analysing of previous systems and suggesting viable changes increasing their functionality becomes of paramount importance. The article discusses issues related to telematic security systems. It particularly focuses on visual monitoring systems.

2 Characteristics of Visual Monitoring Systems Used in Transport Telematic Systems

Visual monitoring systems (CCTV - *Closed Circuit Television*) could be used as part of a transport telematic system. It is a system of hardware and software measures intended for observing, detecting, registering and signalling irregularities indicative of potential threats. They comprise (depending on the configuration) the following basic devices [1, 3, 5, 6, 17]:

- cameras equipped with adequate lenses,
- transmission media with interfaces,
- recording devices,
- monitors,
- power supply,
- other (e.g. additional illumination, uninterruptible power supplies, processing devices).

Fig. 2 illustrates a typical system composed of a maximum of 16 video cameras, a digital video recorder and two monitors (a multi viewer previews all 16 video cameras - 4-split, 9-split and 16-split screen options available, the other feeds the picture from a motion detection camera which detected a movement). The digital video recorder acts as a sequential switcher, a video splitter and a recording system with movement detection. This enables optimisation of the recorded image. It can also be equipped with an Ethernet card enabling a remote control via internet (intranet) over TCP/IP protocol. Such solutions are used at toll booths along toll roads.

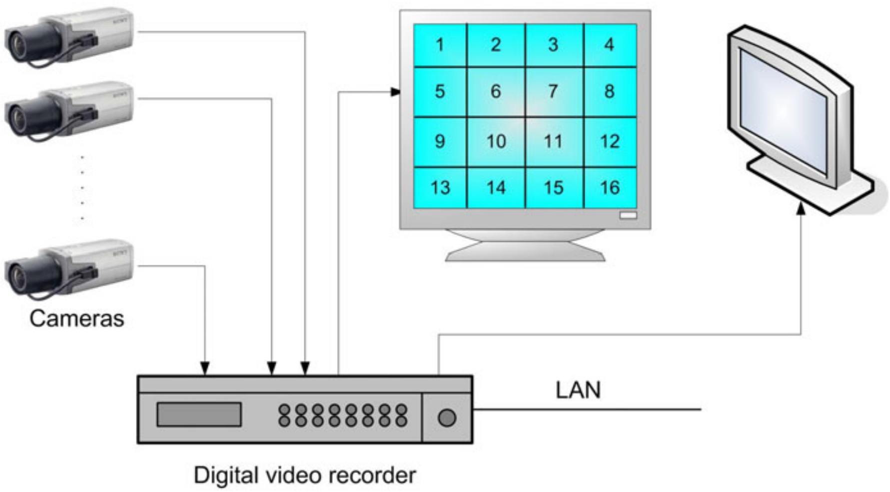


Fig. 2. System including 16 video cameras and a digital video recorder

Digital video recorders from different toll plazas can be connected via optical fibres with the central highway agency control centre. Advantages of this solution are:

low attenuation, resilience to external magnetic fields, confined energy radiation. It is used at facilities located at a considerable distance from each another. It requires converter devices to transform electric signals into optical and reverse.

A rapid IT development [7] made possible to connect directly video cameras with computer networks. Hence, images from cameras installed at different protected locations could be previewed using a computer with appropriate software connected to a computer network. The computer is equipped with built-in video server and has its own IP address. Small and local visual monitoring systems (due to high costs) will continue to use systems based on digital video recorders using coaxial cables. Scattered complexes and buildings equipped with modern telecommunication installations more often use visual monitoring system over TCP/IP. This protocol is called CCTV IP (*Closed Circuit Television Internet Protocol*). Image processing devices have their own IP addresses. One could use off-the-shelf, standard analogue elements such as video cameras thanks to video servers. The signal could be recorded through traditional video recorders and network video recorders.

3 CCTV IP Visual Monitoring System in Highway Telematic Systems

Permanent transport structures (e.g. toll plazas) which comprise one or more buildings could be potentially equipped with several solutions of visual monitoring systems. They are usually small structures (already built) which do not require many video cameras. Thus digital video recorders operating a network interface can be used as well as analogue cameras with TCP/IP network connectivity via a video server. Then such system can be connected to the surveillance centre. Should the structures be large or new-built already installed with adequate telecommunications infrastructure, the CCTV IP is the preferred solution. Fig. 3 illustrates schematically a CCTV IP system.

The surveillance centre plays a superior role to local operators' centres. It manages the entire CCTV IP system (e.g. installing new IP cameras and video servers including configuration) and also stores device-specific information from IP cameras (and/or video servers). Additional recording capabilities are generated by network video recorders.

The image from CCTV IP cameras is recorded via dedicated software. It is usually intended for intelligent video monitoring. Owing to its capabilities, video surveillance centres can be created anywhere, regardless of camera locations (both network and analogue via a video server). The movement and objects can be detected (e.g. missing object detection capability, object tracking capability etc.) thanks to intelligent image processing and advanced filters. That software usually enables the PTZ control (Pan/Tilt/Zoom), provided the camera has adequate hardware capabilities. The built-in privacy zone masking enables to cover unwanted or private areas of the image.

A CCTV IP visual monitoring system offers capabilities unavailable with analogue counterparts:

- video transmission over long distances has no impact on the image quality,
- multiple recording and copying of the same footage does not damage the quality of recorded image,
- video signal recorded once by the camera can be used any number of times regardless of location,
- bi-directional communication with the camera enables not only its configuration, but also other functions (e.g. intelligent cameras),
- motion detection and alarm capabilities of the camera itself (intelligent cameras),
- data buffering of image recorded by the camera,
- *Power over Ethernet* capabilities.

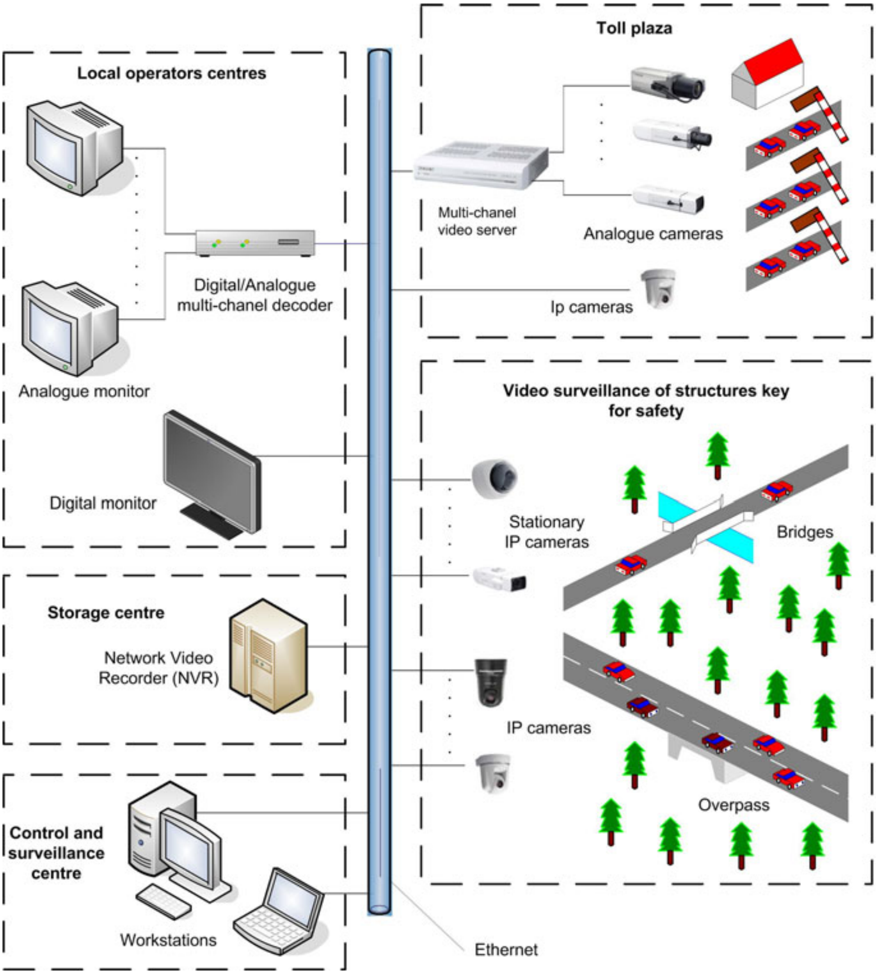


Fig. 3. Schematically presented example of a CCTV IP system

When designing modern visual monitoring systems used as part of transport telematic systems one should consider the following:

- video streaming directly from the camera,
- image resolution in megapixels,
- cameras with thermovision capabilities (thermovision cameras),
- bulk memory functionality available in cameras and video servers,
- Power over Ethernet (PoE),
- quality of service (QoS),
- software with the image preview functionality for mobile devices (e.g. smartphone),
- ONVIF-compliant (Open Network Video Interface Forum).

Those guidelines mean that the designed visual monitoring system will have better operational parameters and it would be a better platform for future updates (particularly so if the equipment used was ONVIF-compliant).

4 Conclusion

The issues discussed in this paper concerned visual monitoring systems used in transport telematic systems (especially in highway telematics) and their potential to increase the travel safety. At present, the developing and deploying of these or similar solutions is of paramount importance to Poland. Single implementations are insufficient. The European and global trends also have to be followed as they aim to create integrated control and surveillance centres of transport telematic systems, particularly so as far as telematic security systems are concerned. This approach will help to install systems with high operational efficiency. Of course, one should always bear in mind emergency intervention services and operational procedures should any danger arise (especially in case of accidents and terrorist attacks).

As part of further research, the authors plan to conduct a reliability analysis of CCTV IP visual monitoring systems. They plan to put emphasis on bulk memories in cameras and video servers. The other research direction will concern electromagnetic compatibility issues related to those systems.

References

1. Dunstone, T., Yager, N.: Biometric System and Data Analysis. Springer (2009)
2. Dyduch, J., Paś, J., Rosiński, A.: The basic of the exploitation of transport electronic systems. Technical University of Radom, Radom (2011)
3. Harwood, E.: DIGITAL CCTV. A Security Professional's Guide. Butterworth Heinemann (2007)
4. Holyst, B.: Terrorism, vol. 1, 2. LexisNexis Publishing House, Warsaw (2011)
5. Kaluzny, P.: Closed Circuit Television. WKiL, Warsaw (2008)
6. PN-EN 50132-7:2003 standard: Alarm systems. CCTV surveillance systems for use in security applications. Application guidelines

7. Collective work: Handbook of data communications, vol. 1, 2, 3. IDG Publishing House, Warsaw (1998, 1999, 2002)
8. Rosiński, A.: Design of Electronic Protection Systems with Utilization of the Method of Analysis of Reliability Structures. In: Nineteenth International Conference on Systems Engineering (ICSEng 2008), Las Vegas, USA (2008)
9. Rosiński, A.: Reliability analysis of the electronic protection systems with mixed m–branches reliability structure. In: International Conference European Safety and Reliability (ESREL 2011), Troyes, France (2011)
10. Rosiński, A.: Reliability Analysis of the Electronic Protection Systems with Mixed – Three Branches Reliability Structure. In: International Conference European Safety and Reliability (ESREL 2009), Prague, Czech Republic (2009)
11. Siergiejczyk, M., Gago, S.: Concept of monitoring and surveillance system at a railway juntion. In: VI International Scientific and Technical Conference LOGITRANS 2009, Szczyrk (2009)
12. Siergiejczyk, M., Rosiński, A.: Optimisation of transport telematics electronic systems operational process. Polish Journal of Environmental Studies. Stud. 20(5A) (2011)
13. Siergiejczyk, M., Rosiński, A.: Reliability Analysis of Electronic Protection Systems Using Optical Links. In: Zamojski, W., Kacprzyk, J., Mazurkiewicz, J., Sugier, J., Walkowiak, T. (eds.) Dependable Computer Systems. AISC, vol. 97, pp. 193–203. Springer, Heidelberg (2011)
14. Siergiejczyk, M., Rosiński, A.: Reliability Analysis of Power Supply Systems for Devices Used in Transport Telematic Systems. In: Mikulski, J. (ed.) TST 2011. CCIS, vol. 239, pp. 314–319. Springer, Heidelberg (2011)
15. Siergiejczyk, M., Rosinski, A.: Selected transport telematics elements in public safety assurance. In: IV International Scientific Conference Public Security, PS 2011, Poznan (2011)
16. Siergiejczyk, M.: Maintenance Effectiveness of Transport Telematics Systems, Scientific Works of the Warsaw University of Technology, Transport Series, No. 67, Warsaw (2009)
17. Tistarelli, M., Li, S.Z., Chellappa, R.: Handbook of Remote Biometrics for Surveillance and Security. Springer (2009)
18. Wawrzynski, W.: Telematics in academic discipline of transport. Typescript. Faculy of Transport of Warsaw University of Technology, Warsaw (2003)
19. Wawrzynski, W.: Transport telematics - conceptual range amd field of application. Communications Digest no. 11, Warsaw (1997)
20. Wawrzyński, W., Siergiejczyk, M., et al.: Final Report on Grant KBN 5T12C 06625. Methods for Using Telematic Measures to Support Realisation of Transport Tasks, Warsaw (2007)
21. Website of U.S. Department of Transportation, Federal Highway Administration, information on ITS systems and control centres, http://www.fhwa.dot.gov

Theoretical and Practical Investigations of Railway Switch Drives

Jakub Młyńczak and Jerzy Łukasik

Silesian University of Technology, Faculty of Transport
Krasinskiego 8, 40-019 Katowice, Poland
{jakub.mlynczak,jerzy.lukasik}@polsl.pl

Abstract. The paper shows an exemplary process of the investigation of a switch drive for the compliance with the Polish requirements. The study is aimed to propose a procedure and clarify the requirements for switch drives. The switch drive is a system which regulates and governs directly the safety of the transportation of people and goods. The lack of explicitly expressed requirements is the source of the problems and the difficulties in the process of the approval to use and the use itself of switch drives. During the research process should also be forecast the application of telematics solutions.

Keywords: approval tests, switch drive, standards.

1 Introduction

The switch drive is a very important unit in the railway traffic control system. However, a huge problem is still the issue of the standardisation of the requirements imposed for the equipment of this type. On the other hand, the lack of the requirements clearly specified by the consumer and the user and also those imposed by certifying bodies leads to a dualism of such requirements. It often turns out that the requirements from a potential user are not in accordance with those presented by the approving body and those with the procedures and standards used by the switch drive manufacturer in his production process. In this paper, an exemplary process of the investigation of the switch drive serves as an example of the attempt to standardise the efforts in this field.

In accordance with the principle adopted by the Office of Rail Transportation, the process of obtaining the Approval Certificate for the newly developed equipment for railway traffic control at the station includes the following stages:

a) check and verification of the results of the testing performed up to now with regard to the compliance with the requirements applicable in the Polish railway system;

b) tests aimed to check:

- correct equipment operation (in this case by the full testing of the drive),
- interoperability of the drive with the railway traffic control systems (in this case by the safety and functionality tests in the laboratory on a representative sample of the systems),

J. Mikulski (Ed.): TST 2012, CCIS 329, pp. 166–174, 2012.

c) use tests in the order to assess the operation and the behaviour of the equipment during its use.

In this paper, an attempt of the assessment of an exemplary switch drive is presented with regard to the possibility of using it in Polish conditions, and by using the resulting data, the development of an opinion, which would constitute one of the formal documents needed to apply for the issue of the timely Approval Certificate by the Office of Rail Transportation for use by Polish railways.

To achieve this goal, an analysis of the drive documentation is necessary, and first of all, the documentation of operation and maintenance, technical conditions of the manufacture and the commissioning, test reports delivered by the manufacturer, and the assessment of the system on the basis of the drive functionality characteristics with regard to the applicable requirements in this field. In modern railway traffic control technology is still missing telematics solutions, particularly in the diagnostics devices. And in the case of switch machine the telematics diagnostic system would be particularly needed.

2 Description and the Use of the Drive

The electrically powered switch drive (hereafter called the "drive") most often is made in two versions: as the "right-hand side" and the "left-hand side". It is intended for the installation on turnouts. The "right-hand side" or "left-hand side" version is determined by the location of the drive on a given side of the turnout when looking from the facing point. An exception is that the drive may be rearranged on site from the "right-hand side" to the "left-hand side", or vice versa.

The drive is intended to be supplied with a voltage of 3 x 400 V AC, 50 Hz. Also, one-phase 230 V AC, 50 Hz drives are used. The slide travel is usually 220 mm, but other slide travel values are also used.

The drive is composed of the subassemblies of shifting, keeping, switching, and control.

The shifting subassembly consists of electrical driving motor, gear transmission (toothed, worm, or hydraulic type), and overload coupling. Depending on the version, the keeping subassembly is of a non self-locking or self-locking type.

The switching subassembly consists of both the mechanical and the electrical parts, which are controlled by a guide disk and a switching off lever.

The drive control subassembly is composed of the system of control slides and levers, mechanically tied with the electrical part of the switching subassembly.

The drives for switches are located outside or inside the track way.

The drive is connected to the movable parts of turnouts and derailing ramps with a turnout switch or derailing ramp-setting rod.

The control of the position of the switch blades is carried out with the control rods connected directly to the movable parts of the turnout, and then with the drive control system.

The non-self-locking drives maintain the movable parts of the turnout in their limit position with the keeping force. These drives provide the possibility of signalling a rail vehicle approaching from the wrong direction.

In Polish conditions, there is also a term of “non-self-locking drive of increased keeping force”.

Self-locking drives maintain the movable parts of the turnout in their limit positions with the force exceeding 25 kN. These drives do not provide the possibility of signalling a rail vehicle approaching from the wrong direction. This signalling is achieved by individual controllers.

In emergency conditions, it is possible to switch the drive with a crank. While switching the drive with a crank, the power supply is temporary disconnected from the motor for the time needed to put the crank into the drive.

3 Assessment of the Fulfilment of the Technical Requirements by the Switch Drive

The technical assessment of the switch drive is normally performed on the basis of the documentation and the technical descriptions, the test programmes, and the test reports delivered by the manufacturer. The analysis and the assessment of the construction/functional characteristics and safety should be carried out with reference to the user’s requirements (e.g. PKP PLK S.A. [Polish State Railways Polish Railway Lines, joint stock company]) specified in standards, requirements, and norms listed in the literature [1] through [13].

The verification of the completeness of the qualification tests for the switch drive is conducted on the basis of the scope of investigations arising from the Regulation [9].

3.1 Functionality Assessment

After analysing the technical documentation of the drive, it is necessary to find out whether the drive performs all the functions related to the switch, such as:

- execution of a specified shifting of the switch blades (derailing ramp, movable frog point) within a given time,
- keeping the blades in their limit position with a specified force,
- providing the control of the limit position,
- signalling the trailing point movement in the non-self-locking drive by the loss of the control of the limit position.

The force of keeping of the switch drive in the self-locking version is at least 25 kN. This force is maintained mechanically and it does not depend on the drive-setting force. This guarantees the stability and keeping this parameter all the time.

On the basis of the above factors, it is normally found that the drive in the self-locking version with the keeping force of 25 kN can be used on lines of high speeds (up to 250 km/h) on the turnouts appropriately adapted to those speeds.

Besides, the drive is adapted for heavy-duty operations.

Correct drive operation is guaranteed at the supply voltage: with 3-phase motor, 3x400 V (+10%, - 30% of the rated value).

The full assessment of the drive functionality will be possible after a trial use in actual operating conditions in the PKP PLK S.A. railway network.

3.2 Assessment of the Resistance of the Drive to Environmental Exposure

The drive documentation made available by the drive manufacturer allows to assess the impact of the environmental exposure on the operation of the drive. The analysis of the documentation has shown that the drive meets the PN-EN 50125-3 environmental standards and the PN-EN 60068 standard series for appropriate tests.

The investigation of the drive for resistance to environmental exposure included the following tests:

- Checking of the resistance to cold. The test has been conducted at a temperature of -40° C.
- Checking of the resistance to dry heat. The test has been conducted at a temperature of +80° C.
- Checking of the resistance to the cyclic damp heat.
- Checking of the emission of electromagnetic interferences.
- Checking of the vibration strength.
- Checking of the insulation resistivity has been carried out with an instrument of the test voltage = 1000 VDC.
- Checking of the electrical insulation resistance has been carried out with an instrument of the test voltage = 2000 V/50Hz.
- Checking of the housing IP. The housing provides the international protection equivalent to IP 54 according to PN-EN 60529.
- Checking of durability.

On the basis of the above factors, the switch drive has been found as fulfilling appropriate safety levels in the railway traffic as well as the standards and requirements for this type of the equipment within the resistance to the environmental exposure.

3.3 Safety Assessment

The investigation included checking the possibility of the operation of the drive together with typical control-setting systems.

The check procedure consisted in the introduction of potential single and double damages to the tested drive. The damages simulated defects in the limit switches inside the drive occurring during switching. Then the response to these defects in the control-setting circuit was observed. As damages, the following events have been adopted:

- failing in the functioning of the system, forcing the change of the status of limit switches,
- jamming of the limit switches,
- breaks and shorting of cable wires.

On the basis of the investigations carried out on a representative sample of the control systems, it has been found that the drive can operate together with the setting systems used in the PKP PLK S.A. railway network within the systems covered with the investigations carried out.

4 The Assessment of the Fulfilment of the Formal Requirements with Regard to the Scope of the Investigations

In accordance to the purpose of its use, the electrical switch drive is subjected to the criteria described in § 1, item 2, letter a) of the Regulation of the Minister of Infrastructure of 26 September 2003 [12] on the equipment for the control of the railway traffic and the equipment for the control of driving trains and railway signalling.

On account of the above the requirements for the investigations necessary for obtaining the certificate of approval for use presented in the § 3 of the Minister of Infrastructure of 12 October 2005 [9] are applicable.

These requirements include particularly:

(a) investigations of the compliance with both appropriate Polish and European Standards,
(b) investigation of the compliance with the appropriate European technical specifications with regard to the interoperability (TSI), or in case of lacking such compliance, with the appropriate national technical specifications and standards, the list of which is established by the Office of Rail Transportation and published in the Official Journal and notified to the European Commission,
(c) tests of the functionality of the equipment and software in regular conditions,
(d) tests of the functionality in the conditions of a single damage interaction, in a specified range, including but not limited to the interaction of single damages, their detection and safe response to them, the impact of multiple damages on the drive operation, and also the same in the conditions of operation disturbances (variable power supply parameters, frequency, etc.),
(e) tests of the functionality in variable external circumstances, as climate conditions, mechanical interactions (including those of strength), electrical interactions (including electromagnetic compatibility - EMC), and the protection against unauthorised access,
(f) safety investigation for the operation in the use conditions,
(g) qualification tests for the functionality, reliability, and safety.

The overall compilation of the tests carried out is presented in the Table 1. In the first column, the requirement according to the § 3, item 1 of the Regulation [9] is defined; in the second column, the necessity of the use of an appropriate testing is pointed out with YES or NOT; in the third column, the performance of an appropriate testing is indicated with YES or NOT; in the fourth one, "positive" or "negative" result is presented, and finally in the fifth column, some notes on the testing procedure are shown.

Table 1. The overall compilation of the results of the testing and the switch drive tests

1	2	3	4	5
The scope of investigation arising from § 3, item 1 of the Regulation [43] "Equipment for railway traffic control" Item 1 - The scope of type investigation of the control and protection equipment of the railway traffic	Does it concern the drive?	Has the investigation been carried out?	Result	NOTES
a) investigation of compliance with the requirements specified in the appropriate Polish and European Standards for a given type of the equipment	YES	YES	Positive	
b) investigation of the compliance with the technical conditions and the requirements given in the appropriate national technical specifications and standardisation documents referred to in the Art. 25d of the Act on Railway Transport	YES	YES	Positive	
c) test of the functionality in normal conditions	YES	YES	Positive	
d) test of the functionality in the conditions of interaction of damages at:				
- interaction of single damages	YES	YES	Positive	
- independence of hardware elements	NO	-	-	N/A for the switch drive
- detection of single damages	YES	YES	Positive	
- response to a detected single damage and the ability to remain in a safe state	YES	YES	Positive	
- interaction of multiple damages	NO	-	-	N/A for the switch drive
- variable voltage, current, and frequency parameters	YES	YES	Positive	
e) investigation of the functionality in the external interaction conditions carried out by:				
- electrical testing	YES	YES	Positive	
- climate testing	YES	YES	Positive	
- mechanical testing	YES	YES	Positive	

Table 1. (*continued*)

- tests for the electrical interaction, including electromagnetic compatibility - EMC	YES	YES	Positive	
- strength tests	YES	YES	Positive	
- testing of the IP provided by the housing	YES	YES	Positive	
- protection against an unauthorised access	YES	YES	Positive	
f) investigation of the operation safety in the conditions of real use of the equipment	YES	YES	Positive	On-site within the field tests
g) following qualification tests				
- functionality tests	YES	NO	-	on site within the field tests
- reliability tests	YES	NO	-	as above
- safety tests	YES	NO	-	as above

The scope and the methodology of the tests and investigations carried out and delivered by the manufacturer is substantially correct. On the basis of the positive results of the tests and the investigations as well as the performed analyses, it can be stated that the requirements contained in [9] for the necessary testing listed in letters (a), (b), (c), (d), and (e) for the obtaining of the timely certification of approval for use have been fulfilled.

The investigations and tests listed in letters (f) and (g) have to be confirmed and carried out during the use tests according to the "The Program of the Use Testing of the Switch Drive".

5 Conclusion

On the basis of the positive investigation, the results and the analyses of the documentation delivered by the Client, it can be stated that:

1. The switch drive is featured with the required functionality and safety, and that the formal requirements contained in the Regulation of the Minister of Infrastructure of 12 October 2005 on "The scope of the testing necessary for the obtaining of the certificate of the approval for the use of equipment intended for the control of the railway traffic" for the obtaining of the timely certification of the approval for use have been fulfilled.

2. The switch drive intended for switching and keeping of the railway switches and the movable frog points with external locks, and the derailing ramps in their limit position can be approved for the use tests in the PKP PLK S.A. railway network.

3. The switch drive in the self-locking (non-trailable) version with the keeping force of 100 kN can be used on lines for high speeds (up to 250 km/h) in the turnouts appropriately adapted for such speeds.

4. The switch drive can operate together with 3-phase setting systems listed in the opinion (developed after investigations) and used in the PKP PLK S.A. railway network.

5. The switch drive meets the safety requirements and the requirements of the CENELEC standards with regard to the environment and electromagnetic compatibility.

As it can be noted, the process of the switch drive investigation requires a great commitment in conducting it. The switch drive is a device which operates in variable, often adverse environmental conditions. Also, the interoperation of the drive with the turnout is not always correct due to bad technical conditions of the latter. The requirements imposed for switch drives in the standards [24] do not define precisely normative requirements, but only a part of technical requirements. On one hand, such an approach is safe for those who develop norms since they do not become outdated as quickly as standards, but on the other hand, some problems of the lack of clarity of the normative requirements appear.

References

1. Elektryczne napędy zwrotnicowe, Dokument normatywny nr 86/98, CNTK, Warszawa, marzec (2000)
2. Instrukcja WTB-E10 – Wytyczne techniczne budowy urządzeń sterowania ruchem kolejowym w przedsiębiorstwie PKP stanowiące załącznik do zarządzenia nr 43 Zarządu PKP z dnia (September 9,1996)
3. Norma PN-EN 50121-4: 2002 (U) Zastosowania kolejowe. Kompatybilność elektromagnetyczna. Część 4: Emisja i odporność na zakłócenia urządzeń sygnalizacji i telekomunikacji (2002)
4. Norma PN-EN 50125-3: 2003 (U) Zastosowania kolejowe. Warunki środowiskowe stawiane urządzeniom. Część 3: Wyposażenie dla sygnalizacji i telekomunikacji (2003)
5. Norma PN-EN 50126: 2002 (U) Zastosowania kolejowe. Specyfikacja niezawodności, dostępności, podatności utrzymaniowej i bezpieczeństwa (2002)
6. Normy PN-EN 60068:2002(U) Badania środowiskowe (2002)
7. Norma PN-EN 60529:2003 Stopnie ochrony zapewnianej przez obudowę (Kod IP) (2003)
8. Obwieszczenie Prezesa Urzędu Transportu Kolejowego z dnia 08 sierpnia 2005 r. w sprawie ustalenia listy właściwych krajowych specyfikacji technicznych i dokumentów normalizacyjnych, których zastosowanie umożliwi spełnienie zasadniczych wymagań dotyczących interoperacyjności kolei (2005)
9. Rozporządzenie Ministra Infrastruktury z dnia 12 października 2005 r. w sprawie zakresu badań koniecznych do uzyskania świadectw dopuszczenia do eksploatacji typów budowli i urządzeń przeznaczonych do prowadzenia ruchu kolejowego oraz typów pojazdów kolejowych. Dziennik Ustaw Nr 212 Poz. 1772 (2005)
10. Rozporządzenie Ministra Infrastruktury z dnia 18 lipca 2005 r. sprawie warunków prowadzenia ruchu i sygnalizacji (Dz. U. nr 172 z 2005r, poz. 1444) (2005)
11. Ustawa o transporcie kolejowym (Dz. U. nr 86 z dnia, March 28 (2003))

12. Rozporządzenie Ministra Infrastruktury z dnia 26 września 2003r. w sprawie typów budowli i urządzeń przeznaczonych do prowadzenia ruchu kolejowego oraz typów pojazdów kolejowych, na które wydawane są świadectwa dopuszczenia do eksploatacji. (Dziennik Ustaw 2003 Nr 175 Poz. 1706) (2003)
13. Rozporządzenie Ministra Transportu i Budownictwa z dnia 23 grudnia 2005r. zmieniające rozporządzenie w sprawie świadectw dopuszczenia do eksploatacji typu budowli i urządzeń przeznaczonych do prowadzenia ruchu kolejowego oraz typów pojazdów kolejowych. (Dziennik Ustaw 2006 Nr 2 Poz. 23) (2005)
14. Czech, P.: Diagnosing of Disturbances in the Ignition System by Vibroacoustic Signals and Radial Basis Function – Preliminary Research. In: Mikulski, J. (ed.) TST 2011. CCIS, vol. 239, pp. 110–117. Springer, Heidelberg (2011)

Determination of the Course of Pressure in an Internal Combustion Engine Cylinder with the Use of Vibration Effects and Radial Basis Function – Preliminary Research

Piotr Czech

Silesian University of Technology, Faculty of Transport
Krasinskiego 8 Street, 40-019 Katowice, Poland
piotr.czech@polsl.pl

Abstract. A huge development of new technologies applied in the automotive industry has been observed in recent years. It is visible both in designing, manufacturing and operating new means of transport. All over the world, a number of research institutions deal with such issues. Currently, there are many tests in progress on the development of internal combustion engines which are all mainly conditioned by ecological aspects. All that is aimed at designs which are the most eco-friendly and at the same time which have the best technological parameters. To maintain a proper functioning the presently produced engines are controlled with a lot of electronic sensors installed in the vehicle. In the test, an attempt was made to determine the course of internal combustion pressure in a cylinder on the basis of registered vibration signals. Radial neural networks taught using the data achieved from the decomposition with the application of a discrete wavelet transform were used in this test.

Keywords: internal combustion engines, pressure, vibration, artificial neural networks, diagnostics.

1 Introduction

An intense development of technologies connected with electronics, information technology and related fields which has been observed in recent years has led to a significant increase in their practical applications in the automotive industry. The vehicles of today are full of most modern systems which serve to increase safety and comfort and to reduce the negative impact on the environment. At the same time the work is continued on further development of such technologies, which can be noticed in various scientific papers – such as [7, 8, 9]. The development of branches connected with state-of-the-art technologies does not limit the performance of basic research work leading to similar measurable effects such as, for example, the increase in safety by increasing the durability of the power transmission system elements [2, 4].

J. Mikulski (Ed.): TST 2012, CCIS 329, pp. 175–182, 2012.

Modern internal combustion engines of cars use integrated ignition-injection systems. By choosing the right value of the ignition advance angle the criterion of maximum turning moment should be taken into account together with a simultaneous low level of harmful substances emission in the car exhaust gas. The value of the ignition advance angle has direct influence on the course of fuel blend peak firing pressure in cylinder.

A correct process of combustion in ZI engines starts at the moment of fuel-air blend ignition. It is initiated with an ignition spark which is formed between the electrodes of an ignition plug. The created flame disperses on the whole charge. The process begins in a laminar way and then continues in a turbulent way. The speed of flame dispersion is dependent on the type of fuel, its composition, temperature and the speed of blend movement.

Three stages can be distinguished during the process:

- stage one – from the moment the ignition spark appears to the moment of pressure increase caused by the reaction of oxidation;
- stage two – from the moment of pressure increase as a result of combustion to the moment when the maximum pressure value in a cylinder occurs;
- stage three – from the moment of maximum pressure value occurrence to the termination of giving off the heat energy.

The first stage includes the ignition delay and the creation of the flame range area. Next the process of fuel oxidation may develop without the input of energy from the outside. The time length of this stage depends strictly on the properties of the blend. It is longer with the decrease of compression degree and ignition energy.

In the second stage the heat is created rapidly and in the form of a flame spreading all over the combustible blend. The time length of this stage depends mainly on:

- ingredients of the blend;
- ignition advance angle;
- degree of compression,
- geometry of the combustion chamber;
- location of the ignition plug;
- rotational speed of the engine;
- degree of blend swirl.

In the third stage, at a high temperature, the fuel after-burns in the combustion chamber.

Detailed information connected with the combustion process in internal combustion can be found in [3, 5, 13, 14].

The heat, created as a result of combustion of a part of blend, fed to the not burned part may lead to creation of next ignition centres. From the created centres the flame may spread with a speed much higher than that in normal conditions (20-40 m/s).

The effect of this is the occurrence of a disadvantageous phenomenon called engine knocking (pinging).

The reason of such phenomenon is:

- too high compression degree;
- too soon ignition of the blend;
- too low octane number of the fuel;
- overheating of the engine;
- too small degree of blend swirl in the combustion chamber;
- too big filling of the cylinder.

The intensity of engine knocking (pinging) depends on:

- degree of compression;
- shape of the compression chamber;
- location of the ignition plug;
- ignition advance angle;
- engine load.

The tendency for engine knocking (pinging) increases particularly with the increase in the engine walls temperature.

It results from the fact that such process occurs mainly during a rapid acceleration of the engine and by big loads of the engine. The engine knocking (pinging) causes an increase in dynamic loads of crankshaft system. Such loads may eventually lead to a damage of bearings. An increased heat exchange between the working gas and the walls of the combustion chamber causes the increase in the temperature of head, valves and pistons which leads to shortening of the life of the whole internal combustion engine.

The limitation of the ignition advance angle is applied to prevent the engine knocking (pinging). To limit the probability of engine knocking occurrence the following methods are used:

- use of fuel with the right octane number;
- properly chosen compression degree;
- properly designed shape of the combustion chamber.

In practice two different methods are used to identify the process of combustion. The first of them assumes a direct measurement of the pressure in a cylinder. However, this method is expensive and requires changes in the construction of the motor head. Also in this case a proper shape of the sensor tip is required as well as the choice of the location of the sensor and the maintenance of the combustion chamber geometry. A direct measurement of the pressure is the advantage of this method allowing an easy identification of the combustion process occurring in the engine.

The second method assumes the use of the vibroacoustic effects, which occur during engine knocking and which are transmitted by the engine elements. In this case the vibration signals measured on the head or on the engine frame may be used and on their basis the occurrence of engine knocking identification. The method, in which the signals are measured is easier and cheaper here, but requires the use of complicated mathematical analyses to detect the engine knocking. Undoubtedly, a huge signal deformation occurs here together with the influence of other engine parts on the shape of the registered signal.

The description of the notions connected with the identification of engine knocking in internal combustion engines can be found in [3, 5, 7, 14].

2 Description of Experiment

In the experiment an attempt was made to determine initially the course of pressure in the combustion process on the basis of registered vibration signals.

The object of test was a 4-cylinder internal combustion engine ZI (installed in a vehicle) with capacity of 1.6 dm^3 fuelled with the LPG gas.

Tests were conducted on a Bosch FLA 203 engine test bench, which made testing of vehicles with a single axle drive possible.

The following aspects were registered during tests:

- signal of vibration acceleration in a direction parallel to the cylinder axis;
- signal of vibration acceleration in a direction perpendicular to the cylinder axis;
- signal of the combustion pressure;
- signal of the crankshaft rotation angle with a marker of upper dead centre of piston position.

Vibration signals were registered with the use of piezoelectric sensors placed on the frame in the area of the fourth cylinder.

The signal of combustion pressure was measured on the fourth cylinder. To perform this a piezoelectric quartz pressure sensor type 6121 was used with a Kisler 5011 charge amplifier. The applied converter enables the measurements in particularly difficult conditions, which are to be found in the combustion chamber. It has a measurement range from 0-25MPa and may work in a temperature range from -50 do +350°C without the need of additional cooling application.

Signals were registered with the use of multi-channel measuring device which enabled synchronous sampling with high frequency. The device worked with an application created in the LabView environment.

Tests were conducted for the engine rotational speed of 1500 rpm at full load.

The registered vibration signals were decomposed with the use of a discrete wavelet transform. Tests were conducted for ten levels of decomposition.

The wavelet analysis consists of the signal decomposition and presenting it in the form of a linear combination of basic functions, called wavelets. The feature which distinguishes this method of signal analysis from other methods is a multi-stage signal decomposition, a variable resolution in time and frequency domain and a possibility to use basic functions other than harmonic functions [6, 11].

A Discrete Wavelet Transform of the signal $x(t)$ is marked as a scalar product $x(t)$ and sequence of basic functions $\psi(t)$:

$$DWT = \int_{-\infty}^{+\infty} \psi(t) \cdot x(t) dt \tag{1}$$

As a result of multi-level signal decomposition, the signal approximation is achieved on a given level a_k and a detail sum on the next levels d_l:

$$x(t) = a_k(t) + \sum_{l=1}^{k} d_l(t) \tag{2}$$

where:

d_l – a detail of signal, a high-frequency signal element,
a_k – the signal approximation, a low-frequency signal representation.

Together with the increase in the signal decomposition level the participation of details decreases, which causes that with the decrease of resolution, the contents of details in the signal approximation decrease.

A Discrete Wavelet Transform gives an opportunity of decomposition and selective reconstruction (synthesis) of the signal in the whole range of the analysis. It can be equalled to a signal filtration with a constant relative bandwidth [6, 11].

The tests assumed that the signal of combustion pressure will be determined with the use of signals achieved from vibration signals decomposition on ten levels. However, because the tests registered combustion pressure only for one cylinder the attempt of combustion process identification was performed only in a chosen range of the crankshaft angle corresponding with the occurrence of the combustion process in a given cylinder.

The tests assumed to use the neural networks taught on the data originating from vibration signals to determine the combustion pressure. Therefore, the registered signals were properly processed with the use of normalisation and scaling processes.

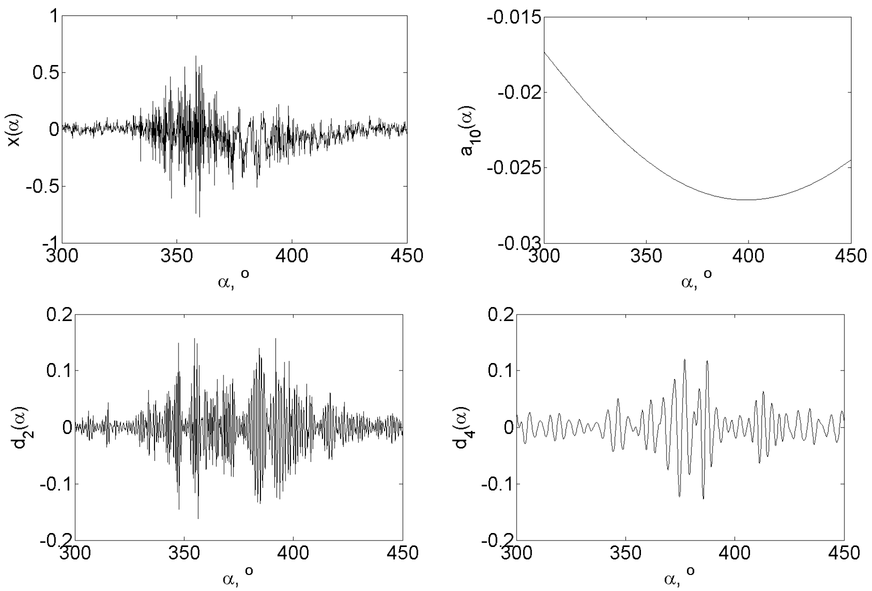


Fig. 1. The course of vibration signal decomposed with the use of discrete wavelet transform

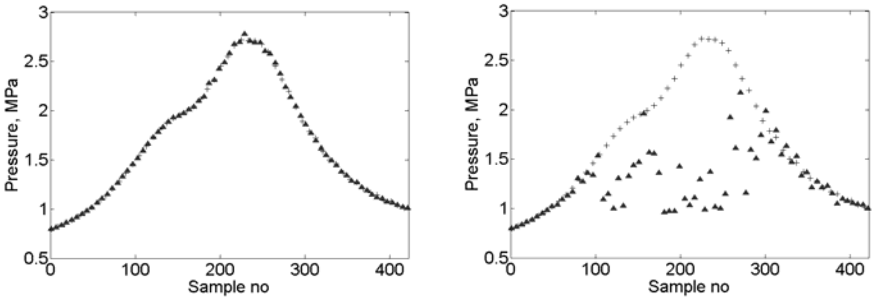


Fig. 2. Influence of coefficient γ on the identified combustion pressure value, the value chosen a) correctly b) incorrectly

An example of vibration signal which was decomposed with the use of discrete wavelet transform is shown in Fig. 1.

Signal values achieved for ten decomposition levels were the input data for the neural networks (radial basis function). The values of combustion pressure were expected to be achieved on the network output.

While using such network type, proper smoothening coefficient γ should be selected. It represents the radial deviation of Gauss functions and is a measure of the range of neurons in the hidden layer. This value, when too low, causes the loss of knowledge generalizing property by the network and if too high, prevents the correct description of details. Similarly to the radial networks, the value of γ coefficient is determined experimentally [1, 10, 12]. The performance of the network for 86 various values of γ coefficient was checked in the experiments.

Example influence of the parameter γ values on the correctness of combustion pressure determination is shown in Fig. 2.

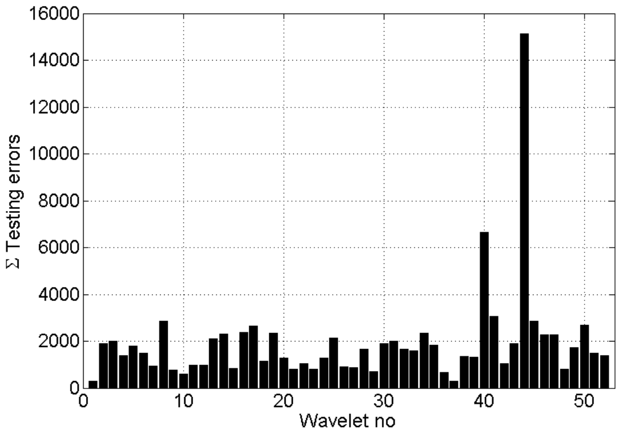


Fig. 3. Influence of basic wavelet on the correctness of the combustion pressure value for the vibration signal measured in a direction parallel to the cylinder axis

Experiments were conducted for vibration signals measured in a direction parallel and perpendicular to the cylinder axis and for 52 basic wavelets. The influence of basic wavelet choice on the correctness of combustion pressure value determination for the signal measured in a direction parallel (a_x) and perpendicular (a_y) to the cylinder axis is shown in Fig. 3 and 4.

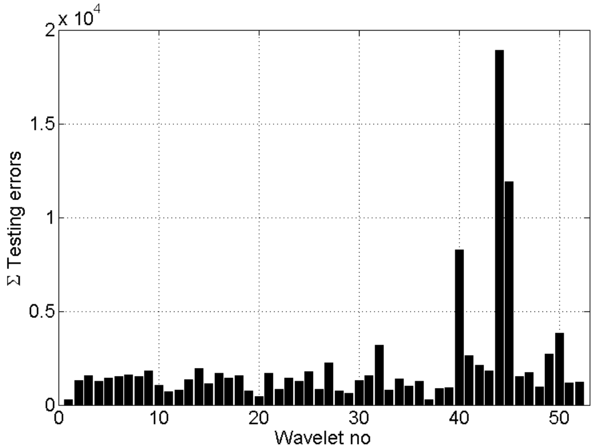


Fig. 4. Influence of basic wavelet on the correctness of the combustion pressure value for the vibration signal measured in a direction perpendicular to the cylinder axis

3 Conclusion

The article presents the results of preliminary tests aimed at the determination of combustion pressure in a cylinder of internal combustion engine with the use of neural networks, which were taught on the data coming from measured vibration signals. The results achieved in the experiment are on a satisfactory level and are the basis to continue further experiments.

The conducted experiments have shown a huge dependence of correct identification of combustion pressure on the choice of the coefficient γ for the radial neural network (radial basis function).

It was also noticed during the tests that the choice of the right basic wavelet used during the signal decomposition with the use of a discrete wavelet transform has a big influence on the determined value of combustion pressure. The best results were achieved for basic wavelets of haar and reverse biorthogonal type.

References

1. Czech, P., Łazarz, B., Wojnar, G.: Detection of local defects of gear teeth using artificial neural networks and genetic algorithms. ITE, Radom (2007)

2. Grega, R., Krajňák, J., Kaššay, P.: Impact of pneumatic flexible coupling on effective vibration value within the mechanical systems. Transactions of the Universities of Košice 3 (2009)
3. Heywood, J.B.: Internal combustion engines fundamentals. McGraw Hill Inc. (1988)
4. Homišin, J.: Tuning methods of mechanical systems by means of torsional oscillation tuner application. Pneumatyka 61(6) (2006)
5. Kneba, Z., Makowski, S.: Power supply and steering of the engines. WKiŁ, Warsaw (2004)
6. Liu, B.: Selection of wavelet packet basis for rotating machinery fault diagnosis. Journal of Sound and Vibration 284 (2005)
7. Madej, H.: Diagnosing the mechanical defects in internal combustion engines masked by electronic steering devices. ITeE, Radom (2009)
8. Mikulski, J.: The Possibility of Using Telematics in Urban Transportation. In: Mikulski, J. (ed.) TST 2011. CCIS, vol. 239, pp. 54–69. Springer, Heidelberg (2011)
9. Młyńczak, J.: Analysis of Intelligent Transport Systems (ITS) in Public Transport of Upper Silesia. In: Mikulski, J. (ed.) TST 2011. CCIS, vol. 239, pp. 164–171. Springer, Heidelberg (2011)
10. Osowski, St.: Neural networks for information processing. House of University of Technology in Warsaw (2000)
11. Peng, Z.K., Chu, F.L.: Application of the wavelet transform in machine condition monitoring and fault diagnostics: a review with bibliography. Mechanical Systems and Signal Processing 18 (2004)
12. Tadeusiewicz, R., Lula, P.: Introductin to neural networks. StatSoft, Cracow (2001)
13. Wajand, J.A.: Piston internal combustion engines of average rotational speed and high-speed. WNT, Warsaw (2005)
14. Wendeker, M.: Steering of ignition in car engine. On-board diagnostic systems of car vehicles. Lublin (1999)

Problems of ITS Architecture Development and ITS Implementation in Upper-Silesian Conurbation in Poland

Grzegorz Karoń and Jerzy Mikulski

Silesian University of Technology, Faculty of Transport
Krasińskiego 8, 40-019 Katowice, Poland
{grzegorz.karon,jerzy.mikulski}@polsl.pl

Abstract. The paper presents a review of the current state of ITS implementation in the Upper-Silesian Conurbation in Poland. A methodology for ITS architecture preparing and implementing for the Upper-Silesian Conurbation, necessary for integrated ITS planning and designing and for the conurbation transportation systems development planning, has also been proposed. Special attention was paid to the implementation problems and barriers, both of individual ITS systems as well as of their whole architecture for the Upper-Silesian Conurbation in Poland.

Keywords: ITS Intelligent Transportation Systems, Upper-Silesian Conurbation, ITS architecture, ITS Services.

1 Introduction

A turbulent development of telematic and telecommunication technologies and also of the automation and control creates unprecedented favourable conditions for *Intelligent Transport Systems* (*ITS*) designing and implementing. However, to make such activities effective and efficient, a systemic and comprehensive approach is required, ensured by so-called *ITS architecture*.

The paper is focused on a proposal of methodology for the Upper-Silesian Conurbation ITS architecture preparing, implementing and using. The analysis of basic strategic documents and of transport development programs for the Upper-Silesian Conurbation shows that now there are no assumptions or plans for such architecture development. Instead, individual cities of the conurbation, recognising the potential of ITS solutions, prepare and slowly implement partial ITS solutions, which cannot be consistent for the whole conurbation as they are aimed at resolving immediate transportation problems of individual cities.

Making an effort to meet the need of developing a comprehensive solution and using the first in the world American ITS architecture[1] as a model, the paper presents

[1] The National ITS Architecture provides a common framework for planning, defining, and integrating intelligent transportation systems. It is a mature product that reflects the contributions of a broad cross-section of the ITS community (transportation practitioners, systems engineers, system developers, technology specialists, consultants, etc.). Source: [12].

J. Mikulski (Ed.): TST 2012, CCIS 329, pp. 183–198, 2012.

– following [9, 12] – basic stages and actions related to the preparation, implementation, use and maintenance of an ITS architecture allowing to introduce such ITS solutions into the conurbation cities, which will be mutually consistent and introduced in the order appropriate to the needs – both local needs in individual cities and also the needs of the whole conurbation.

2 Review of ITS Implementation in Upper-Silesia Conurbation

The Upper-Silesian Conurbation[2] in Poland consists of a group of connected cities, which occupy the area from the city of Dąbrowa Górnicza to the city of Gliwice (see Fig. 1). The Upper-Silesian Conurbation has a strong businesses concentration which covers 18 % of the Silesian Voivodeship (Province) area (1200 km^2).

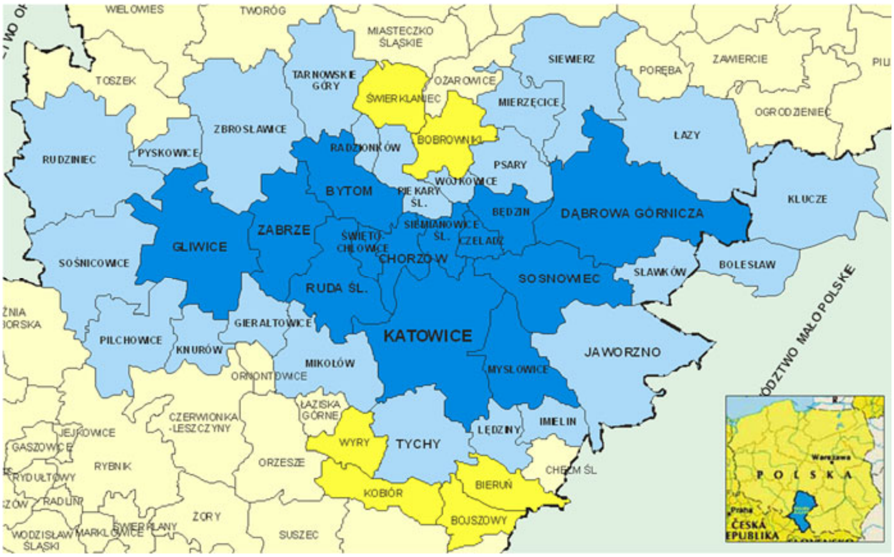


Fig. 1. Location of the Upper-Silesian Conurbation in Poland. Source: [13].

The area is inhabited by over 1.8 million residents what is approx. 38.3 % of the Silesian Voivodeship population. However, it should be emphasised that at the same time the Silesian Voivodeship is one of the biggest provinces in Poland in terms of population – 4.7 million people, that is 12.3 % of the population of Poland. The population density in the modelled area amounts to 1780 persons per 1 km^2 and is several times higher than in the Silesian Voivodeship – 379 persons per 1 km^2. With

[2] A *conurbation* is a region comprising a number of cities, large towns, and other urban areas that, through population growth and physical expansion, have merged to form one continuous urban and industrially developed area. In most cases, a conurbation is a polycentric urban agglomeration, in which transportation has developed to link areas to create a single urban labour market or travel to work area.

regard to the area size, the modelled area is 38.1 % of the Silesian Voivodeship area, which in turn is 3.9 % of the area of Poland.

Two of four *transport corridors* run through Poland (of 10 set Trans European Transport Network TEN-T): corridor III: Berlin – Wrocław – Katowice – Kraków – Lvov and corridor VI: Gdańsk – Katowice – Žilina. Historically formed important railway routes, heading north-south and east-west, are within the area of the Silesian Voivodeship, of which the three main railway lines are a part of the international E and C-E (AGC and AGTC) network: E30 and C-E30: Dresden – Zgorzelec – Wrocław – Katowice – Kraków – Medyka – Lvov – Kiev – Moscow, E65 and C-E65 Gdynia – Warsaw – Katowice – Zebrzydowice – Ostrava – Vienna. The municipal transport network in the Silesian Voivodeship is the most developed in Poland, which is evidenced by, among others, over 20-percent share in the length of all the transport lines in Poland. There are 363 km of tram tracks here (whereas in the Łódzkie Voivodeship there are 207 km and in the Mazowieckie Voivodeship – 241).

The following existing situation may be presented on the basis of strategic documents review and of experts reports in the field of transportation development and ITS development plans for the Upper-Silesian Conurbation [1, 2, 3, 10].

The deployment of intelligent transport systems (ITS) in the Silesian Voivodeships, like throughout Poland, is only at an initial phase. Substantial delay in those technologies implementation may be noticed as compared with the Western Europe countries. Also the issues related to the implementation of modern ITS solutions are not regulated in the Polish legislation. The *Directive 2010/40/EU of the European Parliament and of The Council of 7 July 2010 on the framework for the deployment of Intelligent Transport Systems in the field of road transport and for interfaces with other modes of transport* was passed on the 7th July 2010, however it has not been transposed to the national legislation.

The Silesian Voivodeship and the Upper-Silesian Conurbation have a huge potential and needs in the field of projects implementing the ITS solutions. However, the ensuring of implemented solutions interoperability becomes indispensable to increase their effectiveness. This will enable the interoperation of the equipment of various types and installed in various municipalities.

The following ITS projects are carried out now in the Silesian Voivodeship [1]:

- The City of Gliwice in the Upper-Silesian Conurbation has been implementing a project named "*The expansion of the detection system in the area of the city of Gliwice together with the modernisation of selected traffic lights, stage I*", which is aimed at "*Increasing the traffic smoothness and relieving the traffic jams and as a result the reduction of the amount of pollution emitted by vehicles and of the noise intensity*". The project scope includes the assembling/implementation of the following elements:
  - *Traffic Measurement Systems*,
  - *Automatic Collection of Traffic Data Systems*,
  - modernization of intersections with traffic lights,
  - *Traffic Management Centre*,
  - *Adaptive Traffic Lights*,
  - *Priority for buses in corridors*.

- In the cities of the Upper-Silesian Conurbation the public transport organiser (KZK GOP) has been implementing a project named "*A Dynamic Passenger Information System in the area of KZK GOP activities*" aimed at "*Implementation of an IT system improving the process of public road transport management by the use of solutions from the field of intelligent transport systems.*" The project is related to a pilot implementation of a traffic monitoring system on key transportation routes of the Upper-Silesian Conurbation together with providing the information about the current traffic situation. The project scope comprises the assembling/implementation of the following elements:
  - *Dynamic Passenger Information Management Centre*,
  - the communication infrastructure with a two-way data and control signals transmission between individual system components,
  - *Dynamic Passenger Information* displays together with the necessary stop equipment,
  - the application software for the system together with licences, including a specialised software supporting the traffic analyses in the area of the KZK GOP.
- The City of Katowice in the Upper-Silesian Conurbation has been considering the preparation of a preliminary feasibility study for ITS solutions [1],
- The public transport organiser (KZK GOP) has been considering a solution named "*Silesian Public Services Card – ŚKUP*", which basic services comprise:
  - an *e-ticket* in a "check in – check out" system, due to which the passenger pays only the actually travelled distance and the transport operator will acquire the information on the volumes of passenger flows in the public transport lines; this will be a common ticket for buses, trams, trolleybuses and trains,
  - an *e-wallet*,
  - a *user identifier* (for personalised cards),
  - an *e-signature* carrier.

3 Services and Packages of ITS Architecture

Following [9, p.1] *Intelligent Transportation Systems* (*ITS*) have been defined as: "the application of advanced sensor, computer, electronics, and communication technologies and management strategies – in an integrated manner – to improve the safety and efficiency of the surface transportation system" [3]. ITS are interrelated systems that work together to deliver transportation services. The integration of these

[3] "*Intelligent Transport Systems (ITS)* are advanced applications which without embodying intelligence as such aim to provide innovative services relating to different modes of transport and traffic management and enable various users to be better informed and make safer, more coordinated and 'smarter' use of transport networks.". '*Intelligent Transport Systems*' or '*ITS*' means systems in which information and communication technologies are applied in the field of road transport, including infrastructure, vehicles and users, and in traffic management and mobility management, as well as for interfaces with other modes of transport; – source: *Directive 2010/40/EU.*

systems requires an architecture to illustrate and gain consensus on the approach to be taken by a group of stakeholders regarding their particular systems. An ITS architecture defines a framework within which a system can be built. It functionally defines what the elements of the system do and the information that is exchanged between them, as:

Table 1. *ITS Services* based on U.S. National ITS Architecture [9, 12]

Category of ITS Service	ITS Services
Travel And Traffic Management	▪ *Pre-trip Travel Information,* ▪ *En-route Driver Information,* ▪ *Route Guidance,* ▪ *Ride Matching And Reservation,* ▪ *Traveller Services Information,* ▪ *Traffic Control,* ▪ *Incident Management,* ▪ *Travel Demand Management,* ▪ *Emissions Testing And Mitigation,* ▪ *Highway-Rail Intersection,*
Public Transportation Management	▪ *Public Transportation Management,* ▪ *En-route Transit Information,* ▪ *Personalised Public Transit,* ▪ *Public Travel Security,*
Electronic Payment	▪ *Electronic Payment Services,*
Commercial Vehicle Operations	▪ *Commercial Vehicle Electronic Clearance,* ▪ *Automated Roadside Safety Inspection,* ▪ *On-board Safety And Security Monitoring,* ▪ *Commercial Vehicle Administrative Processes,* ▪ *Hazardous Materials Security And Incident Response,* ▪ *Freight Mobility,*
Emergency Management	▪ *Emergency Notification And Personal Security,* ▪ *Emergency Vehicle Management* ▪ *Disaster Response And Evacuation.*
Advanced Vehicle Safety Systems	▪ *Longitudinal Collision Avoidance,* ▪ *Lateral Collision Avoidance,* ▪ *Intersection Collision Avoidance,* ▪ *Vision Enhancement For Crash Avoidance,* ▪ *Safety Readiness,* ▪ *Pre-crash Restraint Deployment* ▪ *Automated Vehicle Operation,*
Information Management	▪ *Archived Data*
Maintenance and Construction Management	▪ *Maintenance And Construction Operations*

– the functions (e.g., gather traffic information or request a route) that are required for ITS,
– the physical entities or subsystems where these functions reside (e.g., the field or the vehicle) [8],
– the information flows and data flows that connect these functions and physical subsystems together into an integrated system.

The *main components* of an ITS architecture are *subsystems* and *information flows* [9, p3]. *Subsystems* are individual elements of the ITS that perform particular functions such as providing traveler information, managing traffic, etc. Subsystems include center systems, field components, vehicle and infrastructure equipment, and traveler devices that participate in ITS. *Information flows* depict ITS integration by illustrating the information links between subsystems. They define information that is exchanged between subsystems such as traffic information, incident information, or surveillance and sensor control data. ITS integration is not only technical but institutional as well.

ITS offer the potential to improve traffic safety, to reduce traffic congestion, and to increase economic productivity, by means of the *ITS Services* (see Tab.1) and *Service Packages* (see Tab.2).

As mentioned in American National ITS many of the user services are too broad in scope to be convenient in planning actual deployments. They often don't translate easily into existing institutional environments and don't distinguish between major levels of functionality. In order to address these concerns and to support the creation of service based regional ITS architectures, e.g. ITS architecture for conurbation, finer grained set of deployment-oriented ITS service building blocks, called *Service Packages*, were defined from the original user services [6, 12] (see Tab.2).

Table 2. *ITS Service Packages* based on U.S. National ITS Architecture [6, 7, 12]

Group of ITS Service Package	ITS Service Packages
Traffic Management	▪ *Network Surveillance,* ▪ *Probe Surveillance,* ▪ *Traffic Signal Control,* ▪ *Traffic Metering* ▪ *HOV (High Occupancy Vehicle) Lane Management,* ▪ *Traffic Information Dissemination,* ▪ *Regional Traffic Management,* ▪ *Traffic Incident Management System,* ▪ *Traffic Decision Support and Demand Management,* ▪ *Electronic Toll Collection,* ▪ *Emissions Monitoring and Management,* ▪ *Roadside Lighting System Control,* ▪ *Standard Railroad Grade Crossing,* ▪ *Advanced Railroad Grade Crossing,* ▪ *Railroad Operations Coordination,* ▪ *Parking Facility Management,* ▪ *Regional Parking Management,*

	▪ *Reversible Lane Management,* ▪ *Speed Warning and Enforcement,* ▪ *Drawbridge Management,* ▪ *Roadway Closure Management,* ▪ *Variable Speed Limits,* ▪ *Dynamic Lane Management and Shoulder Use,* ▪ *Dynamic Roadway Warning,* ▪ *VMT (Vehicle Miles Traveled Tax) Road User Payment,* ▪ *Mixed Use Warning Systems.*
Public Transportation	▪ *Transit Vehicle Tracking,* ▪ *Transit Fixed-Route Operations,* ▪ *Demand Response Transit Operations,* ▪ *Transit Fare Collection Management,* ▪ *Transit Security,* ▪ *Transit Fleet Management,* ▪ *Multi-modal Coordination,* ▪ *Transit Traveler Information,* ▪ *Transit Signal Priority,* ▪ *Transit Passenger Counting,* ▪ *Multimodal Connection Protection.*
Traveler Information	▪ *Broadcast Traveler Information,* ▪ *Interactive Traveler Information,* ▪ *Autonomous Route Guidance,* ▪ *Dynamic Route Guidance,* ▪ *ISP (Information Service Provider) Based Trip Planning and Route Guidance,* ▪ *Transportation Operations Data Sharing,* ▪ *Travel Services Information and Reservation,* ▪ *Dynamic Ridesharing,* ▪ *In-Vehicle Signing,* ▪ *Short Range Communications Traveler Information.*
Advanced Safety Systems	▪ *Vehicle Safety Monitoring,* ▪ *Driver Safety Monitoring,* ▪ *Longitudinal Safety Warning,* ▪ *Lateral Safety Warning,* ▪ *Intersection Safety Warning,* ▪ *Pre-Crash Restraint Deployment,* ▪ *Driver Visibility Improvement,* ▪ *Advanced Vehicle Longitudinal Control,* ▪ *Advanced Vehicle Lateral Control,* ▪ *Intersection Collision Avoidance,* ▪ *Automated Vehicle Operations,* ▪ *Cooperative Vehicle Safety Systems.*

Commercial Vehicle Operations (CVO)	▪ *Carrier Operations and Fleet Management,* ▪ *Freight Administration,* ▪ *Electronic Clearance,* ▪ *CV (Commercial Vehicle) Administrative Processes,* ▪ *International Border Electronic Clearance,* ▪ *Weigh-In-Motion,* ▪ *Roadside CVO Safety,* ▪ *On-board CVO Safety,* ▪ *CVO Fleet Maintenance,* ▪ *HAZMAT (Hazardous Materials) Management,* ▪ *Roadside HAZMAT Security Detection and Mitigation,* ▪ *CV Driver Security Authentication,* ▪ *Freight Assignment Tracking.*
Emergency Management	▪ *Emergency Call-Taking and Dispatch,* ▪ *Emergency Routing,* ▪ *Mayday (SOS) and Alarms Support,* ▪ *Roadway Service Patrols,* ▪ *Transportation Infrastructure Protection,* ▪ *Wide-Area Alert,* ▪ *Early Warning System,* ▪ *Disaster Response and Recovery,* ▪ *Evacuation and Reentry Management,* ▪ *Disaster Traveler Information.*
Archived Data	▪ *ITS Data Mart,* ▪ *ITS Data Warehouse,* ▪ *ITS Virtual Data Warehouse.*
Maintenance and Construction Operations	▪ *Maintenance & Construction Vehicle and Equipment,* ▪ *Tracking,* ▪ *Maintenance & Construction Vehicle Maintenance,* ▪ *Road Weather Data Collection,* ▪ *Weather Information Processing and Distribution,* ▪ *Roadway Automated Treatment,* ▪ *Winter Maintenance,* ▪ *Roadway Maintenance and Construction,* ▪ *Work Zone Management,* ▪ *Work Zone Safety Monitoring,* ▪ *Maintenance & Construction Activity Coordination,* ▪ *Environmental Probe Surveillance,* ▪ *Infrastructure Monitoring.*

4 Development of ITS Architecture for Upper-Silesian Conurbation

An architecture is important because it allows the integration of options to be considered prior to investment in the design and development of the elements of the system. It defines the systems and the interconnections and information exchanges between these systems. It is functionally oriented and not technology specific, which allows the architecture to remain effective over time. It defines “what” must be done, not “how” it will be done. The functions the system performs remain the same while technology evolves.

Because of the spatial scale of the Upper-Silesian Conurbation and the lack of a single authority managing its development, including the transport development as well, separate ITS systems are operating or are being designed and under construction in individual cities of the conurbation. However, there is no single consistent concept of their architecture, which would allow providing *appropriate ITS services*, adapted to individual cities specific situation and to their local needs as well as to the specific nature and needs of the whole conurbation area and its links with other conurbations and with the national transport system.

A comprehensive approach to the ITS architecture creation for the conurbation requires the application of *Systems Engineering*[4], because it reduces the risk of schedule and cost overruns and increases the likelihood that the implementation will meet the user’s needs. Also Systems Engineering benefits include: an improved stakeholder participation, more adaptable, resilient systems, a verified functionality and fewer defects, a higher level of reuse from one project to the next, and a better documentation. In this context the *ITS project* term may be defined as follows: “*ITS Project means any project that in whole or in part funds the acquisition of technologies or systems of technologies that provide or significantly contribute to the provision of one or more ITS user services as defined in the conurbation ITS architecture*.” This broad definition, adopted following [11], comprises both smaller ITS projects limited to the purchase and installation of field equipment – controllers, signals, etc., and also larger ITS projects which support integration of multiple systems and development of custom software – for example, transportation management centres and traveller information systems.

The development of an ITS architecture for the Upper-Silesian Conurbation will enable its use by the stakeholders (cities of the conurbation) to prepare specific ITS

---

[4] *Systems Engineering* (International Council of Systems Engineering - http://www.incose.org/) is an interdisciplinary approach and means to enable the realization of successful systems. It focuses on defining customer needs and required functionality early in the development cycle, documenting requirements, then proceeding with design synthesis and system validation while considering the complete problem. Systems Engineering integrates all the disciplines and specialty groups into a team effort forming a structured development process that proceeds from concept to production to operation. Systems Engineering considers both the business and the technical needs of all customers with the goal of providing a quality product that meets the user needs.

projects, meeting the regional objectives and also having a dynamic traffic management in view [4, 5]. In addition, together with the operating ITS systems offering inter alia services of the Electronic Payment System, it is possible to apply specific *Demand-Side Strategies*, like [4]:

- *Financial Incentives: Tax Incentives, Parking Cash-Out, Parking Pricing, Variable Pricing, Preferential Parking, Distance-Based Pricing,*
- *Incentive Rewards Programs, Travel Time Incentives, Transit / HOV Signal Priority Systems & Queue Jumps, High-Occupancy Vehicle (HOV) Lanes, High-Occupancy Toll (HOT) Lanes,*
- *Targeted Strategies: Mode Strategies, GRH - Guaranteed Ride Home, Transit Pass Program, Shared Vehicles,*
- *Departure-Time Strategies, Worksite Flextime, Coordinated Event or Shift Scheduling,*
- *Route Strategies: Real-Time Travel Route Information, In-Vehicle Navigation Systems, Web-Based Route-Planning Tools,*
- *Trip Reduction Strategies: Employer Telework Programs & Policies, Employer Compressed Work Week Programs & Policies.*

The conurbation ITS architecture may be an important tool for use in the transportation planning, programming, and project implementation. It can be used to maximize appropriate integration of projects identified by the planning process. The planning process and related outputs also refine the architecture over time as the feedback is incorporated as part of architecture maintenance. And therefore the relationship between the conurbation ITS architecture and transportation planning is on-going and iterative.The conurbation ITS architecture provides information for updating the long range plan, the transportation improvement program and other plans. It will also provide information for use in other planning studies and activities, including the congestion management plans, corridor and sub-area studies, performance-monitoring activities, transit development plans, and other locally defined studies or plans [5].

Table 3. Steps of development of ITS architecture based on U.S. National ITS Architecture [9]

Steps of ITS architecture development	**Description of objectives and results**
Step #1 Identification of conurbation and needs of ITS architecture	Step #1 Defines *who will be involved* with (and served by) the architecture and *how the conurbation ITS architecture* development will be *structured*.

Table 3. (*continued*)

Step #2 Data collection for ITS architecture	At step #2 the existing and planned ITS systems in the conurbation are inventoried, *the roles and responsibilities of each stakeholder in developing, operating, and maintaining these ITS systems are defined*, the *ITS services* that should be provided in the conurbation *are identified*, and the contribution (in terms of functionality) that each system will make to provide these ITS services is documented.
Step #3 Definition of interface between ITS systems	At step #3 the existing and planned *interfaces between ITS systems are defined*. First, the connections between systems are identified, and then the information that will be exchanged on each of the interfaces is defined.
Step #4 Implementation of ITS architecture	Step #4 include a *sequence of projects*, a *list of needed agency agreements*, and a *list of standards* that can be considered for project implementation.
Step #5 Use of conurbation ITS architecture	The ITS services can provide the basis for *operational strategies* that can be used to *improve the transportation system* to meet the *conurbation's vision* and *goals*. The conurbation ITS architecture in step #5 is an important *tool* for use in transportation *planning*, *programming*, and *project implementation*. The conurbation ITS architecture can be used to support evaluation and prioritization of strategies in two ways: - the 1st is through the definition in the architecture of *archiving and data collection systems* that *support collecting the data needed for evaluation*, - the 2nd is through the *detailed definition of ITS projects* and their *sequencing* that can be used to *support prioritization efforts*.
Step #6 Maintain the conurbation ITS architecture	A maintenance plan is used to guide *controlled updates to the conurbation ITS architecture* baseline so that it continues to accurately reflect the region's existing ITS capabilities and future plans. As ITS projects are implemented, new ITS priorities and strategies emerge through the transportation planning process, and the scope of ITS expands and evolves to incorporate new ideas, the conurbation ITS architecture will need to be updated.

Source: steps based on [9]

Basic stages of the ITS architecture for the Upper-Silesian Conurbation development and expansion may be formulated so as presented in Table 3.

Objectives, sources and results of each steps are presented in Fig. 2÷7 (based on [9]).

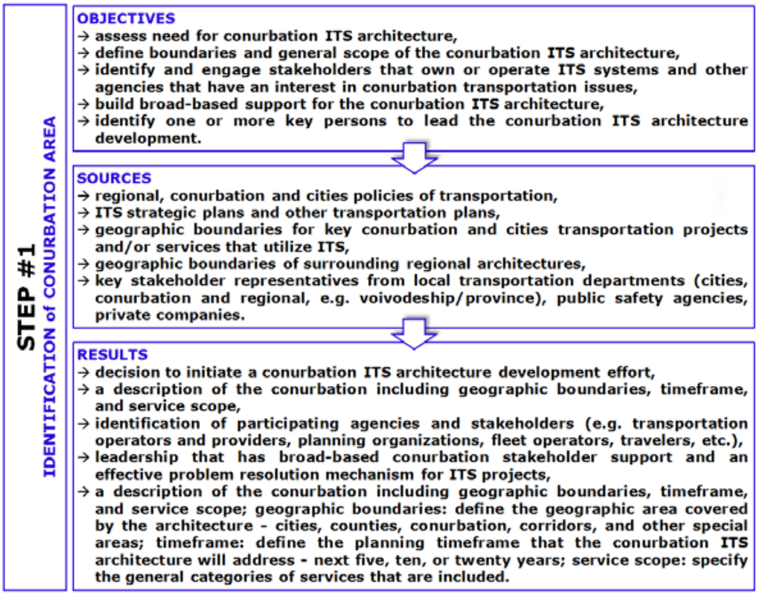


Fig. 2. Step #1 – Identification of conurbation and needs. Source: based on [9].

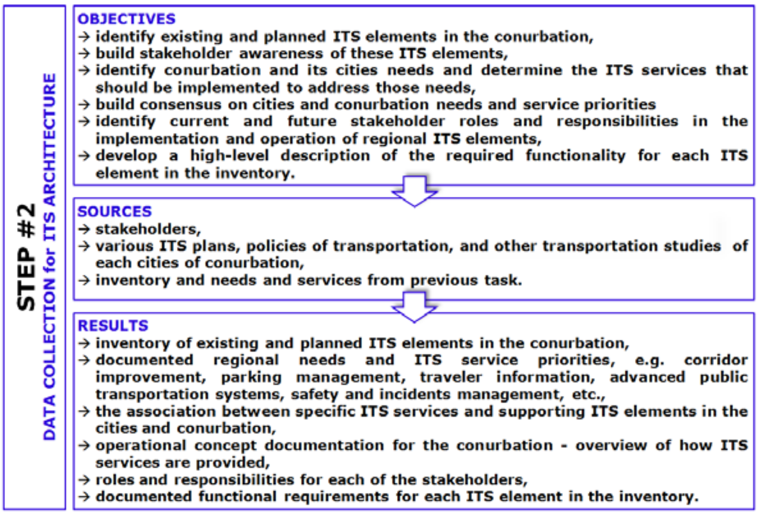


Fig. 3. Step #2 – Data collection for ITS architecture. Source: based on [9].

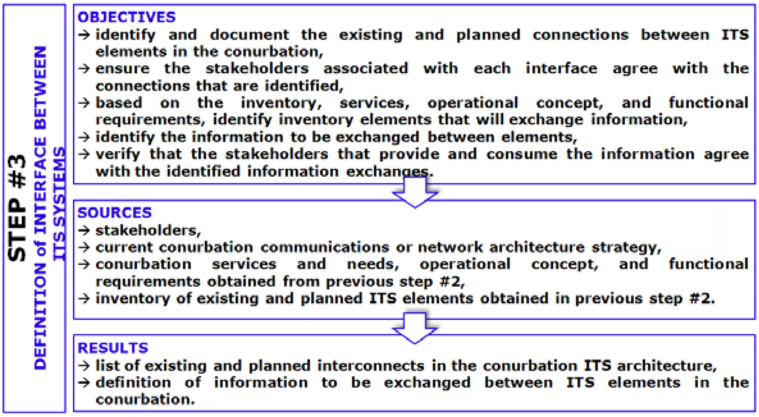


Fig. 4. Step #3 – Definition of interface between ITS systems. Source: based on [9].

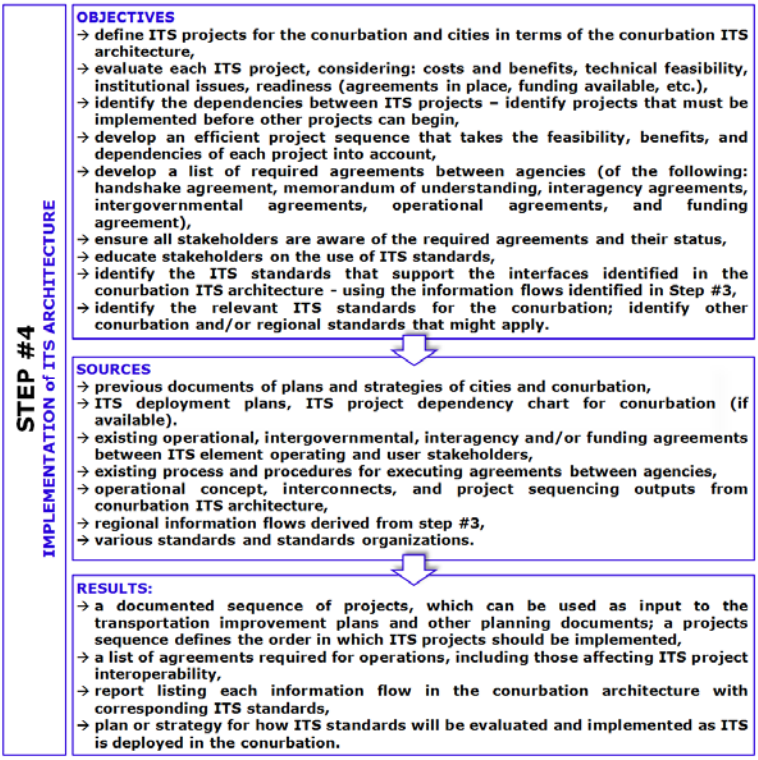


Fig. 5. Step #4 – Implementation of ITS architecture. Source: based on [9]

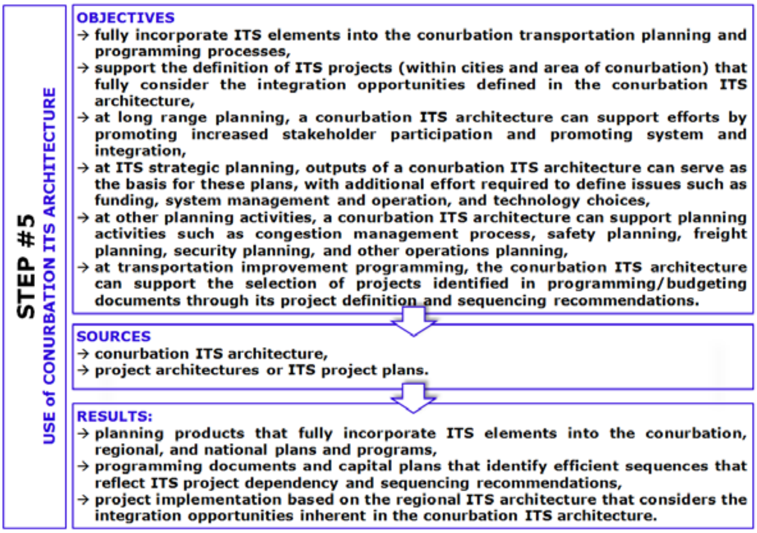


Fig. 6. Step #5 – Use of conurbation ITS architecture. Source: based on [9].

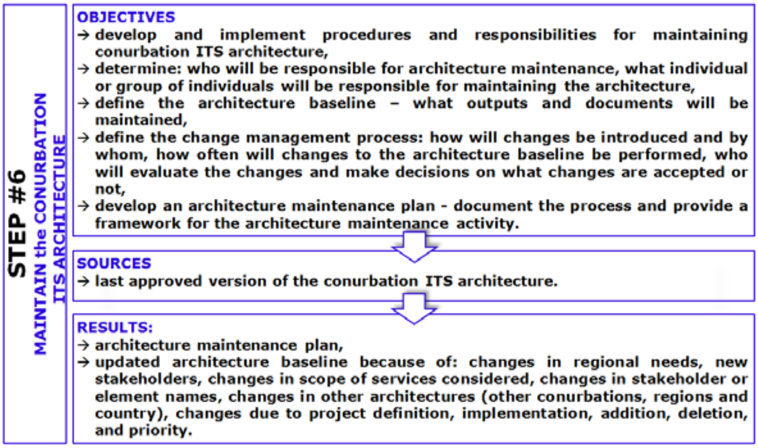


Fig. 7. Step #6 – Maintain the conurbation ITS architecture. Source: based on [9].

5 Conclusion

The most popular ITS solutions being implemented now in the Upper-Silesian Conurbation include: *Adaptive Traffic Lights*, *Public Transport Priority and Supervision*, *Video Surveillance* with the signal transferred to selected institutions, i.e. police, rescue centre, fire brigade / municipal police and also Dy*namic Passenger Information*, *Parking Guidance and Information*, *VMS for Route Guidance*.

The following issues are mentioned among main barriers of ITS solutions implementation in the Upper-Silesian Conurbation [1, 2, 3, 10]: formal-legal problems and procedures and also limited funds. The formal-legal barriers are related among other things to the lack of ITS strategy implementation on the national level and to insufficient adaptation of the Polish legislation to faster and faster developing intelligent technologies. A great amount of time needed for a project preparation as compared with a quick development of ITS technology frequently makes that the equipment planned to be deployed – due to a long time of necessary permits obtaining as well as of design documentation preparation – at the moment of investment implementation uses a technology considered obsolete on the market.

Moreover, the lack of common transportation policy of the cities of the Upper-Silesian Conurbation is a significant barrier. Because of the legislation, each city on its own takes care of the road infrastructure passing its territory and hence also pursues its own policy relating to ITS solutions implementation. This results in the individualisation of cities activities, in the lack of information exchange between them and in the impossibility to start common actions.

The specified problems substantially restrict a possibility of creating integrated ITS systems in the Upper-Silesian Conurbation and delay the development of their appropriate architecture – especially under conditions of missing national ITS architecture.

References

1. Analiza możliwości zastosowania zintegrowanych rozwiązań w systemie transportu miejskiego w województwie śląskim – Raport końcowy opracowany przez Deloitte Polska, Katowice. Marshal's Office Silesian Voivodeship, Katowice, Poland (December 2011)
2. Assumptions of the transport system development strategy for Silesian Voivodeship (Założenia strategii rozwoju systemu transportu województwa śląskiego). Marshal's Office Silesian Voivodeship, Katowice, Poland (April 2012)
3. Diagnoza stanu systemu transportu województwa śląskiego. Marshal's Office Silesian Voivodeship, Katowice, Poland (March 2012)
4. Karoń, G., Żochowska, R., Sobota, A.: Dynamiczne zarządzanie ruchem w aglomeracji górnośląskiej z wykorzystaniem ITS. In: VIII Konferencja Naukowo-Techniczna z cyklu: Problemy komunikacyjne miast w warunkach zatłoczenia motoryzacyjnego. Wyd. SITK Oddział w Poznaniu, pp. 373-398. Poznań (2011)
5. Karoń, G., Mikulski, J.: Transportation Systems Modelling as Planning, Organisation and Management for Solutions Created with ITS. In: Mikulski, J. (ed.) TST 2011. CCIS, vol. 239, pp. 277–290. Springer, Heidelberg (2011)
6. Key Concepts of the National ITS Architecture, Version 7.0 (2012), http://www.iteris.com/itsarch/index.htm (last updated March 28, 2012)
7. National ITS Architecture. Service Packages. Architecture Development Team for Research and Innovation Technology Administration (RITA) US Department of Transportation Washington D.C. (2012)
8. National ITS Architecture. Physical Architecture. Architecture Development Team for Federal Highway Administration, US Department of Transportation (2007)

9. Regional ITS Architecture Guidance (Version 2.0 July 2006). Developing, Using and Maintaining an ITS Architecture for Your Region. U.S. Department of Transportation, FHWA-HOP-06-112 EDL-14317 (2006)
10. Strategy of transport system development for Silesian Voivodeship. (Strategia rozwoju systemu transportu województwa śląskiego). Marshal's Office Silesian Voivodeship, Katowice, Poland (February 2010)
11. Systems Engineering for Intelligent Transportation Systems. U.S. Department of Transportation FHWA-HOP-07-069 EDL-14340 (2007)
12. U.S. National ITS Architecture, Version 7.0 (2010) http://www.iteris.com/itsarch/index.htm (last updated March 28, 2012)
13. Plan zagospodarowania Województwa Śląskiego, http://www.silesia-region.pl

Can the Increasing of Energy Consumption of Information Interchange Be a Factor That Reduces the Total Energy Consumption of a Logistic Warehouse System?

Andrzej Bujak[1] and Paweł Zając[2]

[1] Wroclaw School of Banking,
29-31 Fabryczna St, 53609 Wrocław, Poland
andrzej.bujak@interia.pl
[2] Wroclaw University of Technology, Faculty of Mechanical Engineering,
5 Łukasiewicza St, 50371 Wrocław, Poland
pawel.zajac@pwr.wroc.pl

Abstract. The studying of relationships between subsystems, creating a logistic warehouse system, is one of the most interesting issues in terms of designing and optimization of logistic warehouse systems. It is important to use of these relationships in practice. The authors propose to assume the energy consumption of the system (which was calculated according to the described model), as an index that allows to compare individual options. Thanks to this model it is possible to match the system of information interchange to the given equipment, and vice versa. Three warehouse options were analyzed from the given warehouse system point of view. Three options of a warehouse system of information interchange were tested from the given warehouse equipment point of view.

Keywords: optimization of logistics system, energy consumption, bar codes.

1 Introduction

Nowadays, many requirements which were established for the logistics warehouse systems have changed. Energy consumption is very often one of the most important elements in the processes of the planning, the estimation of efficiency or the revitalization. Energy is defined as a balance of energy that was delivered to the logistics system. This energy can be for example connected with the rotation of unit loads. It is worth to emphasize, that there is no universal index, thanks to which it would be possible to evaluate logistic warehouse systems, taking the energy consumption into consideration. In the literature there are some attempts of creating this kind of index, for example: Directive of the European Parliament and of the Council 2006/32/WE of April 5, 2006 on energy end-use efficiency and energy services and repealing Council Directive 93/76/EEC; renewed Lisbon Strategy and National Cohesion Strategy (2007-2013).

J. Mikulski (Ed.): TST 2012, CCIS 329, pp. 199–210, 2012.

"United Nations Framework Convention", colloquially called The Kyoto Protocol (December, 1997 and it has been in force since 16 February, 2005). As a result of agreements, a decision was made according to which the "Index EEI" was established as a system of energy efficiency classification. This index describes the efficiency of devices. Within the index there is a scale (from A to G), where "A" is the best class and it means that the efficiency of the device is the highest and "G" is the poorest class, and it means that the efficiency of the device is the lowest. There are also some subclasses, for example A+ or A++. These classes are calculated from the proportion of a year-long energy consumption of a given device to the standard energy consumption of the device, which is defined by specific regulations.

The information connected with this subject matter was collected on the basis of reviews of the research database: EBSCO, COMPENDEX, PROQUEST, Baztech, e-Journals: [1-13].

89% of transport energy in Poland is consumed in the road transport and 5.6% of energy is consumed in the rail transport. The rest of energy is consumed in the air transport and minute quantities in inland navigation and cabotage.

A method of calculating the energy consumption of logistic warehouse systems will be presented in this paper. This method can be used to verify the system of information interchange in logistic systems. Because they have no tools, in some cases designers are acting according to their intuition and experience in these types of decisions. Apart from these, there are no analytical tools, which can help in the selection of warehouse system elements [5, 2].

In the second subsection of this paper a method of calculating the energy consumption will be described. Then, example calculations will be presented and results will be described. Conclusions will end the paper. The author provides a wide range of literature.

2 Digitisation and Calculation of Energy in the Model

The level of the terminal floor, in front of the entrance to the warehouse (where the reloading takes place) was assumed to be the reference level – the energy of a pallet that is received to the warehouse amounts to 0 (in warehouses without temperature regimes – this problem is going to be touched on), so every movement of the pallet with the material generates energy changes. There is an assumption, that the energy is equal to the work that has to be done to move the pallet [3].

$$E_E = E_P + E_S + E_Z + E_C + E_W \tag{1}$$

Where: E_P - energy, which is necessary to lift the pallet; E_S - energy, which is necessary to lift empty forks; E_Z - energy which is necessary to transport the pallet; E_C - energy which is necessary for a truck travel without a load; E_P - energy which is necessary to execute tasks connected with the use of forks.

The potential energy can be calculated from the well-known formula E=mgh, but the kinetic energy has to be calculated on the basis of driving characteristic of conveyors and vehicles.

There is no formula in the literature, thanks to which it would be possible to calculate the energy of a forklift truck. So for the purposes of this research, the formula was created on the basis of the simplest case.

$$F = m \cdot a + m \cdot g \cdot \mu \tag{2}$$

where $\mu = f(mat1, mat2)$. The equation of the movement of the truck can be presented by means of relationship (3):

$$F(V) = m \cdot a + W_t \tag{3}$$

where: $m \cdot a$ - inertial force; $W_t = f_t \cdot Q$ - resistances to motion; f_t - rolling resistance coefficient; $F(V)$ - driving forces.

Resistances to motion can be divided into basic and additional. The rolling resistance, suspension dumping, toe-in resistance and aerodynamic drag are included in the first group. The grade resistance, and resistances connected with the drawbar pull, the start-up and with the turn are included in the second group. The equation of movement has a form:

$$\xi \cdot m \cdot \frac{dV}{dt} + W_{r.coal} = F_N(V) \tag{4}$$

where: ξ - ratio of rotor. Possibilities of energy recovery, short time of storing this energy and the conversion of this energy into different forms – all these factors are taken into consideration in this method.

The organization of the warehouse transport system: organization of the warehouse, of the work, of devices.

An assumption was made that there are: 1,2,…, m devices; docks (receiving docks and shipping docks) 1,2,3,…, s; storage areas (rows of racks/intermediate storage area) 1,2,3,…, k. The process of warehousing consists of receiving, storing, picking and giving the loads out. This model assumes that pallet unit loads will be used in the picking process there. In the process of unit load moving the following stages may be distinguished, starting with setting the unit load next to the receiving dock, through the process of storing and giving it out: receipt of a pallet (i); transport of the pallet (j); putting the pallet back on the chosen place in the intermediate storage area (l).

In some cases during the process, there are changes of devices between the stages i,j,l. Each stage is executed by a device, so the storing of pallet (a movement of the unit load in the transport-warehouse system) can be presented as three numbers (i,j,l), that present the number of the device, which is carrying out the warehousing.

The process of storing can be divided into single operations: o_1 - travelling of the empty forklift truck to the place, where the unit load is placed, o_2 - executing the task which is connected with the use of forks, o_3- delivering the pallet to the PP_1, o_4 - return of the empty forklift truck to point M (readiness point), o_5 - reloading, o_6 - transport from the PP_1 to the PP_2, o_7 - reloading, o_8 - taking the pallet to the rack, o_9 - executing the task which is connected with the use of forks, o_{10} - return of the empty forklift truck to point M (readiness point).

The data characterising each device is also given here: a single energy demand in the horizontal direction; a single energy demand in the vertical direction; the data is written as matrix ZE. The ZE is an m x 2 matrix.

$$ZE = \begin{matrix} 1 \\ 2 \\ . \\ . \\ m \end{matrix} \begin{bmatrix} ze_{11} & ze_{12} \\ ze_{21} & ze_{22} \\ . & . \\ . & . \\ ze_{m1} & ze_{m2} \end{bmatrix} \tag{5}$$

Unloading, introduction of the unit load to the warehouse.

There is a set of n pallets, which are located in the particular dock. There is also information about these pallets:

The mass of the pallet.

The address of storing place within the warehouse.

The creation of the transport parameters matrix PT is the next step. This is an n x 5 matrix, where: PT_{i1} = the mass of the pallet I; PT_{i2} = the track of the unloading stage of the pallet I; PT_{i3} = the track of the transport stage of the pallet i; PT_{i4} = the track of the locating stage of pallet I; PT_{i5} = the height of the storing of pallet I.

$$PT = \begin{matrix} 1 \\ 2 \\ 3 \\ . \\ n \end{matrix} \begin{bmatrix} pt_{11} & pt_{12} & pt_{13} & pt_{14} & pt_{15} \\ pt_{21} & \dots & \cdots & . & pt_{25} \\ . & . & . & . & . \\ . & . & . & . & . \\ pt_{n1} & . & . & . & pt_{n5} \end{bmatrix} \tag{6}$$

It is necessary to define the technology of unloading for each pallet (unit load). Because there is a possibility of reloading from one device to the material handling device between some stages of the process, it is necessary to determine a single energy consumption for these operations. The matrix dimensions are m x m and there are three numbers (i,j,k), and each number executes a given stage of unloading. For example, the compositions of numbers (3,5,1) means that the first stage of the process is performed by device number 3, the second stage by device number 5 and the third stage by device number 1. The collection of these technologies for all pallets creates matrix T, where t_{ij} – indicates the number of the device executing stage j of pallet I unloading process (i=1, 2, 3, ..., m; j=1, 2, 3)

$$T = \begin{matrix} t_{11} & t_{12} & t_{13} \\ t_{21} & t_{22} & t_{23} \\ t_{n1} & t_{n2} & t_{n3} \end{matrix} \tag{7}$$

Matrix Z is defined in the following way: Z_{ij} = *>0* when the reloading from one device i to the second device j is possible and if it is equal to the single energy consumption of this operation there is no reloading

$$Z = \begin{bmatrix} Z_{11} & Z_{12} & \cdots & Z_{1M} \\ Z_{21} & Z_{22} & \cdots & Z_{2M} \\ \cdots & \cdots & \cdots & \cdots \\ Z_{M1} & \cdots & \cdots & Z_{MM} \end{bmatrix} \tag{8}$$

For each pallet and taking into consideration technology (T), it is easy to determine an elementary energy consumption. This energy consumption is going to be indicated as an n x 10 matrix O. Elements of this matrix are defined: $O_{i1} = ZE(T(i,1),1)$, $O_{i2} = ZE(T(i,1),2)$, $O_{i3} = 0_{i1}$, $O_{i4} = 0_{i1}$, $O_{i5} = ZE(T(i,1),T(i,2))$, $O_{i6} = ZE(T(i,2),1)$, $O_{i7} = Z(T(i,2),T(i,3))$, $O_{i8} = ZE(T(i,3),1)$, $O_{i9} = ZE(T(i,3),2)$, $O_{i10} = 0_{i8}$.

To obtain the energy consumption for each operation, it is necessary to correct elements of the matrix, taking the track, the load and the height into consideration. Coefficients of corrections are determined on the basis of the PT matrix. We obtain a *WK (n x 10)* matrix, which is defined as a matrix of the coefficients of correction of energy consumption for each operation and for each pallet.

After the operation of matrices WK and O multiplication, we will obtain matrix E, where the elements of this matrix correspond to the energy consumption for each operation and for each pallet.

The matrix multiplication is defined as:

$$WK_{n\times10} \otimes O_{n\times10} = E_{n\times10} \tag{9}$$

where:

$$e_{ij} = WK_{ij} \cdot O_{ij} \tag{10}$$

The total energy consumption is:

$$E_C = \sum_{i=1}^{n}\sum_{j=1}^{10} e_{ij} \tag{11}$$

The thermal energy is considered in the case of analysing warehouse systems with the temperature regime. In these cases it is important to take the heat exchange into consideration – the heat that penetrates walls, the ceiling and the floor of the warehouse chamber; the heat that was abstracted from cooled loads on the pallet; the heat which was given back by the air, which was unintentionally introduced to the chamber; the heat connected with the work of the radiator fan; the heat which is generated by people; and other kinds of heat, according to the equation.

Modern devices in warehouses are very often equipped with technical systems, which allow energy recovering, energy storing and converting it into different forms (by using heat accumulators, storage cells, pressure accumulators). What is more, there is also a possibility of analysing the utilization of solar energy (water heating, generation of electricity). The model was not tested in regard to taking a geothermal source of energy into consideration.

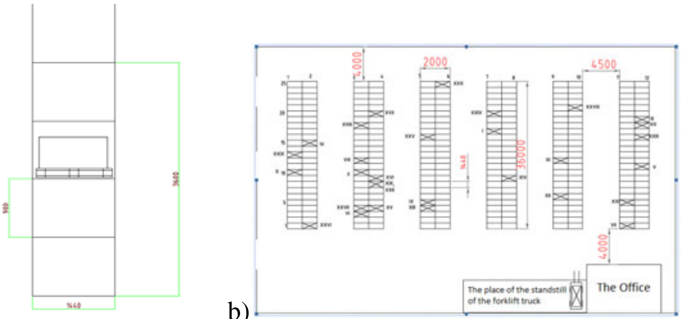
The flow of information in logistics system is described by the quantity and the cost of equipment which is needed to guarantee the flow of information. The energy which is necessary to the information interchange on the accurate level to guarantee the required warehouse turnover was also taken into consideration. . The energy of the system of the information interchange in the logistics warehouse system can be divided into: the energy of identification of elements and the energy of information transfer (in some cases the data transfer). In the paper, this problem was taken into consideration. Functionality of RFID marks which relies on the affiliation both to the first and to the second group was taken into consideration. These marks can be used to recording of actual positions of elements of the logistics system, in the context of minimization simple cycles of their work.

3 The Application of the Model to the Given Case

For the purpose of examination of the above cited concept, calculations were made for a given warehouse. The system of information interchange was matched with a given layout of racks and aisles. There were three options of information interchange: a) the printed information, 2) a mobile terminal with a docking station, c) a radio terminal. Values of several parameters, which describe the logistics warehouse system, were assumed for the purpose of making calculations easier:

PR – annual turnover of the warehouse (the number of unit loads – euro pallets – that move through the warehouse during a year) = 400000; *PD* – daily turnover of the warehouse (the number of unit loads – euro pallets – that move through the warehouse during a day) = 1231 (325 working days for a year); *Ph* – average number of unit loads that are serviced during one hour = 77; reloading of 1096 pallets per day, so there are 1500 rack slots and the weight of a pallet is 800 kg. All forklift trucks working in this warehouse are equal. Forklift trucks and all characteristics that are connected with them, are based on the characteristic of a JUNGHEINRICH EFG 316 forklift truck. There are 5 forklift trucks in the warehouse, but to make calculations clearer, there is an assumption that all tasks are executed sequentially (only one truck is working at the given moment). Forklift trucks move along straight sections and the truck speed amounts to 13 km/h. Rack slots are denoted (it is the number of the lane, the number of the place in that aisle and the number of the level, where the pallet with the load is placed). In the case of the warehouse, two docking terminals are placed on the sides of the building, and in the middle of the lanes. To make the calculations clearer, the authors make an assumption that the weight of each unit load in this warehouse is 800 kg.

The test consists in the simulation of transportation of 30 pallets (with the assumption that unit loads shipments are being delivered at the terminal). The mechanical energy which is necessary to execute all tasks is described by means of previous model. The coefficient of resistance to motion amounts to 0.1. Intermediate storage areas were generated with the use of Random Number Generator from the Windows Microsoft Office Package, where: (number of activity, number of door, number of lane, number of place, number of level): (1,3,7,17,5), (2,2,3,10,5), (3,1,12,18,3), (4,1,2,15,5), (5,3,2,11,2), (6,1,3,3,5), (7,3,11,1,1), (8,3,3,12,2), (9,2,5,5,1), (10,2,1,10,3), (11,1,9,12,4), (12,1,9,6,3), (13,1,5,4,4), (14,2,8,9,4), (15,3,4,4,1), (16,1,4,9,1), (17,2,4,20,1), (18,1,3,18,4), (19,3,4,8,2), (20,2,12,18,5), (21,3,11,5,4), (22,3,4,8,3), (23,2,12,16,3), (24,1,7,20,4), (25,1,5,16,4), (26,3,2,1,4), (27,3,3,4,4), (28, 3,0,21,1), (29,1,1,13,4), (30,2,6,25,5). The data is the same for each of three variants of the warehouse. There are five levels of racks in the warehouse (Fig.1).



a) b)

c)

Rack slots	Number of racks	Number of levels	Number of lanes	Number of slots in one lane
1500	6	5	12	25

Fig. 1. Test warehouse: a) pallet rack, b) functional scheme of warehouse, c) parameters of action

3.1 The Determination of Pallet Movement Tracks on the Horizontal Surface

Unit loads are moved by means of forklift trucks. A forklift truck starts from the point called “standstill” and then it moves to the particular location to take a unit load. The forklift truck holds it and takes it away to the door. Then, depending on the variant: “1” it returns to the “standstill” point, “2” it returns to the docking station (which are placed in the middle of lanes), “3” it starts a new activity.

The scheme of tracks is presented here for variant 1. To illustrate the transport process, the tracks for the firsts 5 orders are presented. These are results obtained from formula (12), for variants 1,2,3.

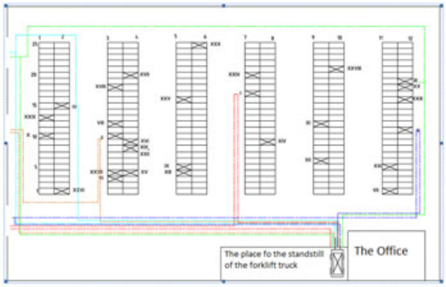


Fig. 2. The scheme of the tracks for the forklift truck in the warehouse number 1 (tracks 1-5)

$$\Delta s = K + 4{,}5 \cdot L_x + 2 \cdot NR \quad (12)$$

Where: *K – Constant length of the track covered by a forklift truck (drive from door nr 3, pulling out and return to the "standstill" point, then travel and return to the edge of the racks); Lx – Distance between racks (from the axis of the middle of an aisle, to the axis of the middle of the next aisle); NR – Distance between the edge of the rack to the storage place of a pallet – product of the slot width and the number of the slot* $NR = 1440 \cdot n$*, n – the slot number.*

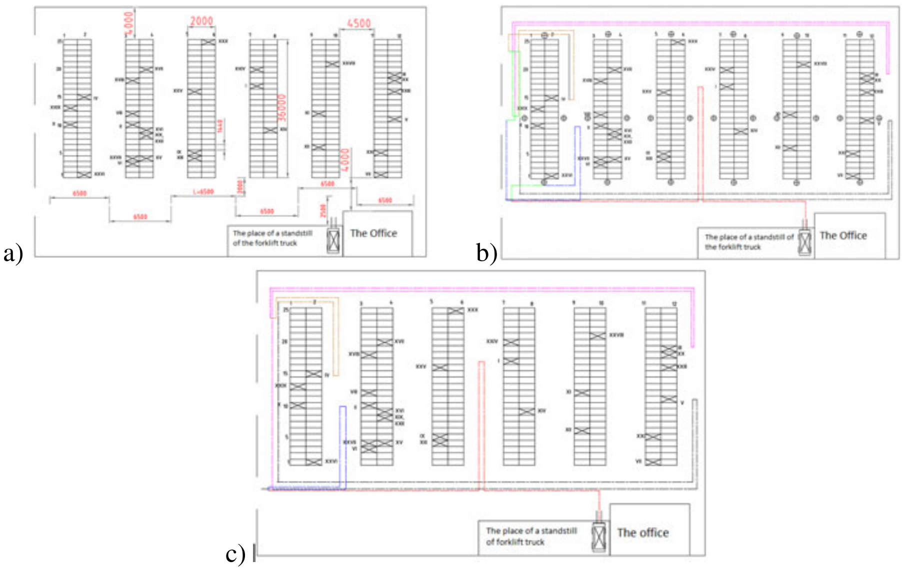


Fig. 3. Distances covered by a forklift truck in a) variant 1, b) variant 2, c) variant 3

3.2 The Work Connected with the Displacement of Pallets

The quantity of energy that is necessary to move 30 pallets by a forklift truck is connected with the characteristic of the handling equipment. In this example, three

Jungheinrich EFG 316 forklift trucks were used, and their driving characteristic is presented on the picture below.

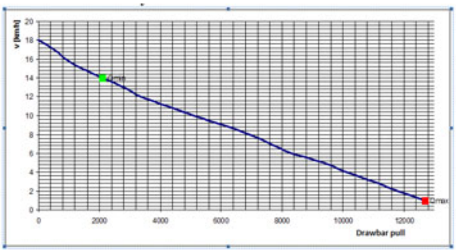


Fig. 4. Driving characteristic of a Jungheinrich EFG 316 forklift truck

Taking into consideration that $W = F \cdot \Delta s = E_K$, the work executed by a forklift truck can be calculated from formula (3).

It results from the assumptions that the forklift truck moves with the speed of *V = 13 km/h*. During turning the speed decreases to 5 km/h. Authors make an assumption that the track that is necessary to turn, to brake and to speed up, amounts to 6m (straight line). The calculation of loss resulting from braking and accelerating gives the difference of given terminal speeds. The statement of calculation results gives the driving force.

The table shows that for the speed of v = 18 km/h, the driving force decreases below 0. It results from the fact that the forklift truck (for which calculations are made), can achieve a speed of v = 17 km/h. It is also important to remember about making turns at the speed of v = 5 km/h, manoeuvres connected with the pallet picking from the storage, travel to the axis of track and manoeuvres connected with the unloading next to the receiving dock. One manoeuvre is braking from the speed of v = 5 km/h to v = 0 km/h and then accelerating on a 2 meters long track. Statement of the tracks, which are covered by forklift trucks in the warehouse, are presented below.

Table 1. Statement of tracks for warehouses 1-3

Warehouse nr 1				
Δs [m]	Number of turns	Track with turns 5 km/h [m] Δ_{sz1}	Track with straight segments [m] Δ_{sp1}	Maneuvers [m] Δ_{sm1}
3637,1	171	1266,0	2371,1	240
Warehouse nr 2				
Δs [m]	Number of turns	Track with turns 5 km/h [m] Δ_{sz2}	Track with straight segments [m] Δ_{sp2}	Maneuvers [m] Δ_{sm2}
2599,51	119	954	1645,51	240
Warehouse nr 3				
Δs [m]	Number of turns	Track with turns 5 km/h [m] Δ_{sz3}	Track with straight segments [m] Δ_{sp3}	Manuvers [m] Δ_{sm3}
2339,76	116	936	1403,76	240

Results of kinetic energy calculations for individual warehouses are presented below.

Table 2. Results of kinetic energy calculations

	Warehouse nr 1	Warehouse nr 2	Warehouse nr 3
Ekp	5712451,16	3964410,798	3381979,63
Ekp5h	4177209,76	3147755,224	3088363,616
Ekp5r	4177209,76	3147755,224	3088363,616
Ek0h	1080995,61	1080995,61	1080995,61
Ek0r	1080995,61	1080995,61	1080995,61
Sum [J]	16228861,92	12421912,48	11720698,09

3.3 Pallet Movement Tracks on the Vertical Surface

The potential energy exists in the transport process during the lifting and lowering of the transported pallet. According to the authors' assumption, the pallet lowering is realized due to the gravitation, and that is why it is also necessary to take into consideration a small amount of energy to prevent the forks striking into the floor. It has a small influence on the result of the calculations, so in this case this energy amounts to 20% of energy that is necessary to lift the forks (with or without the load). During the reloading of the pallet, the potential energy exists, when there is a movement of forks: first, when the truck receives the pallet from the rack slot, and second, when the pallet is delivered next to the receiving dock. The value of the energy in the first case is the function of the storage height, and in the second case it is constant because the pallet is placed on the one-meter platform next to the receiving dock. Each of cases occurs once during one reloading cycle. In the case of warehouse 1, 2 and 3, the first and the second type of reloading occurs only once (the weight of the transporter system amounts to 700 kg).

We multiply the lifting height (1 m) by 30 and it is the same for each warehouse. The value of the energy can be calculated from the formula $E_p = m \cdot g \cdot h$.

Table 3. Potential energy used in the process of moving 30 pallets [3], [4]

Potential energy			
Warehouses			
$E_{p1'}$	414080,1	E_{p1}	496896,12
$E_{p2'}$	441450	E_{p2}	529740
E_p [J]		1026636,12	

The total energy values necessary to move unit loads are presented below.

Table 4. Calculations of the total energy of each warehouse

Total energy			
	Variant 1	Variant 2	Variant 3
E_k+E_p [J]	17255498,04	13448548,6	12747334,21

4 The Analysis of the Calculation Results

Using the method from the second subsection, it was possible to calculate the energy consumption of moving the unit loads for three different variants. This model calculates the energy consumption taking into consideration not only the energy of forklift trucks, but also the energy consumption of the information interchange systems (matrix “O” – single energy consumption). The way of information interchange between the workers and the manager is the factor affecting the energy consumption. In variant 1, the forklift truck operator is given a printed order. He executes the task and then, after each completion he has to inform the manager about this task and then he has to get a new order. In variants 2 and 3 there is an automated information interchange. Each of these methods has a particular energy consumption that influences the energy consumption of the whole logistic system. Results are true for universal electric forklift trucks and typical operator terminals.

Also the making of the storage height higher will result in changing the energy consumption of the information interchange system, because it will be necessary to cover the additional space of the warehouse with radio-waves.

5 Conclusion

Several conclusions can be drawn from the calculations: the knowledge connected with the energy consumption of a warehouse system taking into consideration elementary energy consumptions (for example the system of information interchange) may be used when a decision is to be made about the choice of the information interchange system. Any simplest operations can decrease the energy consumption of logistic a warehouse system by 5%. The application of topical methodology of energy calculation allows performing a detailed analysis, not only in the context of the information interchange system.

References

1. Bartoldi, J.J., Platzman, L.K.: An O(nlogn) planar travelling salesman heuristic based on spacefilling curves. Operations Research Letters (2003)
2. Bozer, Y.A., Schorn, E.C., Sharp, G.P.: Geometric approaches to solve the Chebyshev salesman problem. IIE Trans. (2000)
3. Gudehus, T.: Principles of order picking: operations in distribution and warehousing systems. Essen-Germany (2005)
4. Hwang, H., Moon, S., Gen, M.: An integrated model for the design of end-of-aisle order picking system and the determination of unit load sizes of AGVs. Comput. Ind. Eng. (2002)
5. Kwaśniowski, S., Zając, P.: The Automatic Identification in Logistics Systems. Publishing house of Wroclaw University of Technology, Wroclaw (2004)
6. Kwaśniowski, S., Nowakowski, T., Zając, M.: Intermodal Transport in Logistics Networks. Publishing house of Wroclaw University of Technology, Wroclaw (2008)
7. Makris, P.A., Giakoumakis, I.G.: Interchange heuristic as an optimization procedure for material handling applications (2003)
8. Munzebrock, A., Persico, G.: Regalbedingeraete: Optimierung der Antriebssteuerung. F-H Foerdern Heben (1994)

9. Rouwenhorst, B., Reuter, B., Stockrahm, V., van Houtum, G.J., Mantel, R.J., Zijm, W.H.M.: Warehouse design and control: framework and literature review. Eur. J. Operat. Res. (2000)
10. Zając, P.: Electronic Data Interchange in Logistics Systems. Publishing house of Wroclaw University of Technology, Wroclaw (2010)
11. Zając, P.: Transport and storage system optimization in terms of energy. Research brochures of the Warsaw University of Technology, Warsaw (2010)
12. Zając, P.: Concept of model of estimate of - store house system transport. Total Logistic Management, Gliwice (2009)

Attenuation Measurements of Overvoltages on Contact Line

Andrzej Białoń[1,2]

[1] Silesian University of Technology, Faculty of Transport
Krasińskiego 8, PL 40-019 Katowice, Poland
andrzej.bialon@polsl.pl
[2] Railway Institute
Chłopickiego 50, PL 04-275 Warsaw, Poland
abialon@ikolej.pl

Abstract. In paper, the problems of measurements overvoltage attenuation in a contact line have been described. The investigations have been conducted with telematics synchronization measurements on the test track of the Railway Institute and the Railway Institute Laboratory in Warsaw. The investigations consisted in the introduction of voltage kicks directly to the contact line or its laboratory model with the use of a kick generator. In this publication, the measurement systems used in both the field and the laboratory investigations have been shown. The publication contains also the recorded results of the investigation. The courses of both the voltage and current impulses have been analysed. The results of this analysis allowed to elaborate the conception of the varistor protection of the contact line against overvoltages.

Keywords: measurements overvoltage, contact line.

1 Introduction

The occurrence of overvoltages in contact lines is caused mostly by commutation processes and atmospheric discharges. The first factor arises as the result of switching over in the circuits of traction substations and in the main circuits of electric locomotives and electric traction units. The second one results from a direct lightning strike in the contact line or by the induction of overvoltages caused by lightning strikes relatively close to the contact line.

In both cases, the overvoltage wave propagates in two directions from its source and it is gradually attenuated. It is known from the studies of the available literature that the contact line elementary parameters are featured with a significant non-homogeneity. Therefore, the propagating overvoltage wave will reflect from the discontinuities, such as the anchoring of the contact line, network turnouts, sectional cabin connections, etc. It should be noted that the cases presented in the literature do not describe these phenomena in the explicit and unequivocal way, and particularly the degree of the attenuation of overvoltages by the contact line versus the distance from the source.

J. Mikulski (Ed.): TST 2012, CCIS 329, pp. 211–220, 2012.

One of the main problems in studying the overvoltages propagation on the contact lines is the synchronization of measurements at various points distant from each other even up to several kilometers. Usage one of the basic tools used in telematics, namely the so-called "time stamp" made it possible to realize such measurements in the field.

Investigations were carried out within a section of the contact line of the test centre of the Railway Institute near Żmigród, Poland and in the Railway Institutive Laboratory in Warsaw. The investigations were aimed to determine the overvoltage attenuation in the contact line as the function of the distance from the source. It was the first time when such investigations were carried out in Poland.

2 Investigation Methodology

The length of the contact line section of the RI test track, where the investigations were carried out, was 7,760 m (two contact line types: 2C120-2C-1 - 9 sections 7.28 tkm with an insulation of 25 kV; YC150-2C150 - one section of 1.09 tkm with an insulation of 25kV). The measurements of the overvoltage attenuation in the contact line were carried out for two arrangements: 1) contact line section without load, and 2) contact line loaded with the resistance of 240 Ω.

The testing consisted in the introduction of electric kicks from an impulse wave generator into the contact line and then recording of both the voltage and the current impulse amplitudes at specific sites between the generator and the measurement section end (every one kilometre). The output resistance of the used generator was 2 Ω, and the maximum level of the voltage kick 6.9 kV; standard kick was 1.2/50 μs.

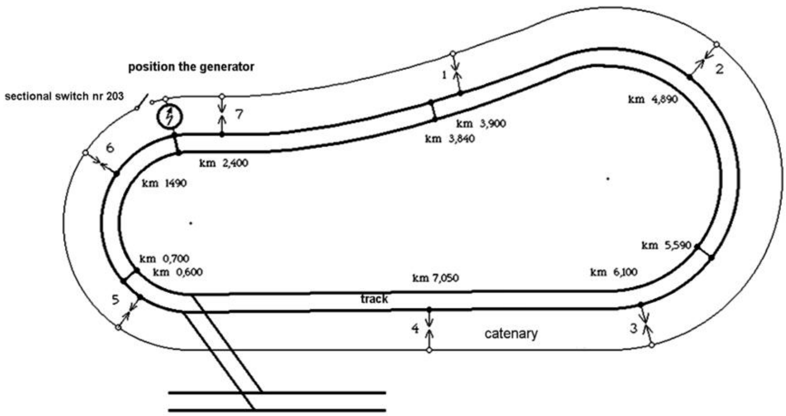


Fig. 1. The scheme of contact line of the test centre of the Railway Institute

In Fig. 1, the position of the generator in the test centre of the Railway Institute is shown. During the measurements, the contact line was open out with the sectionalising breaker No. 203, close to which the stationary station with the kick generator was located.The voltage and current impulses at the beginning of the tested contact line section were recorded with an oscilloscope, whereas at the other side of the 203 breaker, the station for the recording of the voltage impulses on the end of contact line section was located. The measurement systems are shown in Fig. 2 to 4, and the acronyms have the meaning: TEK - two-channel measuring oscilloscope, CR - Rogowski coil with the voltage-current ratio 400 mV/A, W203 - sectional breaker No. 203.

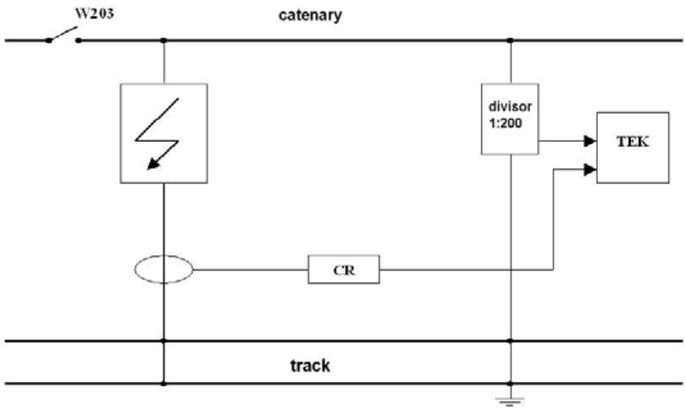


Fig. 2. Measurement system on the kick generator station

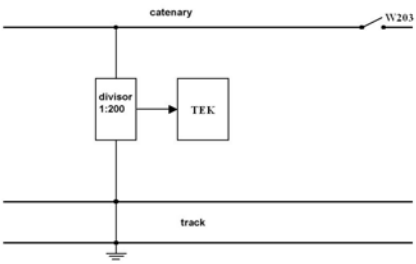


Fig. 3. Measurement system at the end of the contact line w/o load

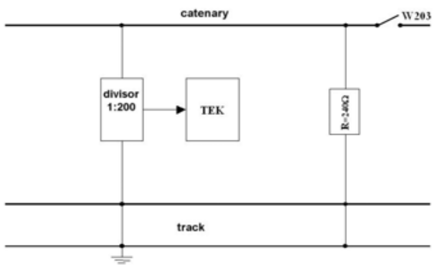


Fig. 4. Measurement system at the end of the contact line loaded with a resistor

The voltage of the electric kick for the above-mentioned arrangements of the contact line section (w/o load and with a loading resistor) was recorded every one kilometre starting from the generator up to the end of the line section. Measurements synchronization was carried out using the so-called "time stamp". This is made possible to illustrate the kick shape and its amplitude in the function of the distance from the kick generator. By using these records, the characteristics of the attenuation of the kick amplitudes on the tested contact line section versus the distance were developed.

Besides, the voltage amplitude course at the beginning of the measurement section and the surge current at the end of the contact line section loaded with the resistance of 240Ω was recorded This allowed to determine the delay of the current wave in respect to the surge voltage wave. Here the measurements were synchronized with the "time stamp" too.

Because of the necessity to reduce the field testing costs, it was decided to carry out a part of the investigations in laboratory conditions by the creation of a model corresponding with the best approximation to the contact line on the RI test track.

The following assumptions for the laboratory contact line model have been made:

- the model imitation of the contact line should be created with the elements of lumped constants,
- the model with regard to the electrical parameters should meet the asymmetric, homogenous long-line requirements,
- the contact line model should be composed of separate elements representing one-kilometre sections of the contact line,
- the resistance of the induction elements and connections within the model should be equal to the resistance of the contact line section of the test track,
- electric kicks (comparable as for their shape to the impulses occurring in actual conditions) of the maximum amplitude of 6.9 kV and the duration 1.2/50 μs and 10/700 μs will be introduced to the model circuits,
- the used elements have to meet the insulation criterion required for the interoperability with the kick generator,
- the induction and capacitance elements of the model should be arranged one to another in such a way that the possibility of arising inductive and capacitive couplings is eliminated,
- the parameters of the model components have to strictly correspond to the elementary parameters of one kilometre of the contact line,
- when introducing the kicks with an amplitude of 6.9 kV into the model, all safety conditions have to be assured for the operating teams and the measuring apparatus.

Basing on these assumptions, a laboratory model has been created, which was composed of 8 inductive elements and 8 earth capacitances. The subassemblies imitating the contact line section of 1 km in length have been carried out so that electrical parameters corresponded to the elementary parameters of the contact line used in the Żmigród test track. The following parameters of the one-kilometre module of the contact line have been assumed:

$$L = 466 \pm 5\ \mu H, \qquad R = 0.062\ \Omega, \qquad C = 10.5\ nF \pm 5\ nF.$$

The inductive element imitating the one-kilometre contact line section has been made in the form of a coil wound on an open ferrite core of a high magnetic permeability. The coil has been wound with insulated copper wire of 2 mm diameter. The inductivity of the wound coil was about 466 ± 5 μH and the resistance was 0.035 Ω. The capacitor representing the ground capacitance has been made with two metal plates separated with a dielectric. The capacitance of the capacitor made in this way was about 10.5 nF ± 5 nF and it was adjustable.

The electrical diagram of the laboratory model together with the values of the used elements is presented in Fig. 5.

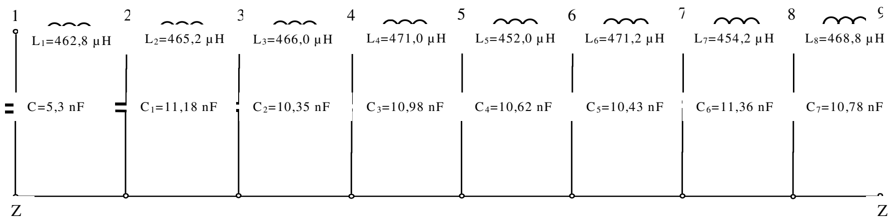


Fig. 5. Electrical system of the 8-km section of the contact line

In the investigations, the HAEFELY generator generating impulses of 1.2/50 μs (8/20 μs) and 10/700 μs shape with the maximum amplitude of the kick amounting to 6.9 kV was used.

To achieve the aim of the testing, it was necessary to establish the time relations and find out the phenomena occurring in the electrical circuit imitating the contact line as a long-line.

The laboratory investigation has been carried out on the created model, which imitated the contact line section of the length of about 7,760 km.

The kick amplitudes were recorded concurrently with two TEKTRONIX oscilloscopes of THS720 type. Fig 3 shows the basic measurement system. The measurements were carried out for two cases:

- with no load on the model output,
- with load with a resistance of 240 Ω.

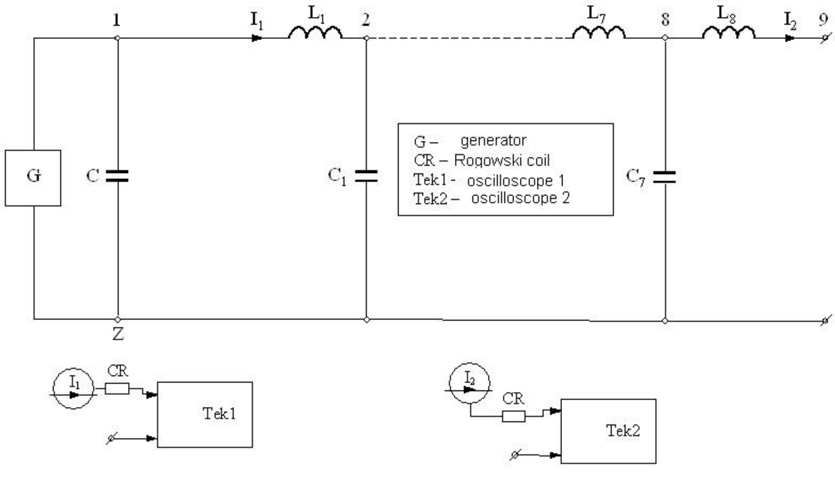


Fig. 6. Electrical measurement system used in the laboratory model of the concerned contact line section

For the creation of the model and further testing, the data related to the elementary parameters of the contact line obtained from the earlier testing [3] conducted on the Żmigród test track were used. The 240 Ω resistance corresponds approximately to the wave impedance of this contact line section.

3 Investigation Results

Fig. 7 shows the course of the recorded voltage impulse for the open-out contact line, without any load on the beginning of the line section (kick generator station), and Fig. 8 is the same, but on the end of the section. The voltage of the introduced kick was 5,900 V.

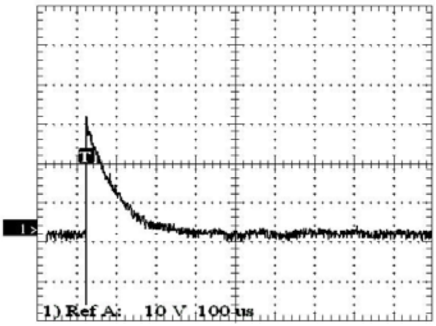


Fig. 7. The course of the voltage wave (the beginning of the open-out contact line section)

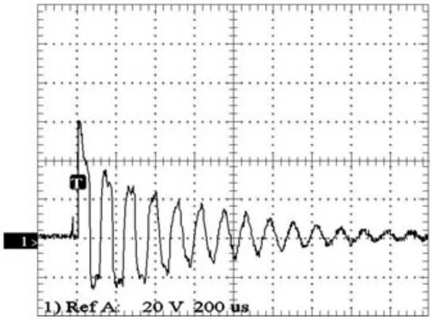


Fig. 8. The course of the voltage wave (the end of the open-out contact line section)

When comparing the amplitudes of the voltage waves presented in the above figures, it can be concluded that the tested contact line section behaves as an unloaded long-line. The amplitude of the impulse at the end of the measurement section increased from 5.6 kV up to 12.16 kV, whereas the impulse current on the kick generator output was 130 A, which proves the significant capacitance of the contact line in relation to the rails. The inductance and the capacitance of the tested contact line section formed the attenuated oscillations of the frequency of about 8.5 kHz.

In the second measurement variant (contact line loaded with a resistor), the impact of the resistance connected at the end of the measurement section on the kick amplitude measured on this resistance has been checked. The Fig. 9 shows the results of the voltage wave amplitude recording at the beginning and at the end of the section (Fig. 10) of the contact line loaded with the resistance of 240 Ω, respectively.

For the 240 Ω load resistance, the voltage wave amplitude has been measured and it amounted to 6.4 kV with the wave amplitude on the input amounting to 6.16 kV (Fig. 9 and Fig. 10, respectively). By comparing these voltage wave amplitudes, it can be stated that the 240 Ω load resistance is close to the wave impedance value for this contact line section.

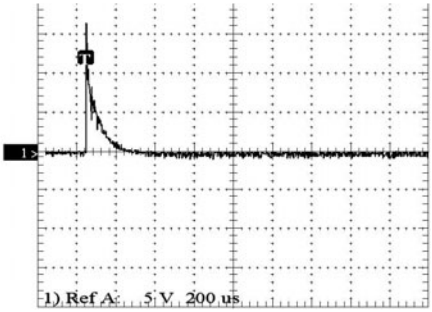


Fig. 9. The course of the voltage wave (the beginning of the contact line section loaded with the resistance of 240 Ω)

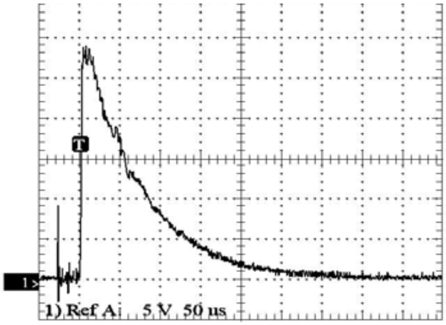


Fig. 10. The course of the voltage wave (the end of the contact line section loaded with the resistance of 240 Ω)

Within the second measurement variant, the input current amplitude values and the values of the amplitudes of the current flowing through the 240 Ω load resistance were also recorded. The amplitude of the input current was 140 A, and with the 240 Ω load resistance, it was 25 A. So significant differences between the amplitudes of the input and output currents are caused by the leakance through the capacitances between the contact line and the rails along the measurement section concerned.

4 Characteristics of the Overvoltage Attenuation on the Contact Line versus the Distance from Their Source

To select the layout of the elements of the contact line surge protection, it is necessary to know the distribution of the kick amplitudes in the function of the distance from the source of these kicks and the impact of the heterogeneity of the tested line section, such as contact line anchoring, network turnouts and viaducts, on the distribution of these kicks. Because of that, the kick amplitude distribution characteristics have been developed on the basis of the results of the investigations carried out earlier. In Fig. 11, the characteristic for the tested section of the contact line without load is presented, whereas Fig. 12 shows the same, but for the section with the load of 240 Ω.

When comparing the oscillograms, it was found that the kick amplitudes on the laboratory model output were as much as twice of the ones on the input. This was due to the reflection of the voltage wave from the open out long-line at the end. With no element constituting the load of the long-line, certain oscillations appeared on its output. Similar oscillations and the doubling increase of the amplitude were observed at the field testing on an unloaded contact line in Żmigród. This attests of the comparable electrical properties of both the laboratory model and the actual contact line section. It should be emphasised that with the developed model, similar current wave delays in relation to the voltage wave were obtained, as for the contact line section in the Żmigród test track (about 30µs).

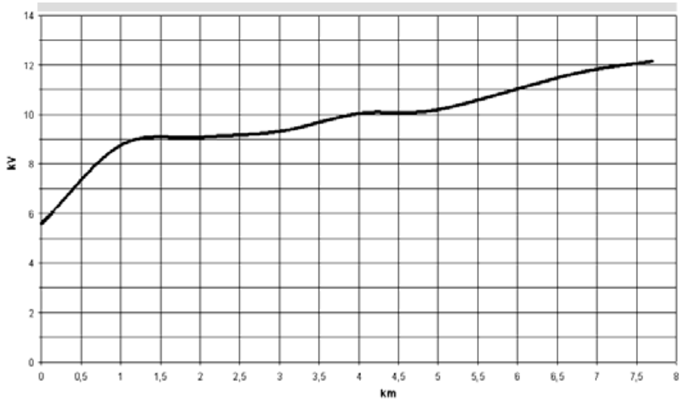


Fig. 11. The distribution of kick amplitudes versus the distance for the contact line section without load

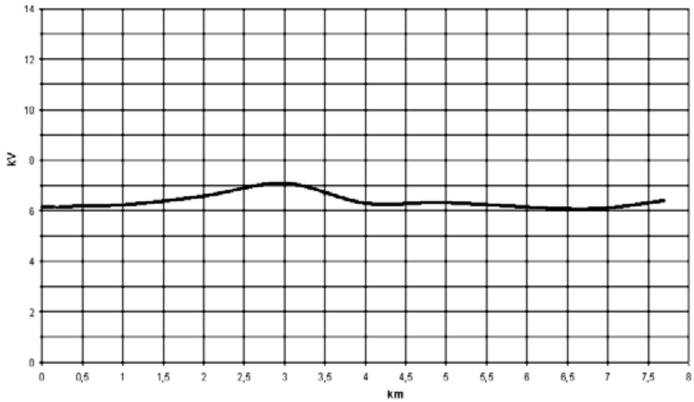


Fig. 12. The distribution of kick amplitudes versus the distance for the contact line section with 240 Ω load

Investigation on the laboratory model output loaded with the resistance of 240 Ω has also been performed. The investigation results show that no differences between the measured attenuation on the laboratory long-line model and the actual section of the Żmigród contact line were found.

The characteristics illustrating the distribution of the kick amplitudes versus the distance from the kick generator are shown in Fig. 13 and 14.

5 Conclusion

The considered section of the open-out contact line is characterised approximately with the properties identical to the homogeneous long-line. Due to the interference between the voltage impulse reflected from the unloaded end of the line section and the generator impulse, the kick amplitudes increase with the increase of the distance.

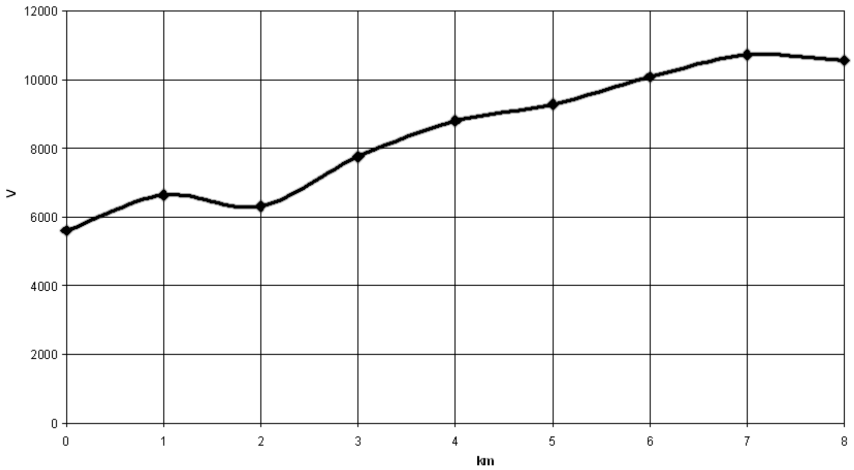


Fig. 13. Kick amplitude (1,2/50 μs) distrubution versus the distance from the generator for the model of the contact line w/o load

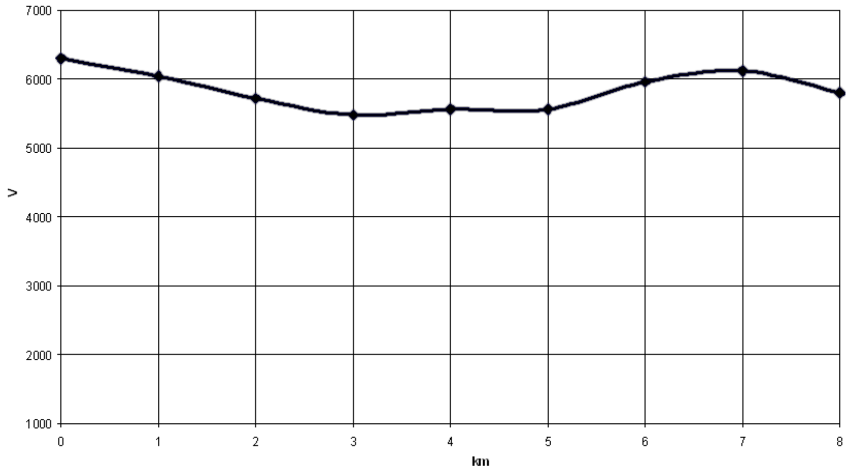


Fig. 14. Kick amplitude (1,2/50 μs) distrubution versus the distance from the generator for the model of the contact line with the load of 240 Ω

Analysing the courses, it should be noted that the characteristics of the distribution of kick amplitudes versus the distance for the contact line loaded with the resistance of 240 Ω have two specific places. In the first place (about 3 km from the generator), the voltage kick amplitutude increases. This is caused by the presence of a viaduct, which in turn contributes to the local increase of the contact line ground capacitance, and indirectly in relation to the railway rail, which causes a non-homogeneity. In the second place (at the end of the measurement section), the wave amplitude increases.

This is connected with maladjustment of the load resistance to the wave impedance value of this contact line section.

The presented results allow to state that the homogeneous contact line (i.e. without intersections, turnouts, branches, etc.) attenuates the voltage kicks to an insignificant degree only.

Carrying out field measurements were possible only by using one of the telematics tools, namely the use of so-called "time stamp to measurements synchronization".

The developed and created laboratory model of the section of the contact line as a long-line behaves within the electrical parameters as an actual section of the contact line (the section of the contact line in the test track centre in Żmigród).

The laboratory model of the contact line section has allowed to continue the investigations, which were necessary to establish the optimum distances between the varistor overvoltage limiters protecting the insulation of the contact line against overvoltages that may occur in it (including atmospheric discharges). During the laboratory investigations, the interaction of the varistor limiters at impulse waves acting on the contact line has been determined. Also, the interaction relationship between the horn gap and the varistor overvoltage limiter has been identified.

References

1. Praca CNTK 4291/10: Opracowanie nowego systemu ochrony sieci trakcyjnej przed przepięciami, badania eksploatacyjne nowego systemu, określenie lokalizacji podłączenia ochrony od urządzeń sterowania trakcja i urządzeń sterowania ruche. Etap1, Warszawa (2007)
2. Praca CNTK 3889/10: Budowa i poligonowe badania prototypowego systemu ochrony przed przepięciami z ogranicznikami warystorowym. Etap1, Warszawa (2011)
3. Praca IK: Budowa i poligonowe badania prototypowego systemu ochrony przed przepięciami z ogranicznikami warystorowymi. Etap1 Zbadanie i określenie na drodze pomiarowej tłumienności przepięć przez sieć trakcyjną w funkcji drogi. Część 2. Badania laboratoryjne (December 2011)
4. Praca CNTK nr 6915/23: Opracowanie dopuszczalnych parametrów zakłóceń dla urządzeń srk, łączności i pojazdów trakcyjnych. Warszawa (1999)
5. Mikulski J., Młyńczak J.: Pomiary parametrów powrotnej sieci trakcyjnej. In: Materiały XI Międzynarodowej Konferencji Komputerowe Systemy Wspomagania Nauki, Przemysłu i Transportu' TRANSCOMP, Zakopane (2007)
6. Białoń A., Furman J., Kazimierczak A.: Badania uszynień grupowych z wydłużoną długością sekcji. In: XIV Międzynarodowa Konferencja "Komputerowe systemy wspomagania nauki, przemysłu i transportu TRANSCOMP 2010" Zakopane (December 6-9, 2010)
7. Adamski, D., Białoń, A., Furman, J., Kazimierczak, A., Laskowski, M., Zawadka, Ł.: Problematyka tłumienności przepięć w sieci trakcyjnej 3kV DC. In: IX Konferencja Naukowo-Techniczna Logistyka, Systemy Transportowe Bezpieczeństwo W Transporcie, Szczyrk (April 17-20, 2012)

Geometrical Determinants of Car Equivalents for Heavy Vehicles Crossing Circular Intersections

Elżbieta Macioszek

Silesian University of Technology, Faculty of Transport
Krasinskiego 8, 40-019 Katowice, Poland
elzbieta.macioszek@polsl.pl

Abstract. Individual groups of vehicles affect the conditions of traffic streams to different extent. For a given group of vehicles, this effect should be approached as a variable determined by some geometrical and traffic properties. This impact can be estimated by using one of the models used for the purpose. One of such models is a model based on comparisons speed of group of vehicles in traffic stream. This paper presents the results of the measurements of the spatial average velocity of vehicles that cross circular intersections (usually roundabouts) and the analyses of the effect of selected geometrical parameters on the velocity of vehicles crossing an intersection and consequently the car equivalents for heavy vehicles.

Keywords: circular intersections, roundabouts, passenger car equivalents.

1 Introduction

The conditions in traffic streams of vehicles largely depend on the types of cars moving within this stream. In the most of actual traffic cases, traffic conditions are far from ideal, mainly due to the varied structure of the traffic stream. Proportion of heavy vehicles in traffic impacts on the quality of road traffic conditions. In analyses, presence of heavy vehicles in the traffic stream is predominantly modeled by means of car equivalents for individual types of vehicles. Equivalents for individual groups of vehicles refer to the number of passenger cars which have similar effect on conditions in the traffic stream as the analyzed vehicles. For comparison purposes, the mixed traffic stream is calculated from actual units [P/h] into equivalent [E/h] units, using individual equivalent values [E_i] for each "*i*" particular group of vehicles. In the most of modern guidelines (e.g. American guidelines HCM 2000 [3], Polish guidelines [12], German guidelines HBS 2001 [4], and Dutch guidelines [13]), car equivalents for heavy vehicles adopt constant values. There have also been other publications in the literature which demonstrated that the values of car equivalents for the types of vehicles should be approached as random variables [6, 11].

2 Determinants of Car Equivalents for Heavy Vehicles

Few studies have examined the effect of traffic conditions on the value of car equivalents in roundabouts to date (e.g. [7]). The most of the studies have focused on

J. Mikulski (Ed.): TST 2012, CCIS 329, pp. 221–228, 2012.

other elements of road infrastructure, such as internodal sections [11], intersections with traffic lights [1] or the areas of road and street networks. Differentiation between the type-related groups of vehicles can be obtained through field research on selected elements of transport network. Global identification of traffic flow with modal division of traffic is possible through e.g. application of modern technologies, such as GIS, GPS and GSM (for more details see e.g. [9, 10]).

The value of car equivalents in roundabouts can be affected by a number of factors, which in particular include geometrical properties (e.g. outside diameter, width of the circular roadway and the width of the main lane), traffic properties (e.g. traffic intensity in the circular roadway and at individual inlets, degree of traffic load) and other factors (e.g. location of the roundabout, atmospheric and psychophysical traits of a driver etc.) The effect of these factors can be the most noticeable under conditions of unlimited traffic with high values of traffic stream intensities. The results of the following studies on the values of car equivalents are presented below:

- the investigations of the effect of the width of the lane on two-lane two-way roads on the values of car equivalents were presented in the study [2]. The study found that the effect of the lane width in the mixed traffic stream is more noticeable when the vehicles do not move one after another in a column. It was also observed that the values of equivalents rise linearly with an increase in the lane width,
- the investigations of the effect of the defects in the surface of the inlets to the roundabouts on the value of car equivalents were presented in the study [8]. It was found that the better the quality of the surface in the inlet the higher the values of car equivalents. For example, the car equivalent for buses on the surface without defects amounted to $E_A = 5.40$ [-], whereas the same equivalent for the surface with defects was reduced to $E_A = 5.26$ [-]. Different values of car equivalents can be explained by different speeds of vehicles between driving on good quality and damaged roads,
- the investigations of the effect of directional traffic distribution in the intermodal sections on the values of car equivalents were presented in the study [8]. It was found that the values of car equivalents for heavy vehicles are lower than in the cases when directional traffic distribution is different than 50%/50%,
- the investigations of the traffic load on the value of car equivalents for heavy vehicles on roundabouts were presented in the study [6]. It was demonstrated that the changes in the values of traffic intensity on the roundabouts impact on the values of car equivalents. The car equivalents adopt higher values at higher traffic intensities, whereas their values stabilize with higher loads in circular roadways in roundabouts,
- in the study [1], the value of car equivalent for heavy vehicles was made dependent on: traffic intensity, type of the heavy vehicle and the proportion of heavy vehicles in traffic. The value of car equivalent for a particular type of heavy vehicle rose with an increase in traffic intensity and a proportion of heavy vehicles in the traffic and it ranged from $E_{SCi} = 1.0$ to 2.18 [-],
- the study [14] demonstrated that the value of car equivalent for the heavy vehicles in motorways rose with an increase in traffic intensity and decreased with an increase in proportion of heavy vehicles and the number of lanes.

3 The Effect of Selected Geometrical Properties on Car Equivalent

The author of the present study carried out the investigations in the roundabouts in order to determine whether the vehicle speed in a roundabout (and consequently the value of car equivalent for heavy vehicles) is affected by such geometrical properties as width of inlets, width of circular roadway and outside diameter. Because of the adopted measuring technique (measurement of the value of time intervals between the preset vehicle sequences) it was found that the values of time intervals are strongly correlated with the velocities of the vehicles. Therefore, the time of vehicles' crossing intersections was measured during the first stage of the study at each research plot. The time of passing was recorded with the following measurement cross-sections:

- cross-sections at the section before the roundabout: measurement was carried out at the section where the cars did not decelerate significantly yet before they approached the line where the cars are obliged to yield the right of way,
- cross-sections at the inlet to the roundabout: inlets are the locations were a significant decrease in the vehicle velocity can be observed, caused by approaching the line where the cars are obliged to yield the right of way,
- cross-section after leaving the circular roadway of the roundabout: the measurement concerned the time for the section from the moment of crossing the line of yielding the right of way to the moment leaving the roundabouts at the selected outlet,
- cross-section at the outlet of roundabout: the measurement concerned the time after leaving the circular roadway by the vehicle.

The analysis focused only on those vehicles which did not stop at the inlet in order to yield the right of way to the vehicles which were going on the roundabout's circular roadway.

After the determination of the time when vehicles went on the defined sections on roundabouts (the approach road before the roundabout, inlet, circular roadway, outlet), and, given the length of the road covered by the vehicles, it was possible to determine the speed of the vehicles passing the roundabouts for the stream which is homogeneous and mixed in terms of the composition (Fig. 1 and Fig. 2).

Based on Figures 1 and 2, it can be concluded that the passenger cars cross the area of intersections at the speed which is by 2 to 9 km/h higher with respect to the speed of heavy vehicles. It was verified based on the test for two variances whether the differences between mean speeds of passenger cars' and heavy vehicles' passing are statistically significant.

When setting the level of significance at $\alpha = 0.05$ yielded some grounds for rejecting the zero hypothesis and adopting the alternative hypothesis that suggested higher mean velocities when crossing the intersection with respect to mean passing velocities of heavy vehicles.

The observations demonstrated that the higher the widths of the lane at the inlet and at the circular roadway, the faster the vehicles cross the roundabout. However, the analysis of the results of the investigations did not support these observations. The

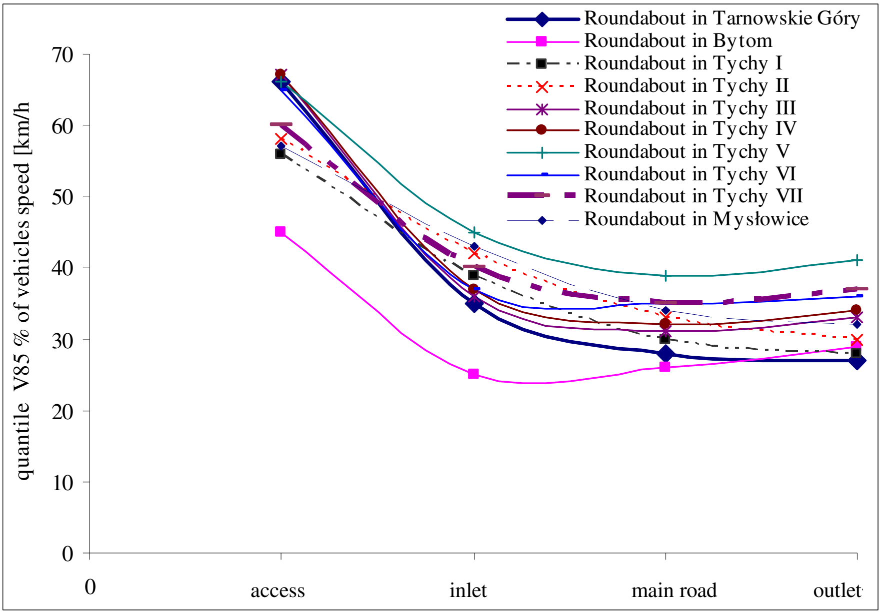


Fig. 1. Results of vehicles speed researches (quantile V_{85} % of vehicles speed) for traffic stream consisting only from passenger cars. Source: own.

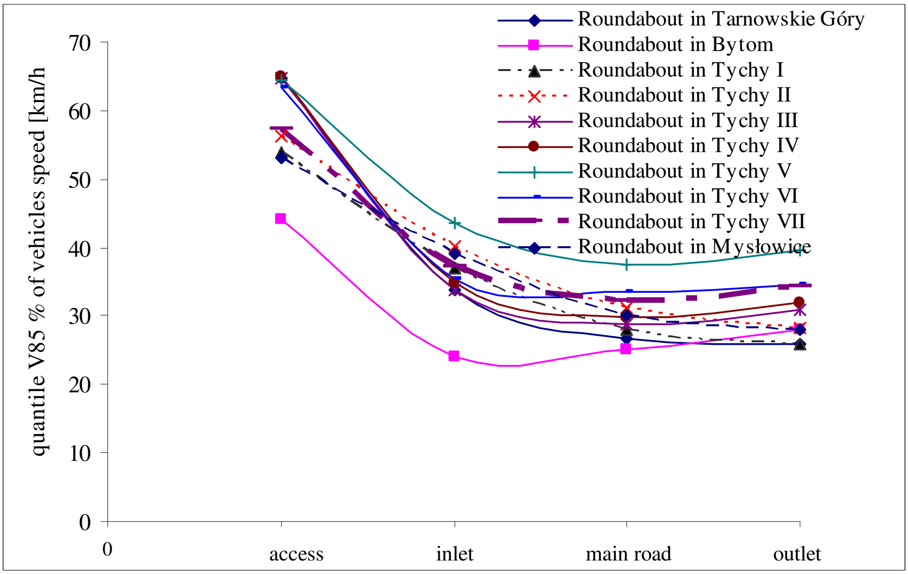


Fig. 2. Results of vehicles researches (quantile V_{85} % of vehicles speed) for traffic stream consisting from different type of vehicles. Source: own.

ranges in the quantile samples V_{85} % of the velocity of crossing at the inlets and on the circular roadway were poorly correlated with both width of inlets (R^2 =0.38 [-]) and the width of circular roadways (R^2=0.34 [-]) (Fig. 3 and 4). Therefore, it can be concluded that the speed of vehicles' crossing the inlet and the circular roadway does not significantly depend on the width of inlet and the width of the circular roadway in the roundabout.

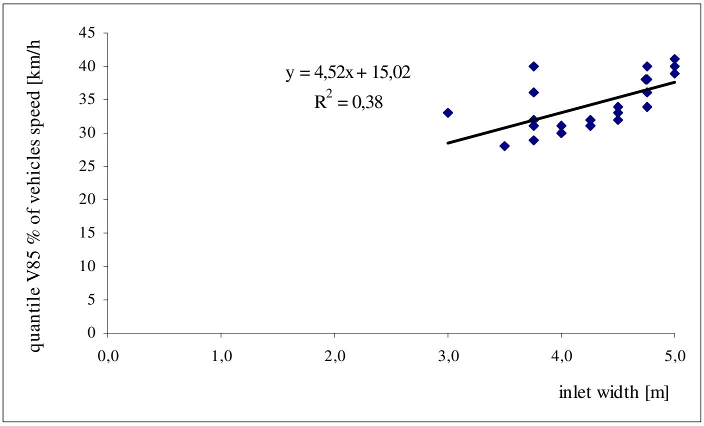


Fig. 3. The influence of inlet width on value of quantile V_{85} % of vehicles speed. Source: own.

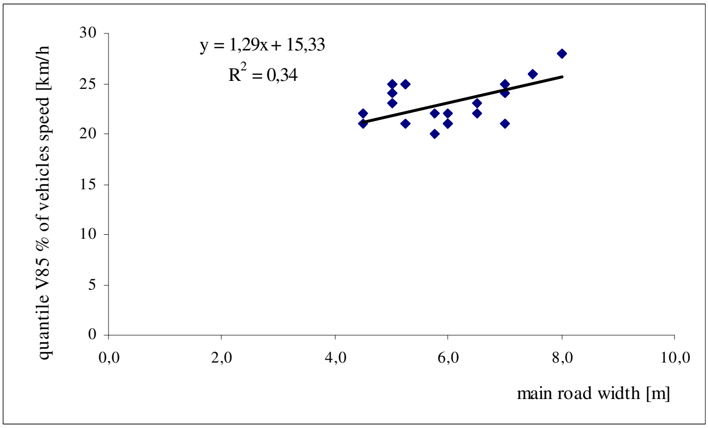


Fig. 4. The influence of main road width on value of quantile V_{85} % of vehicles speed. Source: own.

The next stage analyzed the effect of the external diameter on the speed of vehicles' crossing the inlet, circular roadway and the outlet of the roundabout in the traffic streams which were homogeneous and mixed in terms of their composition (Fig. 5-7). The results of the initial analysis show a comparable and statistically significant effect of the value of outside diameter of the roundabout on the value of quantile V_{85} % of the velocity of vehicles at the inlet, circular roadway and the outlet from the roundabout. The coefficient of determination for the homogeneous stream when crossing the inlet, circular roadway and the outlet, ranges from 0.59 to 0.66 [-], whereas this value for the mixed traffic stream ranges from 0.67 to 0.71 [-]. The velocity of crossing the roundabout rises with an increase in the outside diameter of the roundabout. It can be concluded that the geometrical properties of the roundabout determine to some extent the velocity of vehicles when crossing the roundabout. The

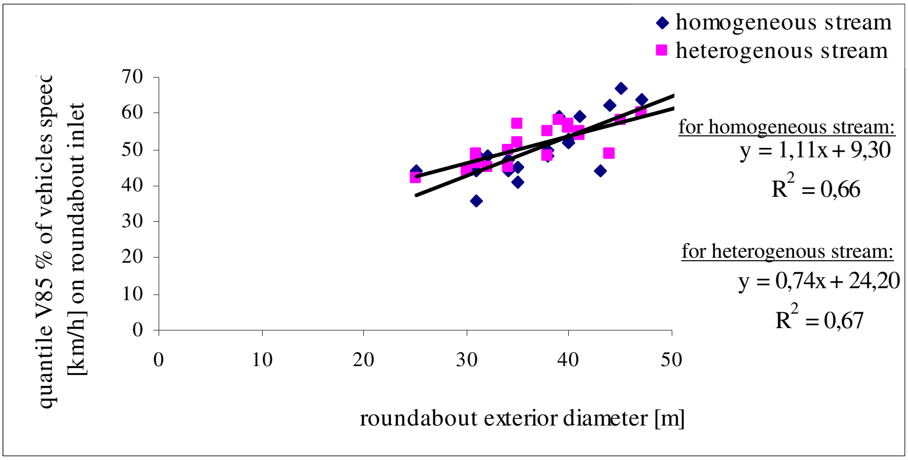


Fig. 5. Relationship between quantile V_{85} % of vehicles speed on roundabout and exterior diameter. Source: own.

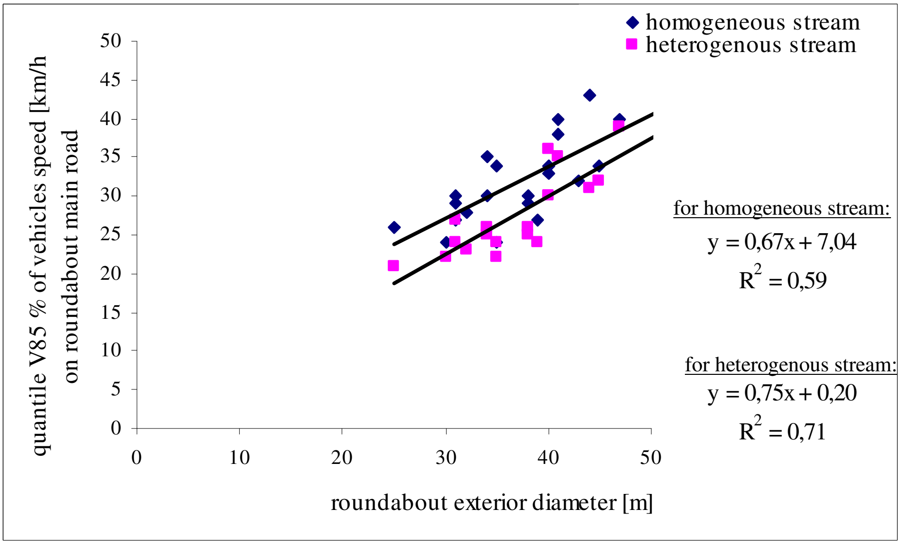


Fig. 6. Relationship between quantile V_{85} % of vehicles speed on main road of roundabout and exterior diameter. Source: own.

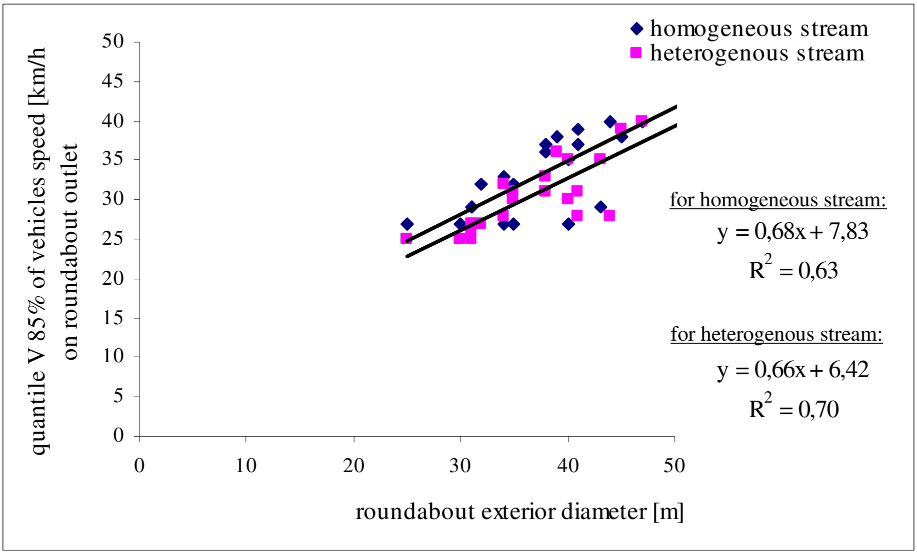


Fig. 7. Relationship between quantile V_{85} % of vehicles speed on roundabout outlet and exterior diamater. Source: own.

velocity of crossing can also be affected by other factors, not analyzed in this study, and by overlapping of several determining factors.

The last stage of the analysis was determination of the car equivalent for the heavy vehicles at the given quantiles V_{85} % of the velocity of vehicles in traffic stream (Fig. 8). The equivalents were determined for the vehicles moving on the circular roadway only for these values of velocity quantiles where the measurement data were available. It can be concluded based on the obtained results that the value of car equivalent changes with an increase in the velocity of vehicles that move on the circular roadway and ranges from 1.92 to 2.25 [-].

4 Conclusion

The following conclusions can be drawn from the investigations and the analyses presented in this paper:

– the results of the presented study demonstrate the need for viewing car equivalents as the variables dependent on some geometrical and traffic-related properties of the intersection. The presented correlations are only approximate. Determination of the definite relationships will be possible after conducting some more research,
– the results of preliminary analysis demonstrate a comparable effect of the outside diameter of the roundabout on quantile V_{85} % of the velocity of vehicles at the inlet of the circular roadway and the outlet from the roundabout. The coefficient of determination for the homogeneous stream (when crossing the inlet, circular roadway and the outlet) range from 0.59 to0.66 [-], whereas this value for the

mixed traffic stream amounts to from 0.67 to 0.71 [-]. The velocity of vehicles' crossing the roundabout increases with an increase in the outside diameter of the roundabout,

- the velocity of the vehicles' crossing the inlet and at the circular roadway showed no significant correlation to the inlet width (R^2 = 0.38 [-]) and the width of the circular roadway (R^2= 0.34 [-]),
- geometrical properties of the roundabout are to some extent determined by the velocity of crossing the roundabout. Nevertheless, other factors, not analyzed in this study, or a combination of several determinant factors might also have an effect.

References

1. Benekohal, R., Zhao, W.: Delay-Based Passenger Car Equivalents for Trucks at Signalized Intersections. Transportation Research Part A 34, 437–457 (2000)
2. Chandra, S., Kumar, U.: Effect of Lane Width on Capacity Under Mixed Traffic Conditions in India. Journal of Transportation Engineering 129(2), 155–160 (2003)
3. Highway Capacity Manual 2000. TRB, Highway Capacity Committee, Washington (2000)
4. HBS. Handbuch fur die Bemessung von StraBenverkehrsanlageb. Forschungsgesellschaft fur StraBen und Verkehrswesen e.V., Koln (2001)
5. Logghe, S., Immers, L.H.: Heterogeneous Traffic Flow Modelling with the LWR Model Using Passenger Car Equivalents. In: Proceedings of the 10^{th} World Congress on ITS, Madrid (2003)
6. Macioszek, E.: Analiza Wpływu Stopnia Obciążenia Ruchem na Wartość Współczynnika Przeliczeniowego dla Pojazdów Ciężkich na Skrzyżowaniach Typu Rondo. Logistyka 6, 2071–2079 (2010)
7. Macioszek, E.: The Passenger Car Equivalent Factor for Heavy Vehicles on Roundabouts. In: Janecki, R., Sierpiński, G. (eds.) Contemporary Transportation Systems. Selected Theoretical and Practical Problems. The Development of Transportation Systems. Monograph, vol. 256, pp. 127–137. Wydawnictwo Politechniki Śląskiej, Gliwice (2010)
8. Sachdeva, S.: Speed-Flow Relationship and Capacity Analysis for an Identified Intercity Road Network. Kurukshetra University, India (2004)
9. Sierpiński, G.: Travel Behaviour and Alternative Modes of Transportation. In: Mikulski, J. (ed.) TST 2011. CCIS, vol. 239, pp. 86–93. Springer, Heidelberg (2011)
10. Sierpiński, G., Celiński, I.: Koncepcja Użycia Technologii Sieci GSM do Identyfikacji Przemieszczeń w Miastach. Logistyka – Nauka 6, 2663–2669 (2011)
11. Stanley, J.: Passenger Car Equivalents of Trucks Under Lane Restriction and Differential Speed Limit Policies on Four Lane Freeways. University of Kerala, Kerala (2009)
12. Tracz, M., Chodur, J.: Metoda Obliczania Przepustowości Rond – Instrukcja Obliczania. GDDKiA, Warszawa (2004)
13. Verweij, C., et al.: Roundabouts-Application and Design. A Practical Manual. Ministry of Transport, Netherlands (2009)
14. Webster, N., Elefteriadou, L.: A Simulation Study of Passenger Car Equivalents (PCE) on Basic Freeway Sections. Transportation Research Part B 33, 323–336 (1999)

Development of Telematic System Solution for Public Transport Sustainability

Antons Patlins and Nadezhda Kunicina

Riga Technical University, Faculty of Power and Electrical Engineering
Riga, Latvia
{antons.patlins,nadezda.kunicina}@rtu.lv

Abstract. The development of telematic system for public transport control is based on the analysis of sustainable development principles, taking into account a list of impacts. The control of several transport modes, based on common principles is discussed. The business case of application is analyzed for the city of Riga. The telematic system solution offered in the current research is able to provide a common technological platform for a sustainable development of the transport system, including an embedded solution for the environmental impact assessment. The legal and historical requirements for statistical data, measurements and information flow are concentrated within one platform. A multi agent platform approach allows using this platform for the on-line transport system control, including the transport system dispatching, the changing of signalling plan, the managing of transport flow, preventing of accidents, and fast solutions for registered accidents.

Keywords: Public transport, telematic system, system sustainability, centralized control system solution, large system control, transport system, monitoring.

1 Introduction

The development of a public transport control system in cities requires considering a huge amount of factors. The maintaining of transport system sustainability requires developing transport system control procedures as well as strong measuring and decision making tools. The decision making in a large system requires statistical data. The list of statistical data in a transport system up to now is not standardised.

The International Association of Public Transport (UITP) has some useful research and opinions, like integrating public transport & urban planning: a virtuous circle published in [14], however the development of a sustainable transport system in the city is still a challenge.

There are several studies of the state of the art in the passengers transportation carried out recently in European cities in a wide range from passengers behaviour to technical solutions and implementation of electro-technical devices.

The analysis and comparison of transport systems in different cities as well as the analysing of future requirements of the transportation system is the topic of many

J. Mikulski (Ed.): TST 2012, CCIS 329, pp. 229–239, 2012.

research in several European projects within the CIVITAS Initiative under FP7 (CIVITAS MIMOSA, CIVITAS VANGUARD, etc.).

Analyzing various projects and literature references it is possible to resume the main unsolved questions in the subject.

There are a lot of definitions of sustainability, but they are scattered in different sources of information. In the literature or in the standardisation documentation there are no fully defined principles of public transport system sustainable development. There are also no common opinions about the urban public transport safety systems and about controlling the public transport system as a whole.

2 Legislative and Natural Requirements of the Parameters Monitoring in the Public Transport

The main issue related to the passengers service is still not formalized, there are Draft TSI of Telematic Applications for Passenger services (Developed by the European railway agency). This recommendation will be mandatory for the EU Member States after the positive consultations with actors involved and the EU acceptance procedure. These recommendations are mainly related to long–distance transportation at the moment (as a railway transport), but in many cases it is important to take up the same principles, to define a route for special categories of passengers, like persons with reduced mobility. Another factor, which should be taken into account, are the Noise directive requirements, such as noise maps [19]. The national metrology centers established in Member States are already responsible for the collection and processing of environmental, metrological, and other information. As an example, they carry out environmental monitoring and inform the society on the environmental situation, ensure the geologic supervision of earth mineral resources and of rational use of them, and realize the state policies in the spheres of geology, meteorology, climatology, hydrology, air quality, and cross-border air pollution influence [15]. At the moment there are no common requirements for a sustainable development of a system, as well as criteria and indicators of sustainability are not standardized.

3 System Sustainability

The definition of sustainable transport indicators – see [16].

The sustainable transport indicators (STIs) are regularly updated performance measures that help transport planners and managers to take into account a full range of economic, social and environmental impacts of their decisions.

According to [17] the reporting is based on formal DPSIR approach, using selected (out of a total of 40) TERM ("indicators of the transport and environment report mechanism") indicators developed by the [17] for the transport sector.

An unequivocal allocation of the indicators to only one of the five groups is not always possible, but the DPSIR-approach can be seen as a useful model, but the technical solution to take up all requirements, using the same technical platform, is not still developed.

Table 1. DPSIR classification of selected TERM indicators [17]

Number	Indicator	Category (within DPSIR)
TERM 01	Transport final energy consumption by mode	D
TERM 02	Transport emissions of greenhouse gases	P
TERM 06	Fragmentation of ecosystems and habitats by transport infrastructure	S
TERM 09	Transport accident fatalities	I
TERM 37	National monitoring systems	R

4 Control of the Large Dynamic System

The control of transport system, taking into account a large system control methodology, is used to define a wireless telematic system solution for a sustainable public transport. The control of large scale dynamical systems is one of the biggest challenges facing control engineers today. Large scale systems are common in applications such as the chemical process control, the power generation and distribution and the highway and air traffic control, among others. Multi-agent, scarce resource problems are of particular interest, where a large number of agents, equipped with communication and control capabilities have to make efficient use of a scarce resource. The size and complexity of these systems makes it difficult to approach those using tools from the classical, central control literature. To manage the system complexity, designers are often forced to use hybrid control schemes, i.e. designs that make use of both discrete and continuous controllers [1].

The latest achievements [3] allow analysing a transport system as an active system – a model of organizational system, which takes into account the existence of divergent interests among the subjects of management (agents) and the active behaviour, that is providing information to the Governing Body (centre) and the choice of actions based on their own interests. The main method of research is mathematical modelling (systems analysis, decision theory, operations research) of procedures for making the management decisions.

The theory of active systems associated with the mechanism of open management of active systems proposed in 1969 by V.N. Burkov, in which the conditions for the maximum of the objective functions of agents (the conditions of perfect agreement) were used as the constraints of the problem of choosing the optimal system plan. The interests of agents guarantee the accuracy of information provided.

In some tasks a transport system could be defined as a linear system, in some as a nonlinear (amount of vehicles). The main issue is to build such a control system, to achieve its optimal condition via critical parameters. The adaptive management application for a transport system will bring benefits in the control task [4].

Let's define the transport system as a dynamic system [5]. Let's assume X and Y as input and output signals/data. So we can also talk about the changes of the system time. Such systems are characterized by changes in time. The system can be described at different time moments and can define the necessary parameters, or to perform calculations continuously at all times. Such control is called a dynamic process control.

For dynamic systems, knowing the set of input parameters X1 at time t1, it is possible to develop such control C at time t2, that it would be an output parameters set Y2, so we can make the technological system regulation by following condition.

$$S(t_1) \xrightarrow{C^1} S(t_2) \tag{1}$$

4.1 Transport Systems

A public transport system could be formally defined as a large system with many different situations during the operation. Such system could be formally analysed as a large system under conditions of risk and uncertainty.

The transport system is usually divided for simplicity of analysis into different subsystems by transport modes or by transported issues: road transport, railway transport, maritime transport, air transport, river transport as well as passenger and freight transport. There are special subsystems, for example an emergency transport (police etc.) as well as special transport categories (military, dangerous goods ((Directive 2008/68/EC - inland transport of dangerous goods of the European Parliament and of the Council of 24 September 2008 on the inland transport of dangerous goods. This Directive replaces Council Directive 94/55/EC, Council Directive 96/49/EC and Council Directive 96/35/EC.)).

All this different categories should be dynamically analysed by different actors of the system, many measuring systems are defined to formalise the categories of data requested.

Most frequently signals and impulses are characterized by a mathematical model S(t), U(t), F(t) etc. The transportation system management is provided with a control signal. Functions describing the signals can be both real and imaginary. If the signal is described by some function of time, it is a one-dimensional function. If the output signals are characterised by a number of time functions of the given sequence, then these are multidimensional functions.

The notion of transport systems suggests the idea that different elements are linked and interdependent, forming a system, "a combination of interacting elements organized to achieve one or more stated purposes" [6]. In the case of transport, this idea expresses the interaction of traffic, infrastructures, mobility, and other components, to produce transport (and traffic).

Mobility and accessibility are two aspects, which attract particular attention in the transport literature. The implementation of sustainable development demands the design of systems consistent with ecological principles, namely, an economic development subject to constraints imposed by the natural systems. The basic criteria to design and construct an environmentally sustainable transport system are: environmental stewardship by a co-ordinated multidisciplinary teamwork, the implementation of best practices and policies and the development of environmental performance indicators [7]. The combination of provision of accessible and effective mobility while improving the natural, built and social environment is the essence of sustainable transport. For these reasons, engineers are increasingly undertaking design problems in which a holistic understanding of natural systems in needed. Transport projects can be planned, designed, built, operated and maintained in such a way that when assessed, on an overall basis, they represent a net positive impact on the environment.

4.2 Sustainable Development

A comprehensive definition of a sustainable transport and of a sustainable transport system as well as principles of sustainable transport are partly defined in [8].

The World Commission on Environment and Development, called Brundtland Commission, is usually credited with one of the first definitions of what is understood today by sustainable development. In Our Common Future, the Commission defined the sustainable development as the "development that meets the needs of the present without compromising the ability of future generations to meet their own needs". The thesis continues: "It contains within it two key concepts:

- The concept of "needs", in particular the essential needs of the world's poor and future generations, to which an overriding priority should be given, and
- The idea of limitations imposed by the state of technology and social organisation on the environment's ability to meet present and future needs".

Analysing the literature, it is possible to formalize a public transport system sustainable development:

$$PTSSD(\Delta T) = \langle E1, Qs, E2, G \rangle(\Delta T) \quad (2)$$

Where: PTSSD(ΔT) - public transport system sustainable development in time period (ΔT); E1 – Environmental sustainability; Qs – Quality of service for passengers, including safety, security and comfort level; E2 – Economical sustainability aspects of transportation; G – Influence of Governance; (ΔT) – Time period.

5 Three-Level Control for Public Transport System Sustainable Development

It is offered to create a three-level centralized control system of passenger transport of the city of Riga. The first level comprises the vehicle equipment and processes control. At the second level the control of each type of transport will be carried out. The third level will allow the supervision of the system as a whole.

According to the control theory and system theory fundamentals, a scheme is offered for a global control of public transport system.

St – transport system; {St1-Stn} 0 St – subsystems of the whole transport system; {Tr1-Trn} – kinds of public transport; {V1-Vn} – vehicles of public transport; {I1-In} – infrastructure objects in the public transport system; {P1-Pn} – passengers; {Cr1-Crn} – the list of sustainable development criteria for the public transport system: (environmental sustainability, passenger safety, comfort level, economic aspects, influence of governance).

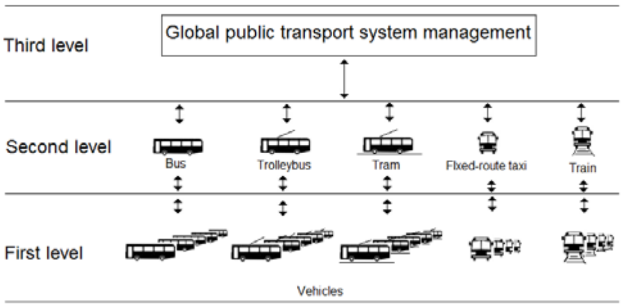


Fig. 1. Three-level public transport system management concept

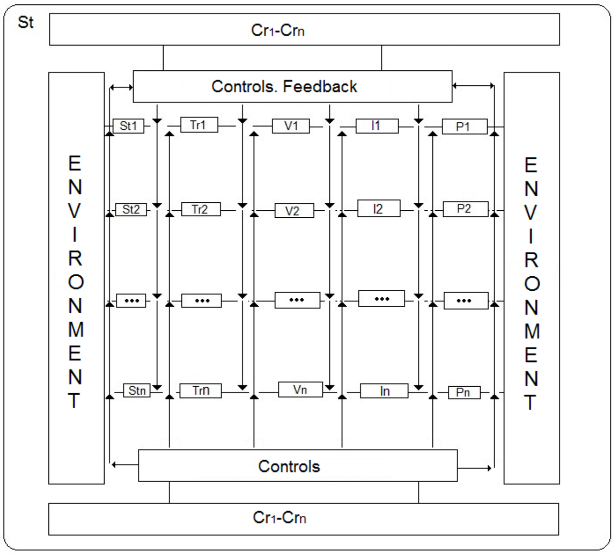


Fig. 2. Transport control system structure. Third (global) control level.

St – transport system; W – Influence of environment; {St1, St2, St3 } 0 St;

St1 – trams system; St2 – trolleybus system; St3- bus system;

{St11-St1n } 0 St1 - trams.{St21-St2n } 0 St2 - trolleybuses. {St31-St3n } 0 St3 - buses.

X – input signals/parameters ; X={X1, X2, …Xn}

Y – output signals/parameters; Y={Y1, Y2, …Yn}

C – control for all the system; C={C1, C2, …Cn}

C1 – tram system controls; C1={C11, C12, …C1n}

C2 – trolleybus system controls; C2={C21, C22, …C2n}

C3 – bus system controls; C3={C31, C32, …C3n}

C1v – trams controls; C1v={C1v1, C1v2, …C1vn}

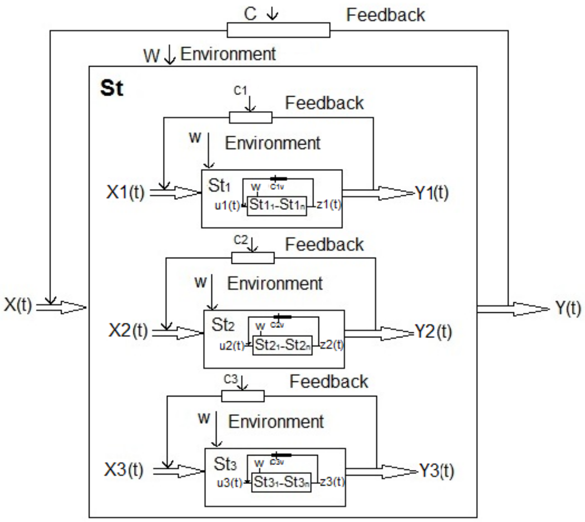


Fig. 3. Transport control system structure from the system theory point of view, using a 3-level transport control system concept. The second control level.

C2v – trolleybuses controls; C2v={C2v1, C2v2, …C2vn}

C3v – buses controls; C3v={C3v1, C3v2, …C3vn}.

This system has many variables and a lot of regulators. Such systems are called multi-circuit control systems.

The bus management system in the city of Riga is an example how the second level of public transport system control can work. See Figure 4.

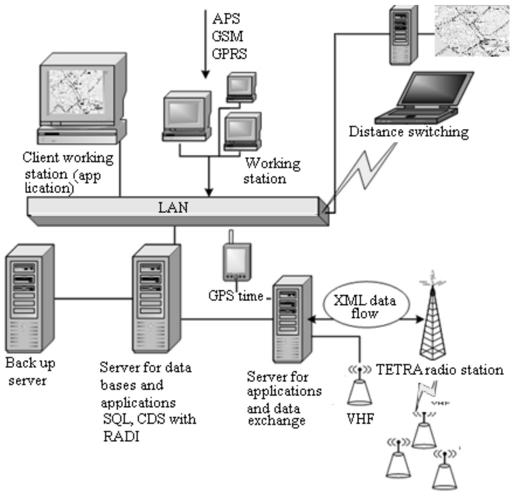


Fig. 4. An example of second level public transport control system – structure of ASOS bus management system in the city of Riga

The public transport control system requires having control procedures, which should:

- carry out the control over the movement of city passenger transport with a possibility of the operational control of movement,
- provide a possibility of carrying out voice communication sessions between dispatchers and mobile objects in certain places of a city, for example on movement terminal points. The given function allows to regulate manually an exchange order, in case of any non-standard situation,
- provide support of an emergency communication in case of breakage of vehicles within all the city territory. A vehicle driver receives a direct communication channel with emergency services or communicates through the dispatching node. The emergency communication can be made active in any point of the city,
- give a possibility at any moment to call any system subscriber from the dispatching centre,
- provide the protection against offenders in case of an attack, and also a security system of the vehicles fleet within all the city territory. A vehicle driver receives a direct channel to "Ambulance" and "Militias" services to transfer a voice message. An alarm signal giving by "a latent button" is possible,
- provide gathering of the information from on board equipment (sensors, the measuring equipment, buttons) and fixing sensors on the base equipment,
- provide an information transfer on stationary display equipment (for example digital boards at stops),
- support the possibility of the emergency notification of the population through a complex infrastructure. The notification through on board sets (a circular call on all vehicles) and stationary equipment (boards at stops).
- support the possibility of traffic regulation by the public transport. Regulations can be carried out in an automatic mode displaying the amendments to movement on the display of the on board set,
- provide uniform time for all system objects, which is displayed on the on board sets and which is corrected automatically from the dispatching centre,
- provide possibilities of self-diagnostics and the visualization of devices and components condition [13].

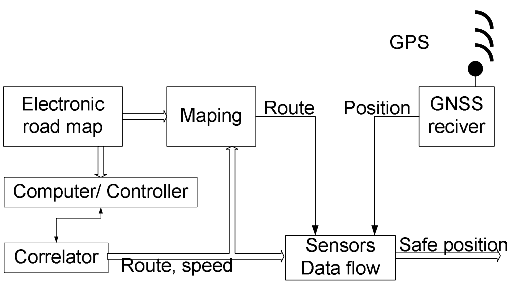


Fig. 5. An example of second level public transport control system – structure of ASOS bus management system in Riga city

The modules of monitoring containing a GPS-receiver, a microcontroller, a flash-memory and a GSM GPRS modem are mounted in vehicles. Apart from GPS and GSM aerials, the GPS monitoring module is probably connected with sensors of the fuel level, sensors of the fuel consumption, sensors of mechanisms work, temperature sensors, sensors of the passenger traffic volume, of ignition, of oil pressure and other sensors. It allows collecting in real-time the information significant for further analysis [13].

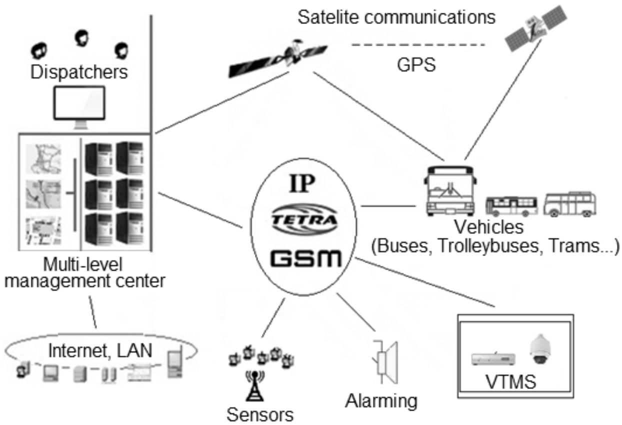


Fig. 6. Wireless communications between all system objects

6 Quality of Service for Passengers

The persons with reduced mobility (PRM) at the moment use the passenger transport to a very limited extent. The EU community has carried out many studies and has developed strategies to make the public transport easily accessible to PRM. This is shown by such documents as EU directive 2008/164/EC, research projects as ACCESS 2 ALL, Isemoa etc. The requirements of EU directives are mandatory only for new infrastructure objects; the reconstruction of existing stations and providing the access for all kinds of passengers is costly and takes time. They are changing the infrastructure and such a 3 – level control system allows increasing the comfort of PRM passengers. The real-time vehicle tracking allows planning the optimal route for PRMs.

A bi-directional communication with the operator will help persons to achieve the transport and to get the information about technical possibilities of stops from the operator via any available communication channel.

7 Conclusion

The telematic system solution is able to provide a common technological platform for a sustainable development of transport system. The legal and historical requirements for statistical data, measurements and flow information are concentrated within one platform. The multi agent platform approach allows to use this platform for on line transport system control, including the dispatcher centre of transport systems, the changing of signalling plan, managing the transport flows, preventing accidents, and fast solutions for registered accidents.

References

1. Lygeros, J.: Hierarchical, Hybrid Control of Large Scale Systems. UCB-ITS-PRR-96-23. California PATH Research (1996)
2. Deshpande, A., Godbole, D., Gollu, A., Semenzato, L., Sengupta, R., Swaroop, D., Varaiya, P.: Automated highway system tool interchange format. (preprint) California PATH Technical Report, Institute of Transportation Studies, University of California, Berkeley (1996)
3. Institute of Control Sciences V.A. Trapeznikov Academy of Sciences (2012), `http://www.ipu.ru/node/11925` (date of access: March 15, 2012)
4. Institute of Control Sciences V.A. Trapeznikov Academy of Sciences, `http://www.ipu.ru/node/11920` (date of access: March 15, 2012)
5. Greivulis, J., Kuņicina, N.: Analogo Elektronisko shēmu analīzes un aprēķina metodes. RTU izdevniecība (2009)
6. International Council on Systems Engineering (INCOSE), Systems Engineering. Handbook: A Guide for System Life Cycle Processes and Activities, Version 3. International Council on Systems Engineering, San Diego, CA (2006)
7. Kehagia, F.: The implementation of sustainability in highway projects. Int. J. of Sustainable Development Planning 4(1), 1–9 (2009)
8. Hall, R.P.: Understanding and Applying the Concept of Sustainable Development to Transportation Planning and Decision-Making in the U.S. PhD thesis, Massachusetts Institute of Technology, Cambridge, USA (2006)
9. WCED. Report of the World Commission on Environment and Development, United Nations General Assembly, A/RES/42/187 (December 11, 1987), `http://www.un.org/documents/ga/res/42/ares42-187.htm` (date of access: March 15, 2012)
10. Matveev, A.: Development of Methods, Algorithms and Software for Optimal Design of Switched Reluctance Drives, 248 p (2006)
11. Shetty, D., Kolk, R.A.: Mechatronics System Design, 2 edn., Cengage Learning, USA, 60 p (1997)
12. Patļins, A., Galkina, A., Kuņicina, N., Čaiko, J.: Design Alarm Processing Techniques and Tools for On-Line Monitoring of Mechatronic Systems. In: MSM 2011: 7th International Conference Mechatronic Systems and Materials, Lithuania, Kaunas (2011)
13. Patlins, A., Kunicina, N., Galkina, A., Ribickis, L.: Development passengers transfer procedure for city transport in Riga. In: 17 ITS World Congress Busan, October 25-29, BEXCO, Korea, Busan (2010)
14. International Association of Public Transport (UITP), `http://www.uitp.org/advocacy/positions.cfm` (date of access: March 15, 2012)
15. Latvian Environment, Geology and Meteorology Centre, `http://www.meteo.lv`

16. Lee, R.W., Wack, P., Jud, E.: Toward Sustainable Transportation Indicators for California. MTI REPORT 02-05, Mineta Transportation Institute, San José State University, San Jose, Ca, USA (2003)
17. EEA. Transport at a crossroads – TERM 2008: indicators tracking transport and environment in the European Union. EEA Report No 3/2009, European Environment Agency, Copenhagen, 52 p (2009b)
18. Joumard, R., Gudmundsson, H. (eds.): Indicators of Environmental Sustainability in Transport. An Interdisciplinary Approach to Methods, COST 356 final report. Institut national de recherche sur les transports et leur sécurité – INRETS, France, 422 p (2010)
19. Environmental Noise Directive (Directive 2002/49/EC)

A Proposed Model of Extended Coordination for the Traffic Light Control at Intersections

Filip Tresler

VŠB-Technical University of Ostrava, Institute of Transport
17.listopadu 15/2172, 708 33 Ostrava-Poruba, Czech Republic
filip.tresler@email.cz

Abstract. The contemporary method for the coordination of urban intersections design for the same speed of approach of around 45 km·h^{-1} does not allow changes in the coordination speed based on faster changes of the required intersection capacity resulting from changed intensity of traffic during a day. However, the speed of traffic flow changes significantly with the road saturation. The coordination is often ignored in off-peak intensities, when the minimisation of the waiting times decreases the cycle length of light signaling equipment (LSE). However, this leads to changed green waves (in time), which means that the coordinated groups of vehicles pass the intersection at different times. Because of fixed constant distances between intersections, the preservation of the green wave might not be possible and even at a lower traffic intensity the number of cars stopping can increase, the waiting times prolong and therefore the traffic smoothness can be deteriorated. The model of extended coordination defines the dependence of duration of traffic lights controlled intersections cycle on the intensity of the coordinated entry, the speed of approach and the distance between the coordinated intersections. The length of the cycle defined this way improves the traffic smoothness by the coordination of traffic light controlled intersections also in the cases of unfavourable conditions for bidirectional coordination.

Keywords: model, coordination, traffic lights controlled intersections.

1 Introduction

The principle of traffic lights control at an intersection alternates the allocation of preference to particular traffic flows, which are in collision to each other. After the introduction of traffic light control, the arrangement is such, compared to a non-controlled intersection, that the main flow is periodically stopped on the red signal and enables therefore smooth passage of vehicles from side directions on the green signal. The alternation of these phases can dispatch, upon sufficient capacity of the design signal plan, all vehicles arriving at the intersection both during the red signal, and during the green signal. If the traffic lights controlled intersections are in close proximity, it is obvious that it will be suitable to allow the created queue passing one intersection to pass also through the next intersection, appropriately through more

J. Mikulski (Ed.): TST 2012, CCIS 329, pp. 240–248, 2012.

intersections in a formation of several intersections (controllers). The implementation of intersections coordination on the main roads with priority results in the reduction of fuel consumption and therefore of exhaust gas emissions, the noise levels are improved, and thanks to the increased speed homogeneity, also the traffic safety improves and the death toll reduces, and even the impact of stress on drivers in unsmooth traffic is inconsiderable.

Various systems for the optimisation of vehicles passage in a road network equipped with LSE are already sufficiently known, proven and deployed in practice, whether it involves autonomous groups of controllers, or systems with super ordinate control level. Their purpose is to control the involved traffic area so that to reduce as much as possible the average travel time or the average delay, namely in the peak intensities during a day. By now, however, the control of passage in off-peak intensity has been neglected, when the LSE length of the cycle is reduced to minimise the waiting times. Because of the given constant distances between intersections, therefore maintaining of the green wave might not sometimes be possible at all and as a paradox, even at a low traffic intensity, the number of vehicles stopping and waiting times might increase and therefore the traffic smoothness can be limited. The purpose of this article is to present a proposal of a model for the control of line coordination, which will increase the traffic smoothness also in the above mentioned conditions, unfavourable for usual coordination. The output from the model of so-called "*extended coordination*" is a suitable length of the LSE cycle based on the intensity of coordinated inflow considering the total intersection capacity, the speed of approach coordination and the distance of coordinated intersections.

For the proposal of the model of the extended coordination, it is necessary to understand the basic coordination project specified in Technical Conditions 81 (TC – regulations for designers). Therefore the following chapter will cover the basic introduction into the theme of coordination of traffic lights controlled intersections according to the standard proposal.

2 Line Coordination - Green Wave Principle

Utilization of this control method is suitable for the coordination of line roads and areas with homogenous course of traffic intensity during the day. In line intersections coordination (green wave), the calculated coordination parameter is used as the basis, so called "*offset*". It defines, based on distances between the intersections, and based on the speed of cars cluster, the time interval of green signals initiation on the signal devices in the coordinated direction. The operation conditions of line coordination require the same length of the cycle at all intersections, a real speed of approach (verified by measurements), the knowledge of time variations of intensity etc.

The above mentioned speed of approach (velocity) V_P is therefore defined as the speed of the vehicles queue (green line) passing through the formation of intersections without stopping. It is usually marked in the distance-time diagram by the angle of the axis line of the green line towards the time axis. The choice of the speed of approach V_P is recommended according to the following rule [1]:

$$0,85 \cdot V \leq V_p \leq V \tag{1}$$

Where V........ speed limit (km per hour).

At high traffic intensities (the used capacity level min. 0.8), the required speed of approach in fact is not usually maintained due to the traffic flow density [1], [3]. The use of a low speed of approach also is usually not successful because drivers do not maintain an overly low speed. It has been proven in practice that the proposed coordination speed on urban roads should not be under 40 $km \cdot h^{-1}$ and on the other hand, it is not possible to use a higher speed than the speed limit on the involved road section (typically 50 $km \cdot h^{-1}$).

The law [2] provides that it is possible to increase the speed limit by a local provision, however max. by 30 $km \cdot h^{-1}$. This provides space for utilization of variable traffic signs – the speed limit. If the experience with influencing the traffic flow is transferred from fast roads to coordinated LSE controlled intersections, operated exclusively in the urban area at speeds given by the law on roads operation, it is possible to suppose that a small increase in the speed limit and its strong presentation, for example through elements used in systems of traffic line control, could result in improved comfort of the road network users. An increased capacity of roads and improved smoothness of vehicles movement between LSE controlled intersections should be the result. Considering the projected extension of the possible speed of approach of coordinated vehicles flow, the presented coordination model has been named "*model of extended coordination*".

3 The Model of Extended Coordination

Before determination of the final model of extended coordination it is necessary to present the basic initial formulas.

3.1 Dependence of the Capacity of the Entry (Intersection) on the Length of the Cycle of the Signal Plan

Any coordination of intersections needs to be examined with regard to their capacity. The relation defining the capacity of the entry into the intersection is given by the term [1]:

$$c = S \cdot \frac{g}{C} \tag{2}$$

Where c .... capacity of the entry lane (vehicles per hour)
S .... basic saturated flow of the entry lane (vehicles per hour)
g .... length of the green time (s)
C .... length of the cycle (s).

In the practice, however, you cannot consider only the duration of the green light (capacity) of one entry regardless the total required capacity of the intersection. To identify the dependence of an intersection capacity or the examined entry to the intersection on the cycle length we suppose that in the cycle length the length of the green of the most important (usually the main) direction is decisive. If we state that the total LSE cycle length is given by the sum of the time of green for the main (coordinated) direction, the time of green for the other (side) entries and the sum of the decisive inter-green times in sequential phases, it is possible to complete the formula defining the capacity of the entry to the intersection:

$$c_C = S \cdot \left(1 - \frac{\sum g_m}{C}\right) - S \cdot \frac{\sum t_i}{C} \tag{3}$$

Where c_C....... capacity of the coordinated entry (vehicles per hour)
Σg_m.... duration of the green light of other side entries (s)
Σt_i...... sum of decisive inter-green times between particular phases per .... cycle (s).

The term in brackets will be named "*coefficient of utilization of reduced cycle duration by the coordinated direction*" and we will mark it n_c.

3.2 Dependence of the Length of the Cycle on the Coordinated Intersections Distance and the Speed of Approach

The proposal of the green signal in the coordinated direction is limited by peripheral conditions, when one of them is the mutual time position of the coordinated lines which is influenced by the intersections distance (unchanging condition) and further on by the speed of approach (variable condition). In the case of bidirectional green waves there is a characteristic point on the distance-time diagram, so called "*dividing point*", which is the crossing point of the central lines of two inverse green flows. The distance of the points can also influence the intersection capacity, because for the phases division certain basic structures result from the distance of the intersection from the dividing point. For the bidirectional roads, it is ideal for the green wave, if the distances of the dividing points are identical with the distances between the intersections centres. However, it is not always possible to achieve equality of the dividing points and the position of intersections. We will mark the distance between the dividing points as d_{DP}. The following relation is valid between the time of the cycle, the speed of approach identical in both directions and the distance of the dividing points:

$$C = \frac{7{,}2 \cdot d_{DP}}{V_P} \tag{4}$$

To finally determine the dependence of the cycle length on the intensity of the coordinated entry, the speed of approach and the distance between the coordinated intersections, we will use the previous formulas (3) and (4).

The number of vehicles, that can really cross the stop-line profile per one hour is the capacity of the entry into a controlled intersection and it depends [4] on the ratio of the green signal in the control cycle and the saturated flow (2). Therefore, to define the dependence of the cycle length on the intensity and the speed of the coordinated traffic flow in formula (3), we will replace the saturated flow with the formula defining the dependence of the traffic flow intensity on the speed with appropriate parameters. A macroscopic speed-intensity model [3] is chosen, where variables – the maximal density K_j and the speed on a free road V_f – represent parameters of the concrete traffic flow on the concrete road. In formula (3), where the substitution of the macroscopic model of the traffic flow for the saturated flow, after the adjustment with the use of formula (4), gives us the resulting formula defining the cycle length versus the intensity, the speed of approach and the distance between the coordinated intersections:

$$C = \frac{K_j \cdot \frac{7{,}2 \cdot d_{DP}}{C} \cdot \left(V_f - \frac{7{,}2 \cdot d_{DP}}{C} \right) \cdot \sum t_i}{K_j \cdot \frac{7{,}2 \cdot d_{DP}}{C} \cdot n_C \cdot \left(V_f - \frac{7{,}2 \cdot d_{DP}}{C} \right) - c_C \cdot V_f} \tag{5}$$

This therefore represents the final formula of the extended coordination model designed for improved smoothness of traffic in coordination of LSE controlled intersections. In the formula, the required length of the cycle C is on the left and also on the right side. After adjustments, we get a polynomial of the third degree. The result therefore includes three roots of possible lengths of the cycle dependent on the required intensity of the coordinated entry, the speed of approach and the distance between coordinated intersections. Usually only one of the roots fulfils the condition of usability in a practical proposal – i.e. the minimal or maximal length of the cycle applicable for real traffic control at a traffic light controlled intersection. By the proposal of the extended coordination of intersections, one works always with the most loaded intersection from the chosen line of intersections. Other intersections must "adapt" their capacity through the same length of the cycle and therefore the waiting time at their side entries might increase. Certainly, this problem concerns generally any proposal of the coordination and it is the necessary tax for passage smoothness in the preferred direction.

4 The Model Verification by Simulation

To verify the functionality of the extended coordination proposal, a micro-simulation in PTV VISSIM software has been applied. The purpose of the simulation is to find the travel times for a simple model construction of 2 traffic lights controlled intersections (No 1 and No 2 – see Fig. 1.). For this simple traffic network, a proposal of the coordination has been elaborated according to TC 81 and according to the newly proposed model of extended coordination and the mutual comparison of both models is carried out.

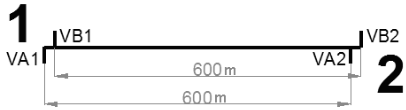


Fig. 1. Scheme of model formation of intersections for experiment purposes

The formation of two model traffic lights controlled intersections contains typical parameters from the real traffic. The communication between the intersections is designed as bidirectional, two lanes with each lane 3.5 m wide without a possibility of overtaking. The mutual distance between the intersections is determined as 600 m. The intensity of the coordinated entry is 500 vehicles per hour for one direction with the coefficient of utilization of reduced cycle duration by the coordinated direction at the amount of 0.7 identical for both intersections. The sum of the decisive inter-green times is 30 s equal for both model intersections. The calculation based on these input parameters (3) gives the minimal length of the cycle 70.84 s, rounded to integer 70 s. The calculation based on the mentioned input parameters then comes to duration of the green light in the coordinated direction for 20 s. All simulated cases of the proposal of coordination are depicted in the distance-time diagrams in Fig. 2.

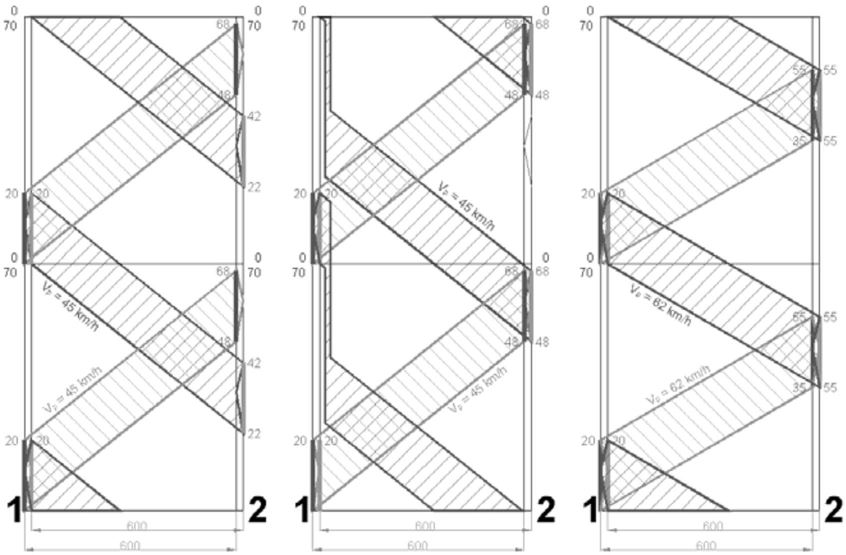


Fig. 2. Distance – time diagrams for the proposals of the coordination according to TC 81 (left and middle) and for the proposal according to the model of the extended coordination (right)

In intersection No 1, we will complete the signal plan so that the passage of both coordinated entries was at the same time of the cycle. The coordination proposal according to TC 81 for intersection No 2 will not result in simultaneous passage of coordinated flows (Fig. 2 - left). At the same time, this results in the capacity insufficiency of side entries into intersection No 2, because the coefficient of utilization of reduced cycle duration by the coordinated direction in the required value is not preserved. The second possibility of the proposal of the coordination according to TC 81 offers a simultaneous passage of coordinated lines in both intersections, however with preserved speed of approach only for the passage in one direction (Fig. 2 – middle). Only the coordination proposal of the fictive intersections according to the model of extended coordination enables a preserved required capacity of all entries and at the same time a passage in both directions at the speed of approach (Fig. 2 - right).

For the simulation of the coordination proposal according to TC 81, the priority has been chosen to preserve the capacity of all entries (Fig. 2 – middle) and this case is simulated below in the Vissim software.

4.1 Output Values of Micro-simulation

The functionality "*Travel Time Sections*" in the Vissim software records travel times including delays, it means the total travel time for all vehicles in the specific section. The resulting average travel times of vehicles between the intersections obtained from the calculation and simulation for both proposals of the coordination designs are generally recorded in Table 1 below.

Table 1. Comparative table of travel times of the particular directions in a fictive coordination obtained from the calculation and simulation

	Travel time	**Direction 1→2**	**Direction 2→1**
Coordination according TC 81	Travel time set by calculation (s)	48.00	92.00
	Average travel time set by simulation (s)	55.33	92.61
Model of extended coordination	Travel time set by calculation (s)	35.00	35.00
	Average travel time set by simulation (s)	40.83	40.01
Percentage difference of travel times set by simulation (in favour of extended coordination)		-26.2%	-56.8%

The unambiguous conclusion from results of travel times received from the calculation and namely from more realistic values from the simulation is a shorter travel time through a formation of intersections when the model of extended coordination is proposed compared to the proposal of the coordination according to TC 81.

5 Conclusion

The preservation of coordination in both traffic directions of the roads is the most required proposal of traffic lights controlled intersections coordination. The contemporary method for the design coordination of intersections (according to the Technical Conditions 81) for the constant speed of approach of approx. 45 $km \cdot h^{-1}$ does not enable a change of the coordination speed upon a more significant change of the required intersection capacity (required length of the cycle of the signal plan) caused for example by a changed traffic intensity during a day, while the speed of the traffic flow changes strongly with the given road saturation. If the recommendation is the proposed coordination speed of approx. 85% of the maximum speed limit (in towns 50 $km \cdot h^{-1}$) and simultaneously it has been proven in practice that the speed of approach below 40 $km \cdot h^{-1}$ is not maintained, it means an obvious conflict and these two limits give us a very narrow range of usable speeds of approach. The model of extended coordination with utilization of variable traffic signs eliminates this failure.

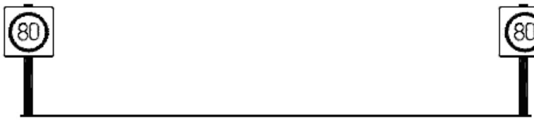


Fig. 3. Minimal configuration of variable traffic signs for application of the extended coordination model

In the case of a practical proposal, a model of extended coordination for a real area of intersections, it is necessary to keep a set of certain technical conditions and mainly the requirements on the traffic safety. For example, variable traffic signs regulating the speed limit must be behind each intersection, their minimum configuration is shown in Fig. 3. In each coordinated direction, there must be more than one lane enabling to overtake slower vehicles. A valid ban on stopping must be effective on the road, because parking vehicles worsen the functionality of a green wave and increase the risk of collision from behind. Moreover, not controlled pedestrian crossings are not permissible and the speed limit must be clearly presented when the variable traffic signs are off. Controllers must enable setup changes for yellow signal duration.

The model has not been yet deployed in practice, however, based on several micro-simulations on real locations and their results, it is possible to consider that the extended coordination model is a model improving traffic smoothness in line coordination of lights controlled intersections.

References

1. Technické podmínky TP81 – Navrhování světelných signalizačních zařízení pro řízení provozu na pozemních komunikacích. II. Vydání. Zpracovalo Centrum dopravního výzkumu Brno. Praha: Ministerstvo dopravy, květen, 124 s (2006) ISBN 80-86502-30-9

2. Zákon o provozu na pozemních komunikacích (online). Ostrava: Nakladatelství Sagit, a.s. Sbírka zákonů ročník 2000, částka 98, ze dne, October 19, 2000, účinnost dnem 1. ledna 2001, Dostupné z, http://www.sagit.cz
3. Gerlough, L.D., Huber, J.M.: Traffic Flow Theory: A Monograph TRB Special Report 165. Transportation Research Board, Washington DC (1975)
4. Roess, R.P., Prassas, E.S., McShane, W.R.: Traffic Engineering. Third Edition. Pearson Education, Inc., Upper Saddle River (2003) ISBN 0-13-191877-X

How the Ionosphere Affects Positioning Solution Using Terrestrial and Satellite Navigation Systems?

Jacek Januszewski

Gdynia Maritime University, Faculty of Navigation
al. Jana Pawla II 3, 81–345 Gdynia, Poland
jacekjot@am.gdynia.pl

Abstract. The Earth's ionosphere, a dispersive medium, is an atmospheric layer that lies typically between 50 and 1,000 km of altitude. The physical parameters of ionosphere have a direct influence on propagation delay (Satellite Navigation Systems – SNS) and on a radio waves propagation (Terrestrial Radionavigation Systems – TRNS), that both cause significant user's position error and additionally in the case of TRNS the lack of the position sometimes. The most frequently used solutions to this problem, as the dual–frequency receivers which permit the calculation of ionosphere-free pseudorange, model of the ionosphere with eight coefficients transmitted in navigation message or differential mode with pseudorange corrections (SNS) and different corrections to the measurements as SWC and ASF in Loran C system (TRNS) are described in this paper. The detailed relations and equations concerning ionospheric delay are presented also.

Keywords: ionosphere, ionospheric refraction, ionospheric delays, ionospheric scintillations, satellite navigation systems.

1 Introduction

Telematics can be defined as any integrated use of telecommunications and informatics. Hence the application of telematics is with among other things any of the following:

- the technology of sending, receiving and storing information via telecommunication devices in conjunction with affecting control on remote objects.
- telematics includes but is not limited to satellite navigation systems (SNS) as GPS or GLONASS technology integrated with computers and mobile communications technology in automotive navigation system.

Telecommunication is the transmission of information over significant distances to communicate. Nowadays telecommunications include the use of the orbiting satellites. That's why telematics involves the impact of the ionosphere on the terrestrial and satellite signals propagation and the determination of the user's position.

J. Mikulski (Ed.): TST 2012, CCIS 329, pp. 249–257, 2012.

2 Ionosphere

The structure of the Earth's atmosphere can be described, for most practical purposes, as asset of concentric spherical shells with different physical and chemical properties. Various subdivisions are possible, but with respect to signal propagation a subdivision into ionosphere and troposphere is advisable. The ionosphere, is the region of Earth's atmosphere of ionised gases (free electrons and ions). The upper limit of the inospheric region is not clearly defined [1 - 5]. The ionisation is caused by the Sun's radiation, and the state of the ionosphere is determined primarily by the intensity of the solar activity [4]. The height of the lower and upper part of Earth's ionosphere can be different, e.g. 75 km and 500 km [6] or 70 km and 1000 km [1, 7], respectively.

Ionospheric region is divided into subregions, or layers (D, E, F1 and F2), according to the electron density. The heights of these layers and the electron density of ionized particles are not constant, both change with the time of the year, seasons, and time of day. Generally, the region between 250 and 400 kilometers above the Earth's surface, known as the F–region, contains the greatest concentration of free electrons. This region affects the satellite signals propagation considerably.

3 Effects of Ionosphere on Terrestrial Radionavigation Systems

The signals transmitted by the stations of terrestrial radionavigation systems (TRNS) can arrive at user's receiver as groundwave or skywave or both groundwave and skywave. It depends on the frequency carrier, radio range, daytime, propagation conditions etc. The user's position with the lowest error can be obtained if the signals arrive directly from all used stations as groundwave without any refraction. If at least from one station signals arrive as skywave it means that:

- in TRNS based on phase or difference phases measurements the user's position cannot be determined. That's why at nighttime both the range of the station and total coverage area of every chain of this system are less than during daytime considerably,
- in TRNS based on time or difference time measurements the user's position can be obtained only if the special corrections were used, e.g. SWC (Sky Wave Correction) in Loran C system. The station range on skywave is greater than on groundwave, e.g. in Loran C it is 2300 Nm or more and 1200 Nm respectively, but position's accuracy is in the first case less considerably.

As the refraction index of the ionosphere depends on the varying electron and ion concentrations, although distinct boundaries do not exist, reflection and refraction take place in a more or less layered medium. Thus signals at a low frequency are reflected at low altitudes, while high-frequency signals are only refracted, but they may develop an increased angle of incidence so that they are reflected at higher layer [8]. Frequencies below about 500 kHz (among them once Decca Navigator system and since 1977 Loran C system) are reflected in the D layer, 0.5 – 2 MHz in the E layer and 5 – 30 MHz in the F layer, whereas frequencies above 30 MHz are not reflected in the ionosphere.

4 Effects of Ionosphere on Satellite Navigation Systems

As Medium Earth Orbits (MEO) with height about 20000 km are used for constellations of navigation satellites such as GPS, GLONASS, Galileo, Compass, the signal transmitted from satellite passes through the ionosphere.

4.1 Ionospheric Refraction

As ionised gas is a dispersive medium for radio waves the ionosphere affects the propagation of satellite signals. The speed of these signals depends upon the number of free electrons in the path of a signal, defined as the TEC (Total Electron Content), i.e. the number of electrons in a tube of 1 m^2 cross section extending from the user's receiver to the satellite. TEC is measured in units of TEC Units – TECU, defined as 10^{16} electrons/m^2. As the path length through the ionosphere is shortest in the zenith direction, TEC in the vertical direction (TECV) is the lowest. TECV typically varies between 1 and 150 TECU [4] or between 1 and 1000 [1]. The state of the ionosphere is described by the electron density N_e with the unit [number of electrons/m^3] or [number of electrons/cm^3]. For a given place and time, vertical TEC can vary 20–25% from its monthly average [1].

The main contribution of the inospheric refraction corresponds to the first order ionospheric term, which is inversely proportional to the square of the frequency. That's why the refraction coefficient describing the propagation of phase n_p and coefficient of the group delay n_g can be expressed by the approximate relations [7]:

$$n_p = 1 - \frac{40.3N_e}{f^2} \qquad n_g = 1 + \frac{40.3N_e}{f^2} \tag{1}$$

If N_e is the number of electrons per cubic meter, numerator 40.3 N_e is expressed in Hz^2, f is frequency carrier. In ionosphere, a dispersive medium, the carrier and the modulation travel at different speeds, respectively with phase velocity v_p and group velocity v_g. These velocities can be defined as follows:

$$v_p = \frac{c}{1-\frac{40.3N_e}{f^2}} \qquad v_g = \frac{c}{1+\frac{40.3N_e}{f^2}} \tag{2}$$

where c is the speed of light in vacuum.

As a consequence of the different velocities, a phase advance and group delay occur. Therefore SNS carrier phases are advanced and ranging codes are delayed. That's why the measured carrier phase pseudoranges are too short and the measured code pseudoranges are too long compared to the geometric range between the satellite and user's receiver. The amount of the difference is the same in both cases.

The ionospheric group delay for different frequencies carrier and the number TEC if the satellite elevation angle is 90^O is presented in the table 1, for different angles and different heights of the ionosphere with maximal electron density if frequency carrier is equal 1.6 GHz and 1.2 GHz in the table 2. The difference Δ_{iono} between measured and geometric range called ionospheric refraction for a phase Δ_{ionop} and group delay Δ_{ionog} can be written as

$$\Delta_{ionop} = -\frac{40.3TEC}{\sqrt{1-\frac{R_z\cos(h_t)}{R_z+h_{j\max}}}}\cdot\frac{1}{f^2} \qquad \Delta_{ionog} = \frac{40.3TEC}{\sqrt{1-\frac{R_z\cos(h_t)}{R_z+h_{j\max}}}}\cdot\frac{1}{f^2} \tag{3}$$

where R_z is radius of the Earth, h_t satellite elevation, h_{jmax} the height of the ionosphere with maximal electron density.

Table 1. Ionospheric group delay [m] for different frequencies and TEC, elevation angle 90O

Frequency carrier	TEC [el/m^2] 10^{16}	10^{18}	$5 \cdot 10^{18}$	10^{19}
150 MHz	17.9	1 790	8 950	17 900
400 MHz	2.52	252	1 260	2 520
1.2 GHz	0.28	28	140	280
1.6 GHz	0.16	15.8	79	158
10 GHz	0.004	0.4	2	4

Table 2. Ionospheric group delay [m], for different satellite elevation ht, for different frequencies carrier L and different heights of the ionosphere with maximal electron density hmax, TEC = 1018

L [GHz]	h_{max} [km]	Satellite elevation h_t [^{O}] 0	5	10	15	20	25	30	45	60	75	90
1.6	250	81.0	77.3	68.8	59.3	50.9	44.0	38.6	27.8	21.9	18.2	15.8
	300	74.2	71.4	64.5	56.5	49.2	42.9	37.9	27.6	21.8	18.2	
	350	69.0	66.7	61.1	54.2	47.6	41.9	37.2	27.4	21.7	18.1	
	400	64.8	62.9	58.1	52.1	46.3	41.0	36.6	27.2	21.6	18.1	
1.2	250	144.0	137.4	122.3	105.4	90.5	78.2	68.6	49.4	38.9	32.4	28
	300	131.9	126.9	114.7	100.5	87.5	76.3	67.4	49.1	38.8	32.4	
	350	122.7	118.6	108.6	96.4	84.6	74.5	66.1	48.7	38.6	32.3	
	400	115.2	111.9	103.4	92.7	82.3	72.9	65.1	48.4	38.5	32.3	

4.2 Corrections of Ionospheric Delays

As at given moment the values of TEC and h_{jmax} are for almost all users, except for the scientific centres, known, but with too low accuracy or, in major cases, completely unknown, four different methods of the corrections of ionospheric delays were defined. These methods can be used in particular by the users determining own positions in real time.

The most frequently used solutions to this problem are: differential mode with pseudorange correction, satellite based augmentation system with geostationary satellites, the dual–frequency receivers which permit the calculation of ionosphere-free pseudorange, model of the ionosphere with eight coefficients transmitted in navigation message used in single-frequency receivers.

Differential Mode. The fundamental principle of differential mode of SNS is the comparison of the accurately measured position of a fixed point, referred to as the terrestrial reference station, with position obtained from a stand-alone SNS receiver (without any corrections) at that point. By comparison of measured and calculated results at this position, corrections, called Pseudo Range Corrections (PRC), are produce for distribution to all users within the coverage area. This technique can be used and works well only if an essential portion of the measurement errors is caused by factor outside the receiver which at the same time are correlated in the area. The four most important of these errors are the following:

- ionospherical delays of the signals,
- tropospheric time delays,
- ephemeris errors,
- satellite clock errors.

As the ionospheric group delay, the tropospheric delay and both last errors are always positive, PRC is always negative, and its value depends on satellite elevation and daytime among other things. For signal arriving at vertical incidence (90^{O}) PRC ranges from about 3 m at night to as much as 15 m during the day. At low satellite viewing angles (less than 10 degrees) PRC can range from 9 m at night up to 150 m during the day [7]. Actual values of PRC for all satellites used for the positioning are signalised by receiver.

The number of beacons (stations) transmitting DGPS corrections has been increased in 10 last years considerably. In 2002 there were in the world 162 stations with status operational, 62 with status on trial and 20 planned, in 2006 these numbers were 235, 57 and 11 respectively. At present (2012) 281 operational stations are localized in 37 countries, the greatest numbers of these stations are in USA (38), Japan (27), China (21) and India (19). In Poland there are two stations, Dziwnow and Rozewie. Since few years there are two operational stations, both in Ukraine (Yenikal'skiy Lt and Zmeiniy Lt), transmitting DGLONASS corrections (in RTCM protocol 6 message types, numbers 31–36) also. At present in Russia 4 DGPS/DGLONASS stations (1 Baltic Sea, 3 Black Sea) have the status on trial, next 12 (4 Arctic Coast, 8 Pacific Coast) are planned [9].

The user must be in coverage of least one reference station and must be equipped with DGPS receiver; the signals arrive from GPS satellites and from PRC station.

The Satellite Based Augmentation Systems (SBAS). The Satellite Based Augmentation Systems (SBAS) use a network of terrestrial monitoring stations to perform SNS ranging measurements. SBAS provides integrity data and correction parameters for the satellite orbits, the satellite clocks, and the ionosphere influence using geostationary satellites (GEO) as the communication path. SBAS message differentiates between 64 types, type 18 provides ionosphere grid point masks, type 26 the L1–only user with vertical ionospheric delay values over a grid of locations with predefined latitude and longitude values [7]. The words L1–only user mean that this ionospheric corrections only apply for single-frequency SNS receivers, at the time of this writing (2012), GPS receivers only.

The SBAS models the vertical delays of the ionosphere at the inospheric grid points (IGP), which commonly span a regular raster of 5^{O} x 5^{O}. The user's receiver estimates the ionospheric delays for every satellite in four step process.

Actually some SBAS are fully operational (WAAS within the United States and Canada, EGNOS within Europe, MSAS within Japan and Southeast Asia), some are under construction (GAGAN within India, SDCM within Russia).

The user must be in coverage of least one SBAS GEO satellite and must be equipped with GPS/SBAS receiver; the signal from GPS satellites and correction parameters from GEO satellite.

Single-Frequency Receiver Model of the Ionosphere. In most commercial SNS receivers only one frequency (L1 in the case GPS system) is available. That's why the pseurorange measurements on two frequencies cannot be made. Consequently, models of the ionosphere are employed to correct for the ionopsheric delay. One important example is the Klobuchar model, which assumes that the vertical ionospheric delay can be approximate by half a cosine function of the local time during daytime and by a constant level during nightime. In the case of GPS system the ionospheric data collected from subframe 4 of navigation message NAV can be used to reduce the ionospheric effect. In page 18 of this subframe there are eight ionospheric data: α_0 (69 – 76), α_1 (77 – 84), α_2 (91 – 98), α_3 (99 – 106), β_0 (107 – 114), β_1 (121 – 128), β_2 (129 – 136), β_3 (137 – 144). These eight coefficients are computed from a global, empirical model and updated every 10 days (every 5 days if the Sun is particularly active). In order to use this model the user's single-frequency receiver has to compute the obliqueness of the signal from the satellite, i.e. the position where the signal path intersects the average ionosphere altitude. This altitude has been set to 350 km. In addition, the receiver has to compute his own geomagnetic latitude (as the ionosphere is a function of the geomagnetic latitude). Using this model one can reduce the user root mean square (rms) position error caused by ionosphere effect at least by 50% percent [7, 10] or 55 – 60% [8].

Dual-Frequency Receiver. The difference between the refraction index (i.e. the phase velocity) in a vacuum and in the ionosphere is consequently inversely proportional to the square of the frequency. That's why the satellites of all SNS transmit the signals on at least two frequencies, which makes it possible to determine the constant of proportionality and, consequently, remove the influence of the ionosphere to a large extent. Differencing pseudorange measurements ρ_1 and ρ_2 made on f_1 and f_2 frequencies, respectively, enables the estimation of both the N_1 and N_2 delays, respectively, neglecting multipath and receiver noise errors. Delay N_1 becomes

$$N_1 = (\rho_2 - \rho_1) \cdot \frac{f_2^2}{f_1^2 - f_2^2} \tag{3}$$

therefore the estimate of ionosphere-free pseudorange measurement at f_1 ρ_{1F} can be presented as

$$\rho_{1F} = \frac{f_1^2}{f_1^2 - f_2^2} \cdot \rho_1 - \frac{f_2^2}{f_1^2 - f_2^2} \cdot \rho_2 = a \cdot \rho_1 - b \cdot \rho_2 \tag{4}$$

where a and b are coefficients which depend on carrier frequencies f_1 and f_2 only. The values of these coefficients for all possible combinations of two frequencies of GPS and Galileo systems are presented in the table 3.

Table 3. Dual frequency measurements, estimation of ionosphere-free pseudorange measurements, coefficient values for GPS system and Galileo system

System	Frequency carrier [MHz]		Coefficient	
	f_1	f_2	a	b
GPS, Galileo	1 575.42	1 176.45	2.261	1.261
GPS	1 575.42	1 227.60	2.546	1.546
	1 227.60	1 176.45	12.26	11.26
Galileo	1 575.42	1 278.75	2.931	1.931
	1 575.42	1 207.14	2.422	1.422
	1 278.75	1 207.14	9.186	8.186
	1 278.75	1 176.45	6.527	5.527
	1 207.14	1 176.45	19.920	18.920

The ionospheric group delay is in direct proportion to TEC along the propagation path of a satellite signal and in inverse proportion to frequency carrier. That's why the stand-alone SNS receiver one frequency uses always the highest frequency of given SNS, e.g. 1575.42 MHz in the case of GPS system and Galileo system.

Ionospheric Scintillations. At times, the mentioned above F–region of the ionosphere, becomes disturbed, and small-scale irregularities develop. When sufficiently intense, these irregularities scatter radio waves and generate rapid fluctuations, called scintillations, in the amplitude and phase of radio signals. Scintillation activity varies with operating frequency, geographic location, latitude in particular, local time, season, magnetic activity, and 11–year solar cycle. Scintillation may accompany ionospheric behaviour that causes changes in the measured range between the user's receiver and the satellite. Its activity is most severe and frequent in around the equatorial regions, particularly in the hours just after sunset. In high latitude regions, scintillation is frequent but less severe in magnitude than that of the low latitude. In the mid-latitude regions (Europe) scintillation is rarely experienced [7, 11].

5 Conclusion

- the increasing use of satellite navigation systems for precise position measurements has provoked an increased interest in ionosphere behaviour, particularly propagation properties as functions of time
- the atmosphere around the Earth, ionosphere in particular, affects the travelling speed of the SNS signals and causes measurement error
- to eliminate ionospheric error single frequency SNS receivers apply SBAS ionospheric corrections, while dual frequency SNS receivers use the measurements on two transmitted satellite frequencies
- for precise local differential positioning PRC can be used if the distance between the user and reference station is less than 10 km, if this distance is greater dual frequency receiver is required
- ionospheric refraction affects the stability of satellite clocks using satellite navigation system time transfer
- ionospheric scintillation, the rapid fluctuations in the amplitude and phase of radio signals caused by small-scale irregularities in the ionosphere, is one of the most potentially significant threats for all SNS and SBAS. This phenomenon can lead to a receiver being unable to track one or more visible satellites for short periods of time

References

1. Seeber, G.: Satellite Geodesy. Walter de Gruyter, Berlin-New York (2003)
2. El-Rabbany, A.: Introduction to GPS the Global Positioning System. Artech House, Boston/London (2002)
3. Groves, P.D.: Principles of GNSS, Inertial, and multisensor integrated navigation systems. Artech House, Boston/London (2008)
4. Misra, P., Enge, P.: Global Positioning System Signals, Measurements, and Performance. Ganga-Jamuna Press, Lincoln (2006)
5. Samana, N.: Global Positioning Technologies and Performance. John Wiley & Sons, New Jersey (2008)

6. Prasad, R., Ruggieri, M.: Applied Satellite Navigation Using GPS, Galileo, and Augmentation Systems. Artech House, Boston/London (2005)
7. Kaplan, E.D., Hegarty, C.J.: Understanding GPS Principles and Applications. Artech House, Boston/London (2006)
8. Forssell, B.: Radionavigation systems. Artech House, Boston/London (2008)
9. Admiralty List of Radio Signals, The United Kingdom Hydrographic Office, vol. 2, NP 282 (2011/2012)
10. Bao-Yen Tsui, J.: Fundamentals of Global Positioning System Receivers. Wiley Interscience, Hoboken (2005)
11. Langley, R.: Ionospheric Scintillations How Irregularities in Electron Density Perturb Satellite Navigation Systems. GPS World 23(4), 44–90 (2012)

The Benefits of Cloud Computing in the Maritime Transport

Jolanta Joszczuk–Januszewska

Gdynia Maritime University, Faculty of Navigation
al. Jana Pawla II 3, 81345 Gdynia, Poland
jolajj@am.gdynia.pl

Abstract. Maritime transport is the prominent mode of trade between European Union (EU) and non-EU member countries. As a result, EU ports and government administrations are under constant pressure to increase the efficiency and quality of their services, recognising the critical role of Information and Communication Technologies (ICT). Cloud computing (CC) as significant advance in ICT development is expected to dominate the maritime transport development landscape in the coming years. The advent of CC where applications are served with data that is stored on the Internet "in the clouds" and can be accessed and shared by the parties involved in the supply chain operation has evoked various degree of interest. It can be explained some of the benefits and challenges of using cloud services in maritime transport which are presented in this paper, particularly the first benefits of some Maritime Transport Single Window implementation initiatives already making use of cloud technology. Also are described the benefits of using CC in maritime transport concerning its technical fleet management operations.

Keywords: Cloud Computing, Maritime Transport, Information and Communication Technologies.

1 Introduction

Maritime transport is the prominent mode of trade between European Union (EU) and non-EU member countries. As a result, EU ports and government administrations are under constant pressure to increase the efficiency and quality of their services.

The EU's member states represent a diverse and fragmented network of trade stakeholders, ranging from ship owners, shipping agents, and freight forwarders, to port authorities and terminal operators, further complicated by required adherence to both national and EU-level trade policies and regulations. Companies and government authorities have traditionally faced challenges in providing the technology needed to ensure interoperability, standardization and availability of information between trade stakeholders.

In the short to medium term, the most promising development for maritime transport is e-Maritime, which is becoming the focus for the simplification and cohesion of administrative requirements and procedures, with a spill-over into commercial

J. Mikulski (Ed.): TST 2012, CCIS 329, pp. 258–266, 2012.

applications [1, 2]. The EU e-Maritime initiative, is seen as a cornerstone for the achievement of the strategic goals of the EU Maritime Transport Strategy 2018, recognising the critical role of Information and Communication Technologies (ICT) for productivity and innovation.

Cloud computing (CC) as significant advance in ICT development is expected to dominate the maritime transport development landscape in the coming years.

The advent of CC where applications are served with data that is stored on the Internet “in the clouds” and can be accessed and shared by the parties involved in the supply chain operation has evoked various degree of interest. It can be explained some of the benefits and challenges of using cloud services in maritime transport which are presented in this paper, particularly the first benefits of some Maritime Transport Single Window implementation initiatives already making use of cloud technology. An example is SOGET operator which announced development of its e-Maritime Port Single Window project on Microsoft Cloud Technology. Also are described the benefits of using cloud computing in maritime transport concerning its technical fleet management operations, on the example MESPAS software.

2 The Term “Cloud Computing”

The term “cloud computing” refers to several different computing paradigms, not all of which are completely new [3]. For example, as the U. S. National Institute of Standards and Technology (NIST) has explained, cloud computing has three service models [4]:

- **Software as a Service** (SaaS) through which applications are provided in the cloud;
- **Platform as a Service** (PaaS), through which a cloud provider permits users to create or run applications using languages and tools supported by the provider while the provider delivers the underlying infrastructure such as servers, operating systems, or storage; and,
- **Infrastructure as a Service** (IaaS), through which a customer can deploy a computing infrastructure similar to a virtualized environment.

The essential characteristics of all three models include self-service (a customer can access new capabilities), shared resources, and rapid elasticity (i.e., as a business grows, it can rapidly add additional processing power and storage).

Additionally, cloud computing provides IT resources, as a service, in a dynamic and scalable manner over a network. It has five essential characteristics:

- on-demand self-service;
- broad network access;
- resource pooling;
- rapid elasticity;
- measured service.

The following list identifies the two main deployment models:

- **Private cloud**. The cloud infrastructure is operated within a single organization. In this case, internal groups such as business units consume resources and services provided by a single internal (i.e. the IT department) or external cloud computing provider;
- **Public cloud**. The cloud infrastructure and services are available to the general public or a large industry group and is owned by an organization selling cloud services.

3 Cloud Computing in Maritime Transport Single Window Initiatives

3.1 The Initial Idea behind the Single Window Concept

Global trade expanded rapidly during the 1980s and 1990s. The resulting complexity and speed of the modern supply chain and the number of parties involved greatly increased the requirements for information controlling the flow of goods. But despite the breakneck developments in information and communications technologies (ICT) and trade data-exchange standards during the same time, trade documentation exchanges remained mostly paper-based. However, in the modern trade environment such paper-based exchanges cannot satisfy the need for efficiency and security.

One “omnibus” means of addressing this problem that has gained considerable momentum over the past 10 years is the so-called “ Single Window” [5].

3.2 An Introduction to Single Window

In 2005 The Recommendation and Guidelines on establishing a Single Window, United Nations (UN) Economic Commission for Europe United Nations Centre for Trade Facilitation and Electronic Business (UN/CEFACT) Recommendation 33, was published. This recommendation defines an Single Window as a facility that allows parties involved in trade and transport to lodge standardized information and documents with a single entry point to fulfil all import, export, and transit-related regulatory requirements [6].

If information is electronic, then individual data elements should only be submitted once. In practical terms, the single window aims to expedite and simplify information flows between trade and government and bring meaningful gains to all parties involved in cross-border trade. The single window is generally managed centrally by a lead agency, enabling the appropriate governmental authorities and agencies to receive or have access to the information relevant for their purpose. In addition, participating authorities and agencies should co-ordinate their controls. In some cases, the single window may provide facilities for payment of relevant duties, taxes and fees.

A single window does not necessarily imply the implementation and use of ICT, although facilitation can often be greatly enhanced if governments identify and adopt relevant ICT for a single window.

This paper addresses the enhanced single window through the use of some trend in the use of ICT for Single Windows in maritime transport. Cloud computing as significant advance in ICT development is expected to dominate the Single Window development landscape in the coming years.

3.3 Examples of Single Widow

Some Single Window initiatives already make use of cloud technology in maritime transport. Microsoft applies cloud computing principles to single window.

Trans-Kalahari Corridor Regional Single Window. On May 11, 2011, the World Customs Organization held its annual IT Conference and Exhibition, featuring the theme of cloud computing – a technology expected to revolutionize trade and customs operations. At the conference, representatives from United States Agency for International Development (USAID) and the government of Namibia unveiled plans to create the Trans-Kalahari Corridor Regional Single Window (RSW). A partnership between USAID, Microsoft, and the Customs Agencies of Namibia and Botswana, the RSW will be a cloud computing-based trade application – built on Microsoft technology – that will link customs processes between the governments of Namibia, Botswana, and South Africa [7].

A single window provides a platform creating a single entrance point for all data and documents necessary to import or export goods – with the goal of creating a more efficient and secure customs process. The Trans-Kalahari Corridor RSW will automate customs connectivity, reduce trade costs, and increase compliance between the customs authorities of all three countries. By enabling faster and more secure trade, this initiative aims to facilitate economic growth and development in the region.

With the RSW, the three countries will be able to leverage their existing customs platforms, while making their data available in the cloud for the first time. The RSW also allows for the integration of additional border agencies, enabling future development of an Integrated Border Management solution in the future.

Cloud computing is where the future of customs technology lies because it transcends national boundaries, creates greater IT efficiencies and helps trade stakeholders such as importers, exporters, border agencies access distant markets. It offers scalability, flexibility and unprecedented interoperability which will help trade grow as never before. Microsoft is excited to work with USAID and its partners to develop the Trans-Kalahari Corridor RSW.

SOGET's e-Maritime Port Single Window Project. An example is one of different forms of Single Window - Port Single Windows and Port Community Systems. Other authorities with a substantial role in trade—such as the Port Authority—have established a limited, port-centric, "Single Window", commonly referred to as either a "Port Single Window" or a "Port Community System" [8].

Headquartered in Le Havre, France, since 1983, SOGET is the leading Port Community System (PCS) operator in France and a public-private partnership between the Port Community of Le Havre, Port of Le Havre Authority and French

Customs. SOGET PCS is in operation throughout French ports as well as overseas. SOGET also partners with Bureau Veritas, a world leader in conformity assessment and certification services, to offer a concession service to implement and operate Port Community Systems.

This project, originally launched in June 2010, as the next generation of Port single window solutions, is supported by European Regional Development Fund (ERDF) and is addressing the European Union's "Maritime Transport Strategy 2018".

The e-Maritime Port Single Window provides a single sign-on allowing all major maritime, port, logistics, and government trade stakeholders to access real-time data application modules. This solution also aims to improve the efficiency of business processes, by standardizing exchanges and providing interoperability with European and international port and government single window systems.

Based on Microsoft technology, SOGET is providing port communities with the building blocks of e-maritime transactions and processes. The solution leverages several Microsoft technologies including BizTalk 2010, SQL server 2011 and Hyper V 2008, creating an open, integrated and innovative solution specifically tailored for the e-maritime industry.

The e-Maritime Port Single Window improves the efficiency of business processes, by standardizing exchanges and providing interoperability with European and international port and government single window systems.

Using the Port Community System, it takes 6 minutes and 19 seconds on average to gain customs clearance of goods.

SOGET announced development of its e-Maritime platform on Microsoft cloud technology in May 2011 and is currently under development.

Bureau International Maritime. Belgian-based, Microsoft partner, Bureau International Maritime (BIM) creates tailor-made solutions for the public and private sector using new technologies in the areas of training, certification, security, cargo tracking and single window [9].

With several training centres located in Belgium, France and Congo, and an international network of partners cross Europe, Asia, America and Africa, BIM benefits from localized expertise.

In 2005, BIM began research on the Guichet Unique Maritime (GUMAR) Single Window application, and in 2009 the development of the Project began. Based on Microsoft technology, including Microsoft SharePoint Server and Microsoft BizTalk Server, the single window solution significantly improves efficiency by allowing transactions to become paperless and simultaneous.

Every participant in the trading process, from buyer to receiver to forwarder to maritime actor, is able to access the GUMAR Single Window application and work in parallel, speeding up the process enormously. The solution provides a single entry point for standardized information and documents, which all parties involved in trade and transport can access and therefore fulfil all import, export, and transit related regulatory requirements.

Users can access the GUMAR Single Window application in a public single window office, on their own PC or on their company's computer running the application securely through the interface system. For convenience, users are notified by SMS when their files have been sent and received.

Average time (5–7 days) required for one container to gain clearance in Congo with GUMAR, reduced from 15–17 days before.

4 Benefits of Mespas R5 Cloud Computing in Terms of Fleet Management in Maritime Transport

The benefits of using cloud computing in maritime transport concerning its technical fleet management operations can be described on the example MESPAS software.

Today, MESPAS (located in Switzerland) is the world's leading fleet management software provider based on cloud computing and the fastest growing software as a service (SaaS) company in the maritime transport.

Mespas R5 is a cloud-based software system and allows you to access and operate data, applications and IT infrastructure as a service via the Internet, the so-called 'cloud'. It's the most efficient, scalable and innovative way of running technical fleet management software – and presents a low-cost, flexible alternative to the have your own software approach.

The following list identifies the essential benefits [10]:

- **Keep Your Costs in Check.** Using the mespas R5 software means you don't need to run your own server infrastructure for your technical fleet management system. This reduces your capital and operational expenses. Since the software is run on a subscription basis (software as a service), your expenses become transparent and predictable. The costs for running the central hardware is spread across all users of the system, making the mespas R5 system far more cost-efficient than maintaining your own server infrastructure.
- **Improve Efficiency.** All relevant people have access to the same up-to-date information (technical information, work instructions, manuals, supplier infos, etc.), and the software is easy to use. This improves efficiency by leaps.
- **Reduce Time and Costs Spent on IT.** The software is maintained and updated centrally at MESPAS. You don't need to worry about central hardware (server), security, backups, and IT personnel for the technical management software.
- **Speed of Deployment.** Adding or removing vessels is a matter of days or weeks, and not months. Save time while scaling up, and save money by scaling down when additional resources are no longer needed. Your company becomes more agile and flexible.
- **Central Data Storage, Accessible by all Relevant People.** Master data (OEM data, manuals, spare parts, etc.) is provided by MESPAS on the central server. Your own business data is kept securely on the central server. It's accessible by all relevant stakeholders via the Internet. You own your data, but you don't need to own the central server. All you need is standard IT equipment, an Internet connection (office) or ship/shore connection (vessel), as well as the Mespas

Cube. The Cube was developed to enable the secure operation of the software aboard the vessel, and to enable multi-user functionality for the crews.

- **Generate Your Own Key Performance Indicator (KPI) Reports or Choose from Standard Reports.** Since data is stored in one secure place, it becomes comparable: across your products, vessels, or across the entire company. With real-time data at your fingertips you are able to monitor KPIs across the fleet, and thus make informed decisions.

Benefits of mespas cloud solution involved from comparison to other solutions, according to concept and costs, are given in the Table 1.

Table 1. MESPAS cloud solution wins over inhouse-installed software [10]

MESPAS Cloud Benefits	**CONCEPT**	Traditional / inhouse-installed software
Cloud-based system 1 off-site central server infrastructure, i.e. software and central server hardware accessible via standard IT equipment from anywhere with Internet access State of the art / future proof solution	Technology	Traditional software Clients run multiple servers, maintained and hosted on-site. Software installed on local computers; changes and enhancements to be implemented individually. Cost and time intensive
Multi tenancy to share hardware, memory and software among multiple tenants (clients) Software can easily be updated for all users via central server	Multi tenancy vs. single tenancy	Single tenancy means the system houses data and software for each company or vessel on a separate server. Time and infrastructure consuming; difficult to compare data and to ensure that users run the same version of the software
Available 24/7 via secure internet connection; server-farm that meets highest standards re: server redundancy, power backup, data backup, internet connection, fire protection, environmental control	Availability, security and disaster recovery	Local solution features significant lower physical and digital security. Higher disaster recovery costs
MESPAS software is constantly being enhanced and advanced. Improvements and additions are made available to customers free of charge, and deployed with little or no client involvement.	Software development	Software updates incur costs such as upgrade fees to software provider, deployment costs, as well as considerable costs for testing in own environment.
IT is core competence of MESPAS	Status of IT	IT is not core competence of company

Table 2. (continued)

	COSTS	
- No CapEx - capital expenditures (only standard IT equipment needed) - Predictable OpEx - operating expenditures. according to actual usage	Cost types	- Large initial CapEx - CapEx for upgrades and adding new users/vessels - Large and unpredictable OpEx

5 Conclusion

Cloud computing services play now and will play an increasingly important role in the maritime transport in the future.

In this paper presented some of the benefits and challenges of using CC in maritime transport. Cloud Computing is enabling the next generation of customs solutions, expanding global trade, economic growth and opportunities.

These benefits, based on examples, are connected with application of CC in Maritime Transport Single Window initiatives and fleet management operations.

Main characteristics of SaaS can be found in mespas R5 solution also. In terms of fleet management in maritime transport, there are two distinctive features the SaaS software possess: high quality data management and easy to manage offline availability [11].

In terms of data management, the SaaS provider manages – on behalf of customers – the data and documents that are shared by all users of the application. Since this is done centrally and for all clients, the costs for this are very low for each client, i.e. definitely much lower than if each customer would have to do this on his own.

Having offline availability means that office staff access the central database in real-time through a secure internet connection – no matter when and where they are located. Crews aboard the vessels work offline by accessing the database on board, which is mirrored to the central server's database ashore. Regular synchronization as part of the standard sync schedule of each vessel ensures that both ship and shore work with the same up-to-date information.

References

1. Lynch, G.: e-Maritime Overview. Periodic Study, SKEMA Interactive Knowledge Platform for Maritime Transport and Logistics (2010)
2. McLaughlin, H.: e-Maritime Periodic Study Summary. SKEMA Interactive Knowledge Platform for Maritime Transport and Logistics (2010)
3. Franclin, J. C., Chee, B.J.S.: Cloud Computing. CRC Press Inc., Taylor & Francis Group, Boca Raton (2010)

4. Babcock, Ch.: Management Strategies for the Cloud Revolution. McGraw-Hill, New York (2010)
5. Koh, J., Tsen, T.: Ten years of Single Window Implementation: Lessons learned for the future. Global Trade Facilitation Conference, Connecting International Trade: Single Windows and Supply Chains in the Next Decade (2011)
6. e-Freight D3.2 – Reference Solutions for Next Generation National Single Windows (2011)
7. Trans-Kalahari Corridor regional Single Window, http://www.oecd.org/dataoecd/37/56/48367286.pdf
8. SOGET's e-Maritime Port Single Window Project, http://www.microsoft.com/publicsector/ww/international-organizations/projects/Pages/enabling-maritime-trade-automation.aspx
9. Bureau International Maritime, http://www.bimv.com/en/SingleWindow/
10. MESPAS, http://www.mespas.com/
11. Thoma, Ch.: Software as a Service - revolutionizing fleet management. Digital Ship 42(1), 40–42 (2011)

The FTTD Method Application to the Safety Analysis of Changeable Block Distance System

Jan Magott[1], Andrzej Lewiński[2], Piotr Skrobanek[1], and Andrzej Toruń[3]

[1] Wroclaw University of Technology, Institute of Computer Engineering, Control and Robotics
Wybrzeże Wyspiańskiego 27, 50-370 Wrocław, Poland
{jan.magott,pawel.skrobanek}@pwr.wroc.pl
[2]Technical University of Radom, Faculty of Transport & Electrical Engineering
Malczewskiego 29, 26-600 Radom, Poland
a.lewinski@pr.radom.pl
[3] Railway Institute
J. Chłopickiego 50, 04-275 Warsaw, Poland
atorun@ikolej.pl

Abstract. The paper deals with the safety of a new train control method based on the Changeable Block Distance. Such method applies the open radio transmission standards including GPS and GSM-R for trains positioning, monitoring and controlling in a typical dispatcher area. The CBD method may be treated as a new technological overlay on classical train control systems, because it can improve the capacity of a railway line assuring the same, high level of safety (SIL4). The applied method of a Fault Tree with Time Dependencies allows to regard the time parameters of transmission connected with delays, interferences and interrupts together with reactions of the dispatcher system and of the railway staff. The results of the FTTD analysis confirm the assumed measures and suggest the trends in future applications of open transmission wireless standards in the railway control and management.

Keywords: Changeable Block Distance, Fault Tree with Time Dependencies, Safety of Railway Control and Management System.

1 Introduction

Traditional train monitoring system used FBD (Fixed Block Distance) rule of traffic control corresponding to sequential occupancy of insulated rail sections [4]. The FBD method requires the highest level of safety (SIL4) but has important limitation of railway line capacity. It is related to assurance of minimum length of block section necessary to breaking distance .(Technical implementation of the FBD rule does not require any additional track-train transmission systems (TTT class). This method allows the drive a train with maximum running speed to 160 km/h corresponding to correct interpretation of railway signaling system signals by the train driver. Therefore, higher running speed than 160 km/h require the use of additional TTT class systems, in which the information about actual allowed speed profile and free route

J. Mikulski (Ed.): TST 2012, CCIS 329, pp. 267–275, 2012.

distance is passed directly to the driver`s cab as an MA (Movement Authority (– permission for a train to run to a specific location within the constrains of the infrastructure) on the DMI (Driver Machine Interface – device to enable communications between TTT system and the train driver).

For this purpose, a control systems TTT class: ATP – Automatic Train Protection, ATC - Automatic Train Control, ATO – Automatic Train Operation, assigned to safety SIL-4 level [7] for which according to the railway safety standards (EN-50129) must be analyzed with respect risk analysis. For this type of TTT class systems the mathods such FTA (Fault Tree Analysis) and FMEA (Failure Mode and Effect Analysis) are recommended as standard methods of risk analysis [1, 2].

Currently used TTT class systems oin Europe Railways, allow the safety drive a train with speed above 160 km/h, but still existing the FBD rule decreases the efficient management of trains. To improve the capacity of the railway line the CBD (Changeable Block Distance) rule can be used as an alternative method of safety time distance control base on the existing signaling track equipment . Implementation of the CBD rule proposed in the presented concept assumes data transmission between RBC (Radio Block Centre) and trains via open public radio network standards (GSM-R, WiFi, WiMax). The typical risk aspects such: delays, interferences, interrupts or integrity code faults in data transmission or unauthorized attempts of reception are required in safety data transmission procedures between RBC and communicating trains.

In this paper the influence of the mentioned transmission aspects for proper operation of CBD system is analyzed using the FTTD (Fault Tree with Time Dependencies) method.

2 Fault Tree with Time Dependencies

Fault Tree Analysis (FTA) is one of the most commonly used techniques for an analysis of safety systems. The technique involves an identification of dangerous situations, so called faults, and causes (events) that may lead to its occurrence. The faults are situations which directly, or through initiating sequence of events, may lead to an accident. FTA is a top-down approach [1, 2]. First, dangerous situations (faults) are identified in a system. Next, one fault tree (FT) for each of faults is constructed. Fault is the top-event. A next level of the FT is constructed using direct causes of the top-event. The top-event and its causes are connected using a logical gate, as will be shown. Then, in the same way, a next level of the FT is constructed (for causes of the causes of the top-event). The FT is finished on a specific level of details or on event for which we cannot or do not need to specify the reasons.

In contrast to the classical FTA, the Fault Tree with Time Dependencies Analysis (FTTDA) allow, additionally, for analysis of timing relationships between events, as will be shown in Chapter 4.

The time parameters determined usually by experts on basis of system construction (minimal and maximal period of time for opening a valve, for message sending etc.)

or the physical laws (e.g. speed of sound in air) or other aspects (e.g. software implementation).

To avoid ambiguity, a notation of events and gates in Fault Tree with Time Dependencies (FTTD) was formalized. In this chapter, only definitions of selected gates (which was used in FTTDA in section 4) will be shown. More information about FTTD and FTTDA can be found in [4, 5].

An example of the formal notation of events and gates is given in Fig. 1.

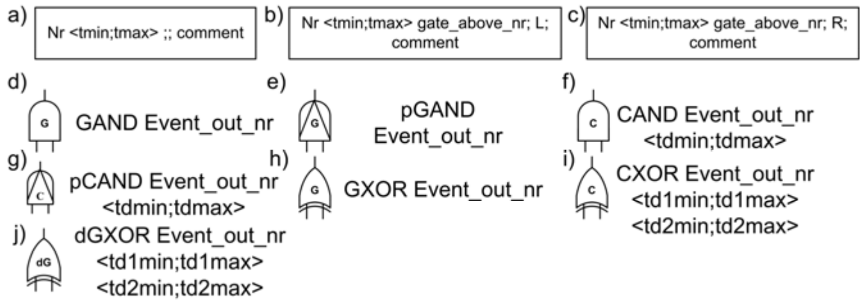


Fig. 1. The notation of: a) the top-event (fault), b) the left entering event, c) the right entering event d) the generalization AND gate, e) the priority generalization AND gate, f) the causal AND gate, g) the priority causal AND gate, h) the generalization XOR gate, i) the causal XOR gate, j) the generalization XOR gate with start delay

The events are described by a unique number *Nr*, the time parameters <tmin, tmax> - respectively the minimum and maximum duration time of the events (in particular, the minimum can be equal 0 and the maximum 0 or ∞). In addition (except the top-event), each event contains information about the gate and the input of the gate with which is connected (L - left, R - right). The last element is description. The time parameters for generalization event are not obligatory – they are counted during the analysis based on the parameters of input events.

A notation of the gates includes information as follows: a type of gate (e.g. generalization XOR, causal AND). For causal gate and generalization gate with start delay the minimal and maximal time delay are given

Let us assume the following designations for any event *e*: $\tau(es)$ – the time instant, start of the event *e*, $\tau(ee)$ – the time instant, end of the event *e*. In contrast to the parameters given in the Fault Tree, $\tau(es)$ and $\tau(ee)$ are relative to a contractual "0" time instant and are counted into analysis process. The "0" time instant is usually the start of the top-event.

Calculation of the time conditions for the start and the end of events allows for recognition of time relation between the events that can lead to the dangerous situation, as will be shown.

A formal definition of gates used in the next section is as follows:

Causal XOR

$$\text{occur}(z) \Rightarrow ((\text{occur}(x) \wedge \text{duration}(x)>td1min \wedge \tau(xs)+td1min \le \tau(zs) \le \tau(xs) + td1max \wedge \tau(zs) \le \tau(xe)) \oplus (\text{occur}(y) \wedge \text{duration}(y)>td2min \wedge \tau(ys) + td2min \le \tau(zs) \le \tau(ys) + td2max \wedge \tau(zs) \le \tau(ye))) \quad (1)$$

where:

- occur(*event*) – predicate, event *event* had occur, *event* ∈ {x,y,z},
- duration(*event*) – a duration time of the *event*, *event* ∈ {x,y},
- td1min, td1max, td2min, td2max, respectively, minimal and maxima delay time between a cause *x* (*y*) and effect *z*.
- ⊕ - "exclusive OR" operator.

This formula has the following meaning: if the event *z* (effect) has occurred then event *x* (*y*) must had occurred earlier. Additionally, start of the event *z* had to be no sooner than *td1min* (*td2min*) and no later than *td1max* (*td2max*) counted from τ(xs) (τ(ys)) –start of the event *x*. Some examples of the time relations between events for causal XOR gate are given in Fig. 2.

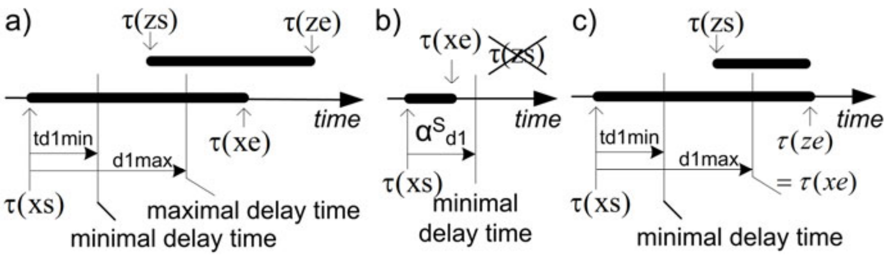


Fig. 2. Examples of time relation between causes x and effect z for a) causal XOR gate (event z occur), b) causal XOR gate (event z not occur), c) generalization XOR gate with delay start (event z occur),

Generalization XOR with Delayed Start (dsGXOR)

$$\text{occur}(z) \Rightarrow ((\text{occur}(x) \wedge \text{duration}(x)>td1min \wedge \tau(xs)+td1min \le \tau(zs) \le \tau(xs) + td1max \wedge \tau(zs)=\tau(xe)) \oplus (\text{occur}(y) \wedge \text{duration}(y)>td2min \wedge \tau(ys) + td1min \le \tau(zs) \le \tau(ys) + td2max \wedge \tau(zs)=\tau(ye))) \quad (2)$$

The main difference between this gate and causal XOR gate is as follows: in causal XOR gate the end of effect not depend on end of cause, in this gate the end of effect is equal the end of cause. The FTTDA involves the following operations:

- establishing the „0" time instant (usually the top-event start),
- calculating time parameters for top-event,
- calculating the time parameters for other events.

Some discussion the FTTDA can be found in [4, 5].

3 Absolute Changeable Block Distance - Description

The CBD conception assumes the time spacing control used dynamic block section between trains. The length of block section may change in time according to the current traffic situation in the controlled section of railway line. This means that the distance between trains is changeable and its actual length is the result of processing information about position and speed of all following supervised trains in the given railway line. [4]. The rule of the Absolute Changeable Block Distance $(CDB)_A$ presented in the Fig.3 bases on the virtual block section with not fixed reference points (connected with distance between station) but flexible modified corresponding to given traffic situation. This method may be compared with "electronic visibility" when the actual speed depends on position and speed of a previous train.

It is the Absolute Changeable Block Distance $(CDB)_A$ when the second train receives the permission of drive (ZNJ) to the place nearer than calculated on last report about position (LR – location report) of a previous train. (It is connected with assumption about "zero distance" stop immediately after sending the location report (RoP_1) and total breaking way with overlap of second train.)

The distance control process is shown at the Fig. 3. According to CBD_A rule the "*train 2*" receives movement authority (ZNJ_2) about precise position with respect to calculated localisation of "*train 1*". If the case of calculated distance between train 2 and 1 is shorter "*train2*" (KP_1) receives from Radio Block Centre new permission (ZNJ). ZNJ (movement authority) includes a collection of information on the validity of the permission, location of the end of movement authority, target speed at the end of authority, temporary speed restriction and static speed profiles. Finally when the driver outrun MA parameters train system automatically start – " service brake – emergency brake".

In functional analysis of the time sparing controls system with CBD_A rule can identify three basic scenarios for service:

1. Full Supervision mode (PN) – the train equipment mode giving full protection against over-speed and over-run. Train goes under the movement authority (ZNJ),
2. No equipped (NW), or system faliture mode - Train goes under the indication of a signals.
3. Staff Responsible (OM) – Mode after stopping a train. This mode allow the driver to move the train under his own responsibility after stop train. Train goes with a maximum speed equal 20 km/h, until to the first signals and next the train goes on NW mode (when train equipment is out of service) or until the recept new movement authority (ZNJ) when train equipment is full of service, then train goes on PN mode)

Linking operational scenarios resulting from the principles of the system, in particular, error detections (e.g. train equipment failure, delays, interferences, interrupts in the public radio-transmission network or errors in GPS reading) or error in the operations of the drivers, as a result of which the train stopped.

In the analysis of the effects of detected errors are defined two group of errors:

- First (critical fault) – the train is breaking implementing immediately,

- Second (non-critical fault) – the train stops in accordance with the movement authorization.

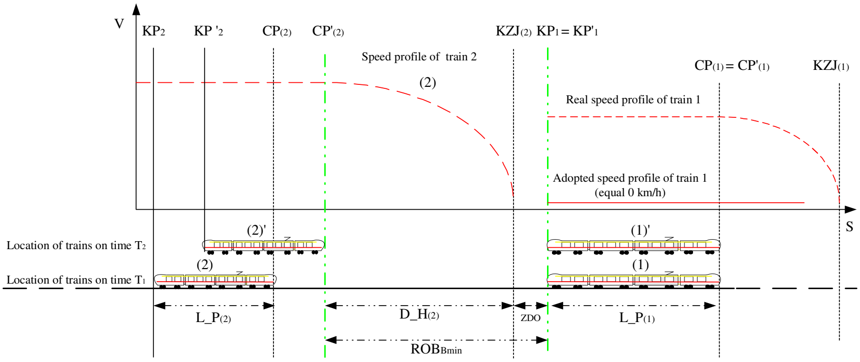


Fig. 3. The rule of Absolute Changeable Block Distance

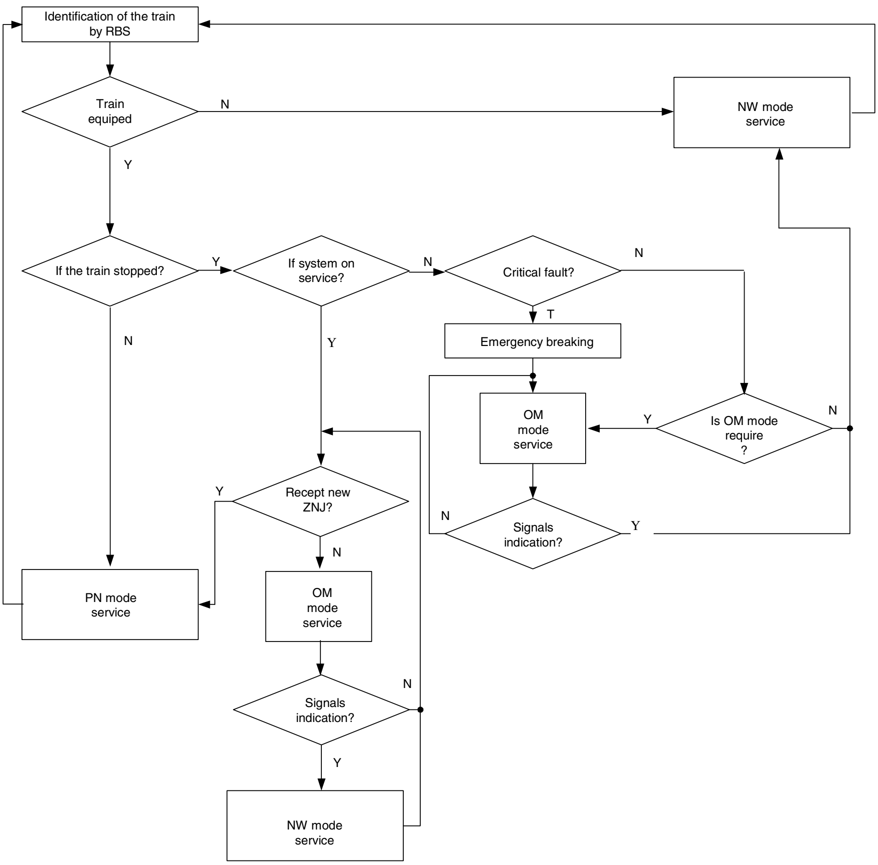


Fig. 4. The relations between operational scenarios

4 Case Study

In the paper, FTTD for top event „hitting the preceding train" will be analysed. In case study, „driving on visibility" mode is not analysed. In order to simplify the case and because of limited room:

- the analysis will be omitted – only results of the analysis will be presented,
- the following assumption is accepted: RSW (supplementary radio system) and CS (control centre) are reliable.

The FTTD is given in Fig. 5.

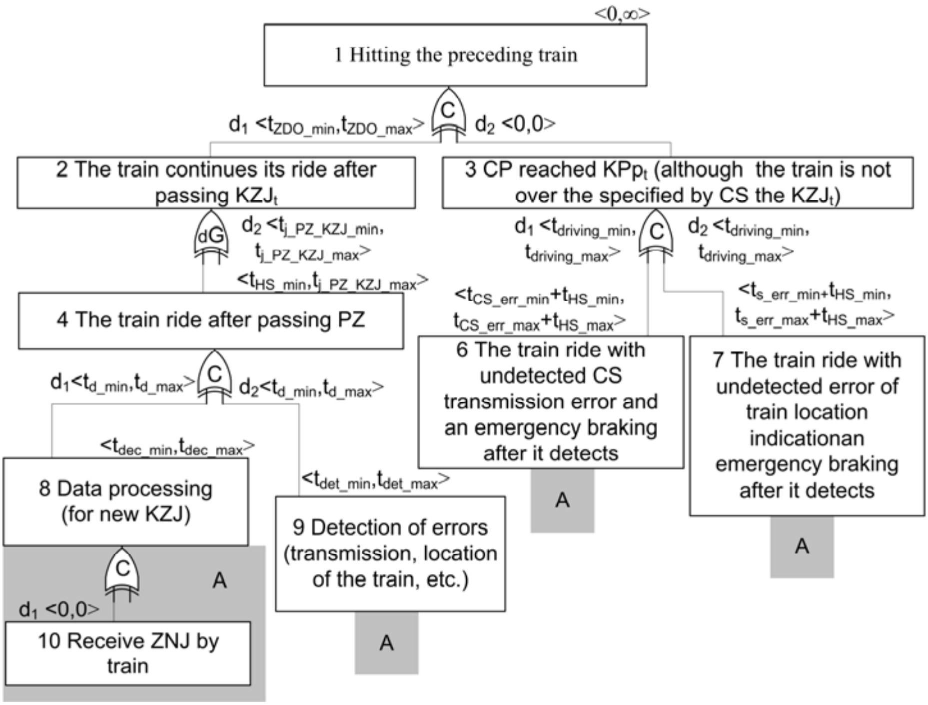


Fig. 5. FTTD for train traffic according to CBD control

Meaning of symbols and time parameters is as follows:

- CP – front of a train, CPp – front of a preceding train,
- KP – end of a train, KPp – end of a preceding train, KPp_t – end of a preceding train in time instant *t* (when the train receives an agreement to drive ZNJ),
- KZJ – end of permission to drive, KZJ_t – end of permission to drive given by CS in instant *t*,
- PZ – point where the train has to start to decelerate (according to speed function),

- t_{ZDO} – driving time through the protection distance ZDO; t_{ZDO_min}, t_{ZDO_max} , respectively, is minimal, maximal driving time through the ZDO (when the train speed is maximal, minimal),
- $t_{driving}$ – driving time when the train reaches the KP of preceding train,
- t_{CS_err} – length of time interval when an transmission error from the CS is not detected (including case when an intruder/a terrorist simulates that he/she is the CS, and this activity is not detected),
- t_{s_err} – length of time interval when there is an error of train location indication,
- t_{HS} – braking time (until the train is stopped),
- t_{dec} – decision time of new KZJ preparation (train location message transmission from train RSW to the CS, CS processing, message transmission from the CS to the train RSW),
- t_{det} –errors detection by RSW and their evaluation time (to determine whether "riding on visibility" is needed.

During a construction of FTTD it has been chosen that faults would be considered as relative to instant when the train receives the ZNJ message (because RSW and CS are reliable and new parameters are calculated using the ZNJ message). In the FTTD (Fig. 5), it is modeled by synchronization of faults by event 10.

In the FTTD, two sub-trees can be extracted.

Left sub-tree (with root event 2) models normal system work. For example, event 8 occurs in normal work mode, but its time parameters are required in order to find bounds on the other time parameters to avoid slowing down the train.

Right sub-tree (with root event 3) models abnormal behaviour. Cases when the train or the CS receives incorrect information are analysed. Such situation can occur if e.g. someone has intercepted a transmission or the train localization was incorrect because of a failure. There is no room to discuss the problems in detail. Timing analysis gives some requirements imposed on the system.

The analysis gives us the following results:

- $t_{CS\_err\_max} < t_{driving\_min} - t_{HS\_max}$ - detection time must be sufficiently short; for example, it is possible to introduce a supplementary analysis of the position of trains in order to verify the accuracy of data transferred from the CS or additional identity verification of CS,
- $_{s\_err\_max} < t_{driving\_min} - t_{HS\_max}$ – detection of train location error must be sufficiently short,
- $t_{dec\_max} < t_{d\_min}$ – time needed to process the data and set new parameters should be sufficiently short so as not to cause the train’s braking,
- $t_{det\_max} < t_{d\_min}$ – analogically as $t_{dec\_max} < t_{d\_min}$,
- $t_{j\_PZ\_KZJ\_min} + t_{ZDO\_min} > t_{HS\_max}$ – stop of the train should be no later than the minimum time after which can create a dangerous situation + time reserve.

5 Conclusion

According to analysis of CDB method of train control the critical failure – the collision of two trains in the same direction is connected with occurrence of several faults simultaneously (transmission system or emergency train braking faults) . The necessary condition of such critical situation is very good explained in FTTD method. This method is more restrictive for safety analysis, because the typical method of safety assessment of railway control systems assumes the scientifically small probability of multiple faults (the intensity of multiple faults corresponds to very short time of fault diagnostics defined in obligatory standard [7]).

The FTTD results of CDB analysis allow to satisfy the fail-safe criteria defined in CBD system design and may be applied as an additional condition for capacity optimization. The results presented in the paper are related tot he first stage of more complex research work connected with fault analysis in failures in other parts of Railway Train Monitoring, Management and Control Centre such Radio Block Centre or Movement Authority.

In the paper only the critical failure of CBD_A is analyzed, the non critical failures generally connected with transmission faults and delays require the similar FTTD analysis. The results may be used for design and implementation of appropriate procedures assuring the fail-safe operation of CBD system.

The FTTD schemes may be treated as a supplementary documentation form of safety analysis necessary in railway control systems in certification process.

References

1. Fault Tree Analysis (FTA), International Standard, IEC 61025 (2006)
2. Fault Tree Handbook, NUREG-0492, U.S. Nuclear Regulatory Commission, Washington D.C., 2055 (1981), `http://www.nrc.gov/reading-rm/doc-collections/nuregs/staff/sr0492/sr0492.pdf`
3. Jureczko, M., Skrobanek, P.: Tool for analysis of the fault tree with time dependencies. Przegląd Elektrotechniczny 86(9), 179–183 (2010)
4. Lewiński, A., Toruń, A.: The Changeable Block Distance System Analysis. In: Mikulski, J. (ed.) TST 2010. CCIS, vol. 104, pp. 67–74. Springer, Heidelberg (2010)
5. Magott, J., Skrobanek, P.: Timing analysis of safety properties using fault trees with time dependencies and timed state-charts. Reliability Engineering & Systems Safety 97(1), 14–26 (2012)
6. Magott, J., Skrobanek, P.: A Method of Analysis of Fault Trees with Time Dependencies. In: Koornneef, F., van der Meulen, M.J.P. (eds.) SAFECOMP 2000. LNCS, vol. 1943, pp. 176–186. Springer, Heidelberg (2000)
7. Railway applications - Communication, signaling and processing systems. Safety-related electronic systems for signaling. International standard PN-EN 50129-2010 (U)
8. Railway applications - Communication, signaling and processing systems. Safety-related communication in transmission systems. International standard PN-EN 50159-2010 (U)

Development of Cellular Automata for Simulation of the Crossroads Model with a Traffic Detection System

Krzysztof Małecki[1] and Stanisław Iwan[2]

[1] West Pomeranian University of Technology, Dept. of Computer Science
49 Żołnierska St., 71-210 Szczecin, Poland
kmalecki@wi.zut.edu.pl
[2] Maritime University of Szczecin
11 Pobożnego St., 70-515 Szczecin, Poland
s.iwan@am.szczecin.pl

Abstract. Cellular automata are considered by some researchers as one of the computer science areas, which is an artificial intelligence. Even though their history goes back to the forties of the last century, much attention is still paid to the use of cellular automata in the process of testing and simulation of different phenomena. The universality of cellular automata led to their application in many areas such as: physics and traffic modeling. The authors present the achievements in the traffic modeling field and formulate a new traffic model. This new model helps to simulate the traffic at the crossroads of two-lane and two-way roads. In addition, the authors present some aspects of the prepared system, which enables to model any area (intersections and roads) and to carry out the traffic simulations in a microscopic scale.

Keywords: cellular automata, crossroads model, simulation, traffic detection.

1 Introduction

In today's world, almost everybody comes across the problem of street congestion – some as drivers, others as pedestrians. For most of them the only solution to this situation is expanding the network of roads and sidewalks. Unfortunately, there are real limitations to such a policy, due to reduction in the number of densely built-up urban areas and the need to protect the rest of the "green" ones, which are necessary to preserve life on our planet.

Therefore, the need to streamline the existing transportation system is being observed. Of course, this can be interpreted in different ways, e.g. by modeling and controlling the traffic or by continuous education of drivers to show them incorrect behaviors on the roads. For this purpose it is useful to have adequate mathematical models and adequately prepared computer applications (software).

Table 1 shows the existing traffic models using cellular automata. Some of them have been implemented and presented in the form of a system available on website [6]. Another system, which deserves attention is the TrafficSim [7]. It includes a few

J. Mikulski (Ed.): TST 2012, CCIS 329, pp. 276–283, 2012.

implemented traffic models [8]. According to the information presented on the project website [7], the authors developed a simple simulation environment for the traffic.

Table 1. The traffic models using cellular automata

The model	Description
The model of Chopard-Luth-Queloz	The elementary cellular automaton, defined by S. Wolfram, as a rule number 184; it submit the traffic on the one-way and single-lane road. Cars move in one direction.
The model of Nagel-Schreckenberg [4]	Model describing the cars' movement on the one-lane and one-way road. The lane is divided into sections (cells) with a length of 7.5 meters. Each vehicle moves at a speed no greater than the maximum speed, but not less than zero. The model has an absorbing boundary conditions. Transition function is divided into 4 stages: acceleration, braking, and shifting a random event.
The model of Chowdhury-Schadschneider [1]	The crossroad's model, where traffic is only right or up. The model has periodic boundary conditions. Transition function as a model NaSch.
The model of Treiber: Ring-Road [6]	Vehicles move on the two-lane and one-way road, represented by an array of two rows and *n*-columns. Transition function as a model NaSch, but the vehicles may change lane.
The model of Treiber: On Ramp [6]	The two-lane and one-way road, with one-way run-up way (the third lane). Transition function consists of 5 stages: a lane change, acceleration, braking, shifting and random events.
The model of Treiber: Lane closing [6]	The two-lane and one-way road. The fragment of this road is closed. Transition function consists of 5 stages: a lane change, acceleration, braking, shifting and random events.
The model of Treiber: Uphil Grade [6]	The two-lane and one-way road. The traffic difficulties are simulated: the slow going vehicles moving on the steep climbs obstruct the traffic.
The model of Treiber: Traffic Lights [6]	The network of cells with two rows and columns to provide a two-lane and one-way road with traffic lights. Transition function consists of 5 stages.
The model of Treiber: Lane Changes [6]	The model identical to Ring-Road model, but a few obstacles are placed on both lanes of traffic, forcing cars to change lanes.
A few models of Bartodziej [8]	Development of several models, including Chowdhury-Schadschneider, Ring-Road and On Ramp.
This article	Extending of crossroads' models presented in [1] and [8]. The authors propose to change the functions of transition - the second stage, by considering the inductive loop, which eliminates the possibility of blocking the crossroads.

Unfortunately, the number of existing models (they do not include all aspects of road traffic) and computer applications (particularly those free for the users) is still insufficient. Therefore, relying on the analysis of the available solutions, the authors developed a crossroads model (presented in this paper). Additionally, the idea and some basic elements of an environment based on cellular automata theory are presented, to analyze drivers' behavior in a given situation in a specific structural area. The application is still being developed.

2 Extending the Existing Model

This paper proposes a thesis that developing a theory and mechanism preventing intersection traffic congestion will increase the motor vehicle traffic capacity of cities.

2.1 The Idea

In Poland, there is an appropriate article in the Road Traffic Code [9, Article 25, para. 4], which provides that "…a vehicle driver is not allowed to enter an intersection if there is no place at it or behind it to continue driving…". Unfortunately, the experience of Polish cities shows that Poles have a problem with respecting this rule, consequently increasing the congestion phenomenon.

The easiest way to verify this thesis is to construct a mathematical model and verify its correctness using the computer application.

2.2 The Model of Crossroads with "Inductive Loop"

In the classic model of Chowdhury-Schadschneider concerning traffic at the crossroads [1] and [8], the situation shown in Figure 1 is possible. In order to avoid this phenomenon, the authors worked out a new model, based on models presented in [1] and [8].

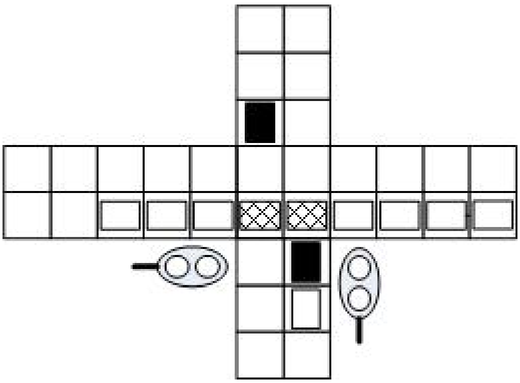


Fig. 1. The models [1] and [8] may lead to the intersection traffic congestion

The model applies to crossroads of two-lane and two-way roads, assuming that the distance between crossroads does not need to be identical. One cell of automata can be free or busy (then it has the numerical value: $V_{j,i} = 0, 1, \ldots, V_{\max}$, where j is the number of ways, i – the number of cell for which the vehicle's speed is set at the road section).

Extending the existing crossroads models concerns the use of the so-called inductive loop behind the crossroads (Fig. 2). The transition function, like in models [1] and [8], consist of 4 stages, but the second one (breaking) was modified; stages: 1, 2 and 4 are consistent with the model [4]:

1. acceleration – vehicles whose speed ($V_{j,i}$) at time t is less than $V_{\max}$, increase it by one:

$$V_{j,i} < V_{\max} \rightarrow V_{j,i} = V_{j,i} + 1\,, \tag{1}$$

   where $V_{j,i}$ – the speed of a vehicle situated on the road number j and the cell number i;
2. braking – if the traffic lights at the next crossroads is on with a color:
   - red and

$$V_{j,i} > \min(d_{j,i}, s_{j,i}) \rightarrow V_{j,i} = \min(d_{j,i}, s_{j,i}), \tag{2}$$

   - or: green and

$$d_{j,i} > s_{j,i} + p_{j,l} \rightarrow V_{j,i} = \min(V_{j,i}, d_{j,i}), \tag{3}$$

   - or: green and

$$d_{j,i} \le s_{j,i} + p_{j,l} \rightarrow V_{j,i} = \min(d_{j,i}, s_{j,i}), \tag{4}$$

   where $d_{j,i}$ – the distance from the vehicle located in the cell number i and road number j to the next vehicle, $s_{j,i}$ – the distance from the vehicle located in the cell number i and road number j to the next traffic lights, $p_{j,l}$ – the distance from the traffic lights on the road number j to the inductive loop placed behind the crossroads;
3. a random event – with a given probability, the cars reduce their speed by one:

$$V_{j,i} > 0 \wedge P < p \rightarrow V_{j,i} = V_{j,i} - 1, \tag{5}$$

   where P – the random variable, p – the probability of appearance of random event;
4. a move – vehicles move through the number of cells as their speed was, the time variable changes the value: t=t+1.

To ensure better intersection traffic capacity in the situation shown in Figure 2, the traffic lights could change to red (according to the formula (4) the vehicle located in the marked field will not enter the crossroads). Then the cars marked in black will be

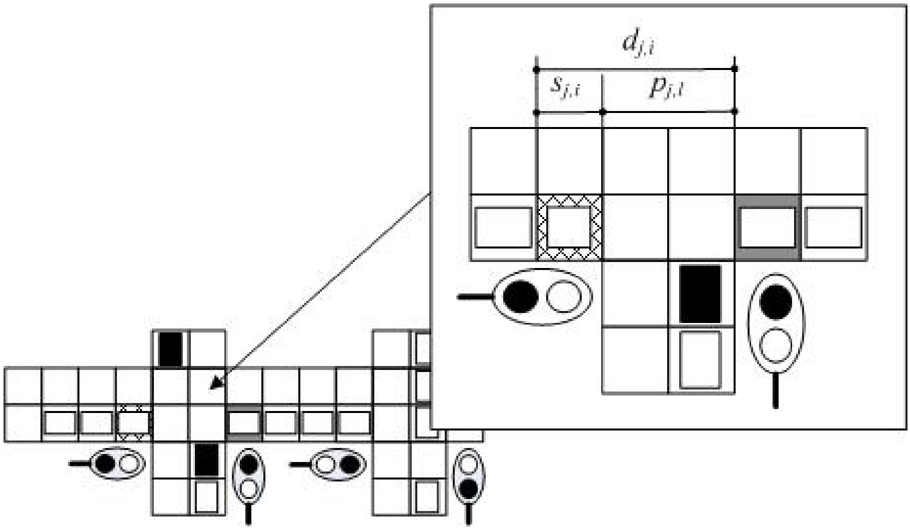


Fig. 2. The models [1] and [8] may lead to the intersection traffic congestion

able to continue their movement a bit earlier. The next change of traffic lights may take place straightaway after the place behind the crossroads is empty (the gray box shown in Fig. 2).

3 The Assumptions of the Traffic Simulation System

The system, which is being developed at the Department of Computer Science at the West Pomeranian University of Technology in Szczecin (Poland), consists of two main applications: TrafficCA (traffic simulation based on cellular automata) and MapGen – the map editor, which allows users to generate any area comprising the roads, traffic lights and surroundings.

The main aim of developing such a system is to provide a tool for analyzing drivers' behavior, depending on the situation on the road, according to the models presented in Table 1. The software will include new models describing other possible events, such as aggression with reference to driving comfort of other traffic participants or the impact of inductive loops on reducing intersection traffic congestion.

The map obtained in the system consists of grid cells with n-rows and m-columns. A single cell may represent a roadway or shoulder (the area closed to vehicles). The pavement cells are used to create paths which cars move along. Each track is a tree structure (Fig. 3), and therefore has one root and n-leaves. The root represents a point at which the vehicles' objects are created (entering the zone shown on the map), the node splits into two subsequent paths (fork of the roads) and the leaves are the points where the cars are "destroyed" (they leave the map). For each node a probability value is assigned. Based on that the vehicle determines the choice of the road (path in the tree). For the root the probability value means to emerge a new car on the map at the point. The leaves are the same facilities as other nodes – they only do not hold a reference to the next node.

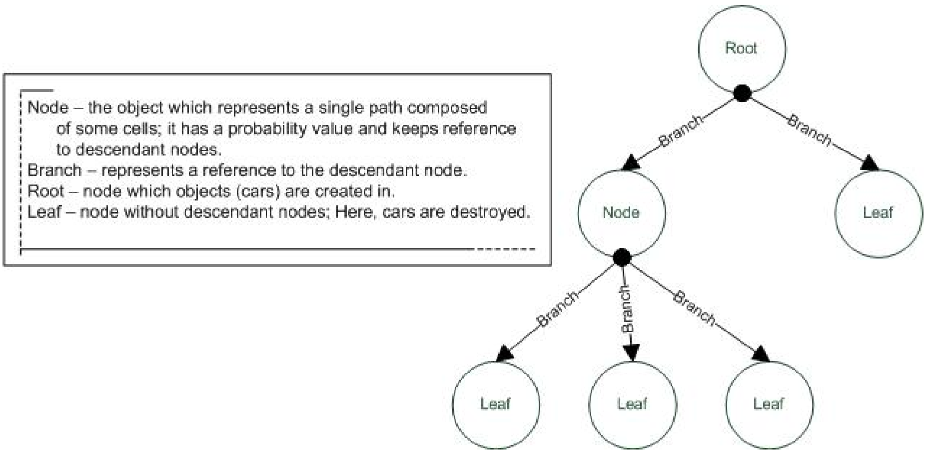


Fig. 3. The tree for the path which the vehicle moves on

Furthermore, any cell typed as "roadway" may contain an object representing the traffic lights. It is defined by three parameters – red and green traffic lights durations and a delay (it is used for proper signals synchronization).

As in the Nagel-Schreckenberg model [4], a simulated car is not bigger than a single cell and keeps references to the cell where it is located and to the cells following it in the path. A car may move to the next cell, if the red traffic light is not on in the current cell and there is not another car in the next cell.

The object "car" holds some values, including information about the current cell and the next cell, the current path (road), the coordinates, speed, maximum speed, acceleration (or delay) and maximum acceleration. Acceleration / delay depends on the distance to another vehicle or traffic lights and random events.

Thanks to the technology and application of new models, the software has great potential to use:

- access through any Web browser;
- map editor that allows to model any area of study;
- a simulation enabling cars differentiation due to the different values of acceleration and maximum speeds, generating random data and modeling of specific situations on the road;
- complex statistics and visualization of data generated by the simulator, such as traffic intensity, average speed, speed and acceleration of vehicles on the road section, the capacity of the area and characteristics of traffic lights;
- and many others.

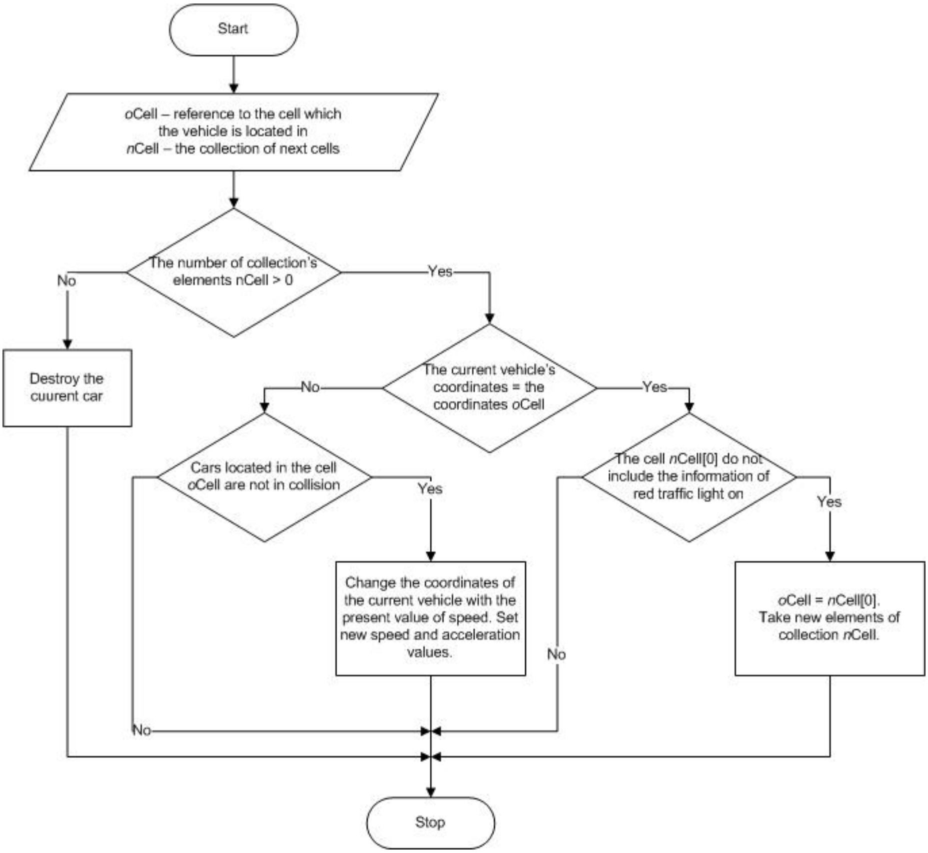


Fig. 4. The algorithm of the vehicle movement

4 Experimental Results

To validate the developed model, the authors implemented two identical crossroads with traffic lights in the simulator. One of the intersections was additionally equipped with a traffic detection system – the induction loops (behind the intersection - to see if the road is passable).

Numerous simulations have shown that the intersection traffic congestion is dependent primarily on:

- –traffic on a given stretch of road,
- –traffic light settings,
- –time periods ("dead zones") between the switching on subsequent traffic lights,
- –the distance between two intersections.

The results of the tests allow to conclude that the detection system (e.g. inductive loops) effectively minimize the risk of conflict situations and prevent blocking of vehicles at crossroads. Obviously, a key aspect is the appropriate location of inductive loops - for this purpose all the above-mentioned factors should be included, and possibly the specific traffic lights durations could be corrected.

5 Conclusion

A new crossroads model which is an expansion of the models presented in [1] and [8], was drawn up. The analysis of the current models has highlighted the phenomenon of intersection traffic congestion, now completely eliminated.

Based on the analysis of the existing traffic models the authors attempted to work out the traffic simulation system and drivers' behavior in various road situations. The basics of the system are presented in this article.

References

1. Chowdhury, D., Schadschneider, A.: Self-organization of traffic jams in cities: effects of stochastic dynamics and signal periods. Physical Review E 59, 1311–1314 (1999)
2. Esser, J., Schreckenberg, M.: Microscopic Simulation of urban traffic based on cellular automata. International Journal of Modern Physics C 8(5), 1025–1036 (1997)
3. Kułakowski, K.: The cellular automatas, AGH-UST (2000) (in Polish)
4. Nagel, K., Schreckenberg, M.: A cellular automata model for freeway traffic. Journal de Physique I 2, 2221–2229 (1992)
5. Wolfram, S.: A New Kind of Science. Wolfram Media (2002)
6. http://www.traffic-simulation.de/ (date of access: January 05, 2010)
7. http://student.agh.edu.pl/~kdzwinel/Projects/TrafficSim/index_pl.html#authors (date of access: November 25, 2009)
8. http://www.ioz.pwr.wroc.pl/pracownicy/weron/prace/Bartodziej07.pdf (date of access: October 11, 2009)
9. The Road Trafic Law, Act introduced on the 20th of June 1997, consolidated text of act. Journal of Law, No 108, Pos. 908 (in Polish)

Marine Integrated Navigational Decision Support System

Zbigniew Pietrzykowski, Piotr Borkowski, and Piotr Wołejsza

Maritime University of Szczecin, Faculty of Navigation,
Waly Chrobrego 1-2, 70-500 Szczecin, Poland
{z.pietrzykowski,p.borkowski,p.wolejsza}@am.szczecin.pl

Abstract. The article presents a concept of an integrated navigational decision support system for vessel traffic control. The system is based on a joint platform for decision support systems on land and ships. The platform construction is recommended to be based on the shipboard navigational decision support system. The necessary scope of work to build such system is defined. The proposed integrated decision support system NAVDEC fits such concepts as e-navigation, e-maritime and intelligent marine transport systems that reflect global trends in the development of navigational systems, widely using modern information communication technologies.

Keywords: sea transport, navigation, decision support system.

1 Introduction

Information technologies available today open increasing possibilities of building systems supporting decision processes in various fields of human activities. This also refers to maritime transport. The use of information technologies in maritime transport is of key importance from the viewpoint of navigational safety enhancement and transport efficiency.

The complexity of decision making processes results from, inter alia, the need to acquire, analyze and process great amounts of data and from time restrictions while choosing a solution (decision making). This, in turn, makes taking decisions by navigators difficult and may lead to dangerous situations, which is confirmed by reports of marine accidents and disasters. Verdicts of maritime courts indicate that wrong decisions – human errors – are among most frequent causes of marine accidents and disasters that have grave consequences: loss of life and health, material losses due to sinking of or damage to the ship and/or cargo, threats and damage to the natural environment.

Information technologies enable creating efficient navigational equipment and systems that, using complex computing algorithms, gather, process and present navigational information to the navigator. For several years such systems have been used on ships and in shore-based centres. Navigational systems in use today mainly execute information functions and to a certain extent assist in safe ship conduct. They automate the information acquisition from various sources and integration, making navigator's decision easier. Concepts such as e-maritime, e-navigation, Maritime ITS

J. Mikulski (Ed.): TST 2012, CCIS 329, pp. 284–292, 2012.

[3, 4, 9, 11] show that Information Communication Technology (ICT) will have increasingly wider applications in sea transport. As a result, there is a trend to transform navigational information systems into navigational decision support systems. Their functions, apart from acquisition, integration, processing and presentation of information, include situation analysis and assessment, generation of solution/s to dangerous situations and the recommendation of these solutions to the navigator as well as shore-based traffic personnel responsible for vessel traffic management.

2 The Navigational Decision Support System NAVDEC

2.1 Description of the System

Human errors should be reduced or eliminated wherever possible, in order to provide for possibly high navigational safety level. This can be achieved only by equipping ships and land based centers with tools that, apart from information functions, will work out solutions to collision situations accompanied by adequate explanations. Primary functions of the navigational decision support system are the following:

- automatic acquisition and distribution of navigational information,
- analysis of a navigational situation ,
- alarming of dangerous situations and indicating the current level of navigational safety, based on criteria commonly used by expert navigators,
- solving collision situations, including automatic determination of an optimal manoeuvre and trajectory in collision situations,
- explanation of the optimal manoeuvre choice,
- interaction with the navigator.

Developed at the Maritime University of Szczecin, the navigational decision support system NAVDEC [8, 10] comes as the first navigational tool worldwide that apart from information performs functions typical of decision support systems. Its innovative functionalities, significantly extending the performance of devices generally carried by ships, have now a status of patent applications.

The NAVDEC system supplements the range of shipboard navigational devices. This real time system handled by the navigator monitors its ship and the environment and records information on the present navigational situation. On this basis the system identifies and assesses the navigational situation (processing) and works out solutions (decisions) assuring safe navigation. The correct functioning of the system requires interoperability with shipboard standard equipment and systems such as: log, gyrocompass, ARPA (Automatic Radar Plotting Aids), GNSS (Global Navigational Satellite System), AIS (Automatic Identification System), ENC (Electronic Navigational Chart), sources of current navigational data (Fig. 1). Like to the ECDIS system (Electronic Chart Display and Information System) the NAVDEC (Fig. 2) it provides on-screen presentation of bathymetric data from an electronic chart, an image of surface situation from a tracking radar, positional information from the AIS and GNSS receivers. It also determines movement parameters of targets in vicinity and presents them the navigator.

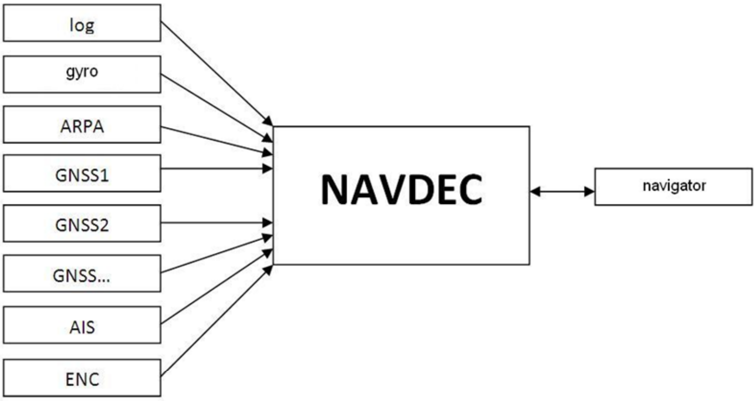


Fig. 1. A diagram of the NAVDEC environment [10]

The data defining navigator's and other ships' (targets) movement parameters make up a basis for effective solutions to collision situations. The accuracy of data presented to navigators is of major importance for their correct situation assessment and decisions to be made. For this reason the NAVDEC system performs the fusion of own ship data (measurements from a number of shipboard GNSS receivers are used) and integration of data on targets from alternative sources (tracking radar, AIS) [1, 10].

Fig. 2. The NAVDEC in use: the m/v Nawigator XXI [10]

Among new features of the NAVDEC is its ability to analyze and assess the navigational situation done in relation to all other or selected targets located within eight nautical miles. This is one of the decision making steps normally taken by the navigator, simply because the situation assessment takes into account relevant regulations. Thanks to the NAVDEC system the navigator is currently advised on the identification of an encounter situation in compliance with the Collision Regulations. This is a considerable aid, particularly in heavy traffic, although collisions of one-to-one ships in the open sea are known (m/v Gotland Carolina and m/v Conti Harmony in 2009).

When a collision situation is developing, it is the navigator who decides on a safe manoeuvre, the one that solves a given situation. The navigator decides on actions to be taken (alter course and/or speed) and the manoeuvre parameters: moment to begin it and values of course and/or speed alteration. The navigator may specify a safe course, on which the target will be passed at a preset range considered as safe. To date, the NAVDEC has been the only tool worldwide capable of performing this function (Fig. 3). The NAVDEC 'knows' the Collision Regulations [2], principles of good sea practice, as well as criteria used by expert navigators [5, 6, 7, 8]. Apart from one specific solution the system submits, alternative solutions complying with the regulations are also determined (possible range of course and/or speed alterations). Moreover, the system explains why a given is proposed, which refers to all targets or selected ones.

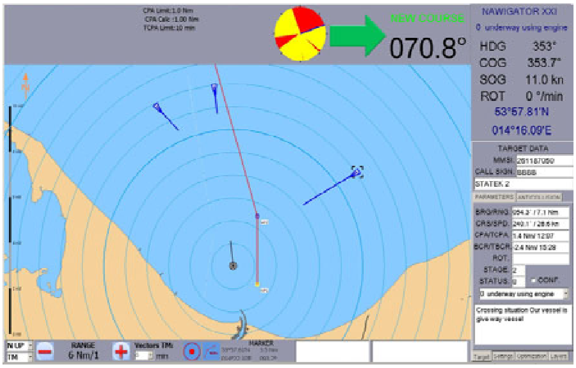


Fig. 3. The NAVDEC interface illustrating a collision situation [10]

Although the system proposes solutions and justifies them, it does not relieve the navigator from responsibility, but make his/her decision much easier to make. However, it is possible to make the ship control automatic by direct connection of the NAVDEC with the autopilot, steering gear, main engine, engine telegraph and controllable pitch propeller.

2.2 Comparative Analysis

Navigational systems installed on ships of the global fleet serve mainly information functions and correspondingly, to some extent provide an aid in safe vessel conduct. However, none of the known systems displays to the navigator ready solutions to a collision situation that are worked out in relation to all vessels in vicinity of their ship. This considerably limits decision support, consequently it limits the effectiveness of collision avoidance.

Today the navigational bridge of a ship is fitted with devices intended to help the navigator to sail safely. The ARPA (Automatic Radar Plotting Aid), a commonly used tracking radar is a marine tool for decision support in collision situations. However, the ARPA is sometimes ineffective due to its numerous limitations.

The NAVDEC has the following advantages in comparison to the tracking radar, presently used on ships for calculating encounter parameters and working out an anti-collision manoeuvre:

- incorporates the Collision Regulations, for both good and poor visibility,
- plans a manoeuvre also for the ship located in the radar blind area,
- the operator is immediately notified about a manoeuvre started by another ship thanks to information on target's rate of turn,
- needs just a few seconds to calculate the encounter parameters. On the other hand, the tracking radar, according to test situations defined by the IMO (International Maritime Organization), after one minute of tracking presents the CPA (Closest Point of Approach) with a one nautical mile accuracy. After three minutes, the maximum allowable error of CPA and TCPA (Time to Closest Point of Approach) calculated by the radar may amount to, respectively, 0.3 Nm and 0.5 minute. These tolerances are given with a 95% probability,
- more accurately calculates the encounter parameters, by:
    - taking account of the ship's size thanks to information on the position of the antenna, received from the AIS (*Automatic Identification System*),
    - use of GPS (*Global Positioning System*) / DGPS (*Differential GPS*) for position determination,
- takes account of the ships' sizes while planning an anti-collision manoeuvre,
- calculates new courses and speeds of own ship, such that other vessels will be passed at a preset CPA.

The ECDIS (Electronic Chart Display and Information System) is another information system commonly used on sea-going ships, often referred to as an electronic chart. Apart from displaying the positions and movement vectors of own ship and targets (vessels in vicinity), it also allows to, inter alia, obtain information on depths, aids to navigation and coastline. However, it is only an information system. The NAVDEC system moves a step forward (Fig. 4). Besides providing the navigator with the same scope of information, necessary for safe navigation, as the ECDIS, NAVDEC recommends ready solutions enabling the ship to safely pass vessels or stationary objects. Moreover, these solutions take into account ship's manoeuvring ability and the present area restrictions, i.e. existing traffic separation schemes, shoreline and safe depths. NAVDEC is, therefore, a very useful decision support system.

NAVDEC was developed to support navigators in decision making on board a sea-going vessel. At present, work is in progress on extending the functionalities of the system. It seems purposeful to undertake work on the introduction of NAVDEC functionalities to Vessel Traffic Services systems (VTS) operating on land.

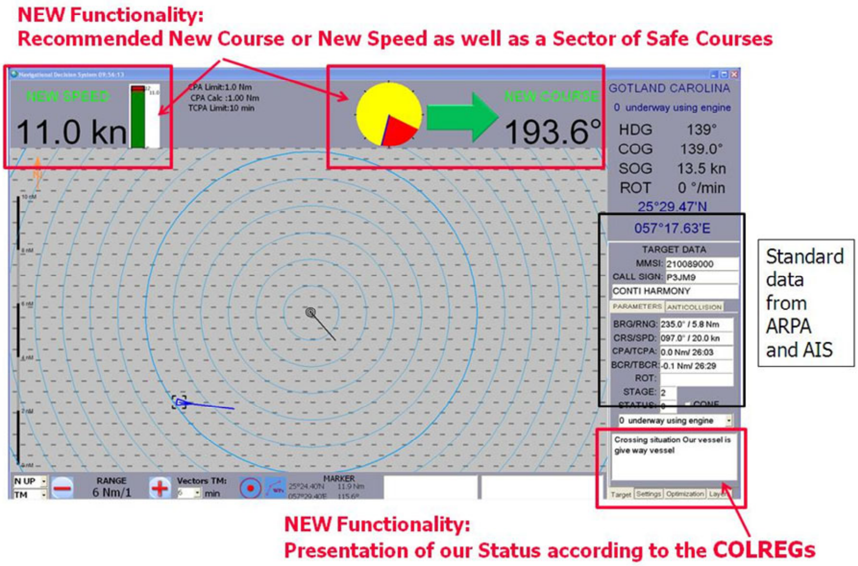


Fig. 4. Innovative functions on the NAVDEC interface [10]

3 A Concept of Integrated Navigational Decision Support System on Ships and in Land-Based Centres

3.1 The Navigational Decision Support System in a Land-Based Centre

VTS systems are information systems installed in centres responsible for vessel traffic monitoring and control. VTS stations were initially equipped with devices such as those on board ships: radars and ARPA systems. Today VTS centres operate specialist systems with such functions as are executed by shipboard systems: ARPA, ECDIS and AIS.

To increase their effectiveness these systems are supplemented with management modules that assist the operator in decision making (VTMS systems). VTMS systems combine three basic services of VTS: information, navigational assistance and vessel traffic management. The extension of system functions mainly refers to the issues of port complex management. Actions are being taken to integrate individual VTS and VTMS systems to enable their co-operation, so that larger areas could be covered within one VTMIS.

The broadening of information systems presently operated in VTS centres by adding the NAVDEC functionalities will enhance the scope of decision support offered by a land-based centre. Such extension will include:

- fusion and integration of navigational data available on board and at a land-based centre,

- analysis and assessment of a navigational situation that account for the Collision Regulations,
- automatic generation of solutions to collision situations, using special computational algorithms, including optimization algorithms,
- explanation of the present navigational situation based on the navigational knowledge base (Collision Regulations, principles of good sea practice, criteria of navigational situation analysis and assessment used by expert navigators),
- justification of the proposed manoeuvre.

This will allow to solve situations involving a group of indicated objects (vessels) in accordance with regulations in force. At the same time, it will be possible to find solutions from the viewpoint suggested by the object operator (Fig. 5), similarly to the NAVDEC system on board vessel.

a) b)

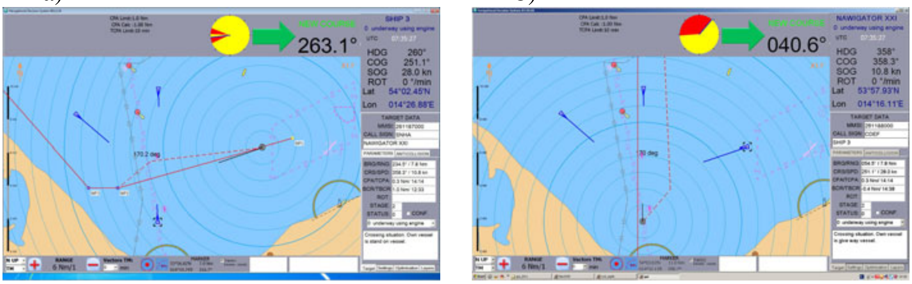


Fig. 5. Presentation of a navigational situation in the NAVDEC system at a land-based entre: a) viewpoint – chosen ship Nawigator XXI; b) viewpoint - chosen ship 3

3.2 The NAVDEC Platform

Safe and effective execution of a transport task involves information exchange and coordination of activities between vessel navigators and land-based centre personnel. There are various consequences of lacking integration of decision support systems on ships and in land-based centres:

1) different scope of available information,
2) possible differences in the identification and assessment of a navigational situation,
3) restricted range of automatic ship-ship and ship-shore-ship information exchange,
4) decision support systems on board and ashore may generate different solutions to the same navigational situation.

All these disadvantages can be eliminated by building a uniform, joint platform of the integrated decision support system NAVDEC, incorporating vessels at sea and land-based centres of vessel traffic management. (Fig. 6).

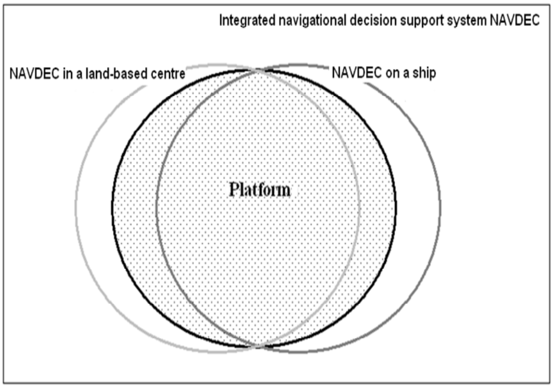


Fig. 6. Integrated navigational decision support system NAVDEC

The platform will include joint modules for shipboard and land-based systems plus tools for the integration of both systems. The joint modules will be, among others, the library of navigational procedures, interpretation of COLREGs, rules and criteria for navigational situation analysis and assessment used by navigators. The platform will also have specialized interfaces that will enable using services that are different for ships and land-based centres. One essential component of the platform should be a subsystem of automatic ship-ship and ship-shore-ship communication. The proposed integrated decision support system NAVDEC will enable automatic communication between users of decision support systems on ships and in VTS centres. It will also incorporate other traffic participants, even if automatic communication with them is not available.

The NAVDEC platform should be equipped with elements that do not exist in either shipboard or land-based decision support system, e.g. interfaces for external systems.

The following actions are required for the development of an integrated navigational decision support system NAVDEC:

- extension and adjustment of the existing NAVDEC modules for shipboard use,
- building of special NAVDEC modules for a land-based centre,
- development of special interfaces enabling using the services different for the ship and land-based centre,
- development of interfaces for external systems,
- development of automatic communication module, including:
  - rules of communication and cooperation between objects (vessels, land-based centres),
  - ontology of navigational information,
  - algorithms of information acquisition and interpretation,
  - algorithms for a dialog between objects,
  - negotiation principles and procedures,

- o methods of ontology representation to fit standards used in software engineering,
- o relevant knowledge bases,
- o operator-system communication model,
- o model of the scope and form of information presented to the operator.

Similarly to information systems of this type, the integrated navigational decision support system NAVDEC has to guarantee data storage and transmission security and varied range of personnel access.

References

1. Borkowski, P.: Algorithm of multi-sensor navigational data fusion – testing of estimation quality. Polish Journal of Environmental Studies 17(3B), 43–47 (2008)
2. COLREGs: Convention on the International Regulations for Preventing Collisions at Sea, International Maritime Organization (1972)
3. European Commission (EC) Communication:Strategic goals and recommendations for the EU's maritime transport policy until 2018, COM, 8 (2009), http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=COM:2009:0400:FIN:en:PDF
4. IMO, NAV 53/13. Development of an E-Navigation Strategy. Report of the Correspondence Group on e-navigation, submitted by the United Kingdom. Sub-Committee on Safety of Navigation, International Maritime Organization, London (April 20, 2007)
5. Magaj, J., Wołejsza, P.: Algorithm of working out anticollision manoeuvre by decision-supporting system. Polish Journal of Environmental Studies 17(4C), 14–18 (2008)
6. Pietrzykowski, Z.: Modeling of decision processes in sea-going ship movement control, Maritime University of Szczecin, Series Study No 43, Szczecin (2004)
7. Pietrzykowski, Z.: Ship's fuzzy domain – a criterion for navigational safety in narrow fairways. The Journal of Navigation (61), 501–514 (2008)
8. Pietrzykowski, Z., et al.: Navigational decision support system for sea-going ships. Maritime University of Szczecin, Szczecin (2011)
9. Pietrzykowski, Z.: Maritime Intelligent Transport Systems. In: Mikulski, J. (ed.) TST 2010. CCIS, vol. 104, pp. 455–462. Springer, Heidelberg (2010)
10. Pietrzykowski, Z., Borkowski, P., Wołejsza, P.: NAVDEC – navigational decision suport system on a sea-going vessel. In: 7th Explo-Ship International Conference, Świnoujście (2012)
11. Weintrit, A.: Telematic Approach to e-Navigation Architecture. In: Mikulski, J. (ed.) TST 2010. CCIS, vol. 104, pp. 1–10. Springer, Heidelberg (2010)

UML – A Part of an Interlocking System Development Process

Karol Rástočný[1] and Karol Rástočný, Jr.[2]

[1] University of Žilina, Faculty of Electrical Engineering
Univerzitná 1, 01026 Žilina, Slovakia
karol.rastocny@fel.uniza.sk
[2] Slovak University of Technology, Faculty Informatics and Information Technologies
Ilkovičova 3, 842 16 Bratislava 4, Slovakia
karol.rastocny@stuba.sk

Abstract. The analysis of functional requirements and their definition presents a very important task in the development of software of an interlocking system. If the UML-based modelling is used, then specific rules and conditions of use need to be defined with respect to the application field, already existing default assumptions, given goals, and target platform properties that will consequently implement the software. This paper deals with the use of a multiple inheritance during the software development with the help of IBM Rational Rhapsody development environment. The specific problems that could lead to ambiguity are explained, since they often lead to systematic failures, and recommendations on how to correctly use the multiple inheritance are proposed.

Keywords: UML, safety, control system, interlocking system.

1 Introduction

A railway interlocking is a special kind of a (safety-critical) control system, whose faulty operation can cause harm to people, significant damage to the equipment or environment, or other unwanted consequences. If the interlocking is expected to have the required safety integrity, then it must fulfil requirements of safety integrity against systematic, as well as random failures.

Systematic failures of an electronic interlocking are closely related to the software. A functional requirements specification is one of the important steps in the initial phase of the interlocking software development.

The customer's requirements are subjected to a thorough analysis, to create a model which allows us to verify the complexity and consistency of the functional specification and to remove potential conflicts in the specification [1]. The creation of the model must be preceded by the understanding of the customer's requirements and based on the proper method supported by an appropriate modelling tool.

A natural language or other informal forms of descriptions have more disadvantages and in general are inappropriate to describe a technical system. When such an informal notation is used, then the transformation of the specification into a software

J. Mikulski (Ed.): TST 2012, CCIS 329, pp. 293–300, 2012.

implementation is usually problematic and rather inaccurate. On the other hand, system models based on semi-formal and formal methods help us to create a comprehensive and unambiguous logical description of the functional properties of the system. The creation of such model is time-consuming, therefore the chosen methods and procedures should allow us to re-use once created model also in the later phases of the system life cycle (e.g. for the partial or even complete code generation).The object-oriented modelling can be successfully employed to achieve this goal. One of many suitable modelling languages that fall into this category is the UML (Unified Modelling Language). The UML is the most used and wide-spread standard of the OOM(Object Oriented Modelling)[2].

Several integrated development environments (IDE) that support a modelling language as well as a programming language (e.g. IBM Rational Rhapsody)are available on the market. Some of such IDEs do not only allow a direct code generation based on a designed model, but also make a dynamic analysis through a dynamic model animation possible. The dynamic analysis of the model lets us to automatically or semi-automatically check the correctness of the model functionality, while significantly decreases abstractness in comparison to the static analysis.

2 UML Employment in the Development of an Interlocking

The currently valid version UML 2.3 is characterised by a rich collection of building blocks (elements, relationships, diagrams, etc.) and rules on how to use these building blocks. However, if the software development process of the interlocking is based on the UML [6], then specific rules that reflect the application field (railway applications), existing default assumptions, given goals, and properties of the hardware and software components of the target platform need to be defined. These rules lead to a degradation of the current UML possibilities with an aim to create a model that is unambiguous, understandable, consistent and verifiable.

One set of rules is related to the management of the model creation (especially, when more persons are involved in the process). These rules serve mainly to ensure the model readability and clarity. Clearly defined rules contribute to the fact, that creators can focus their attention more on the logic underneath the model and less on the formal issues and means of description.

Another set of rules is important from the programmers' point of view – the generated code should be directly applicable in the target application without considerable modifications. The programmers must modify the designed model, so the generated code is executable on the target hardware platform, while the functional algorithms have to remain intact by the modification. Those rules include:

- coding standards (e.g. naming conventions);
- rules of a static structure creation(definition of classes, creation of relations between classes or objects etc.);
- means of message exchange between respective objects;
- ...

These rules are heavily dependent on the planned target software and hardware platform. This paper is focused on one major problem that needs to be solved with the help of mentioned rules – the problem of *multiple inheritance* usage in the process of interlocking system development using an integrated development environment *IBM Rational Rhapsody for C++*with the *UML* used to model a functional behaviour of the system. The problems that can lead to ambiguity and consequently to systematic software failure, are specified in the paper. In addition, recommendations are proposed, that – if properly used – can help analysts and developers to avoid problems related to the inheritance in the UML models.

3 Case Study

3.1 Multiple Inheritance in IBM Rational Rhapsody C++

The main source of equivocality in multiple inheritance is the existence of identically defined methods and attributes of super classes [3]. Other possible sources of equivocality are state machine diagrams (oractivity diagrams) that are defined in super classes, but the IDE IBM Rational Rhapsody does not allow this specific case to happen and when such structure is detected during the model checking, an error is generated: „Error: Multiple inheritance from reactive classes is not supported“.

Various combinations of identically named methods and attributes of the super classes are listed and analysed in the following part of the paper.

3.1.1 Basic Model

The Fig. 1 illustrates a model that consists of the classes *CBase*, *CParentA*, *CParentB*, and *CDaughter*. The class *CBase*is a base class and has a method *BaseOperation()*.

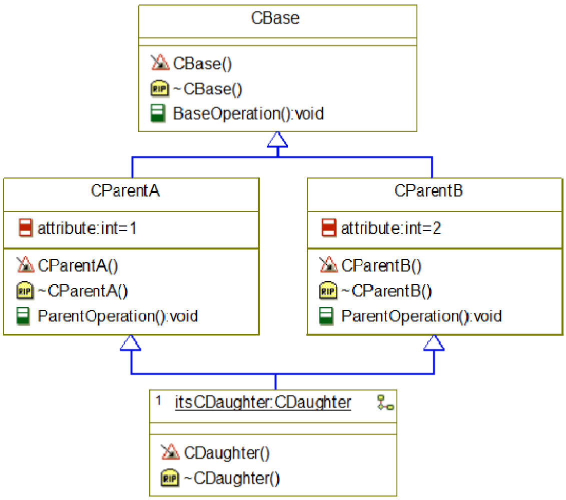


Fig. 1. Basic configuration of the case study

The classes *CParentA* and *CParentB* include an attribute *"attribute"* and a method *ParentOperation()*.

All mentioned methods write on the standard output a text that contains the name of the class in which they are defined, and their name, e.g. *CBase:BaseOperation*. The constructor writes the name of the constructed class and flag of creation, e.g. *CBasecreated*, and the destructor writes also the name of the destructed class and the flag of deletion, e.g. *CBasedeleted*.

The execution of the *CDaughter* class is clearly defined by the activity diagram pictured in Fig. 2. This diagram consists of three activities, in which the instance of the class *CDaughter* has a role of its super classes. This is necessary due to the identical naming of the methods and attributes in the super classes (the compiler needs to know, which method is being called).

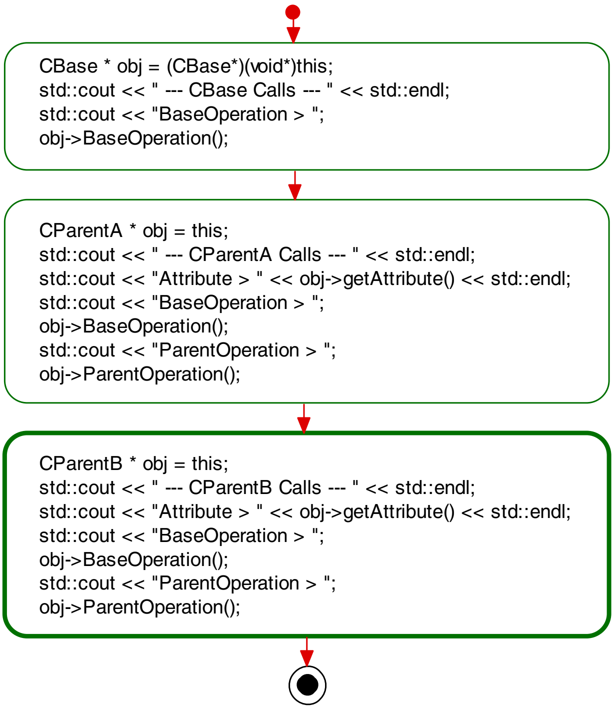


Fig. 2. Activity diagram of the class *CDaughter*

An example of the model output is shown in Fig. 3. It can be clearly seen that the constructor and destructor of the *CBase* class is called twice. The reason for this behaviour is the generalization relationship between class *CBase* and classes *CParentA* and *CParentB*. The sequence of construction and destruction of the instances of the classes *CParentA* and *CParentB* is given by the sequence, in which the generalization relationships have been drawn in the model. The sequence, in which the generalization relationships have been drawn, cannot be inferred from the diagram – that is the source of ambiguity. The output in Fig. 3 also tells us, that the instance of the class *CDaughter* has two identically named attributes *“attribute”* inherited from the classes *CParentA* and *CParentB*. Which attribute is used, and when, depends on the role of the class (another possible source of ambiguity).

CBasecreated CParentBcreated CBasecreated CParentAcreated CDaughtercreated --- CBaseCalls --- BaseOperation>CBase:BaseOperation --- CParentACalls --- Attribute> 1 BaseOperation>CBase:BaseOperation	ParentOperation>CParentA:ParentOperation --- CParentBCalls --- Attribute> 2 BaseOperation>CBase:BaseOperation ParentOperation>ParentB:ParentOperation CDaughterdeleted CParentAdeleted CBasedeleted CParentBdeleted CBasedeleted

Fig. 3. The output of the basic configuration

3.1.2 Simple Declaration of the Method BaseOperation()

If a method is simply defined, then the executed code of the method depends on the role the related object performs (if in the respective class the method is not defined, then the code of the super class that is closest in the inheritance hierarchy is executed). This behaviour of the instance of the class *CDaughter* is explained in Table 1.

Table 1. The behaviour of the instance of the class *CDaughter*in various roles when simply declared method *BaseOperation()* is called

The role	The method BaseOperation() is defined in				
	CBase	CBase CParentA	CBase CParentA CParentB	CBase CParentA CDaughter	CBase CParentA CParentBCDaughter
CBase	CBase	CBase	CBase	CBase	CBase
CParentA	CBase	CParentA	CParentA	CParentA	CParentA
CParentB	CBase	CBase	CParentB	CBase	CParentB
CDaughter	-	-	-	CDaughter	CDaughter

When the instance of the class *CDaughter* calls the method *BaseOperation()*, it cannot have a role of the class *CDaughter*, unless the method *BaseOperation()* is defined in the class *CDaughter*. This effect occurs due to indeterminacy of the method definition source in the class hierarchy.

3.1.3 Virtual Declaration of the Method BaseOperation()

If a virtual method is declared, then a record in the virtual method table is made for each overriding definition of this method. The definition from the youngest subclass in the class hierarchy is used anytime the method is called. If there is no subclass that would possess the definition of this method, then the definition from the nearest parent class is used (Table 2).

Table 2. The behaviour of the instance of the class *CDaughter* in various roles when virtually declared method *BaseOperation()* is called

The role	The method BaseOperation() is defined in				
	CBase	CBase CParentA	CBase CParentA CParentB	CBase CParentA CDaughter	CBase CParentA CParentBCDaughter
CBase	-	-	-	-	-
CParentA	CBase	CParentA	CParentA	CDaughter	CDaughter
CParentB	CBase	CBase	CParentB	CDaughter	CDaughter
CDaughter	-	-	-	CDaughter	CDaughter

Similarly, the instance of the class *CDaughter* cannot have the role of the class *CDaughter* when *BaseOperation()* method is called, unless the method *BaseOperation()* is defined in the class *CDaughter* (just as it was in the simple case of *BaseOperation()* declaration). In the case of the virtual method declaration, the role of the class *CBase* is also forbidden. However, this problem will manifest only during an execution of the application, when the method will be called.

3.1.4 Simple Declaration of the Method ParentOperation()

If the method *ParentOperation()* is simply declared, then the behaviour of the instance of the class *CDaughter* during the call of this method is similar to the simple declaration of the method *BaseOperation()* (Table 3).

Table 3. The behaviour of the instance of the class *CDaughter* in various roles when simply declared method *ParentOperation()* is called

The role	The methodParentOperation() is defined in			
	CParentA	CParentA CParentB	CParentA CDaughter	CParentA CParentBCDaug hter
CParentA	CParentA	CParentA	CParentA	CParentA
CParentB	-	CParentB	-	CParentB
CDaughter	CParentA	-	CDaughter	CDaughter

3.1.5 Virtual Declaration of the Method ParentOperation()

The behaviour of the instance of the class *CDaughter* when the virtually declared method *ParentOperation()* is called is similar to the behaviour of the virtually declared method *BaseOperation()* (Table 4). However, if both classes *CParentA* and

CParentB define the method *ParentOperation()*, but the class *CDaughter* does not, the behaviour is then different and quite expected – the code of the parent is executed, whose role has the instance of the *CDaughter*. One problem arises when the instance of the class *CDaughter* calls a destructor, since the memory heap in which the instance has been created is then damaged.

Table 4. The behaviour of the instance of the class *CDaughterin* different roles during the call of the virtually declared method *ParentOperation()* (*- damage of memory heap)

The role	The methodParentOperation() is defined in			
	CParentA	CParentA CParentB	CParentA CDaughter	CParentA CParentB CDaughter
CParentA	CParentA	CParentA*	CDaughter	CDaughter
CParentB	-	CParentB*	-	CDaughter
CDaughter	CParentA	-	CDaughter	CDaughter

4 Recommendations on Multiple Inheritance Modelling

Following conclusions and recommendations based on the aforementioned case study can be drawn:

- *All destructors shall be declared as virtual* (applicable generally when the inheritance is used). If non-virtual destructors are used, the call of all related destructors (after the instance of the class has been typecasted to one of the parents' type) cannot be ensured.
- *The conflicts of the names of attributes and methods shall be avoided.* This helps to avoid ambiguity caused by rather complicated rules of the object behaviour determination, that depends on the fact if the method is virtual and where is defined (column „*CParentA*, *CParentB*, *CDaughter*" in Table 3 and Table 4). Furthermore, diagrams provide no means of knowing, if the method is virtual or not.
- *The class shall inherit only from one not pure abstract class.* All other parental classes shall be interfaces (pure abstract classes). The problem of indeterminacy in the constructor and destructor call sequence can be avoided via this recommendation.
- *The "diamond" problem shall be avoided* [4], i.e. the case, in which two parental classes inherit from the same class (as it is in the case study model in Fig. 1). The conflict in the names of methods and attributes is inevitably caused.
- *If the behaviour of the subclass needs to be defined by a state machine diagram or an activity diagram, then at least empty diagram shall be defined in the class, which is at the top of the inheritance hierarchy and is not an interface.* So the *OMReactive* class (defined in the OXF framework which is used by IBM Rational Rhapsody) is included in the class hierarchy. From this class all classes with a defined state machine diagram or an activity diagram must inherit. Without this measure, the subclass would inherit directly from the class *OMReactive*.

5 Conclusion

In general, the multiple inheritance can be avoided in the modelling process. On the other hand, multiple inheritance makes the model more abstract and can significantly help to create and to implement reusable software parts and to improve the modularity of the system. That is the reason why the multiple inheritance has its purpose and should not be excluded from the model for no apparent reason – under the assumption that its use is well justified and the recommendations in chapter 4 of this paper are respected.

Such approach can eventually lower the cost and increase the speed of the development of a new software system (through the implementation of reusable elements that have already been tested and their correctness and safety has been validated) and simplify its maintenance (small modules of the correctly designed modular system are easy to implement, to test, and to maintain, so the costs of the software development can be further cut [5]).

Acknowledgments. This paper was supported by the scientific grant agency VEGA, grant No. VEGA-1/0388/12 "Quantitative safety integrity level evaluation of control systems in railway application".

References

1. EN 50 128: Railway applications – Communication, signalling and processing systems – Software for railway control and protection systems (2010)
2. Arlow, J., Neustandt, I.: UML 2 and the unified process: Practical object-oriented analysis and desing. Pearson Education, Inc. (2005)
3. Carré, B., Geib, J.M.: The point of view notion for multiple inheritance. In: European Conference on Object-Oriented Programming on Object-Oriented Programming Systems, Languages, and Applications (OOPSLA/ECOOP 1990), pp. 312–321. ACM Press, New York (1990)
4. Malayeri, D., Aldrich, J.: CZ: multiple inheritance without diamonds. In: 24th ACM SIGPLAN Conference on Object Oriented Programming Systems Languages and Applications (OOPSLA 2009), pp. 21–40. ACM, New York (2009)
5. Baldwin, C.Y., Clark, K.B.: Design Rules: The Power of Modularity. The MIT Press (2000)
6. Rástočný, K., Janota, A., Zahradník, J.: The Use of UML for Development of a Railway Interlocking System. In: Ehrig, H., Damm, W., Desel, J., Große-Rhode, M., Reif, W., Schnieder, E., Westkämper, E. (eds.) INT 2004. LNCS, vol. 3147, pp. 174–198. Springer, Heidelberg (2004)

Approaches to a Solution of Key Management System for Cryptography Communications within Railway Applications

Mária Franeková and Marek Výrostko

University of Žilina, Faculty of Electrical Engineering
Univerzitná 1, 010 26 Žilina, Slovakia
{maria.franekova,marek.vyrostko}@fel.uniza.sk

Abstract. Authors are dealing with the problem of the on-line key management system (KMS) solution, needed to ensure, using cryptographic tools, the authentication of communicated entities of stationary and mobile parts of ETCS (European Train Control System) level 2 (L2). Different approaches to this solution are described on the basis of a symmetric and an asymmetric cryptography. The main part of paper presents a proposal of KMS solution for ETCS L2 based on an asymmetric cryptographic algorithm ECC (Elliptic Curve Cryptography) oriented towards a solution of domain's parameters selection and the identification of particular entities in the train traffic. A practical generation of a pair of keys is carried out via a SW tool Cryptool.

Keywords: key management system, symmetric cryptography, asymmetric cryptography, Advanced Encryption Standard, Elliptic curve cryptography, ETCS, key generation.

1 Introduction

The interlocking systems incompatibility within the railway transport in Europe can be changed using the ETCS (European Train Control System), which is developed in Europe as part of international project ERTMS (European Rail Traffic Management System) [1]. The aim of ERTMS is to create a standardized European railway system, common for all countries of the EU (European Union), which allows transport of trains with the ETCS equipment on all European railway lines. The EU legislation deals with this system integration in the European countries. The Slovak republic starts with building the first level of ETCS, which is in a test operation. The Slovak railway strategy contains also building higher levels of ETCS in the future after an appropriate analysis [2].

According to the track-side equipment of ERTMS/ETCS we can differ three basic application levels L1, L2 and L3 [3]. The GSM-R system (Global System for Mobile Communications – Railway) is the necessary part of solution for ETCS L2 and L3, which executes the radio transmission between stationary and mobile parts of the ETCS system via a secure cryptographic protocol.

J. Mikulski (Ed.): TST 2012, CCIS 329, pp. 301–313, 2012.

The basic function of the stationary part of the ETCS system consists in obtaining the actual data from the relevant interlocking equipment, its processing and transmission to the mobile part of the ETCS system. The main task of the stationary part is to generate an authorization to drive a train and to transmit it with other required data to the mobile part, which is located on the locomotive.

The basic function of the mobile part of the ETCS system is to carry out the train driving monitoring and control on the basis of information exchange between stationary and mobile parts. The main task of the mobile part is to control the non-exceeding of boundary parameters (speed, passed distance) based on the received data from the stationary part of the system.

In ERTMS/ETCS L2 the OBU (on-board unit) in the train and the RBC (Radio Block Central) with connection of the interlocking equipment exchange the information among each other using an un-trusted open transmission system GSM-R (Global System Management for Railway). If an OBU wants to communicate with a given RBC it must be able to verify whether the communication with this RBC is valid or no and vice versa. This procedure is based on concrete cryptographic techniques, which are implemented in the safety layers of Euroradio protocol and use secret keys. It is necessary to underline that these procedures do not offer the tools for cryptographic keys generation, distribution and updating, so called KMS (Key Management System).

At present there are 54 kilometres of railway track in Slovak Republic, between Bratislava - Rača and Nové Mesto nad Váhom. The corridor modernization started already in the year 2000, but the installing and bringing to life of the technological part of ETCS started in the year 2008. ETCS L1 technology was supplied and installed by the Siemens company, especially because of compatibility with the already existing HW and SW components of electronic interlocking systems in the railway section. Railways of the Slovak Republic consider building a GSM-R network in a new project Bratislava - Nové Zámky, subsequently Bratislava - Žilina and the implementation ETCS L2 on the Kúty – Bratislava - Nové Zámky (Štúrovo) track. ETCS L2 is planned also for the Nové Mesto nad Váhom – Košice – Čierna nad Tisou section.

Six corridors were established in the European Union, which have to be equipped with the ERTMS/ETCS system. Corridor E passes through the Slovak Republic on the Kúty – Bratislava – Rusovce – Rajka – Nové Zámky – Komárno – Štúrovo section. Also ŽSR have to follow the decision of the European Commission, that every new build track, or modernized, has to be implemented with the ERTMS/ETCS system. That's why this technology was implemented in the Bratislava – Nové Mesto nad Váhom section and will be implemented along the track of corridor Bratislava – Žilina – Košice – Čierna nad Tisou.

Nowadays the process of KMS development is coordinated by the UNISIG, which is a syndicate of firms as Alstom, Ansaldo STS, AŽD Praha, Bombardier, CAF, Invensys, Memec group, Siemens and Thales. They participate in creating subsets – recommendations for ETCS project, in which KMS is also obtained. At present, subsets recommend solely off-line KMS, what brings errors caused by the human factor. All functions used in the KMS have to be executed in a secure way in accordance with the railway applications standards [3].

The KMS as COTS (commercial off-the-shelf) can be implemented using symmetric and asymmetric key techniques. Advantages from both of the cryptographic solutions are used, by using the public key technique for a safe exchange symmetric key, used for enciphering needed communication between subjects. In COTS technologies standards exist for the key management needs, e.g. [4, 5]. Companies, dealing with the KMS for ETCS systems development, use the existing standards to create the new ones, where the specifics valid for railway control and communication systems have to be added.

2 Functions of Key Management System for Railway Applications

In practice, where the cryptographic systems are used, the key management is considered the hardest and most complex part of cryptographic systems. The key management is a set of processes and mechanisms, which support the key exchanging and considering a permanent relationship between participants. In the case of railway applications, we will call them entities. Under the term of entity we will understand communicating parts in stationary and mobile parts of the ETCS system. The key represents a secret information known only to these entities. By using a computationally safe enciphering algorithm we can ensure the basic assumption, that a sensitive message (e.g. a controlling command) cannot be violated without knowing the key. From practice we know, that a lot of these systems may fail because of errors in the key management. In the key making phase the length of key is very important. This length increase in individual enciphering algorithms can increase the safety of safety related messages. In the case of safety relevant communication in the railway applications the issue is not to secure a large amount of data, like in COTS systems, but to secure short messages, as in the case of failure, there is a risk of damage and threat to health and life of passengers. The following minimal key lengths are recommended now for COTS technologies [6] (Table 1). Table 1 shows recommended minimal lengths of the key for symmetric algorithms (especially DES modification, like 2-DES or 3-DES and AES algorithm) and asymmetric algorithms like RSA (Rivest Shamir Adelman) and ECC for years 2012 to 2040.

In cryptography there are two basic types of key:

Session keys (for symmetric enciphering algorithms, like e.g. 3-DES, AES) have short lifetime (hours, days). They are used for direct data ciphering.

Long-time keys (for asymmetric enciphering algorithms, e.g. RSA) have longer lifetimes (months, years). They are used to protect other, i.e. session, keys or messages of short length.

Table 1. Minimum of recommended key length in bits

Type of algorithm	year 2012	year 2030	year 2040
Symmetric	80	93	101
Asymmetric (RSA)	1464	2493	3214
Asymmetric (ECC)	149	176	191

The method of ensuring that both communicating entities in a railway application are the ones they claim to be, is based on an Identification and Authentication dialogue. In order to ensure a complete protection, this procedure shall take place each time the peer entities effectively start a new communication session between them.

To allow the interoperability between the KMC domain and to perform a secure and consistent inter-KMC key exchange any KMS shall support the following functions:

Generation of the key – creating the key used by the enciphering algorithm or the key hash function MAC (Message Authentication Code), according to the used model of transmission messages for opened transmission systems B0 or B1 (according to EN 50159 [7]). The choice of the enciphering algorithm or hash function is one of the first steps specifying the range of key space. Incorrectly designed key generator can radically reduce the key space and simplify an attack oriented at the key detection (brute force attack).

Storage of the key – this function allows putting away generated keys to the storage before activating them. After acquiring request for the key, the system does not have to wait for the generator, but it will take the key, which has been generated earlier. If longer keys are generated, with the key storage, time can be saved which leads to an increased availability of the system.

Validation of the key – process in which it becomes clear, if the generated key is valid and appropriate for using with the given algorithm. The key validation means, that ciphers like DES, 3-DES have weak and semi-weak keys. The keys validation separates unusable and potentially risky keys.

Distribution of the key – process, in which a secret key became available to two or more entities (e.g. an OBU in a vehicle and a relevant radio block centre - RBC). It has to be guaranteed that the generated and subsequently validated keys will be distributed safely.

Cancelation (deleting) of the key – it is the key deletion process. This request comes from the fact, that all keys are updated in particular time intervals. An attacker must not have an opportunity to get even an inoperative key. This key could read old and used messages. For this reason an old key and all related material (old messages and so) has to be deleted from all subjects owning or using this key.

Updating of the key – the key has a limited validity, to increase the safety of the systems keys have to be updated. It is a process, which includes functions of generating, validating, distributing and deleting the key. This procedure is performed for example when a key was disrupted or lost and when old keys are replaced with new ones.

3 Current state of Key Management System Development

The KMS methodology in the ETCS pilot project still develops and it is directed at the evolution from off-line to on-line systems [8]. Subset [8] is based on the symmetric cryptography and defines the following key hierarchy:

Level 3 – transport keys KTRANS: these keys are used to protect the control communication between the key management centre KMAC and an ERTMS entity RBC (radio block centre), or a track vehicle (OBU) to obtain or renew the authentication keys. Keys used to protect the connection between KMC are known as K-KMC.

Level 2 – authenticate keys KMAC: these keys are used to specify the session keys and to authenticate ETCS entities.

Level 1 – session keys KSMAC: Keys are used to protect the data during transmission between safe entities.

Table 2 presents the keys usage depending on the communicating entities.

Table 2. Keys usage depending on the communicating entities in ETCS

Involved entities	**Key usage**		
	Identification & Authentication	Message Authentication	Enciphering
RBC - OBU	KMAC	KSMAC	-
KMC - RBC/OBU	-	-	-
KMC - KMC	-	K-KMC1	K-KMC2

Since keys are secret, every ERTMS entity (KMC, RBC, OBU) have to include a set of valid authentication keys for domains, in which an OBU is authorized to ride. Subset [8] distinguishes two approaches. The first approach assigns a unique KMAC to every OBU, which will be valid in corresponding domains. The second approach assigns different KMAC keys to OBUs in different domains. Under the name of KMC domain we understand a list of trackside (RBC) and OBU entities, the definition of rules of keys assignment, the identification other domains, where home OBUs are allowed to ride and domains, where OBUs are accepted by that domain.

To ensure safety of KMAC key transactions between domains, K-KMC keys are used. K-KMC keys should be distributed between domains before launching any transactions. They consists of two parts. The first part of key, K-KMAC1, is used to ensure the integrity and authentication of transaction using the CBC-MAC algorithm. The second part of the key, K-KMAC2, is used for ensuring the data confidentiality during the transaction, which means the enciphering of KMAC key.

Subset 038 [8] recommends the use of 3-DES enciphering algorithm. There is an assumption that 3-DES will be satisfactory, from the view of attack length, during the whole ERTMS lifetime. This assumption is correct, but the AES algorithm provides several advantages, resulting in an increased security against attacks.

Authors of [9] are comparing the performance of symmetric ciphers DES, 3-DES, AES and Blowfish. Tests have shown, that in such applications the AES reaches better results than DES or 3-DES. Tests have also shown that the throughput of

3-DES cipher is third compared with the DES algorithm. It is caused by the character of ciphering using this cipher, which means a triple amount of ciphering input.

In [10] and [11] the authors compare besides speed also the effectiveness in relation to the battery life of the equipment, where it is used. In every measurement the AES algorithm had better results than 3-DES.

Paper [12] evaluates DES, 3-DES and AES algorithms in GSM applications. Authors tested all algorithms with brute force attack, where time was measured, which have to be used to breach the algorithm. Tests run for different lengths of key have shown, that the AES algorithm needs the longest time to be breached. With the increasing length of the key, the time needed for a brute force attack on the AES algorithm exponentially raised. It was shown, that this algorithm is many times more resistant than 3-DES, or even DES.

The DES algorithm was designed for the hardware, not software performance. On the contrary, the AES algorithm was constructed for a high performance in both directions. That is why 3-DES algorithm could be better to use, but only if the application has properly adapted the hardware for the cipher use. According to NIST both of the ciphers will coexists until the year 2030, which allows a continuous transition from 3-DES to AES algorithm. The NIST institute consider the AES as a successor of 3-DES algorithm.

Based on these reasons, the use of AES enciphering algorithm will lead to an increased effectiveness, speed and safety at telegrams enciphering and deciphering.

4 Proposal of KMS System Solutions for ETCS L2 Using the Asymmetric Cryptography based on the ECC

Nowadays the RSA is the most used algorithm in cryptographic applications based on the asymmetric cryptography. Although the RSA algorithm is today computationally safe, according to [13] the development of computational possibilities of modern information technology results in growing the length of cryptographic key to keep the required safety integrity. According to GISA (German Information Security Agency) the today's 1024-bits RSA key standard is insufficient and since 2006 a 2048-bit key is recommended. The fact, that the majority of present chips are not able to process the keys longer than 1024 bits, shows the need for alternative solutions. Algorithm ECC (Elliptic Curve Algorithm) is such an alternative in the area of asymmetric cryptography.

The main problem of ECC implementation is now the absence of standardization. Only one possibility of RSA implementation exists but many chances for the implementation of ECC algorithm. Every possibility has an advantage and also a disadvantage, and one of them can be the most effective for certain applications. But this situation creates problems with the interoperability. To enable the communication between all tools, they must support all ECC algorithms. Even when the safety of RSA is sufficient, within a new system solution the designer should consider using also the ECC tools, especially if the system is designed for a long time period.

The efficiency of ciphering algorithms depends on the length of key and necessary computational demand to keep the required safety integrity level. The general advantage of ECC in comparison with RSA is that it requires a much shorter length of key. When we assume that the growth of microcomputers computational ability according to Morrow is low (computing capability is doubled every 18 months), the development of the key length to ensure safety is more effective with ECC than RSA. Table 3 shows the lengths of comparable RSA and ECC key sizes according to [14].

Table 3. Comparable key sizes in bits of RSA and ECC

ECC	RSA
163	1024
233	2240
283	3072
409	7680
571	15360

In the case of proposal of solutions KMS for ETCS L2, it will be needed first to choose the location of KMS (to create the KMC - Key Management Centre), which must be accredited by the certification authority. In the case of Slovak Republic the NBU (National Security Body) is the root certification authority.

The KMC centre should be considered within one domain of ETCS, which would be of a regional nature. Then each country can be considered an independent domain. In this ETCS domain, the control of ETCS must be centred on the highest level. It serves as the central control place, which is fulfilling the function of a dispatcher control, from which it sends commands to hierarchically lower elements. This centre should be established in a building with limited access, for authorized persons only. Then the expected physical and electronic security of the space in the building should be ensured, enabling to place e. g. the key management centrum issuing and handling the keys.

On the basis of aspects described above, the most preferable solution of KMS location is shown in the main control centre of traffic, MCCT (see Fig. 1). The authors see the most important advantage of this solution in the fact, that a secure communication between KMS and MCCT does not need to be resolved, the safety of KMS depends on the personnel only.

For the KMS system solution the authors choose an asymmetric cryptography system based on the very modern principle of elliptic curves ECC (a mathematical description is presented e. g. in papers [13, 14]).

Description of the elements in Figure 1:

NBU (National Safety Body) – introduces the root certification and cryptography authority in the Slovak Republic.

RBC (Radio Block Centre) – introduces the track module of ETCS, which sends messages received from MCCT to trains operated within its own geographical area.

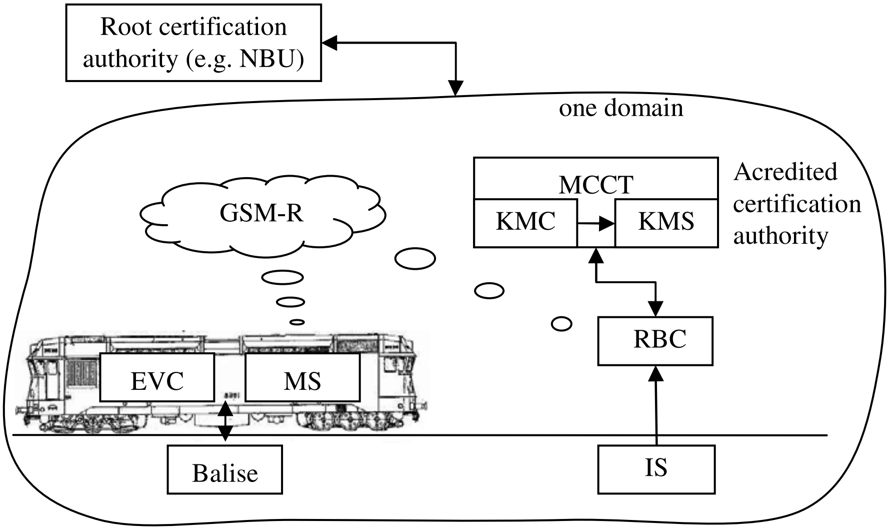


Fig. 1. Solution of KMS architectures for ETCS L2

MCCT (Main Control Centre of Traffic) – its task is to send the information about train traffic via RBC.

EVC (European Vital Computer) – technological PC fulfilling important on board signalization functions.

MS (Mobile Station) – communicates via wireless GSM-R network.

GSM-R – is a wireless network licensed for railways, providing the tracks for communicated trains. It supports the voice and data communication within ETCS.

ER (Euroradio) – is a safety protocol which supports the safety functions and controls the connection of GSM-R equipment

KMS (Key Management System) – module which allows manipulation with cryptography keys from the generation to disturbing of keys.

IS (Interlocking equipment) – sends messages required for train operation.

Balise - introduces a passive equipment within ETCS which provides information about the train position for EVC.

Each KMS domains should contain the following subjects:

- RBC station – serves the communication between the train and MCCT (or by KMS),
- end - RBC station – serves as an RBC station but also as the interface between neighbouring domains,
- train (on board unit) – locomotive,
- MCCT – the main task is to process the driving tracks of particular trains within one domain,
- KMS – executes all operations with cryptography keys (generation, distribution, storage, change, deleting and updating),

- KMS domain manager – a person who controls one domain. Within the solution it is necessary to distinguish three types of domains:
- home domain – is only one for one administrator, other domains appear for the administrator as neighbour or remote.
- neighbour domain – is the domain which has common boundaries with the home domain.
- remote domain – is the domain witch with the home domain dos not have common boundaries.

4.1 Solution of the Process of Particular Entities Identification

Each entity which communicates within the ETCS system must be uniquely identified through identity number ID. This number must be unique within all system domains. That is why each MCCT in a particular domain must contain the database of entities ID for all domains. It is a necessary measure again the ID duplication (e.g. trains). Particular ID numbers for entities of stationary and mobile parts of ETCS will be installed during the execution of KMC centre domain. For this reason we will distinguish the following ID numbers:

- ID_T uniquely identify a train (on-board unit),
- ID_{RBC} uniquely identify an RBC station or end-RBC station,
- ID_{KMS} uniquely identify the KMS,
- ID_{DOM} uniquely identify the KMS domain.

4.2 Solution of Train Tracks Planning

Let us assume that RBC and train know their unique identity numbers (ID_{RBC} and ID_T), according to which they will receive particular messages. Each MCCT centre plans particular train tracks and then informs via messages others MCCTs which participate in these train tracks. This is the case, when the track planning goes across more than one domain. If the train track planning does not interfere with another domain, then interfere MCCT sends to a particular RBC station in its own domain the identity numbers of trains (ID_T), which they store in their database. According to ID_T the RBC stations know, which trains have the authorization to run across a particular area. Then these RBC stations inform particular MCCTs about successful sending of a message by forwarding it. Each train track planned must contain the following identity numbers:

- ID_T which train has been planned for the given track,
- ID_{RBC} sequence particular RBC station with which the train will communicate during its driving.

4.3 Example of Domain's Parameter Option

Let us assume that in all the KMS domain only one setting of elliptic curve parameters will exist. These parameters will be generated in the process of KMC establishment. For the solution of ECC parameters the authors have chosen the Galois field GF(2^m), m = 409 elements long. This type of GF is recommended to keep an optimal safety of 192 bits, using a symmetric AES cipher (Advanced Encryption Standard). We choose a binary elliptic curve and then we generate domain parameters. These parameters are published. Domain's parameter for a binary random curve:

B-409: $m = 409, f(z) = z^{409} + z^{86} + 1, a = 1, h = 2$

S = 0x 4099B5A4 57F9D69F 79213D09 4C4BCD4D 4262210B
b = 0x 0021A5C2 C8EE9FEB 5C4B9A75 3B7B476B 7FD6422E F1F3DD67 4761FA99 D6AC27C8 A9A197B2 72822F6C D57A55AA 4F50AE31 7B13545F
n = 0x 01000000 00000000 00000000 00000000 00000000 00000000 000001E2 AAD6A612 F33307BE 5FA47C3C 9E052F83 8164CD37 D9A21173
x = 0x 015D4860 D088DDB3 496B0C60 64756260 441CDE4A F1771D4D BO1FFE5B 34E59703 DC255A86 8A118051 5603AEAB 60794E54 BB7996A7
y = 0x 0061B1CF AB6BE5F3 2BBFA783 24ED106A 7636B9C5 A7BD198D 0158AA4F 5488D08F 38514F1F DF4B4F40 D2181B36 81C364BA 0273C706

Domain's parameter for a Koblitz curve:

K-409: $m = 409, f(z) = z^{409} + z^{86} + 1, a = 0, b = 1, h = 4$

n = 0x 007FFFFF FFFFFFFF FFFFFFFF FFFFFFFF FFFFFFFF FFFFFFFF FFFFFE5F 83B2D4EA 20400EC4 557D5ED3 E3E7CA5B 4B5C83B8 E01E5FCF
x = 0x 0060F05F 658F49C1 AD3AB189 0F718421 0EFD0987 E307C84C 27ACCFB8 F9F67CC2 C460189E B5AAAA62 EE222EB1 B35540CF E9023746
y = 0x 01E36905 0B7C4E42 ACBA1DAC BF04299C 3460782F 918EA427 E6325165 E9EA10E3 DA5F6C42 E9C55215 AA9CA27A 5863BC48 D8E0286B

Definition of symbols:

m	integral number (or prime number) of the defined finite field GF(2^m),
$f(z)$	irreducible polynomial of degree m defined field GF(2^m) representation
S	160 bits input parameters for hash function SHA-1,
a, b	two elements of the field, which specified the equation of elliptic curve E over GF(2^m),
n	represents the degree of curve (number of points on the curve),
h	co-factor of curve (the biggest prime number in factorization of point number on the curve or its multiple,
x, y	axis of base point on curve $G = (x_G, y_G)$.

Note: The authors choose the domain's parameter for elliptic curve according to recommendation FIPS 186-3 [15] .

4.4 Key Hierarchy Solution

In comparison with the KMS solution on the basis of the symmetric cryptography (described in chapter 3), where in our solution to particular entities we assigned one secret key based on the asymmetric cryptography; within the KMS key pairs must be generated and assigned to all participating entities. Every key pairs will contain a public (Pub) and a private (Pri) key. The message is ciphered using the public key and is deciphered using the private key. Table 4 illustrates the key architecture with the description of key use in the ETCS system.

Table 4. Key hierarchy solution for ETCS L2 using ECC algorithm

Key (Length of key)	**Using**
PubK-AKT (416 bits)	Public key used for message ciphering between two MCCT
PriK-AKT (416 bits)	Private key used for message deciphering between two MCCT
K-TRANS (192 bits)	Private symmetric key used for messages ciphering/deciphering between two entities (the key is created during the exchange of master key)
PubK-MAC (416 bits)	Public key used for message ciphering between – train and RBC station, – MCCT and RBC station RBC (within one domain) – end-RBC station and KMS
PriK-MAC (416 bits)	Private key used for message ciphering between – train and RBC station, – MCCT and RBC station RBC (within one domain) – end-RBC station and KMS

We can use the K-TRANS key during a non-standard communication (not related to the ETCS) between two entities. This key is created during the exchange of keys, when two entities determinate the secret symmetric key K-TRANS, which is then shorted (in an agreed manner) to the required length (e.g. 192 bits for the AES algorithm). The K-TRANS key may be created between any two entities.

Note: the Euroradio protocol must support also the communication through the symmetric cryptography, so that the K-TRANS key would be applied, too.

5 Conclusion

For decades the cryptographic techniques are common safety aids in the area of so-called COTS technologies, e.g. in the financial sector (banking, e-commerce), in the office and in corporate information and communication systems and networks, but their usage is recommended for the safety-related applications only in recent few years.

In railway applications within the ETCS system the L2 requirements for a key management system are still in the development phase. There is a trend from the existing off-line key management system to create an on-line KMS which eliminate human factor mistakes.

Authors summarized in the paper the approaches within the development of key management system presented by the leaders of UNISIG association.

Authors describe the solution on the basis of the asymmetric cryptography in the concrete case of the ECC algorithm use. This way reduces the key length, thereby accelerating and simplifying the transmission of key and this solution is today very advanced and presented as computationally safe. The paper specifies solutions of the process of particular entities identification, train tracks planning and the design of key hierarchy. An example of domain's parameter solution based on the ECC was presented for a key pair generated over the Galois field $GF(2^m)$ for $m = 409$. It is equivalent to a 192-bit symmetric AES algorithm key.

Acknowledgement. This work has been supported by the Educational Grant Agency of the Slovak Republic (KEGA) Number: 024ŽU-4/2012: Modernization of technology and education methods orientated to area of cryptography for safety critical applications.

References

1. Architecture of ERTMS, `http://www.ertms.com/architecture.html` (date of access 2010)
2. Zahradník, J., et al: Application of ETCS and GSM-R in conditions of Slovak Railways. Contract. Customer: Railways of Slovak Republic (2006)
3. Dohnalík, B.: Purposes of Railway of Slovak Republic in area of ERTMS/ETCS applications. In: International Conference Applications of ERTMS/ETCS in Czech Republic, Pardubice (2004)
4. FIPS PUB 171, Key management using ANSI X9.17, National Institute of Standards and Technology (1992)
5. ANSI X9.17 (Revised), American National Standard for Financial Institution Key Management (Wholesale), American Bankers Association (1995)
6. Lenstra, A., Verheul, E.: Selecting Cryptographic Key Sizes in Commercial Applications. PricewaterhouseCoopers Cryptographic Centre of Excellence (CCE) Quarterly Journal (1999)
7. EN 50159: Railway applications. Communication, signalling and processing systems Safety-related communication in transmission systems. CENELEC (2010)
8. UNISIG Subset 038 v2.1.9. Off line key management FIS (2005)
9. Nadeem, A., Col: A Performance Comparison of Data Encryption Algorithms. IEEE (2005)
10. Abdul Kader, H.M., et al.: Performance Evaluation of Symmetric Encryption Algorithms. IJCSNS International Journal of Computer Science and Network Security 8(12) (December 2008)

11. Hirani, S.: Energy Consumption of Encryption Schemes in Wireless Devices. University of Pittsburgh (2003)
12. Sachin, M., Kumar, D.: Implementation and Analysis of AES, DES and Triple DES on GSM Network. IJCSNS International Journal of Computer Science and Network Security 10(1) (2010)
13. Enge, A.: Elliptic Curves and Their Applications to Cryptography – An Introduction. Kluwer Academic Publisher, Boston (2001) ISBN 0-7923-8589-6
14. Cohen, H., Frey, G.: Handbook of Elliptic and Hyperelliptic Curve Cryptography. Chapman & Hall/CRC (2005) ISBN: 1-5848-8518-1
15. FIPS 186-3: Digital Signature Standard (DSS), National Institute of Standards and Technology (2009)

Changes in System of Civil Aviation Protection

Piotr Uchroński

Katowice International Airport,
42-625 Ożarowice, Wolności 90
puchronski@gtl.com.pl

Abstract. For several years, according to the doctrine that the ruling authorities are responsible for ensuring the security of the country, security checks at airports for international traffic was held the Border Guard, you can tell by the Airport Management, which includes the formation of protection - Inner Protection Service, which has the power the security checks and baggage in domestic traffic. It is the Border Guards protect our border checkpoints at airports, carrying out tasks such as the control of persons and luggage, examined formal permission to cross the border. Not without reason, but it uses the past tense when describing above. SG tasks. Well, 18.09.2011r. came into force the Aviation Law, which provides for the resignation of SG on the implementation of activities related to the control of security in favor of the Executive, which is in fact subordinate to him Airport Security Guard, leaving behind only here the powers of supervision.

Keywords: We would like to encourage you to list your keywords in this section.

1 Introduction

Amendments to the Aviation Act, on which the legislator worked for several years, have unfortunately entered the implementation stage at airports. The term "unfortunately" is intentionally used, as the amended Act imposes a number of obligations on Airport Managers that substantially affect their financial liquidity.

This paper will focus exclusively on the change involving the transfer of security control duties in international traffic, previously assigned to the Border Guard, and the consequences of such for Airport Managers. Depending on a size of an airport, the number of aircraft operations (including the number actually performed), airport infrastructure and various other factors, Airport Managers will be required to hire qualified staff (dozens or even hundreds of new employees) responsible for the security checks of passengers and baggage.

1.1 Changes in Civil Aviation Protection

The amended Aviation Act allows for choosing between several organisational models for airport protection. Setting aside economic and social factors, an Airport

J. Mikulski (Ed.): TST 2012, CCIS 329, pp. 314–318, 2012.

Manager may hire staff under various contracts allowed by the Labour Code, sub-contract security control services to an external company that has adequate staffing, or adopt the "mixed model", which assumes cooperation of a manager with an external entity.

When choosing one of these options, one must analyse thoroughly the consequences that a given security protection model may have on the broader security system of the whole airport.

No one may harbour illusions that a change where dozens of new employees are suddenly (18 months being a short period in such instances) brought to work in passenger and baggage control will, at least initially, have no adverse impact on civil aviation protection.

The Manager's role will be to organize the protection system in cooperation with the Border Guard, acting here as a supervisory body.

An indisputable downside and major problem facing civil airports (regardless of the model adopted) is the absolute shortage, or, to put it mildly, insufficient number of workers who are properly qualified to perform security checks.

One should remember that security checks may only be performed by a person who, in addition to a physical protection license, holds the necessary Security Control Operator certificate, issued by the Chairman of the Civil Aviation Office following a number of training courses and examinations.

When analysing specific solutions that a Manager may choose from, it is first crucial to point to the areas that were previously assigned to the Border Guard, now due to be replaced by an Airport Manager.

Existing legal regulations drew a clear line between the competences of the Border Guard (BG) and the Airport Security Guard (ASG), assigning international traffic control to the BG and domestic traffic and staff checkpoints to the ASG.

The international traffic that this paper focuses on includes the following security checkpoints:

- control of security of passengers and hand baggage;
- registered baggage control;
- cargo control.

The number of security control employees (multiplied by the number of scheduled shifts and the required amount of staff per checkpoint) will vary, due to the number of security checkpoints at any given airport.

An Airport Manager hiring security staff internally will increase employment in this group by 100% to 200% on average, which means that an internal security service will be a substantial financial burden for the Manager.

An indisputable advantage of this solution, because of the subordination hierarchy it creates, is the possibility to respond directly to any irregularities detected in security staff work.

This model represents an internal security service reporting directly to the Manager. Thus, the structure of this entity affects the entire airport's protection system.

Despite the obvious advantages for civil aviation protection, such a solution also has a downside, which, whilst not directly affecting airport security, is possibly the principal factor in deciding which the security model should be adopted by an airport.

This may be down to the combination of an economic and a human factor - in the form of workers' organisations allowed by law, commonly known as Trade Unions.

Objectively, and without starting a debate about the activities of the Trade Unions, it should be stated that Airport Managers may see in this scenario the risk of work blockades by security staff who are members of Trade Unions. This point, especially in the light of information provided by both national and foreign media reporting disruptions in the operation of airports, is well-founded and worthy of consideration.

One may claim that the abovementioned security model is only suitable for airports with relatively small international traffic, where a percentage of newly hired security personnel would be acceptable for Managers in both economic and social terms.

Another option allowed by the Aviation Act is the transfer of security control tasks to an external company who provide this type of service. Such an entity will obviously be unable to provide civil aviation protection services without qualified and properly trained personnel. The legislator has decided to meet such organisations halfway, as they are required to guarantee only 10% of the total number of security control operators when soliciting an order. This percentage is unfortunately insufficient to guarantee that the security of passengers and baggage will not be compromised, especially as an airport will need many authorised workers to assume the previous duties of the Border Guard. The small numbers of national centres that provide training for security control operators demand that companies have their own training centres, as well as their own trainers for potential employees.

The argument that speaks in favour of this solution is primarily the European model; in the majority of EU countries, external companies have long provided high quality security services to airports. It should be emphasised that such practices have been in place for over a decade. This would be a revolutionary change to civil aviation protection in Poland.

An external organisation would seem to be the most attractive solution for Airport Managers, both economically and socially. In brief, a contract outsources “security control” to an external entity that provides such, in return for a pre-agreed fee. In this scenario, the airport is not faced with any staff, social or workspace problems.

A well-formulated contract clearly regulates the scope of duties, responsibilities, competences, and even the consequences of defective work. Clearly, it is a convenient solution for Airport Managers’, although it is not fail-safe, e.g. frequent staff rotation has an adverse impact on any airport security system. One may claim that such a model will prove efficient several years hence, when airports and surveillance companies will be appropriately prepared (formally as well as practically) for civil aviation security tasks.

The last variant that may be adopted for security is the “mixed control” model, performed by the Airport Security Guard in cooperation with a surveillance company. Joint control may vary, from strict separation of the ASG’s responsibilities from an external entity in a given area, to tasks performed jointly by the ASG and the external company.

An unquestionable advantage of such a model is its flexibility, for example when a quick response is required in case of staff shortages. Such shortages may be remedied by using ASG staff or, if necessary, by using the resources of external companies, which will ensure smooth control operations at an airport.

When choosing the solution, one should also consider the option whereby a special purpose company is set up directly by an airport, and therefore civil aviation security tasks are fully outsourced whilst the company's policy may be tailored to meet expectations. It should be noted that this option is subject to the restrictions of law, because said entities are required to operate on the market for at least five years.

As with the previous model, this solution requires the Airport Manager and the external entity to put into place a supervision system that will include properly qualified staff with sound theoretical and practical training.

One should remember that responsibility for an airport's protection will however lie with Airport Managers, and that the law does not allow for delegating this responsibility to an external entity performing control services.

Additionally, the amended Act defines in detail the scope of the cooperation of the Border Guard and Civil Aviation Office Chairman in overseeing security controls:

The Border Guard, when performing its oversight duties, is empowered to:

1) monitor and record the performance of a security checkpoint;
2) control the number of security service employees working at a security checkpoint and immediately notify an Airport Manager of any doubts as to their psychological and physical condition;
3) immediately respond to any breaches of civil aviation protection law by airport security service personnel;
4) immediately request that an Airport Manager remedy any serious irregularities detected and notify the Office Chairman of such irregularities;
5) control the certification of airport security staff, attesting to their qualifications;
6) immediately respond to any reports of public order violations at a security checkpoint and in adjacent areas.

Among the three models presented, the last variant, requiring the cooperation of the ASG and an external entity, would seem to be the optimal solution as it reconciles formal and organizational issues.

The security control tasks currently performed by the Border Guard, shortly to be replaced by an Airport Manager, require the intensive efforts of Airport Managers and external entities, who, in order to fill a gap in the market, will be forced to rapidly adapt their organizational structures to the requirements imposed by law.

The problems raised by the transfer of security control to an Airport Manager are both difficult and controversial. Arguments, both for and against each solution, clash against each other. We all became used to the current model, and the protection ensured by the state service makes us feel safer.

There are countries where the so-called "SUFO" successfully perform the security control tasks; in others, the state is responsible for ensuring the safety of passengers. No ideal solution exists, and the economic motives for amending the Act fail to convince me. The whole problem of removing the Border Guard from security checkpoints is narrowed down to the “budget spending cuts”, whereas the burden of ensuring security is shifted towards society as a passenger will now have to pay to feel safe when using air transport services.

References

1. Aviation Act dated 3 July 2002 (Journal of Laws no 130, item 1112)
2. Commission Regulation (EU) No 185/2010 of 4 March 2010 laying down detailed measures for the implementation of common basic standards on aviation security
3. Regulation (EC) No 300/2008 of the European Parliament and the Council of 11 March 2008. on common rules in the field of civil aviation security
4. Council of Ministers of 19 June 2007 on the National Civil Aviation Security Programme to the principles of aviation security (Journal of Laws No. 116, item. 803, as Amended, D.)

Development of the Coastal and Global Ships Traffic Monitoring Systems

Ryszard Wawruch

Gdynia Maritime University
Morska 81-87, 81-225 Gdynia, Poland
wawruch@am.gdynia.pl

Abstract. Paper describes reasons of the introducing of different coastal and global maritime ships traffic monitoring systems. It presents operational possibilities and limitations of these systems and perspectives of their further development. Global systems are divided into Long Range Identification and Tracking System (LRIT) and satellite monitoring systems of the conventional and non-conventional vessels utilizing class A and B ship borne transponders of the Automatic Identification Systems (AIS) working in VHF band.

Keywords: ships traffic monitoring systems, long range identification and tracking, coastal and satellite AIS monitoring systems.

1 Introduction

Basic functions of the maritime ships traffic monitoring system are collecting, storage and dissemination of following information about ships and passengers and cargo carried at sea received mainly from two sources: shore-based stations of the Automatic Identification System (AIS) working in the VHF band and Long Range Identification and Tracking System (LRIT):

- Ships identification, actual positions and parameters of movement;
- Ships destination and planned route;
- Information about passengers on board sea going vessels;
- Detailed data about marine pollutants and dangerous, harmful and noxious substances carried at sea;
- Information about incidents with ships and cargo at sea and in the ports and about violence of regulations concerning: sea safety, protection of environment and ships traffic and security.

Legal aspects relating to introducing and operation of ships traffic monitoring system are defined in:

1. International Convention on Safety of Life at Sea, SOLAS 74, as amended.
2. Regional agreements, as:
   - Declaration on the Safety of Navigation and Emergency Capacity in the Baltic Sea Area (HELCOM Copenhagen Declaration) adopted on 10 September 2001 in Copenhagen by the HELCOM Extraordinary Ministerial Meeting;

J. Mikulski (Ed.): TST 2012, CCIS 329, pp. 319–327, 2012.

- Directive 2002/59/EC of the European Parliament and of the Council of 27 June 2002 establishing a Community vessel traffic monitoring and information system and repealing Council Directive 93/75/EEC; and
- Directive 2009/17/EC of the European Parliament and of the Council of 23 April 2009 amending Directive 2002/59/EC.

3. National regulations in particular countries.

Operational and technical questions connected with AIS and LRIT are described in appropriate recommendations, resolutions and standards issued by International Maritime Organisation (IMO), International Electrotechnical Commission (IEC) and International Association of Marine Aids to Navigation and Lighthouse Authorities (IALA).

2 Monitoring Systems Utilizing Shore Based Stations of the Automatic Identification System

Automatic Identification System (AIS) was introduced principally due to the:

1. Limitations of navigational radar detection at sea, its sensitivity on hydro meteorological disturbances and often insufficient accuracy and time delay of data presented by radar tracking aids mainly during manoeuvres of tracked objects.
2. Problems correlating the identity of a ship detected by radar.

Before AIS introducing, personnel of the ships monitoring systems utilized as a source of information about ships in monitored area visual observation, radars, radar tracking aids and VHF voice communication.

AIS is an autonomous and continuous broadcast system operating in two dedicated channels in the VHF maritime mobile band: 87B (161,975 MHz) and 88B (162,025 MHz). It handles multiple reports at rapid update rates in autonomous and autonomous and continuous or pulling modes. It may use both 25 kHz and 12.5 kHz simplex channel bandwidths. One 25 kHz simplex channel enables to transmit with the rate of 9600 bits per second about 2000 standard position reports per minute. 12.5 kHz channel enables transmission with the rate of 4800 bits per second about 1000 standard position reports per minute. Simultaneous application of two designated VHF channels increases system capacity twofold. Additionally, ship borne AIS equipment has, not used in practice, interface to connect it to the INMARSAT-C satellite communication terminal. According to the IMO recommendation AIS should be capable of operating:

- In the ship-to-ship mode, to assist in collision avoidance;
- As a mean for littoral state to obtain information about ships and their cargo; and
- As a vessel traffic service (VTS) tool to traffic control and management.

All reports are transmitted and received automatically. In the autonomous mode they are transmitted with frequency depending on the vessel speed and stability of its movement. Minimum rate is equal to one position report per six minutes, maximum - one position report per 2 seconds. Additionally, there is a possibility to call for reports in pulling or continuous mode particular types of ships independent on their present positions or ships being actually in defined sea area using digital selective calling (DSC) technique. Ship borne equipment shall report different data, including ship's identity, type, position, course, speed, navigational status and other safety related information like port of destination, route plan, type of hazardous cargoes carried on board, number of persons on board, etc.

There are two classes of ship borne equipment: AIS class A designated for ships listed in SOLAS Convention and class B designated for so called non-SOLAS vessels like yachts, fishing vessels and small merchant ships. The main difference between them includes number of data transmitted by ship and rate of this transmission. AIS class B transmits less frequently smaller number of data than AIS class A. IMO requires AIS class A to be fitted on all maritime passenger ships and cargo vessels of 300 gross tonnage and upwards engaged on international voyages and on cargo ships of 500 gross tonnage and upwards engaged on home trade. It is estimated that more than 40000 ships currently carry AIS class A. In 2007, the new Class B AIS standard was introduced which enabled installation of new generation of the low cost AIS equipment on non-conventional vessels. This has triggered multiple additional national or regional mandates from some countries like EU member states affecting hundreds of thousands of vessels. In 2010, the most commercial vessels operating on the European Union inland waterways were mandated to fit inland modified and approved AIS class A equipment. The fishing vessels with length over 15 meters registered in the EU member states have to install AIS class A until 2014. Additionally, a number of other countries including China, India, the United States, and Singapore, have started AIS mandate programs which require large numbers of vessels to fit an approved AIS device class A or B for safety and national security purposes.

AIS traffic monitoring systems are introduced on the basis of national or regional regulations. All of them utilize connected into net shore based AIS stations to communicate in VHF band with ships equipped with AIS. For instance, according to the requirements of the Directive 2002/59/EC all EU member state concerned built up shore based installations for receiving and utilising AIS data transmitted in the VHF band and established together with the European Commission vessel traffic monitoring and information systems (VTMIS) filling up following functions:

- Receiving reports sent by ships bound for or leaving ports of a member state, their operators, or agents; and
- Monitoring vessels in sea areas of responsibility of the state and in its ports.

Requirements of this Directive were introduced by the end of 2008.

Monitoring system utilizing AIS shore stations has many limitations arising from legal regulations regarding ship borne equipment. Most important of them are:

1. Obligatory AIS installation on part of the maritime ships only. Smaller fishing boats, leisure crafts, warships, naval auxiliaries and ships owned or operated by governments might not be fitted with described equipment.
2. Open format of AIS data transmitted on unrestricted radio frequencies.

3. Legal possibility to switch off the ship borne AIS equipment on all vessels when the master on his professional judgement believes that the continual operation of AIS might comprise the safety or security of his ship.
4. Semiautomatic method of work of the ship equipment which transmits and receives automatically data introduced in automatic or manual manner without checking its quality, reality and accuracy. As a result of this method of work, AIS may transmit false data.
5. Narrow zone of monitored area along the coast limited by the range of radio-communication in the VHF band between ships and shore stations.
6. Expensive shore infrastructure of the system. European monitoring system introduced as required by Directive 2002/59/EC consists of about 730 AIS shore stations connected into national, regional (Baltic HELCOM system) and European nets.

It means information provided by AIS may not create a complete picture of the situation in monitored area and due to that should be appreciated as supplementary to that derived from visual observation and radar equipment.

Main advantage of described type of monitoring system is wide scope of data transmitted and received free of any charge.

More detailed basic information about AIS and monitoring systems utilizing AIS shore stations introduced by maritime administrations may be found in [2,3,4].

Additionally to the monitoring systems introduced by maritime administrations or agencies responsible for safety and security at sea, there are many commercial web sites tracking ships equipped with AIS built by private companies, different in style and offering different services, like AISHub.net, AIS Live and Fleet Mon. Participants of the 79th session of the IMO Maritime Safety Committee (MSC) decided that the publication on the world-wide web or elsewhere of AIS data transmitted by ships could be detrimental to the safety and security of ships and port facilities. The Committee condemned this kind of publication and urged IMO member governments, subject to the provisions of their national laws, to discourage those who make available AIS data to others in that manner from doing so. In addition, the Committee condemned those who irresponsibly publish AIS data transmitted by ships, particularly if they offer services to the shipping and port industries.

3 Long Range Identification and Tracking System

Today fully operational Long Range Identification and Tracking (LRIT) system provides global identification and tracking of suitable equipped merchant ships sailing outside the areas of coverage of the shore based AIS stations working in the VHF band (outside so called Global Maritime Distress and Safety System (GMDSS) A1 area). System has mixed structure comprising international, national, co-operative and regional LRIT data centres. It consists of the:

- Ship borne LRIT equipment;
- Application System Providers (ASP);
- Communication Service Providers (CSP);

- National, regional, co-operative and international data centres;
- International Data Exchange System (IDES) working according to the approved by IMO Data Distribution Plan (DDP); and
- International Mobile Satellite Organisation (IMSO) working as LRIT Co-ordinator monitoring international part of the LRIT system and acting on behalf of all governments participating in the system.

International data centre was introduced by IMO and works under supervision of this organisation. Governments not participating in this centre established national, co-operative or regional LRIT data centres co-operating with the international centre. They are obliged to report updated relevant details about these centres to IMO. Non-international LRIT data centres served as a national, regional or co-operative vessel monitoring systems (VMS), e.g. as mentioned in chapter 2 VTMIS built by the European Union member states and European Commission according to the requirements of the Directive 2002/59/EC. LRIT data centres serving as VMS may require transmission from ships additional information not required by the LRIT system and at different intervals. It may also perform other functions. An international LRIT data centre may, upon request, collect additional information from ships entitled to fly the flag of some states on the basis of specific agreements signed by IMO with the maritime administrations of these states.

EU member states created one of the largest data centre in the LRIT system (EU LRIT DC) monitoring 1/3 of the world's maritime traffic. It handled around 3 million messages about ships on a monthly basis.

The biggest national LRIT data centre collecting information from approximately 8000 ships was built by Panama Maritime Authority.

The LRIT regulation applies to the following types of ships engaged on international voyages outside GMDSS A1 area:

- All passenger ships including passenger high-speed craft;
- Cargo ships, including high-speed craft of 300 gross tonnage and above; and
- Mobile offshore drilling units.

Ships participating in the system are obliged to transmit in coded manner identification of the ship borne LRIT equipment and ship geographical position and its data and time. Data is transmitted in:

- Six hours intervals (as standard) or more often as required by competent authority, but in time intervals not smaller than 15 minutes; and
- Pulling mode on request of the competent authority.

LRIT data may receive maritime administrations or agencies responsible for maritime safety and security of the SOLAS contracting governments about ships:

1. Entitled to fly its flag irrespective of where such ships may be located outside inner waters of another country.
2. Which have indicated their intention to enter a port facility or a place under the jurisdiction of that government, irrespective of where such ships may be located provided they are not located within the inner waters of another country or in territorial waters of the ship flag state.

3. Entitled to fly the flag of other governments, not intending to enter a port facility or a place under the jurisdiction of that government, navigating within a distance not exceeding 1000 nautical miles of its coast provided such ships are not located within the inland waters of another country or territorial waters of the ship flag state.

Governments bear all costs associated with any LRIT data they had requested and received. They are responsible for farther distribution of the LRIT data and shall not impose any charges on ships in relation to the LRIT information they may seek to receive.

Additionally, Search and Rescue (SAR) services of SOLAS contracting governments shall be entitled to receive, free of any charges, LRIT information in relation to the search and rescue of persons in distress at sea.

Basic advantages and disadvantage of the LRIT system are:

1. Advantages: coded transmissions to authorised receivers only and absence of expansive coastal infrastructure.
2. Disadvantages - limited number of data transmitted from ships (identification, position, data and time only).

It should be emphasized that ship borne LRIT equipment may be switched off on all ships when ship flag state takes this decision or ship master on his professional judgement believes that the operation of LRIT might comprise the safety or security of his ship.

More detailed basic information about LRIT system may be found in [5,6,7].

4 Satellite Monitoring System Utilizing Ship AIS Transmissions

In the past years it was confirmed that reception of AIS signals from space is feasible. Space based AIS technology greatly extends range of the monitoring systems utilizing shore base AIS stations working in the VHF band and creates many new application possibilities for competent maritime authorities. While AIS was originally designed for short range operation as a collision avoidance and navigational aid, it is now possible to receive AIS signals by low earth orbit (LEO) satellite. This is becoming known as S-AIS and is different from LRIT system. Single polar orbiting satellite can provide complete global coverage. As the satellite orbits north-south, the Earth turns east-west underneath the orbital plane. Within about 12 hours, the satellite observes every point on the planet. S-AIS like LRIT system receives using satellites and collects messages transmitted by ships and requires on board equipment to be switched on and efficient. Difference between S-AIS and LRIT system consists in the transmission rate of the ship messages containing different number of data. Thus the information collected from S-AIS and LRIT system is mutually complementary.

Indeed, due to the co-channel interference near densely populated or congested sea areas satellites have difficulties in detecting and receiving AIS messages from space in those areas. The AIS VHF channels 87B and 88B are not restricted and used on shore for another purposes. LEO satellite detects ship AIS transmission together with land users' transmissions done in the same band. To solve this problem World Radiocommunication Conference of the International Telecommunication Union

(ITU) allocated in 2012 two additional channels in VHF band (channels 75 and 76) for improved satellite detection and tracking of vessels. This allocation was made on a primary basis for region 2 and on a secondary basis for regions 1 and 3. Introducing of these two new frequencies will require changes in the construction of already existing ship borne equipment class A and B which has not transceivers working in new designated channels. Due to that, there will be discussion at IMO in near future, whether to introduce this new possibility and modernize or change ship equipment now or to consider AIS working in all four channels as an element of the future e-navigation system developed by this organisation.

S-AIS is under development by European Union (EU) and some states outside EU like Canada, Norway and USA. European Maritime Safety Agency (EMSA) begun collaboration with European Space Agency (ESA) in order to create satellite AIS monitoring and start to built S-AIS in 2007. Monitoring has been jointly implemented by EMSA and ESA via the Data Processing Centre (DPC). DPC collects AIS data received through satellite and ancillary data from the ground stations, processes them and delivers through EMSA prime contractor – Collecte Localisation Satellites (CLS) to the users using the SafeSeaNet (SSN) and other interfaces. The service has been operational since October 20th 2011.

In June 2008 the company ORBCOMM begun to launch new LEO satellites equipped with the capability to collect AIS data and became the first commercial satellite AIS data service to qualified government and commercial subscribers. It established terrestrial network of 15 gateway earth stations around the world ensures timely delivery of the satellite AIS data to its subscribers. US Coast Guard announced plans to collect AIS data by satellite from ships up to 2000 nautical miles away from the US coast to give US authorities maximum warning of approaching ships. System is introduced in collaboration with ORBCOMM.

On April 28th 2008, Canadian company COM DEV International became the first company launching a space based AIS nano-satellite designed to detect AIS signals. It is currently developing a full nano-satellite constellation, global ground network and centralized data processing centre in order to offer global AIS data service. The service is operational and available worldwide as of mid-2010 through exactEarth COM DEV's data services subsidiary. ExactEarth uses a patented ground and space based processing technology to minimize interference AIS signals, thereby improving detection compared with all other satellite based systems.

On July 12th 2010, the Norwegian AISSat-1 satellite was successfully launched into polar orbit. The purpose of the satellite was to improve surveillance of maritime activities in the high north area. AISSat-1 is a nano-satellite, measuring 20x20x20 cm and weighing 6 kg, with an AIS receiver made by Kongsberg Seatex. It is used by Norwegian maritime administration to improve safety at sea for vessels in Norwegian waters and to monitor the transport of dangerous goods and cargo in the high north.

5 Conclusion

Satellite monitoring of sea traffic using ships AIS transmission in VHF band is developing very rapidly and independently of terrestrial AIS and LRIT systems.

Companies developing this technology work with international regulatory and expert authorities to develop the appropriate regulatory and data policy framework to ensure required by IMO security of the AIS data. Some states and regional organisations create complex monitoring systems collecting, processing and presenting information on maritime traffic issued from different available source: terrestrial and satellite AIS monitoring systems, LRIT system, vessel traffic services (VTS), ship reporting systems (SRS), coastal radars, agents, harbour masters, etc. Vessel traffic monitoring and information system (VTMIS) created by the European Union and managed by EMSA may be mentioned as a sample of these complex systems. It is developed in order to:

- Gather and provide information about ships and hazardous cargo movements in EU waters (server SafeSeaNet);
- Monitor the position of EU ships worldwide (EU LRIT Data Centre);
- Provide reliable information to port state control officers about ships to be inspected in port (data base and server THETIS); and
- Detect possible oil spills in EU waters through satellite monitoring (server CleanSeaNet).

The Integrated Maritime Data Environment (IMDatE) aims to combine and process data from EMSA's maritime applications (SafeSeaNet, CleanSeaNet, LRIT, THETIS) and other external sources (e.g. S-AIS, coastal radar), enabling all information to be visualised together on one interface. It should provide, by 2012, different services depending on the user access rights for each application. Data may be delivered via a user friendly web interface and distributed automatically to authorised external systems. SafeSeaNet allowed the real time tracking of the approximately 17000 ships which transit in EU waters on a daily basis through the generation of more than 4 million ship position reports per day.

It should be mentioned that there are some fears in relation to dynamic development of so called AIS navigational functions. Many maritime administrations decided to use AIS to transmit application specific messages (ASM), e.g. information on status of real floating aids to navigation, synthetic and virtual aids to navigation, hydrological, meteorological and tidal data, etc. Existing operational purpose of the AIS intended for ships identification and collision avoidance and utilization of the AIS messages transmitted by ships for the sea traffic monitoring purposes would be at risk if significant volumes of this type additional data communication were put on the existing AIS frequencies. The capacity of the AIS VHF data link (VDL) is limited and the main function of AIS must not be jeopardized by transmission of ASMs.

References

1. European Maritime Safety Agency, Annual Report (2010), `http://www.emsa.europa.eu`
2. Wawruch, R.: Uniwersalny statkowy system automatycznej identyfikacji (AIS). Fundacja Rozwoju Wyższej Szkoły Morskiej w Gdyni, Gdynia (2002)
3. Wawruch, R.: Wykorzystanie systemu automatycznej identyfikacji do monitorowania statków morskich. Przegląd Telekomunikacyjny 12, 969–975 (2007)

4. Wawruch, R.: Utilization of the automatic identification system for collision avoidance and traffic monitoring. In: Mikulski, J. (ed.) Advances in Transport Systems Telematics, Katowice, pp. 261–269 (2007)
5. Wawruch, R.: Global ships monitoring system - basic requirements and principle of introduction. Transport Problems 3(2), 59–68 (2008)
6. Wawruch, R.: Global ships monitoring system - basic requirements and possibility of establishment. Polish Journal of Environmental Studies 17(5A), 99–103 (2008)
7. Wawruch, R.: Światowy system identyfikacji i śledzenia statków. Przegląd Telekomunikacyjny i Wiadomości Telekomunikacyjne 1, 16–23 (2009)

Efficiency of Induction Heating of Rails with Oblong Heaters

Leszek Szychta, Elżbieta Szychta, and Kamil Kiraga

Kazimierz Pułaski Technical University of Radom
Faculty of Transport and Electrical Engineering
26-600 Radom, Malczewskiego 29, Poland
{l.szychta,e.szychta,k.kiraga}@pr.radom.pl

Abstract. Results of laboratory testing on two types of oblong heaters applied to induction heating of rails are presented and discussed in this paper. The article continues the research into practical applications of induction heating to turnouts.

Keywords: induction oblong cup and flat heater, induction heating, rails, turnouts.

1 Introduction

In their earlier work (2010-2011), the authors focussed on developing a simulation model of the induction heating process of 60E1 rail in two-dimensional space using FEMM software [1, 4, 5, 8] and methods of determining key electric and magnetic properties of 60E1 rails, whose knowledge proved necessary to develop a method of turnout induction heating [2, 3, 9]. Time has come therefore to verify the issue in an actual laboratory stand. This paper combines key results of the authors' research with regard to three areas - laboratory testing of rail induction heating process, numerical calculations of magnetic field and eddy currents, and measurements of basic electric and magnetic parameters of a selected rail type.

2 Experimental Stand for Testing of Rail Induction Heating

A laboratory stand for rail induction heating shown in Fig. 1.a) and 1.b) is developed on the basis of testing results of rails' electric and magnetic quantities as well as simulation test results of rail induction heating process executed in FEMM software. This stand comprises two inductor types – thus, two different sources for generation of electromagnetic field.

J. Mikulski (Ed.): TST 2012, CCIS 329, pp. 328–333, 2012.

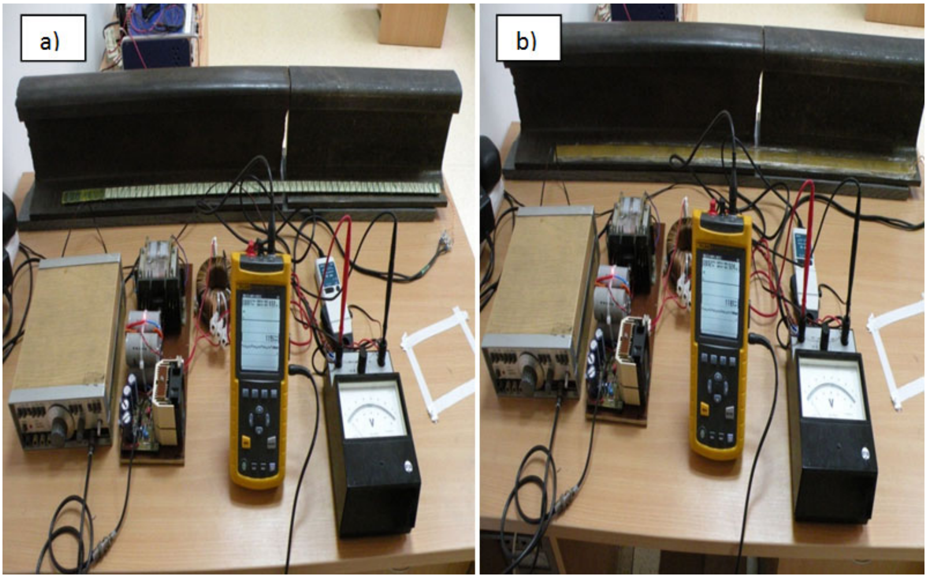


Fig. 1. The laboratory stand with a a) cup, b) flat oblong heater

The heaters used in laboratory testing of rail induction heating are shown in Fig. 2 and Fig. 3.

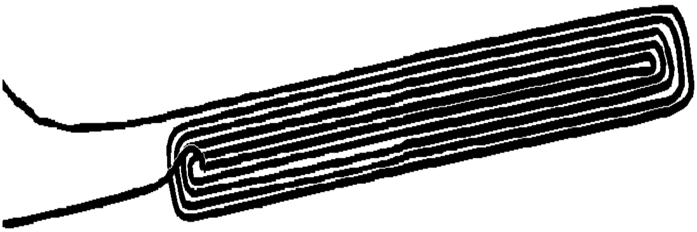


Fig. 2. Oblong flat heater

The laboratory stand under discussion consists of the following elements:

- Two heaters of sizes and dimensions similar to flat-oval heaters applied to electric turnout heating,
- infrared camera,
- functional generator,
- power amplifier,
- analyser of energy quality,
- (mercury and electronic) thermometers,
- personal computer.

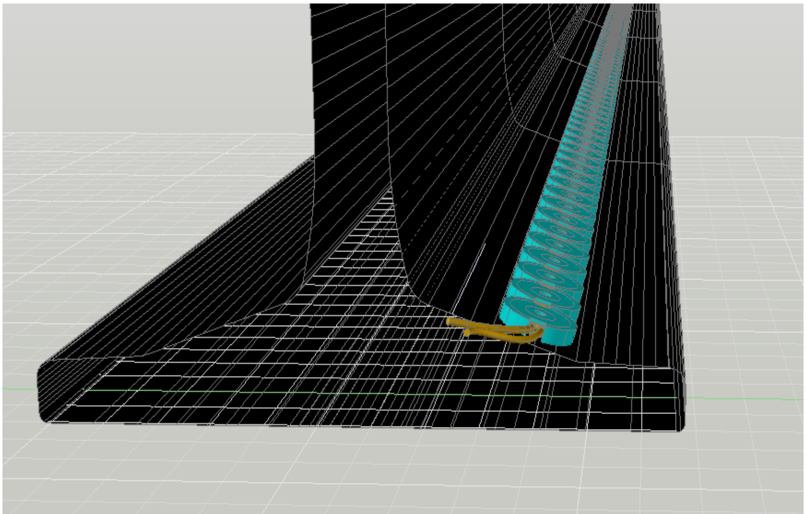


Fig. 3. Oblong cup heater

The magnetic loop of the inductor and receiver consisted of two 1.13m long heaters generating different shapes of the magnetic fields shown in Figures 2 and 3, respectively.

The infrared camera recorded temperature growth in the heated section of rail at selected points in time.

The quality analyzer, equipped with current and voltage measurement probes, tracked current and voltage shapes in the circuit and the angle shift between the tracked courses, thereby the active power [6] lost across a heating element.

The functional generator was designed to generate a sinusoid voltage and supply it to the power amplifier.

A 100W power amplifier served as a source of supply voltage, controlled by the functional generator.

The thermometers read the ambient temperature in the testing room.

The PC recorded measurement data from the experiments.

The following assumptions were adopted for the laboratory testing:

- range of frequency f variations of the supply voltage (0 – 1000)Hz,
- oblong flat and cup heaters are tested at 60W,
- ambient temperature $T_{otoczenia}$ and air temperature $T_{atmosfery}$ in the testing room are constant,
- constant emissivity of the tested material is 0.960 (the rail was covered with a layer of soot) and constant air humidity in the testing room is 16%,
- initial rail temperature $T_{p\text{-}szyny}$ and initial heater temperature $T_{p\text{-}grzałki}$ are constant. After each experiment, the rail and the heater are cooled back to their initial temperatures.

The individual experiments continued for 30 minutes. Images of heated rails were recorded every 30 seconds to capture minimum temperature growths.

3 Distribution of Rail Temperatures in Induction Heating

A series of experiments were conducted on the heaters under discussion in the range of frequency *f* variations of the supply voltage (50 – 1000) Hz. Constant power of 60 W supplied to the heaters was maintained.

The experiments implied the inductors discussed in this paper failed. Temperature recorders showed that only heaters themselves became hot in both cases. Temperature growths ΔT of rail foot were below 3°C and were similar across the entire range of *f* variations.

Results of laboratory tests of rail heating by means of an oblong cup heater for the selected frequency $f = 650$ Hz are illustrated in Figure 4. Results of laboratory tests of rail heating by means of an oblong flat heater for the selected frequency $f = 650$ Hz are illustrated in Figure 5.

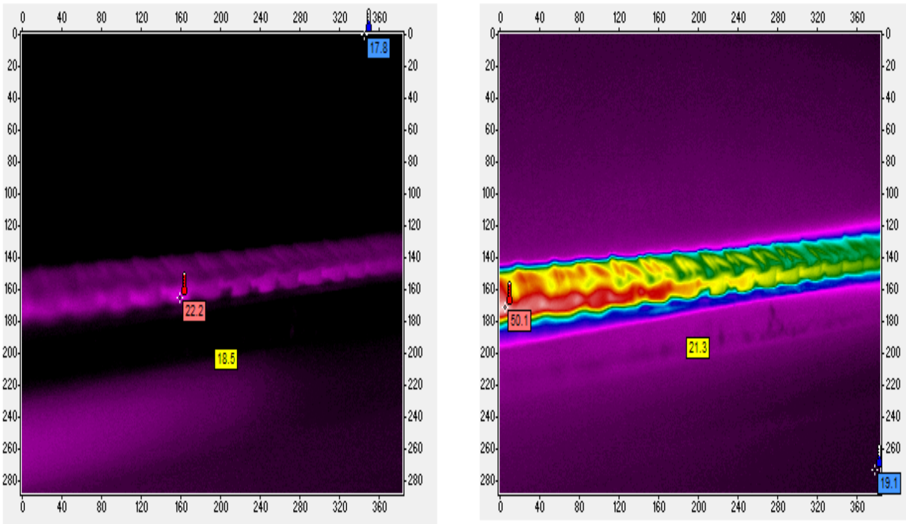


Fig. 4. Image from an infrared camera recording temperature growths ΔT across the rail foot after 30 mins of heating with an oblong cup heater (results of laboratory testing)

For the sake of comparison, Figure 6 illustrates temperature graphs in the rail foot heated with oblong : a) cup and b) flat heaters.

4 Conclusion

The oblong heaters presented in this paper fulfill requirements of resistance heater design in electric turnout heating (requirements of PKP PLK Polish railways) [7] in respect of their geometric dimensions.

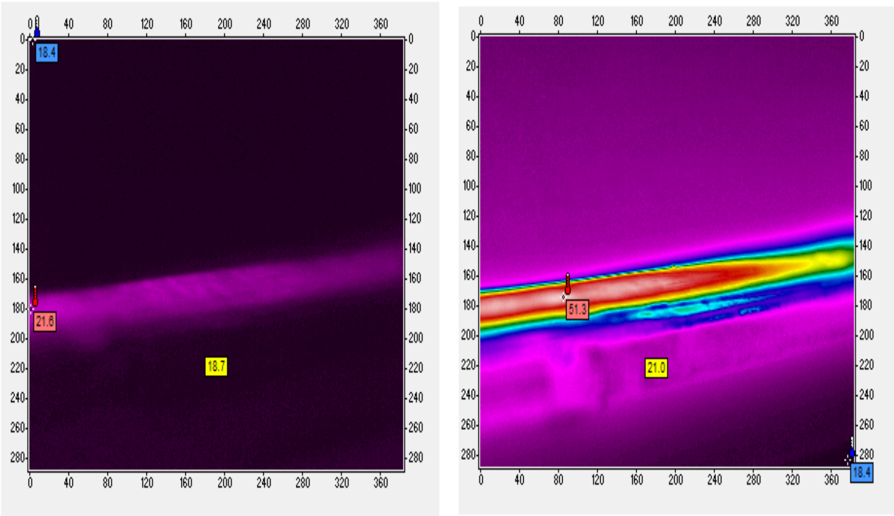


Fig. 5. Image from an infrared camera recording temperature growths ΔT across the rail foot after 30 mins of heating with an oblong flat heater (results of laboratory testing)

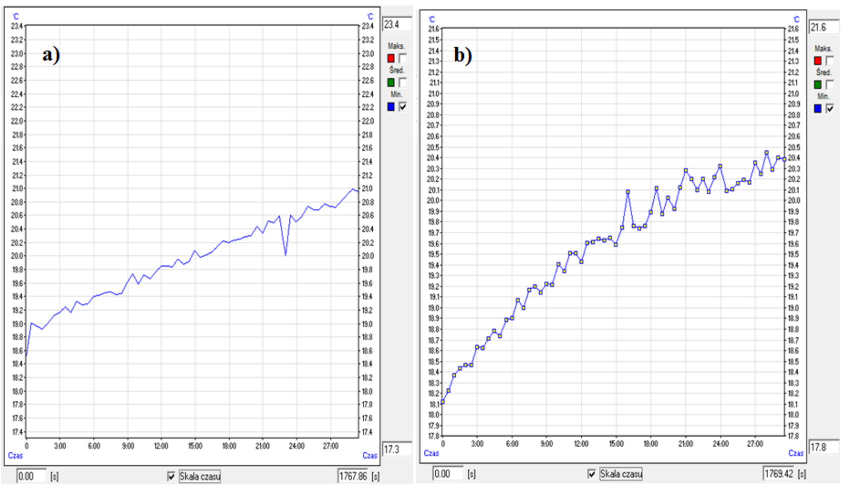


Fig. 6. Graphs of temperature growths ΔT across the rail foot after 30 mins of heating with an oblong a) cup heater and b) flat heater (results of laboratory testing)

Efficiency of rail heating with oblong cup and oblong flat heaters proved very low. Results of the laboratory testing imply that oblong heaters acting as inductors failed to provide a closed magnetic loop for the magnetic flux and the magnetic field they generated was dispersed as a result. The current caused heat to be released across the heater resistance only. The consequent temperature growth across the rail reached a mere few degrees and covered a small portion of the rail immediately adjacent to the heaters. The two design types of heaters presented in this paper will not be applied to induction heating of railroad turnouts, therefore.

Further research should focus on modifying design of inductors employed to heat rails as part of induction heating of railroad turnouts.

References

1. Kiraga, K.: Distribution of magnetic induction and magnetic field intensity in a rail and its environment in femm software, Trencianskie Teplice, Slovakia, September 20-22. Central European School of Doctoral Study, pp. 44–47 (2011)
2. Kiraga, K., Szychta, E.: Badanie wybranych właściwości elektrycznych i magnetycznych szyny kolejowej. In: X Krajowa Konferencja Naukowa Sterowanie w Energoelektronice i Napędzie Elektrycznym SENE 2011, art. 9, Łódź, 16-18 listopada (2011) (in Polish)
3. Kiraga, K., Szychta, E.: Electric and Magnetic Properties of Railway Rail for use in Induction Heating of Rail Turnouts. In: Proceedings of the 2nd International Scientific Conference DESAM 2012 Diagnostics of Electrical Machines and Materials, University of Żilina, Popradno-Povazska Bystrica, Slovakia, February 8-9, pp. 20–26 (2012)
4. Kiraga, K., Szychta, E., Kowol, M.: Distribution of Magnetic Induction and Magnetic Intensity Across A Railway Rail and Its Environment. In: Proceedengs of the 8th International Conference ELEKTRO 2010, Žilina, Slovak Republic, May 24-25 (2010)
5. Kiraga, K., Szychta, E.: Two-dimensional model induction heating of 60E1 rail. Transcom, Žilina, Slovakia, June 27-29, pp. 73–78 (2011)
6. Mikołajuk, K.: Podstawy analizy obwodów energoelektronicznych. Wydawnictwo PWN, Warszawa (1998) (in Polish)
7. PKP Polskie Linie kolejowe S.A. let-5 Wytyczne projektowania urządzeń elektrycznego ogrzewania rozjazdów. Tom 1 Projektowanie instalacji torowych i przytorowych, dobór grzejników, Warszawa (2009) (in Polish)
8. Szychta, E.z., Szychta, L., Kiraga, K.: Analytical Model of a Rail Applied to Induction Heating of Railway Turnouts. In: Mikulski, J. (ed.) TST 2010. CCIS, vol. 104, pp. 96–106. Springer, Heidelberg (2010)
9. Szychta, E., Szychta, L., Luft, M., Kiraga, K.: Application of 3D simulation methods of the process of induction heating of rail turnouts. In: Perpinya, X. (ed.) Infrastructure Design, Signalling and Security in Railway. InTech, Croatia (2012)
10. Białoń, A., Mikulski, J.: Wpływ typu ogrzewania rozjazdów na zużycie energii elektrycznej. Przegląd Elektrotechniczny (9), 37–39 (2009)

Marine and Offshore Telematics Systems

Adam Weintrit

Gdynia Maritime University, Faculty of Navigation
Al. Jana Pawla II 3, 81-345 Gdynia, Poland
weintrit@am.gdynia.pl

Abstract. Throughout the field of transportation and vessel systems control is gaining importance. This paper focuses on the current key problems engineers in this field are facing and highlights some major recent accomplishments. In the fields like marine or offshore systems, the focus of research is on the swarming behaviour of multiple vessels. New sensors and networking will also enable more efficient traffic flow control which will allow for a better use of the resource network capacity. A forecast on future trends in marine and offshore telematics systems including e-Navigation concept is given at the end of the paper.

Keywords: offshore technology, telematics, IT, marine control systems, transport systems, e-navigation.

1 Introduction

The advantage of the latest technical development in the field of automation, electronics, telecommunications, informatics, telematics, geomatics and global position fixing techniques, achievement in maritime and offshore data storing, processing, analysing, transferring and visualisation should be taken into account and applied to the maritime and offshore systems. In the paper the Author defines offshore industry entities: offshore constructions and offshore supply vessels, tries to discuss the key issue and priorities of e-Navigation concept, including hydrographic data, position fixing and communications, as well as satellite navigation systems, the use of telemetry on offshore worksite, autonomous marine control systems installed on board of unmanned surface vehicles (USVs), unmanned underwater vehicles (UUVs) and remotely operated vehicles (ROVs), describes marine and offshore telematics systems which are revolutionizing not only shipping but also the boating industry, and finally governance of offshore IT outsourcing. A forecast on future trends is given at the end of the paper.

2 Offshore Industry Definition

In an industry encompassing functions of drilling, exploration, production and environmental protection competitiveness will be determined less by sheer size and

J. Mikulski (Ed.): TST 2012, CCIS 329, pp. 334–348, 2012.

asset holding and more by agility, speed of response, and capacity for innovation. This can be axiomatically stated of the offshore oil & gas industry today.

2.1 Offshore Constructions

Offshore construction is the installation of structures and facilities in a marine environment, usually for the production and transmission of electricity, oil, gas and other resources.

Construction in the offshore environment is a difficult and dangerous activity. Construction and pre-commissioning is typically performed as much as possible on land or inshore areas. To optimize the costs and risks of installing large offshore platforms, different construction strategies have been developed.

One strategy is to fully construct the offshore facility inshore, and tow the installation to site floating on its own buoyancy. Bottom founded structure are lowered to the seabed by de-ballasting, whilst floating structures are held in position with substantial mooring systems.

The size of offshore lifts can be reduced by making the construction modular, with each module being constructed onshore and then lifted using a crane vessel into place onto the platform. A number of very large crane vessels were built in the 1970s which allow very large single modules weighing up to 14,000 tonnes to be fabricated and then lifted into place.

Specialist floating hotel vessels known as flotels are used to accommodate workers during the construction and hook-up phases. This is a high cost activity due to the limited space and access to materials.

Oil platforms are key fixed installations from which drilling and production activity is carried out. Drilling rigs are either floating vessels for deeper water or jack-up designs which are a barge with liftable legs. Both of these types of vessel are constructed in marine yards but are often involved during the construction phase to pre-drill some production wells. Other key factors in offshore construction are the weather window which defines periods of relatively light weather during which continuous construction or other offshore activity can take place. Safety is another key construction parameter, the main hazard obviously being a fall into the sea from which speedy recovery in cold waters is essential.

The main types of vessels used for pipe laying are the "Derrick Barge (DB)", the "Pipelay Barge (LB)" and the "Derrick/Lay Barge (DLB)" combination. Diving Bells in offshore construction are mainly used in water depths greater than 120 ft, less than that, the divers use a metal basket driven from an "A" frame from the deck. The basket is lowered to the water level, then the divers enter the water from it to a maximum of 120 ft. Bells can go to 1500 ft, but are normally used at 400 to 800 ft.

Offshore construction includes the design, construction, and/or repair of offshore structures, both commercial and military, including:

- subsea oil and gas developments,
- offshore platforms - fixed platforms, semi-submersibles, spars, tension leg platforms (TLPs), Floating Production Storage and Offloading (FPSOs), etc.

– Floating Oil and Gas Platforms - semi-submersibles, spars, TLPs, FPSOs, etc.

2.2 Offshore Supply Vessels

These are cargo vessels that regularly transport goods, supplies or equipment in support of exploration or production of offshore mineral or energy resources. Offshore supply vessels are typically operated by shipowners: either companies set up specifically to own and operate such vessels or companies combine with other vessel operations like salvage, shipping, etc.

The general mission of supporting offshore oil and gas exploration and production can be sub-divided according to specific mission requirements. Also vessels may involve in more than one kind of such specific missions. Like every other design criteria, optimizing based on the priority is done in the design process, and also meeting the safety requirements.

These specific missions include:

– seismic survey to locate oil and gas bearing areas
– towing of rigs and platforms to their location, positioning them, and laying anchoring & mooring equipments
– supplying rigs and platforms with necessary personnel, equipment, stores, provisions, etc.
– subsea operations like ROV operation, diving support, inspection, and maintenance,
– safety standby.

Offshore vessels can be divided into the following types:

– Platform Supply Vessels. These are used for transporting supplies to the rigs and oil platform and return other cargoes to shore. The supplies include fuel, fresh water, equipments, consumables, stores and provision for the operation of the oil platforms and for the personnel onboard. The design and construction of these vessels depend on the working environment - weather conditions, distance from the shore, and hence design varies by location. These ships range from 20 to 100 meters in length and accomplish a variety of tasks. The primary function for most of these vessels is transportation of goods and personnel to and from offshore oil platforms and other offshore structures. In the recent years a new generation of Platform Supply Vessel entered the market, usually equipped with Class 1 or Class 2 Dynamic Positioning System;
– Multi-purpose Supply vessels. These are similar to the Platform Supply Vessels but are fitted with additional systems for example, for subsea inspection, maintenance and repair;
– Anchor Handling vessels. These are used to towing and anchor handling and are fitted;
– Seismic vessels;
– Crew boats;
– Safety / Standby vessel;
– Combination vessels.

Table 1. Modern merchant ships

Dry cargo	Barge · Bulk carrier · Lake freighter · Car float · Coaster · Collier · Container ship · Heavy lift ship · Hopper barge · Lighter aboard ship · Reefer ship · RORO ship · Submarine Cargo Vessel · Train ferry · Livestock carrier
Tankers	Chemical tanker · FPSO unit · LNG carrier · LPG carrier · Oil tanker · Tanker · Gas Carriers
Passenger	Cargo liner · Cruise ferry · Cruise ship · Ferry · HSC · Narrowboat · Ocean liner · RORO ship · Train ferry
Support	Dive support · Fireboat ·Supply ship · Tender · Towboat · Tugboat
Other	Pipe-laying ship · Cable layer · Crane vessel · Dredger · Drillship · Fishing vessel · Icebreaker · Merchant submarine · Narco submarine · Research vessel · Riverboat · Semi-submersible · Snagboat

3 Key Issues and Priorities of e-Navigation

e-Navigation is the future, digital concept for the maritime and offshore sector. Larger and faster ships, greater congestion and reduced manning levels have all provided the impetus for this development [5]. Considering the wide range of options and benefits that could become part of e-Navigation, the primary value of e-Navigation is to join the ship's bridge team and sea traffic monitoring teams to create a unified navigation team that would achieve safer navigation through shared information. For full implementation of such a system it would need to be mandatory for SOLAS vessels and scaleable to all users.

It was suggested that before the primary benefits and value-added services could be realised, an architecture comprising three fundamental elements should first be in place. These are:

- Electronic Navigation Chart (ENC) coverage of all navigational areas (WEND - Worldwide Electronic Navigational Chart Database);
- a robust electronic position-fixing system (EPFS), with redundancy; and
- an agreed infrastructure of communications to link ship and shore.

Specifications for these fundamental elements are contained as follow.

3.1 Hydrographic Data (ENCs)

A full coverage of ENCs for navigational waters will require considerable effort from the world's hydrographic community. It has further been noted that the existence of proprietary updating software in many ECDIS systems has become a key cost issue when implementing ENC data. It is thought that if, through IMO, an open architecture system could be agreed, this would allow a more competitive environment in the purchase, and maintenance of ECDIS systems thus reducing the overall costs of ENC's and increasing the global rate of acceptance. From the seaman‘s point of view there is unsolved question of responsibility for correction of information presented by ECDIS and ENC updating.

3.2 Position Fixing

Electronic position-fixing systems, which could be integrated into e-Navigation, can be divided into Global Navigation Satellite Systems (GNSS), GNSS augmentations, terrestrial radio-navigation systems and non-radio positioning systems. There are two operational GNSS at present (GPS & GLONASS) and two more planned: European - Galileo and Chinese - Compass. It has long been recognized that GNSS require augmentation to achieve the required integrity for safety of life applications and the accuracy needed for specialized navigation and positioning. Augmentation systems fall into two broad categories: Ground Based (GBAS) and Satellite Based (SBAS). GBAS (IALA) maritime beacon system has been the standard GNSS augmentation system for maritime applications. SBAS is based on two operational (WAAS, EGNOS) and two planned public service (MSAS, GAGAN).

There are many high accuracy, local terrestrial radio-positioning systems provided, mostly on a commercial basis, for specialized applications. However, the only terrestrial radio-navigation system with widespread, regional coverage is Loran-C. The Far East Radio-Navigation System (FERNS) is provided under an international agreement between PRC, Russia, Korea and Japan and extends from the Bering Straits to the South China Sea. Saudi Arabia also has a system, covering its own territory and the Arabian Gulf. Non-Radio Positioning Systems is the Inertial Measuring Unit (IMU), usually integrated with GNSS to enhance it and cope with outages.

The problem of fixing position coordinates for navigational needs considered only in terms of measurement error seems to have already been solved in a global scale. Its realization with higher or lower precision is only a function of the technical solution adopted. Therefore, other, equally important, although often omitted, exploitation parameters of navigation systems become crucial. These are: availability, integrity, continuity and also reliability.

The following is a list of key elements required for e-Navigation position fixing:

- appropriate accuracy, availability, continuity, and integrity (alert limit, time to alarm, integrity risk), already included in IMO Resolution A.915(22);
- adequate redundancy;
- compatibility between systems; and
- appropriate datums (vertical and horizontal).

There is also necessity to develop a unified theory of the some navigational criteria (availability, reliability, continuity, and integrity) under consideration and to determine the relations between them, because:

- reliability and availability refers to different functional structures,
- definition of continuity is ambiguous,
- lack of mathematical connection between availability, reliability and continuity,
- vague procedures and methods of determining each of the criteria,
- measurement of the criteria is based only on statistic analysis of empirical measurement data.

These and others methodological problems should be solving as soon as possible, because all fixing systems characteristics have to be considerate in the same standardized way. The next important problems in implementation position systems to e-Navigation are:

- identification of the service provider responsibility (especially for global and wide area positioning systems) for accidents caused by non-operation status,
- to establish international cooperation between GNSS service providers related to others than positioning services (safety of live, commercial, search and rescue, etc.),
- to solve responsibility problem for core navigational system provider and augmentation signal deliverer.

Current GNSS has a common weakness in that they are all subject to accidental or intentional interference. Hence, alternative and independent position fixing capabilities need to be considered. Consideration should be given to independent non-GNSS Electronic Position Fixing System and sensors as a potential component of E-Navigation.

e-Navigation systems should enable the electronic capture of radar ranges, radar and visual bearings, etc. for position fixing.

3.3 Communications

The following is a list of key communication aspects required for e-Navigation, relating to both technical and content:

- autonomous acquisition and mode switching (i.e., minimal mariner involvement needed);
- common messaging formats;
- sufficiently robust (e.g., signal strength, resistance to interference);
- adequate security (e.g., encryption);
- sufficient bandwidth (data capacity);
- growth potential;
- automated report generation;
- global coverage (could be achieved with more than one technology); and
- the use of a single language (English), perhaps with other languages permitted as options.

The following communications issues are among those that will require resolution to achieve the above:

- it seems likely that a satellite broadband link will be required to achieve the above requirements, and consideration must be given to how this will be achieved; and
- the question of cost and who pays for the provision of a satellite broadband link must be resolved early in development of e-Navigation.

The standardization and unambiguous interpretation of information plays an essential role in the appropriate accomplishment of navigational information acquisition and exchange processes in the e-Navigation System. The definition of relevant standards will enable unequivocal interpretation of the information. Measures taken to unify the above mentioned standards are aimed at the development of the navigational information ontology. The starting point for the creation of this ontology is an analysis and classification of navigational information accounting for its kind and range. This will allow to sort out the structure of navigational information, thus the availability and exchange of information will be extended.

4 Satellite Navigation

The provision of satellite navigation services is presently under U.S. monopoly through GPS. GPS was conceived primarily to serve military user requirements and cannot as such satisfy most of civil user expectations. GPS and Glonass, if it matures, require significant adjustments to become civil usable systems. A system satisfying civil user requirements, and particularly transport services, will bear the generic name of GNSS - Global Navigation Satellite System. This situation provides a window of opportunity for Europe to play its cards through Galileo system.

4.1 Civil User Requirements - Diversity

Potential areas of applications are numerous, obviously for all modes of transport, but also for offshore industry, sea bed exploration, geodesy, gravimetry, timing services, hydrography, leisure, ...

Transport Applications:
Today's shipping would be impossible without the extensive use of radio navigation. These maritime and offshore activities could derive numerous benefits from the application of satellite navigation.

Potential benefits for shipping are: low equipment cost; one system world-wide instead of numerous different systems in use today; a high accuracy will enable the use of satellite navigation for harbour approach and in narrow waterways, where radio navigation accuracy was insufficient so far; the integration of satellite positioning with digital map displays simplifies navigation; position reporting systems, particularly for the surveillance of fishing vessels and dangerous goods, can be implemented easily.

Geodesy, Surveying:
Surveyors are a rather small, but important user group. The high cost of maintaining national reference networks as well as the cost of actual surveying tasks can be significantly reduced by the use of satellite positioning. Economic effects in the offshore sector can be significant, as the allocation of revenue from the recoverable deposits within an exploitation block depends on the exact definition of the boundaries of these blocks.

Other Applications:
There are numerous other applications of GNSS, which will become operational once the level of development has reached a mature stage. For instance applications in fish

farming, marine agriculture, where an optimum distribution of organic fish fertilizers can be achieved by linking accurate position data with in situ measurements of yields. Furthermore, the high quality precise time distribution and synchronisation system can be used for timing purposes on the ground. Finally, there is an enormous potential for satellite navigation in leisure activities for hikers, amateur pilots and sailors as the most common examples of user groups for better safety and for pleasure.

GNSS will be the first radio navigation system in history having the potential to fulfil the requirements of practically all areas in which radio positioning and navigation can be applied. Therefore, GNSS is likely to become an indispensable part of the world-wide infrastructure.

Civil User Requirements - Characteristics

The complexity of the present situation can be grasped having in mind the two following fundamental considerations:

- the availability of GNSS serving a public service role and/or commercial service purposes for civil use is foreseeable in the near future but is still not there. GPS give only the flavour of it.
- many categories of users and service providers are dreaming about the potential applications of GNSS as perceived through the experimental use of GPS but their detailed requirements are generally not mature.

Users will increasingly rely on GNSS as it comes to fulfils operational requirements, provided that cost-benefit relation is satisfactory. Operational requirement may comprise compliance of GNSS to specific rules according to:

- the nature of the applications: public service missions, commercial operations, leisure activities,
- the mode of operation: accessory or supplemental or self-sufficient mean of navigation, safety critical or not with respect to human life, vehicles, goods.

These operational requirements can be turned into system performance requirements expressed in terms of the following characteristics:

- accuracy: the system will provide relatively homogeneous accuracy world-wide, some applications may require better precision than could be obtained by technique called "differential". This principle involves a ground reference station (or a network of such stations, depending on the area to be covered), which determines local errors of the satellite positioning and transmits them to the user via a datalink.
- integrity: ability of the system to alert in the appropriate time frame on errors.
- availability: measurement of the amount of hours per year the system will be unusable. This relates to the number of active satellites.
- continuity of service: relates to the commitment required from the service providers to offer perennial service at stable financial conditions. This comprises the institutional issue of who controls the systems as well as the financial issue of the cost recovery mechanism.

5 The Use of Telemetry on Offshore Worksite

Telemetry is a communication technology used for remote sensing and control often used for hydrographical and offshore industry related purposes [3].

The use of telemetry on offshore worksite:

- DGPS shore reference station transmitting differential correction data to all DGPS users on the worksite,
- shore based monitoring station to control and log overall operational data as position off all infield installations like different vessels, information and sensor data from wave height and direction buoy, tidal information from tide gauge,
- information such as position, dredged quantities from dredger to shore and GPS correction data from shore station to dredger,
- position information of survey vessel to shore and DGPS correction data from shore station to vessel,
- position information of stone dumping barge to shore station and DGPS correction data from shore station,
- position information of shallow water survey lounge to shore and DGPS correction data from shore station,
- tidal information from the tide gauge is transmitted via the telemetry of the DGPS shore station to all installations on the work site,
- wave height and direction information transmitted to the shore station and from there available for all users on the site.

5.1 Radio (DGPS) Datalink

The most common type of transmitting corrections is, at the moment, radio transmission. Radio transmission is relatively cheap and provides access to a large group of users. This type of system is also called Local Area Augmentation System (LAAS). Depending on the type of users a frequency is chosen, the most common being: UHF/VHF, MF and HF.

UHF/VHF. These are generally short range systems used in hydrographic surveying. Transmitters are generally portable and can be set up by the user. Special attention should be given to the frequency and power level chosen, since there may be restrictions (like the need to apply for a permit) on either in a specific area / country.

A generally used frequency band for DGPS corrections is the 439 MHz band, which is free for power output levels with a maximum of 0.5 Watt. The range is up to 25 kilometres from the transmitter with these setting. The accuracies achieved with this type of system are in order of 1-2 meters. When a multi reference solution is used the accuracy can be 0.5 meters.

The disadvantage of UHF / VHF is that it is a line of sight system, meaning that obstructions between the transmitter and receiver can block the signal. Another problem with UHF / VHF is the reflection of the signal off islands, structures and buildings, thus creating shadow zones and multipath.

MF. When the survey area is bigger than 20 kilometres, a medium or long range system should be used. Around the world IALA beacons are used to provide range corrections to navigational users of GPS. These corrections can be received using cheap receivers from reference stations placed around most international waters used for shipping. The range of these systems is up to 150 kilometres with an accuracy of about 2-3 meters.

HF. These systems use HF frequencies for transmitting and can cover 100 to 600 kilometres and provide a DGPS system at 1.0 to 5 meters accuracy. In general, these services are not free of charge. The signal will be scrambled by a code, which can be obtained for free from the supplier of the signal. An advantage of these type of systems is the signal monitoring by the supplier and the relatively high accuracy achieved.

Public Radio Broadcast. In some countries signals are broadcasted over public radio as a so-called piggy back signal. In general FM transmissions are used and the accuracies achieved will be in the order of 3 meters.

Eurofix. A specific form of radio telemetry is the Eurofix system. With Eurofix, Loran-C transmitters transmit the DGPS correction signal. Loran-C is a terrestrial radio positioning system that seemed to become a victim to the wide spread use of DGPS. However, since then various reports, Volpe report (United States) and Helios study (Europe), criticized the vulnerability of GPS as a sole means of navigation. Since Loran-C uses terrestrial transmitters with high power signals in a different frequency band it is not vulnerable to the same sources as GPS. As a stand-alone system it is however less precise. To be able to be an independent backup for GPS, Loran is currently being upgraded to enhanced Loran - eLoran.

By combining Loran and GPS, a system is created that is less vulnerable and which can achieve a precision of meters even when the GPS signal is interrupted. The Americans are fully convinced that Loran-C should remain active for this reason and are modernizing their Loran-C transmitters. Some European governments however are less convinced and one of the Loran-C chains, the NELS (Northwest Europe Loran-C System) is reportedly threatened to be shutdown.

The argument often heard is that Galileo will solve the problem mentioned. This is not a valid argument however, since Galileo is vulnerable in much the same way as GPS and/or Glonass since all systems are based on the same principles and frequency bands.

Mobile Telephone. Corrections via mobile telephone are at the moment offered by commercial services only. The advantage of these networks is that they offer a large number of reference stations over a relatively small area. As a result a high precision is achievable without the need to erect an own base station. The main disadvantage is the high cost when performing continuous measurements. As a result these networks are commonly used for calculating base positions. However, with the introduction of UMTS and GPRS techniques costs will probably drop considerably.

Satellite Transmission. As we have seen with radio based telemetry systems, they can cover a couple of hundred kilometres at best. This is not enough from precise positioning in the middle of the Atlantic ocean. When DGPS signals are transmitted via a satellite network, they can cover a much larger area than when transmitted from terrestrial beacons. There are a number of commercial networks based upon satellite telemetry. An example of such a network is the Furgo Omnistar / Starfix network.

A special form of satellite DGPS as WAAS (Wide Area Augmentation System) and its European counterpart EGNOS (European Geostationary Overlay System). With EGNOS the - free - correction signals are transmitted on the same frequency as the GPS signals themselves. As a result no expensive separate DGPS receiver is needed. The GPS receiver however needs to be able to receive the so-called satellite-based augmentation system (SBAS) messages that are transmitted by geo-stationary satellites. Currently three systems are being developed, WAAS (USA), EGNOS (Europe) and MSAS (Japan). EGNOS uses 3 satellites to broadcast the correction messages, two Inmarsat satellites (IOR, AOR) and an ESA Artemis satellite. Next to the correction message, a GPS signal is transmitted as well, so the three augmentation satellites can also be used as an additional GPS satellite.

5.2 Hydroacoustic Telemetry

Another type of telemetry is used for underwater purposes and called hydroacoustic telemetry. This type of underwater telemetry is mostly used in the offshore industry for controlling valves and sensor measurements from subsea installations. For this type of telemetry acoustic underwater hydrophones are used to transmit and receive the signal. Also AUV's (Autonomous Underwater Vehicle) communicate with a baud rate of approximately 2400baud via a underwater acoustic system with mother vessel.

AUV's are used to carry payload like sonar and can carry out bottom profiles during autonomous operation the bottom profile data is logged inside the AUV and can be transferred via a cable when the AUV is on deck of the mother vessel. Only simple commands can be given due to the slow speed of the underwater telemetry.

6 Autonomous Marine Control Systems

During the last decade there has been a rapidly growing interest in the design and development of autonomous marine craft [1, 2, 4]. The craft in question being unmanned underwater vehicles (UUVs) and unmanned surface vehicles (USVs). It is considered that this area of marine control systems design will be one of the key areas for research and application both in the military/naval and industrial sectors for the foreseeable future.

The dynamic characteristics of an UUV present a control system design problem which classical linear design methodologies cannot accommodate easily. Fundamentally, UUV dynamics are nonlinear in nature and are subject to a variety of disturbances such as varying drag forces, vorticity effects and currents Therefore they offer a challenging task in the development of suitable algorithms for motion and position control in the six degrees of freedom in which such craft operate, and are required to be robust in terms of disturbance rejection, varying vehicle speeds and dynamics. It should be noted that the term ''unmanned underwater vehicle'' as used here is a generic expression to describe both an autonomous underwater vehicle

(AUV) and a remotely operated vehicle (ROV). An AUV being a marine craft which fulfils a mission task without being constantly monitored and supervised by a human operator, whilst an ROV is a marine vessel that requires instructions from a human via a tethered cable or an acoustic link.

Although ROVs play an important role in the offshore industry, their operational effectiveness is limited by the tethered cable, and the reliance and cost of some form of support platform. Whilst even though AUVs cannot be considered as being commonplace at this moment, they are thought by many to be the future technology to provide essential platforms for instruments and sensors for various kinds of subsea missions. These missions could include environment forecasting, policing exclusive economic zones and under-ice operations as well as ocean basin monitoring.

Issues surrounding the deployment of AUVs at sea are mainly threefold. The first is a non-technical matter that needs to be resolved as soon as possible and revolves around the legal responsibilities and liabilities for AUVs when working at sea. The second relates to the limited endurance capacity of existing power systems. To overcome this problem, there needs to be a major breakthrough in battery technology and/or a shift to other power sources. The third and final restricting factor is associated with the capabilities of onboard navigation, guidance and control (NGC) systems.

Paramount importance for AUV is the requirement for it to be equipped with a robust navigation subsystem that can accurately predict its current position. High accuracy can be gained by employing costly inertial systems. However, as the popularity and use of AUVs increases so will the demand for continued high navigational accuracy but at low cost. The solution to this ongoing problem lies in the use of inexpensive sensors being used in multi-sensor data fusion (MSDF) algorithms. Without doubt the development of MSDF algorithms is a priority for research. In addition, as the navigation aspects of an AUV improve so must the guidance laws become more sophisticated.

Although AUVs are seen as having great potential, such craft cannot be deployed in shallow or inland waters to undertake, for example, surveying and pollutant tracking tasks. As a result, operational costs are currently high as self contained underwater breathing apparatus (SCUBA) divers or special vessels containing a number of people have to be employed. Hence the interest in providing such services at low cost via USVs which are capable of operating in river systems, and littoral and deep water. USVs can also be usefully commissioned for search and rescue missions, police, and custom and excise operations, and a variety of deterrent, attack and covert military roles.

The dynamic characteristics of USVs will vary depending upon whether it is a mono or twin hull vessel. However, irrespective of the hull configuration, they are all exhibit highly nonlinear behaviour. Further complications arise with these vehicles when attempting to control the surge, sway and yaw modes owing to underactuation. In many cases, underactuated USVs are more easily served using nonlinear control theory. Thus underactuated marine systems and the application of nonlinear control theory in the design of their control systems are considered as a necessary field of continued research.

It may be argued the navigation of an USV is less of a problem than that of an AUV because of the access to GPS information. To a certain degree this may be true, however, they can be required to operate in areas of non-existent/degraded GPS reception. Thus the navigation of USVs is still difficult and thereby the need for intelligent dead reckoning algorithms applies equally to such vessels as well as AUVs.

The navigational aspects of an USV can be further complicated if it operates in a pack with similar vehicles. This can be further exacerbated if the pack or an individual also has to be linked with other air borne and/or subsea autonomous assets.

Hence as USVs can be engaged in multi-entity operations, research into network centric systems is essential along with that for robotic co-operative and swarming behaviour.

7 Marine and Offshore Telematics Onboard of Boats

Marine and offshore telematics systems are revolutionizing not only shipping but also the boating, offshore, recreational fishing and commercial fishing industries. Marine and offshore telematics involves the use of wireless voice and data communication systems which provide vessel tracking, emergency aid, system monitoring, internet access, and other features.

Systems normally consist of a user interface, satellite antenna, and a communication link with the vessel's electronic systems. This technology can be vital to the user since it provides a satellite link to the outside world when other communications may unavailable.

Several companies have entered the marine telematics market for boats. Some of them integrate Global Positioning System (GPS) technologies with electronic mapping and mobile communications to monitor and provide important boat data from port, land or sea. Installed in boats they have received wide acclaim from prominent boat builders and boat owners alike. They offer an array of services that cater to boat security, boater safety, and boating enjoyment, including an advanced emergency messaging system that can pinpoint a distressed boat's location offshore. They can also alert boat owners to potential problems when they are not using their boats, such as low battery voltage or water intrusion in the bilge. In addition, they have a proven record of assisting law enforcement agencies in tracking and recovering stolen boats, offering boat owners a new level of security and safety.

The systems provide emergency communications, detailed messages, including information on GPS position and other critical parameters, between watercraft and the dedicated maritime rescue coordination centre. This data is intended to help rescue agencies locate a vessel and respond appropriately during an emergency.

Other wireless maritime information system companies provide Vessel Monitoring System (VMS) services to commercial fishing vessels. These systems include GPS, two-way communications and e-mail. A network operations centre provides continuous support. The term Vessel Monitoring System, or VMS, refers to a wireless information system that automatically reports fishing vessel position and activity.

8 Governance of Offshore IT Outsourcing

The lack of effective IT governance is widely recognized as a key inhibitor to successful global IT outsourcing relationships. In this study the development and application of a governance framework to improve outsourcing relationships is presented. The approach used to developing an IT governance framework includes a meta model and a customization process to fit the framework to the target organization. The IT governance framework consists of four different elements: (1) organisational structures, (2) joint processes between in- and outsourcer, (3) responsibilities that link roles to processes and (4) a diverse set of control indicators to measure the success of the relationship. The IT governance framework is put in practice in Shell GFIT BAM, a part of Shell that concluded to have a lack of management control over at least one of their outsourcing relationships.

9 Conclusion

The key structural components of a safe and comprehensive e-Navigation policy for marine and offshore purposes are:

- accurate, comprehensive and regularly up-to-dated Electronic Navigational Charts (ENCs), covering the entire geographical area of a vessel's operation;
- accurate and reliable electronic positioning signals, with "fail-safe" performance (probably provided through multiple redundancy, e.g. GPS, GLONASS, Galileo, differential transmitters, Loran C and defaulting receivers or onboard inertial navigation devices);
- provision of information on vessel route, course, manoeuvring parameters and other status items (hydrographic data, ship identification data, passenger details, cargo type, security status etc) in electronic format;
- transmission of positional and navigational information: ship-to-shore, shore-to-ship (e.g. by VTS, coastguard centres, hydrographic offices) and ship-to-ship;
- accurate, clear, integrated, user friendly display of the above information onboard and ashore (e.g. using IBS or INS);
- information prioritisation and alert capability in risk situations (collision, grounding etc), both onboard and ashore; and
- reliable transmission of distress alerts and maritime safety and security information with reduction of current GMDSS requirements by utilizing newly emerged communication technologies.

There is a growing trend towards the use of remote control centres. These require continuous monitoring, with the concept of the digital field wholly dependent on guaranteed secure communications.

References

1. Fossen, T.L.: Marine control systems: Guidance, navigation and control of ships, rigs and underwater vehicles. AS: Marine Cybernetics (2002)
2. Kiencke, U., Nielsen, L., Sutton, R., Schilling, K., Papageorgiou, M., Asama, H.: The impact of automatic control on recent developments in transportation and vehicle systems. Annual Reviews in Control 30(1), 81–89 (2006)
3. Lekkerkerk, H.-J., Theijs, M.J.: Handbook of Offshore Surveying, 2nd edn. Projects, Preparation & Processing, vol. 1. Skilltrade BV, The Netherlands (2011)
4. Roberts, G.N., Sutton, R. (eds.): Advances in unmanned marine vehicles. IEE Control Series. IEE Press (2005)
5. Weintrit, A.: Telematic Approach to e-Navigation Architecture. In: Mikulski, J. (ed.) TST 2010. CCIS, vol. 104, pp. 1–10. Springer, Heidelberg (2010)
6. Weintrit, A., Wawruch, R., Specht, C., Gucma, L., Pietrzykowski, Z.: Polish Approach to e-Navigation Concept. TransNav – the International Journal on Marine Navigation and Safety of Sea Transportation 1(3), 261–269 (2007)
7. Yoon, D., Yi, M., Jeong, J.S., Park, G.K., Son, N.S.: Safety Control of Maritime Traffic Near by Offshore in Time. TransNav – the International Journal on Marine Navigation and Safety of Sea Transportation 5(1), 65–69 (2011)

Some Specific Activities at the Railway Signalling System Development

Karol Rástočný, Juraj Ždánsky, and Peter Nagy

University of Zilina, Faculty of Electrical Engineering
Univerzitná 1, 010 26 Žilina, Slovakia
{karol.rastocny,juraj.zdansky,peter.nagy}@fel.uniza.sk

Abstract. The paper presents authors' opinions on some essential activities related to the interlocking system development. The safety of electronic systems is based on other principles than for the relay systems. That means it is necessary even to change the approach to their safety function definition and safety evaluation. The risk taken by a person using the railway transport should be similar or independent of the country of occurrence. For a signalling system to be acceptable it is not enough to state only the goal fulfilment but it is even mandatory to prove the decrease of the risk to the minimal acceptable level. Objectification of the safety function definition process for signalling systems is very problematic nowadays.

Keywords: signalling system, interlocking system, safety, safety integrity level, development, modelling.

1 Introduction

Like in the case of living organisms, characteristic phases are also identifiable in a life cycle of technical devices (systems).These phases have common or even the same characteristics, regardless of a specific system. The view on a technical system that is based on its life cycle is defined through a list of activities that take place from the initial initiative of the system creation (concept) up to the point in time, when the system is decommissioned. European safety standards for railway applications are based on the life cycle based view of the railway signalling systems and the standard [1] specifies procedures and tasks for each phase of the life cycle that is necessary to perform in order to achieve the system ability to perform its required functions with respect to reliability, availability, maintainability, safety (RAMS) and their mutual interdependencies.

In general, a signalling system is required to fulfil not only control functions, but the safety functions as well. Safety functions of the signalling system are, together with their respective tolerable hazard rate (THR), determined by the risk analysis. They actually present technical measures meant to mitigate the risk related to specific hazards. To create a future signalling system a sole declaration that safety measures have been taken is insufficient. The proof that measures that have been taken decreased the risk at least to an acceptable level is necessary. The task to make the

J. Mikulski (Ed.): TST 2012, CCIS 329, pp. 349–355, 2012.

risk analysis objective is quite problematic nowadays, since it is directly related to the definition of safety functions of the system. Similar problems are identified in the safety assessment of the already developed system. Even though there are European safety standards available, that define rules on how to proceed in a safety functions specification process, those rules are, however, too generally defined.

The aim of this paper is to bring some of the open problems that are related to the development process of a signalling system before the reader. The second aim of the paper is to recommend a process that would eliminate those problems to a great extent. That can be, however, reached only when common rules on the definition of safety functions (application-dependent) and their safety evaluation will be available.

2 Definition of Safety Functions

In the initial phases of the system life cycle a process model need to be created, that allows us to identify hazards related to the railway traffic process, to estimate rates of their occurrence and their possible consequences. The combination of the rate of hazard occurrences and their consequences presents a risk related to the control of a process. If this risk is identified and known, and the tolerable risk is obtained, then safety functions can be defined and their respective THR can be assigned. If the failure of the control system to perform a control function can endanger a train operation, then such function must be considered as a safety function and related safety requirements must be specified.

If the identified hazard can cause a material damage, then the tolerable risk value is heavily dependent on an operator. If the use of a signalling system can somehow cause a human harm (which is usually true), then the operator alone is not allowed to determine the value of a tolerable risk. If the railway transport is discussed globally (at least on a European level), then the risk taken by a user of a railway transport service should be the same, regardless of the country, in which the transit takes place. The achievement of this goal assumes the risk analysis based not only on collective risk principle, but on an individual risk principle as well. Even though the standard [1] provide margins for individual risk analysis (e.g. MEM principle), such an approach to risk analysis is not practically used, because a strict mathematical apparatus is involved in such risk analysis and the results of the analysis are frequently dependent on a specific, time-dependent traffic situation (e.g. traffic volume, duration of the hazard presence, a body position of a user during an accident, immediate train speed).To some extent the historical (statistical) data can help to rate individual risk factors, even though the expert assessment needs to be done at the end. The problem, however, lies also in the fact, that a signalling system is developed and meant to be used in the future, while the statistical data origin naturally in the past. The application of pessimistic approach to the information assessment is an acceptable solution, but this way leads to a considerable exaggeration of safety requirements of the system.

In practice, a collective principle is used in the risk analysis – a matrix method can be employed, for instance, which has a qualitative background and is based on the expert estimation of the hazard rates and their consequences. This approach usually leads to the conclusion, that the failure of any of the elementary safety functions related to the train control can cause a human casualty, which renders the safety requirements on such safety function rather strict, described by the safety integrity level (SIL) 4.

Nowadays there is no binding list of safety functions for different types of signalling systems (station interlocking equipment, line blocking system …).If the risk analysis is not based on the individual risk principle, then the definition of safety functions for each type of signalling system is necessary to make an objective comparison of safety properties of comparable signalling systems manufactured by different manufacturers. The definition of safety functions is nowadays practically based on an agreement between a manufacturer and an operator (or a safety authority). Given the fact that the standard [3] defines the THR on a single safety function, the definition of safety functions alone is also of great interest during the safety evaluation process. Different specification of safety functions can finally lead to different safety analysis results.

A simplified case of execution of a safety function that is performed by separate parts of a system and consequent assignment of the THR to these parts of the system is pictured in the Fig. 1 and Fig. 2.

Let us assume a safety function F, which can be logically decomposed into two partial safety functions $F1$ and $F2$, i.e.:

$$F = F1 \wedge F2. \tag{1}$$

Let the risk analysis be followed by the THR assignment, during which tolerable hazard rate THR_{F1} has been assigned to function $F1$ (THR_{F1} corresponds to the tolerable hazardous failure rate λ_{F1}^{T}). Similarly, tolerable hazard rate THR_{F2} has been assigned to function $F2$ (THR_{F2} corresponds to the tolerable hazardous failure rate λ_{F2}^{T}). Let the equality

$$\lambda_{F}^{T} = \lambda_{F1}^{T} + \lambda_{F2}^{T} \tag{2}$$

be valid, in which λ_{F}^{T} is the tolerable hazardous failure rate assigned to function F.

In the presented example, the signalling system pictured in Fig. 1 fulfils the safety function F. A sensor subsystem (sensor S1 and sensor S2), a subsystem performing logical functions L and an actuator subsystem (actuator A1 and actuator A2) are all together participating in the performance of the safety function F.

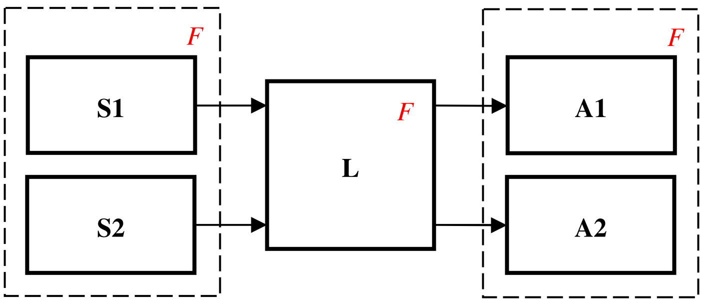


Fig. 1. Block diagram of a system that implements a safety function F

The total hazardous failure rate of the overall system related to the performance of function F can be formulated in the following manner:

$$\lambda_F^H = \lambda_{S1}^H + \lambda_{S2}^H + \lambda_L^H + \lambda_{A1}^H + \lambda_{A2}^H. \tag{3}$$

where λ_i^H is the failure rate of those failures of the i^{th} element (i.e. λ_{S1}^Hdescribes element S1 etc.),whose failures are in relation to function F hazards.

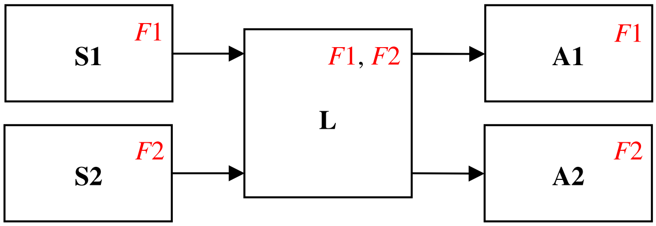


Fig. 2. Block diagram of the system that performs safety functions $F1$and $F2$

Now, let the already described system displayed in Fig. 1 performs two partial safety functions $F1$ and $F2$, as shows the diagram in Fig. 2. The total hazardous failure rate of the system related to individual functions $F1$ and $F2$ can be evaluated:

$$\begin{aligned} \lambda_{F1}^H &= \lambda_{S1}^H + \lambda_L^H + \lambda_{A1}^H, \\ \lambda_{F2}^H &= \lambda_{S2}^H + \lambda_L^H + \lambda_{A2}^H. \end{aligned} \tag{4}$$

When equations (3) and (4) are compared, it is obvious, that

$$\lambda_F^H \neq \lambda_{F1}^H + \lambda_{F2}^H. \tag{5}$$

The statement expressed by equation (5) is in conflict with the assumption expressed by equation (2).

The signalling system can, in general, include parts (modules) that participate in the implementation of more than one safety function. At the same time, some other modules may not participate in the implementation of any safety function. If the module implements no safety function, then no safety requirements are laid on such module and the module is also not subjected to the safety evaluation process.

3 Design of the Signalling System

A signalling system is characterised by a long useful life-time (approx.20 years) and considerable high development costs, in comparison to a number of applications (mainly because of safety requirements). Therefore it is of great importance that such system is not morally outdated even before commissioning or in the first years of its operation (the development and consequent approval process takes years and technological advance is nowadays rather fast, especially in the field of semiconductors). On the other hand, safety requirements force the developers to conservatively use well-known and already proved technology. Only such elements and also tools shall be used, whose properties were already checked in as wide as possible range of not necessarily safety-critical applications, and sufficient references on positive as well as negative experience from their operation is available. The use of modules with exactly defined interfaces in the future can help to solve problems with the use of older, but proved technology. The current trend is to avoid using special parts or system elements, the use of commercial available resources is preferred instead (so-called COTS – Component of the Shelf).

The new generation of signalling systems, based on microprocessors, employs so-called composed safety technique to achieve the fail-safe property. Signalling systems based on composed safety principles use various forms of redundancy to achieve the required safety level. Critical safety functions are implemented at least twice in different manners, while data comparison between individual system channels is used.

The decision on system architecture is one of the most important decisions that must be taken in the development process. A trade-off between the system cost, and the system availability and safety must be usually made, so a great attention should be paid to the decision on the system architecture.

The decision on the system architecture cannot be simply influenced by an intuition of a developer, but a thorough RAMS parameters modelling must be performed in advance instead. Based on the created model an optimal solution must be chosen, so all customer's requirements are respected and fulfilled.

Software is a vital part of electronic programmable signalling systems. The software also affects the safety of the railway traffic process. Systematic failures are typical ones, when speaking about the software safety. An ambiguous or incomplete functional requirements specification is a common source of the systematic software failures. That is the reason why the creation of an analytical model of the developed system is advisable in the early phases of the system's life cycle. Such model should provide means to analyse functional requirements of the system. Suitable methods and tools exist nowadays (e.g. based on unified modelling language - UML), that support not only the analysis, but are able to generate an executable code from the design model[6].

The RUP (Rational Unified Process) method is the method most used in the software development process[4], which can be described by the following characteristics:

- life cycle is divided into four phases (inception, elaboration, construction, transition);
- software product is prone to changes during its life (implementation of new functions, bugs removal, etc.), so some of the phases need to be iterated; the very first run of the life cycle is called initial, each next is then evolutional;
- each phase also consists of more iterations; their number depends on the project complexity;

- different activities are executed in each phase (or iteration), e.g. requirements, analysis and design, implementation, testing; the volume of these activities in each phase or iteration is also project-dependent;
- besides the technical problems, the RUP also brings the tasks of project management in individual phases into focus.

Even though the mentioned method is not entirely consistent with the recommendations of the standard [1] (because the method has been originally developed for the development of a commercial software), it is quite flexible and adjustable to the required application field (railway applications).

The software cannot be safe or dangerous, since in principle the safety is always a system property. The software safety can be assessed only if the specific software is considered as a part of specific signalling system (as an integral part of hardware). The use of signalling systems based on the microprocessor technology brings inarguable benefits (flexibility, modifiability, smaller dimensions, lesser energy consumption ...). On the other hand new problems must be solved – e.g. related to a sequential processor operation, or a large state-space of such system. If the large-scale software is used, its thorough testing that would cover all functional requirements is also demanding, especially to cover it in all possible operational states.

4 Safety Evaluation of the Signalling System

The presentation of a documented proof that the system complies with the specified safety requirements is the necessary condition of a signalling system acceptance [3]. The part of the so-called safety case is also an evidence that the system fulfils the requirements on hardware and software integrity (against random and systematic failures, respectively).

It is necessary to analyse effects of random hardware failures on the safety of a railway signalling system through a suitable qualitative method. State-space based methods (based on Markov chains or Petri nets theory) allow to analyse effects of not only random failures, but also the effects of another safety-affecting factor in a single model. A reflection of specific system architecture, diagnostic coverage, time to detect and negate a failure and of recovery rate is required in such model[7]. The safety integrity of a system is the dominating, but not the only safety attribute of a system. Generally, the safety of a signalling system can be understood as a list of system properties that among others include the safety integrity and availability of the system. System models created by state-of-the-art methods are usually focused on the analysis of a single property. A local optimum can be achieved for each property through those methods. There is a strong interdependency between relevant system properties. It is therefore recommended to try to create such a model with the help of a suitable method that would allow us to find a global optimum with respect to limiting values of chosen system properties and economic acceptability of the proposed solution.

The effects of random hardware failures on the safety of a signalling system can be with more or less success done through a combination of qualitative [5] and quantitative methods of analysis (using the probability-based approach). On the other hand, the effects of software failures (systematic failures caused by either wrong requirements specification or a programmer's mistake) cannot be quantitatively evaluated, but only assessed by qualitative methods. Since in the case of complex software packages the performance of a complete test of the software is virtually impossible, the statement that the software is correct cannot be made even when all tests are done. The only statement that can be made is that the tests did not discover any failure (or if they did, it has been corrected) and the measures that correspond to the required SIL have been taken during the overall software development process [2].

5 Conclusion

The unification of Europe brings with it the unification of basic standards and regulations in the field of railway transport applications. An endeavour to achieve interoperable railways fuels the unification process. The interoperability carries with itself many benefits, such as a decrease of costs of signalling systems development and a gradual uniformity in the views on the signalling systems safety.

The main aim lays in the effort to develop compatible signalling systems that fulfil common European standards and to mutually accept the safety-approval process of individual systems and their parts. A mutual acceptance is possible only when system parts or functions they perform are rather generic and generally valid. The acceptance of the common safety-approval process is virtually impossible when specific system parts or functions are in question, especially when they are based on national-specific standards and regulations. The procedures on how to proceed in the safety approval of the railway signalling systems are specified in the standard [3].

Acknowledgement. This paper was supported by the scientific grant agency VEGA, grant No. VEGA-1/0388/12 "Quantitative safety integrity level evaluation of control systems in railway application".

References

1. EN 50 126-1: Railway applications – The specification and demonstration of Reliability, Availability, Maintainability and Safety (RAMS). - Part 1: Basic requirements and generic process (1999)
2. EN 50 128: Railway applications -Communication, signalling and processing systems – Software for railway control and protection systems (2010)
3. EN 50 129: Railway applications–Safety related electronic systems (2003)
4. Arlow, J., Neustandt, I.: UML 2 and the unified process: Practical object-oriented analysis and design. Pearson Education, Inc. (2005)
5. Białoń, A.: Problem of Railway Traffic Control Devices Disruption. In: Mikulski, J. (ed.) TST 2011. CCIS, vol. 239, pp. 180–191. Springer, Heidelberg (2011)
6. Rástočný, K., Janota, A., Zahradník, J.: The Use of UML for Development of a Railway Interlocking System. In: Ehrig, H., Damm, W., Desel, J., Große-Rhode, M., Reif, W., Schnieder, E., Westkämper, E. (eds.) INT 2004. LNCS, vol. 3147, pp. 174–198. Springer, Heidelberg (2004)
7. Rástočný, K., Ilavský, J.: Quantification of the safety level of a safety-critical control system. In: International Conference on Applied Electronics (AE), pp. 285–289 (2010)

Assessing Transport Telematic Systems in Terms of Data Services Quality

Mirosław Siergiejczyk

Warsaw University of Technology, Faculty of Transport
Koszykowa 75, 00-662 Warsaw, Poland
msi@it.pw.edu.pl

Abstract. Needs signalled by transport telematic system users can be represented by users demand for telematic services delivered by networks transmitting the telematics-based information. Because of internal and external factors that demand depends on a range of events, actions and processes capable of distorting it. This paper aims to qualitatively analyse transport telematic systems in terms of performance and quality of provided services concerning transmission of telematics-based information.

Keywords: transport telematic system, telematic services, data transmission, operational effectiveness, efficiency metrics.

1 Introduction

A rapid development and a growing importance of telematics-based services dedicated for transport requires determining the service level, above all of data transmission services. Needs signalled by users of networks providing telematic services could be represented by the demand for data transmission services provided by a given network. Because of internal and external factors, the satisfying of the demand depends on a range of events, actions and processes capable of creating discrepancies relative to unsatisfied, signalled demand. The quality of data transmission services offered by given telematic system is tightly linked with a set of attributes defined for them. Such set would include the following parameters:

- admissible transmission delay, from the source to the receiver,
- access time (time between service execution and request),
- request time and time of service profile change,
- service execution delay and its fluctuations,
- data transmission time,
- offered bandwidth,
- transmission errors,
- reliability of equipment creating the communication chain.

The analysis of factors influencing the delivery of data transmission services proves that a good practice of describing the behaviour of a transport telematic system could

J. Mikulski (Ed.): TST 2012, CCIS 329, pp. 356–363, 2012.

involve developing models concerning the behaviour of operating systems factoring in both the parameters related to technical properties of network elements key to the service (service availability), and executing tasks ordered by telematic systems and networks (performance). It seems as if the operational effectiveness is a good, multi-faceted characteristics of operational parameters of transport telematic systems.

The assessment of operational effectiveness of transport telematic systems in terms of executing telematic services, bearing in mind characteristic industry demands, is one of fundamental metrics showing the relationship between the needs articulated by the system user and the performance produced by the structure, individual elements' capabilities and resources available in the system. The effectiveness should provide the basic information about system's performance over different periods and under different external and internal conditions. It also very often defined as the relationship between benefits (quantified with revenues or profits) and incurred costs. The effectiveness being a tangible feature of a system should express the relation between set objectives (plans) and met objectives (completed), and also between benefits reaped from delivering on objectives and outlays incurred to meet them. Potential is capable or gives a means to satisfy the needs [2, 9, 6, 10, 12].

2 Transmission of Telematics-Based Information

The information broadcasting and transmitting, i.e. its flow, is one of crucial properties of telematic systems. The distribution of telematic information is strictly linked to telecommunications, i.e. the transmission of information over a distance through different signals – now often electric and optic signals. Telecommunication services can provide multivariate data: alphanumeric data, voice, sounds, motion or still pictures, writing characters and various measuring signals, etc. The information chain is the salient aspect of telematic information distribution in terms of telecommunications. In essence, the chain transmits multivariate messages from a transmitter to a receiver, thus concentrates on two data exchange points. However, the fact that the transmission took place matters, as opposed to the way the information was transmitted. The way of transmitting information is important in the case of a communication chain, which is part of the information chain (Figure 1). In the case of a communication chain, data transmission from the transmitter to the receiver without data identification is important. What is important here is the message conversion to a transmittable signal and the transmission medium. During the transmission, the signal is exposed to interferences, thus often becomes distorted. Hence it is crucial, that the signal transmitted to the receiver is the best reproduction of the original signal. The transmission medium is called a communication channel which is usually a wire (twin-lead, coaxial), an optical fibre or a radio channel.

Factors disturbing the process of telematic data feed have been acknowledged as twofold. The first group includes factors depending on the technical and operational reliability of telematic networks' elements, irregularities in terms of broadly defined

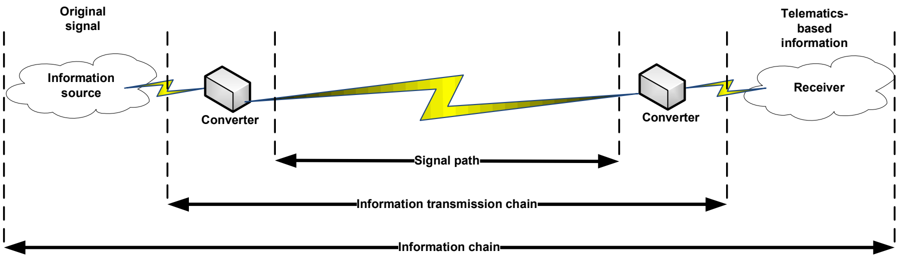


Fig. 1. Telematic information chain and telematic communication chain

electromagnetic compatibility, as well as on the planned and random interference. The performance of a given telematic system can be assessed from these factors' perspective by using its operational characteristics. The second group of factors disturbing the data feed process includes destructive external factors. They are not repeatable during the ICT network's operations and normally they lead to an irreversible ICT hardware damage or long-lasting refurbishment processes taking place at specialised workshops. The destructive impact of disturbing factors is taken into account when estimating the lifetime of ICT networks. Hence it is a noteworthy issue – in this case – to explore whether they have a negative bearing on the operational effectiveness, thus the service from both the standpoint of the user and network operator.

3 Selected Metrics of Effectiveness in Terms of Data Transmission Services

The system effectiveness is probably the most systemic property which factors in the impact of the most important system features and its environment on its performance. It should be informative about the system's capacity and demand for its services or the relationship between benefits (revenue, profit) and costs incurred in order to operate the system. Because of the aforementioned, a single metric could not reliably and in a multi-faceted manner factor in all the requirements. Normally, numerous indicators or sets of indicators would be used, which would give the desired operational characteristics. The following groups are the most popular metrics of system's performance [5]:

- data transmission capacities,
- signal clarity,
- reliability of data transmission,
- availability of data transmission.

3.1 Metrics of Telematic Data Transmission Capacity

Information effectiveness [9]:

$$E_y = \begin{cases} 0 \quad dla\, D_s \le D_s^k \\ ln\, D_s \left[ln\, D_s \dfrac{D_s^g}{D_s^k} \right]^{-1} + ln\, D_s^k \left[\dfrac{D_s^k}{D_s^g} \right]^{-1} \quad dla\, D_s^k < D_s < D_s^g \\ 1 \quad dla\, D_s \ge DD_s^g \end{cases} \tag{1}$$

where: D_s - telematic data transmission capacity of the network;

D_s^k - critical value of D_s when the execution of data transmission tasks is impossible;

D_s^g - boundary value of D_s when the data transmission tasks are fully executed.

***Service availability B_o* -** defined as a steady-state probability of executing the request:

$$B_o = \lim_{T \to \infty} \left\{ \frac{K_1(T)}{K(T)} \right\} \tag{2}$$

where: K_1 - number of successfully completed service requests over observation time T,

$K1$ - number of service requests over observation time T,

Continuous service execution.

That quantity can be defined as the probability of executing order P_{ob}[6]:

$$P_{ob} = (1 - B_o) \cdot (1 - B_p) \tag{3}$$

where: B_p- frequency of disruptions prior to the service completion:

$$B_p = \lim_{T \to \infty} \left\{ \frac{K_p(T)}{K(T)} \right\} \tag{4}$$

where: K_1 - number of requests, whose execution was disrupted over observation time T,

K - number of requests to execute the service over observation time T,

Probability of request execution P_{ru}.

The likelihood of the requested service being executed to the extent specified and continued over a predetermined time.

$$P_{ru} = \lim_{T\to\infty}\left\{\frac{K(T)-K_1(T)-K_p(T)}{K(T)}\right\} = \lim_{T\to\infty}\left\{\frac{K_{ob}(T)}{K(T)}\right\} \tag{5}$$

where: K_{ob} - number of requests processed successfully.

3.2 Metrics of Signal Clarity

The quality of data transmission - widely used in telematic systems - can be given by the metric [8].

Bit Error Rate (BER), S_e

$$S_e = \frac{n_e}{N_e} \tag{6}$$

where n_e - number of bit errors,
N_e - number of analysed bits,

Block Error Rate (BLER), S_b

$$S_b = \frac{n_b}{N_b} \tag{7}$$

where n_b - number of block errors,
N_b - number of analysed blocks.

Character Error Rate P_{bz}:

$$P_{bz} = \lim_{z\to\infty}\left(\frac{Z_b}{Z}\right) \tag{8}$$

where: z_b - number of character errors received by the terminal;
z - number of characters sent from the terminal.

The quality of voice transmission - taking place between the dispatch at surveillance centre and highway emergency system - is determined by the syllable recognition.

Syllable recognition Q_s could be given by the following relationship:

$$Q_s = \lim_{n\to\infty}\left(\frac{n_c}{n}\right) \tag{9}$$

where: n_c - number of syllables correctly recognised by the terminal;
n - total number of syllables sent from the terminal.

3.3 Metrics of Data Transmission Reliability

In telematic systems, the data transmission reliability i.e. the network's capability to establish and maintain a connection between active users over a given time under specific operating conditions, is determined by the reliability of its components and the network's reliability structure.

When evaluating a network's reliability, its "point to point" availability i.e. <a,b> for a,b$\in \Re$, where $\Re$ is a set of network nodes, reliability parameter Ra,b{ φ (x)=1} is limited by the top band determined by the reliability of boundary nodes [82]. Thus:

$$R_{a,b}\{\varphi(x) = 1\} = P\{\varphi(a) = 1\} \cdot P\{\varphi(b) = 1\} \quad (10)$$

where: φ(x) - structural function of telecommunications network's reliability, which is given by the vector x;

φ(a) - structural function of starting node's reliability;

φ(b) - structural function of end node's reliability;

An accurate evaluation of transmission's reliability factors in the impact of reliability structure of bipolar network between a and b nodes, i.e. [6]:

$$R_{a,b}\{\varphi(x) = 1\} = P\{\varphi(a) = 1\} \,.\, P\{\varphi(b) = 1\} \,.\, P_{a,b}\{\varphi(x^*) = 1\} \quad (11)$$

where: x* - is the vector state of a bipolar network without a and b nodes, i.e. the nodes indicated by the network availability criterion.

Given relationships show that requirements important for networks transmitting the telematics-based information are a high reliability of single nodes and a network structure assuring the lowest possible impact of damaged nodes and inter-nodal connections on the system. The limitation of the reliability structure impact can be analysed using two methods.

The first of the methods involves assuring a high line consistency of nodes feeding the telecommunication traffic and keeping the inter-nodal distances low.

The second of the methods involves distributing the allocation of reliability requirements among individual inter-nodal paths by using a graph of the function $R_{a,b}\left\{\varphi(x^*) = 1 = f(\lambda / km\right\}$ where λ / km is an operational failure rate of one kilometre of inter-nodal path.

3.4 Metrics of Data Transmission in Package Networks

The calculation of load availabilities is an entirely different problem [3]. That parameter computing is justified provided some data is sent from a single terminal via multiple connections of different availability. IP networks are a good example. The load availability is a weighted mean of availabilities shown by individual connections and is given by the formula [3, 14]:

$$K_{g_D} = \frac{\sum_i K_{gpi} \times D_{pi}}{\sum_i D_{pi}} \qquad (12)$$

where: K_{gD} - availability of load sent between two nodes via connection *i,*
K_{gpi} – availability rate of the *i-th* connection*,*
D_{pi}, - data rate via the *i-th* connection,
$\sum_i D_{pi}$ - data rate between nodes (provided the connection is unidirectional).

Note, the following relationship always holds: $\min\{K_{gpi}\} \le K_{gD} \le \max\{K_{gpi}\}$.

From the standpoint of the reliability theory, that parameter computing is not justified. However, based on the results [20] this approach might find its application, especially for analysing services provided via converged networks over the IP protocol (*Internetwork Protocol*).

4 Conclusion

The centrepiece of operational effectiveness of transport telematic systems consists of their technical condition, availability, outlays incurred to deliver those tasks and profits made on provided services.

The transport telematics industry (currently almost entirely focused on end user services) offers a range of voice and data services. In order to position oneself within the marketplace, telematic service operators and service providers have to offer a wide array of high quality, rich services. They also have to quickly address the dynamically changing demands in that market. The overarching objective is to provide widely available and functional services of different calibre to the end consumer.

The network's integrity defined as service's independence of access methods and communications protocol is going to be the fundamental evaluation criterion of telematic transport services. Regardless of whether the connection is tunnelled through traditional circuit switching networks or the packet forwarding, a fixed or a mobile network, the service provided has to have constant parameters. Both the service and its level should be the same across different locations - at home, office and during travel in particular.

The offered product portfolio should stay the same regardless of technologies applied. The provided service should be good quality and transparent for the user regardless of whether it is using switched or package data services, and the implementation should be an issue of choosing service provider and/or network operator.

References

1. Amanowicz, M.: Modelling of professional terrestrial mobile radiocommunication systems. WAT, Warsaw (1989)
2. Bobrowski, D.: Effectiveness of technical objects. Praxeology (3-4), 173–181 (1983)

3. Cholda, P., Jajszczyk, A.: Availability assessment of telecommunications networks. Telecommunications Digest and Telecommunications News no. 2-3. Publishing House Sigma NOT, Warsaw (2003)
4. Hardy, W.C.: QoS Measurement and Evaluation of Telecommunications Quality of Service. John Willey & Sons Ltd., England (2001)
5. ITU-T Recommendations E.800: Terms and definitions related to quality of service and network performance including dependability. Approved in 1994.08. Status: In Force
6. Kwestarz, W., Pencak, Z.: Reliability analysis of communications system used in military action. In: Conference Proceedings Military Arms 96, Gdynia (1996)
7. Pencak, Z.: Telecommunications network engineering. Theoretics. WSISiZ scripts, Warsaw (2000)
8. Poręba, J.: Telecommunications metrology. ITA Scripts of Wroclaw University of Technology, Wroclaw (2001)
9. Powierza, L.: Engineering basics of bioagricultural systems. Publishing House of Institute for Sustainable Technologies, Radom (1997)
10. Sienkiewicz, P.: System effectiveness theory. Publishing House Ossolineum, Warsaw (1987)
11. Siergiejczyk, M.: Modelling of Functional Efficiency of Selected Telematics Systems on Highways. Archives of Transport, vol. 19(1-2). Publishing House WDN PAN, Warsaw (2007)
12. Siergiejczyk, M.: Exploitation Effectiveness of Transport Telematics Systems. Scientific Papers of Warsaw University of Technology. Transport Series. Bulletin No. 67. Publishing House OW PW, Warsaw (2009)
13. Stankiewicz, R., Jajszczyk, A.: Methods used to ensure the guaranteed quality of services in IP networks Telecommunications Digest and Telecommunications News no. 2/2002. Publishing House Sigma NOT, Warsaw (2002)
14. Willems, G., Arijs, R., Van Parys, W., Demeester, R.: Capacity vs. Availability Trade-offs in Mesh-Restorable WDM Networks. In: Proceedings of 3rd International Workshop on the Design of Reliable Communication Networks (DRCN 2001), Budapest (2001)

Traffic Flow Analysis Based on the Real Data Using Neural Networks

Teresa Pamuła

Silesian University of Technology, Faculty of Transport
Krasińskiego 8, 40-019 Katowice, Poland
teresa.pamula@polsl.pl

Abstract. The paper presents the analysis of traffic data for determining classes of time series of traffic flow intensity for use in traffic forecasting employing neural networks. Data from traffic detectors on the main access road to the city of Gliwice in the period of past year is the basis for statistical analysis. Four classes of time series are proposed as representative of the traffic flow. The time series map temporarily smoothed detector counts. Different smoothing periods are used to retain the dynamic characteristics of the flows. A neural network is developed to classify incoming traffic data into the proposed time series classes. The specific time series implies a traffic control or management strategy, which indicates the capability of the NN to work out decisions for use in Intelligent Transportation Systems (ITS) applications.

Keywords: traffic flow analysis, traffic flow classification, prediction, time series, neural network.

1 Introduction

The traffic intensity characteristics are prerequisite for identifying trends in the distribution of traffic in the roads' network, for prediction of the road capacity, for determination of the variability of traffic in time.

The traffic description data plays a crucial role in the functioning of Intelligent Transportation Systems (ITS) applications. These provide the road users with e.g. routing information for efficient travelling through the network and enable the traffic supervisors to adapt control strategies to the traffic demand [1].

Neural networks (NN) are proved to be efficient tools for identifying the characteristics of traffic data for ITS [2, 3]. The ability to map historical traffic flow patterns, especially for nonlinear and dynamic evolutions is the essence of applications. NN are used for forecasting traffic flow parameters in different time horizons. Short term prediction is especially useful for determining control decisions at road traffic junctions [4, 5]. NN modelling of traffic flow is incorporated in adaptive signal control algorithms for evaluating performance measures in the optimization horizon [6, 7].

Application of NN for estimating parameters requires careful development of the networks and correct training of their operation. NN are sensitive to "over training"

J. Mikulski (Ed.): TST 2012, CCIS 329, pp. 364–371, 2012.

which means that the set of reference inputs is memorized instead of being characterized by the weights of the network. Minimizing network training errors is a non-convex problem, diverse methods are used to search for the global optimum.

The preparation of training sequences is an important task, in many cases determining the success of accurate modelling of the phenomena - in this instance traffic flow. Removing measurement noise is an important task in adapting measurements for training the NN's [8, 9]. Averaging in moving windows and exponential smoothing are methods commonly used to suppress noise. The size of the window and smoothing factor are the crucial variables which are optimized to achieve the desired robustness of data for NN training and validating.

In [10] it is shown that an aggregation of variables describing temporal characteristics of traffic flows at different time scales, gives an advantage in forecasting. Although variables are all based on the same detector data the temporal filtering reveals characteristics aiding the prediction. Temporal filtering is particularly useful for depicting incidents and overlapping phenomena with variable dynamics.

This paper presents the proposition of combining temporarily filtered data from traffic detectors into classes representing traffic flows requiring distinct actions in the domain of traffic control or management. Four classes of time series, derived on the basis of statistical analysis of detector counts, are proposed.

A neural network is developed to classify incoming traffic data into the proposed time series classes. Prepared sequences of traffic data are used for verification of the performance of the NN.

The main contributions of the paper are the elaborated method of combining detector data and the definition of criteria for this operation. The resulting sequences of data streamline the application of NN for traffic flow prediction.

The first section of the paper presents the detector setup and characteristics of the data set from the detectors. The second section discusses statistical parameters and criteria for classification of the detector data. Some details of the organisation of the developed NN are presented in section 3. These are supplemented with results of training and classification runs. The concluding section summarises the analysis and proposes further investigation topics.

2 Traffic Flow Database

The traffic flow database was prepared using data from traffic detectors placed on the main access road to the city of Gliwice. Map on fig. 1 illustrates the positions of the detectors. These detectors process video streams from cameras mounted above traffic lanes [11, 12]. The reliability of data was confirmed by manual vehicle counts carried out at random intervals during the collection of measurements for the database.

The database contains vehicle counts for 5 minute intervals carried out during the past year from June 2011 till May 2012. Table 1. summarizes the contents. There are in all over 200 thousand vehicle counts from two sets of detectors on traffic lanes covering inbound and outbound traffic to Gliwice.

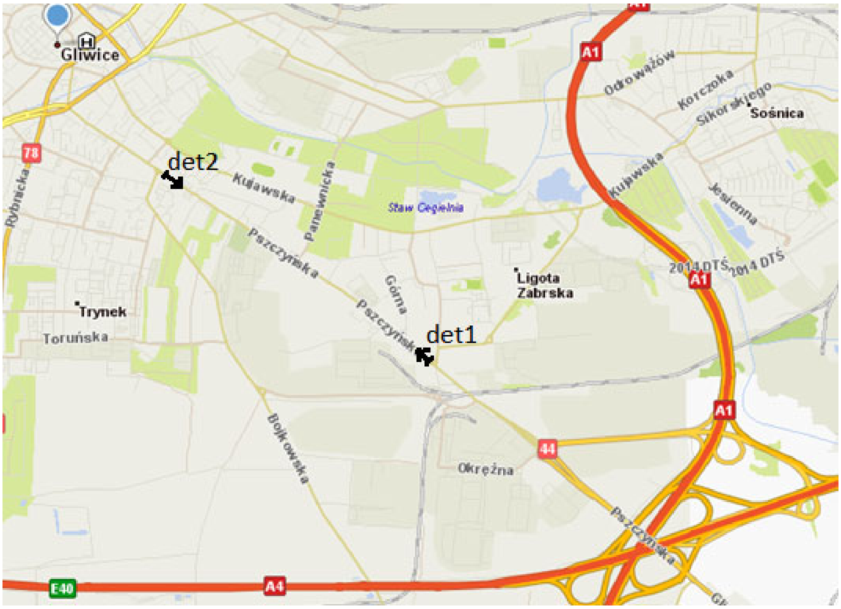


Fig. 1. The placement of the detectors on the Pszczyńska street in Gliwice (source: Targeo)

Table 1. Traffic flow database contents

traffic direction	number of detectors	number of vehicle counts
Inbound	1	101 286
Outbound	1	101 666

The vehicle counts were summed for 15 min intervals and converted to traffic flow values. These values were next scaled to traffic flow intensities per hour.

The urban traffic is characterised by very high intensities at rush hours. Morning peak is generally sharper than the afternoon peak, which is more extended in time. The intensity in the rush hours usually reaches 8 - 12% of daily traffic intensity. Significant fluctuations of traffic, occurring on particular days, strongly depend on the seasonal activities of the drivers. During weekends traffic intensity decreases - with the exception of weekend return hours and some holidays.

The 5 minute counts give a dense highly noisy graph of the traffic flow. These may be suitable as input for local traffic controllers. The 15 minutes interval is the reference period for traffic capacity analysis.

An important element of the traffic flow intensity analysis is the inspection of the traffic fluctuations, related to the following time factors: season, day of week, time of day.

3 Analysis of Traffic Flow Data

The aim of the analysis was to determine the similarity of time series of traffic flow values for different time scales. Season changes, day of week and intraday changes were investigated. The values of mean flow, MAD and MAX flow were chosen as the criteria for combining time series data together.

3.1 Seasonal Fluctuations (Monthly)

Data was analysed on monthly basis for capturing seasonal fluctuations. The average traffic flow and MAD of its values was calculated for every month. The mean value Q_{AV} includes working days and holiday traffic:

$$Q_{AVM} = \frac{1}{k}\left(\sum_{i=1}^{k}\left(q_{M\,\det 1} + q_{M\,\det 2}\right)\right) \tag{1}$$

where:

q_{Mdet1}, q_{Mdet2} – the average daily traffic flow for working days and holidays recorded by detector 1, detector 2, k – number of days in the month.

Table 2. Monthly statistics of traffic flow

Month	The average traffic flow in the working days $Q_{AV(det1+det2)}$ [veh/h]	mean absolute deviation (MAD)
June 2011	918	30%
July 2011	764	8%
August 2011	728	3%
September 2011	780	10%
October 2011	754	7%
November 2011	710	1%
December 2011	656	7%
January 2012	633	10%
February 2012	608	14%
March 2012	669	5%
April 2012	614	13%
May 2012	639	10%

Table 2 presents the values of seasonal fluctuations of traffic flow. The largest difference of mean flows is observed between June and February. The months have scattered values of MAD reflecting a wide spectrum of traffic flow graphs.

Seasonal fluctuations analysed on monthly basis exhibit large variations, which may be accounted for changing number of holidays in the months, unusual weather conditions (harsh winter, rainy summer in 2011, that is conditions when drivers are reluctant to use cars). The combining of monthly traffic flows leads to large discrepancies in representing traffic flows.

3.2 Daily Characteristics

The 3D presentation of traffic flow values in fig. 2 reveals consistently similar values of traffic on week days for particular months. Further investigation proves that some week day time series may be combined together. The intensities of the traffic flow measured in the working days from Monday to Friday are comparable. The largest differences appear on Saturday's and days before holidays, these are caused by the fact that for the majority of the traffic participants these are not working days.

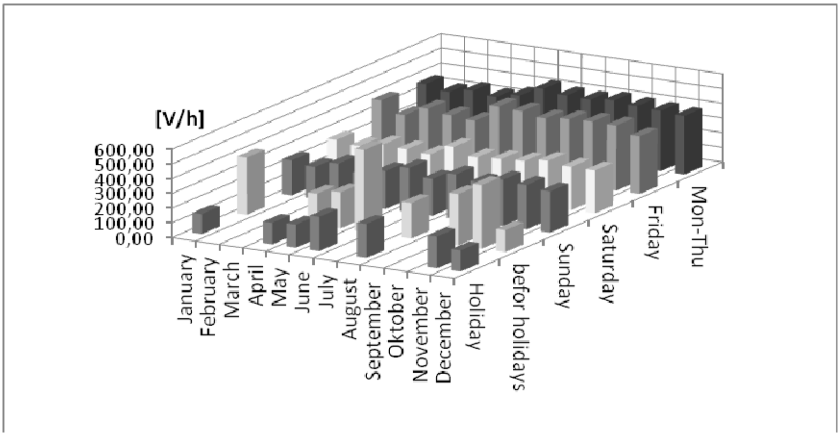


Fig. 2. The average traffic flow for 12 months for selected days

Table 3 presents the values of daily traffic flows. Mean traffic flow for working days Monday to Friday fall within a 1,5 % bracket of values. Corresponding MAX values of the flow in the morning peak confirm the similarity of flows. It is proposed to combine these graphs together forming a class of working day flows.

Table 3. Daily statistics of traffic flow

Day	The average traffic flow [veh/h]	MAX flow in the morning peak 7.00 – 9.00
Monday	393	683
Tuesday	393	676
Wednesday	394	664
Thursday	396	653
Friday	399	634
Saturday	272	390
day before holiday	294	404
Sunday	265	278
Holidays	171	108

Another distinct class is formed by holiday flows which are characterised by low values, less than half of work day flows, and gentle graph slopes. Saturdays and days preceding holidays have larger flows than the previous class and these can also be combined together. The last proposed class is placed in values between Saturdays and holidays with a humpy peak flow in the evening.

Fig. 3 shows the inbound traffic to Gliwice: on the left graphs partly cover each other when the day belongs to the proposed class, on the right combined flows forming the proposed classes of flows are presented.

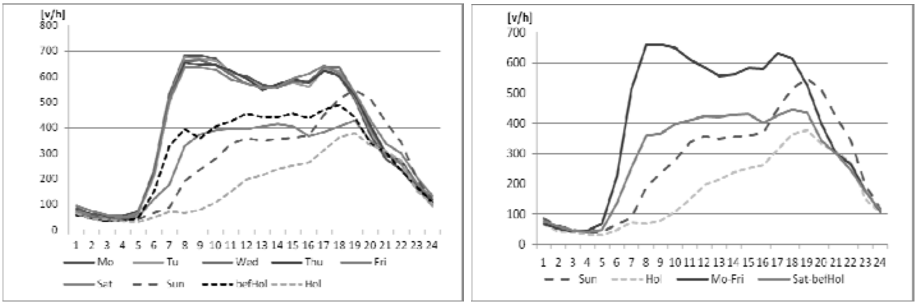


Fig. 3. Traffic time series for all weekdays for vehicles going to the city center

Using elementary statistic parameters it is possible to define distinctly flow classes of daily time series for use with NN applications. These are: workdays Mondays-Fridays, Saturdays and days before holidays, Sundays and holidays.

4 Neural Network

A neural network was developed for the purpose of verification of the capability to distinguish proposed traffic flow classifications [13]. Outputs of the NN are to be used for signalling the need of changing a traffic control or management strategy for use in Intelligent Transportation Systems (ITS) applications. The traffic flow class is recognised basing on values from a window of the time series of traffic flow. The size of the window is minimised just to retain the recognition capability thus reducing at most the signalling delay.

4.1 The Structure of Neural Network

A back-propagation neural network, with six inputs and four outputs and two hidden layers was proposed for investigations as in fig. 4. The number of neurons in the hidden layers were determined experimentally on the basis of authors previous experience in constructing classifying NN's. The proposed network has the structure of 6-30-40-4. The inputs are the values of traffic flow from the analysis window. The window is updated in hourly intervals and contains past 6 values of the traffic flow.

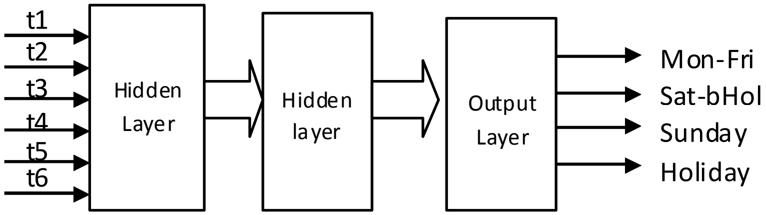


Fig. 4. The structure of neural network for recognizing flow classes

Network outputs correspond to the four proposed classes of traffic flow.

4.2 The Learning Sequences for the Neural Network

The learning sequence consisted of 400 input vectors of 6 elements, 100 for each class. Each input vector comprised of the mean values of traffic flow calculated in hourly periods based on data from the flow database. Output vectors corresponding to the appropriate flow classes were build as sequences of zeros and ones with a 1 at the position of the class.

4.3 Classification Results

The proposed neural network has been implemented and tested. Randomly chosen, from different months and days of week, sequences of traffic flow mean values were used to evaluate the classification performance of the trained NNs.

Preliminary results show that the holiday class and working day class were correctly recognized in 90% and 65% of the tests. Other classes were poorly recognized for morning and evening hours sequences. These sequences contain small values of traffic flow with ambiguous trends of changes.

5 Conclusion

The statistical analysis of traffic flow enabled a distinct classification of traffic flows. Four classes of time series are proposed as representative of the traffic flow. The elaborated method uses basic statistical measures for distinguishing these classes. Different smoothing periods are used to retain the dynamic characteristics of the flows. The chosen criteria for classification are adequate. The developed NN for classification of traffic flow requires further modifications to enhance its recognition rates. An increase of the time resolution of traffic flow values for the NN inputs seems a promising direction for further research.

Acknowledgments. The author wishes to thank ZIR-SSR Bytom for providing video detector data from the Gliwice site.

References

1. Karlaftis, M.G., Vlahogianni, E.I.: Statistical methods versus neural networks in transportation research: differences, similarities and some insights. Transportation Research, Part C. Emerging Technologies 19(3), 387–399 (2011)
2. Pamuła, T.: Road traffic parameters prediction in urban traffic management systems using neutal networks. Transport Problems 6(3), 123–129 (2011)
3. Chrobok, R., Kaumann, O., Wahle, J., Schreckenberg, M.: Different methods of traffic forecast based on real data. European Journal of Operational Research 155(3), 558–568 (2004)
4. Vlahogianni, E.I., Karlaftis, M.G., Golias, J.C.: Optimized and meta-optimized neural networks for short-term traffic flow prediction:agenetic approach. Transportation Research Part C 13, 211–234 (2005)
5. Dia, H.: An object-oriented neural network approach to short-term traffic forecasting. Eur. J. Oper. Res. 131(2), 253–261 (2001)
6. Cai, C., Wong, C.K., Heydecker, B.G.: Adaptive traffic signal control using approximate dynamic programming. Transportation Research Part C 17(5), 456–474 (2009)
7. Srinivasan, D., Choy, M.C., Cheu, R.L.: Neural networks for real-time traffic signal control. IEEE Trans. Intelligent Transportation Systems 7(3), 261–271 (2006)
8. Chen, H., Grant-Muller, S., Mussone, L., Montgomery, F.: A study of hybrid neural network approaches and the effects of missing data on traffic forecasting. Neural Computing and Applications 10, 277–286 (2001)
9. Quek, C., Pasquier, M., Boon, B., Lim, S.: POP-TRAFFIC A Novel Fuzzy Neural Approach to Road Traffic Analysis and Prediction. IEEE Transactions on Intelligent Transportation Systems 7(2), 133–146 (2006)
10. Tan, M.-C., Wong, S.C., Xu, J.-M., Guan, Z.-R., Zhang, P.: An Aggregation Approach to Short-Term Traffic Flow Prediction. IEEE Transactions on Intelligent Transportation Systems 10, 60–69 (2009)
11. Płaczek, B.: A Real Time Vehicle Detection Algorithm for Vision-Based Sensors. In: Bolc, L., Tadeusiewicz, R., Chmielewski, L.J., Wojciechowski, K. (eds.) ICCVG 2010, Part II. LNCS, vol. 6375, pp. 211–218. Springer, Heidelberg (2010)
12. Pamula, W.: Vehicle Detection Algorithm for FPGA Based Implementation. In: Kurzynski, M., Wozniak, M. (eds.) Computer Recognition Systems 3. AISC, vol. 57, pp. 585–592. Springer, Heidelberg (2009)
13. Michalik, K.: Neuronix – Symulator sztucznych sieci neuronowych. Podręcznik użytkownika, Katowice (2006)

Background Suppression for Video Vehicle Tracking Systems with Moving Cameras Using Camera Motion Estimation

Przemysław Mazurek and Krzysztof Okarma

Higher School of Technology and Economics in Szczecin
Faculty of Motor Transport,
Klonowica 14, 71-244 Szczecin, Poland
{mazurek,okarma}@wste.szczecin.pl

Abstract. Camera oscillations and slight movements are typical in the video based parts of the Intelligent Transportation Systems, especially in the cases when the cameras are mounted on the high pylons or pillars, similarly as some street lamps. The influence of strong wind and some vibrations caused by some heavy vehicles may result in some shifts of the images captured as the consecutive video frames. In such situations some typical background estimation and removal algorithms based on the comparison of corresponding pixels of each video frame may lead to significant errors. The influence of such camera motions increases seriously for high focal length corresponding to tracking of distant objects. In order to minimize the influence of such movements the background suppression algorithm using the camera motion estimation is proposed in the paper increasing the stability of the estimated background which is further used in the vehicle tracking algorithm.

Keywords: background suppression, vehicle tracking, motion estimation.

1 Introduction

One of the main purposes of the Intelligent Transportation Systems (ITS) is allowing the improvement of many traffic parameters, such as reduction of the number of traffic jams, shortening the travel time, reduction of toxic emissions by vehicles as well as the reduction of the number of disturbing events (e.g. traffic accidents). Such systems consist of three major subsystems: measuring (acquisition of the traffic data), controlling (containing the traffic control algorithms) and executive (directly affecting the road users). A crucial part is the measuring subsystem containing many road sensors [1]. The most typical ones are the induction loops, WIM (Weight in Motion) sensors or radar sensors (Doppler based) allowing the measurements at a specified fragment of a road regardless of weather conditions. Such measurements can also be conducted at more locations but their cost increases significantly. Furthermore, not all parameters can be revealed using such sensors, so an interesting supplement for them can be video

J. Mikulski (Ed.): TST 2012, CCIS 329, pp. 372–379, 2012.

Fig. 1. The illustration of the image shifts introduced by the camera motion

based measurements. Using the cameras located near the roads many motion and traffic parameters can be measured simultaneously but the working range of video systems and their measurement capabilities strongly depend on weather conditions.

The cameras used in the ITS solutions are typically mounted on dedicated pylons, high pillars, gates or, in some cases, at the buildings' walls. Installations of cameras on the street lamps are rather not used because of their elasticity causing noticeable wavings on the wind affecting the stability of acquired images. In such case the camera's vibrations may significantly influence its working area and range. Such changes are of great importance especially for the cameras with narrow angle lenses, since the acquisition of the video data corresponding with distant objects may be troublesome because of the wind's influence.

Each change of the visible area influences the results of tracking [2], also using the high dynamic range approach [3], especially on the background estimation results, since the commonly used algorithms for such purposes can be applied assuming the constant location of the camera without any motions. Each motion of the camera can be interpreted as the motion of the objects causing the intensity changes of some pixels so further detection is more difficult. Additionally, such motion influences the tracked vehicles' location on the consecutive images acquired by the camera, introducing the changes of objects' velocities both in horizontal and vertical direction. Considering the issues addressed above, a great importance of the image stabilisation leading to the elimination of the wind's influence in the ITS solutions can be noticed. An exemplary illustration of the possible results of the camera motion on the stability of the acquired images is shown in Fig. 1 where the translations can be easily noticed at the image boundaries (only two vehicles are in motion). Fig. 2 contains the results of the background estimation [4,5] for two methods discussed in detail in some earlier papers [6,7,8]: exponential smoothing (with smoothing parameter $\alpha = 0.95$) and median algorithm together with the differences between the "ground truth" background and its estimates. The images have been generated as synthetic using the captured sequence of images for the static camera in order to determine the reference background image. As illustrated both algorithms converge to the reference background image, especially for low density traffic, but the character of the absolute estimation errors during the first phase is different for both methods (global for the exponential smoothing and local for the median algorithm).

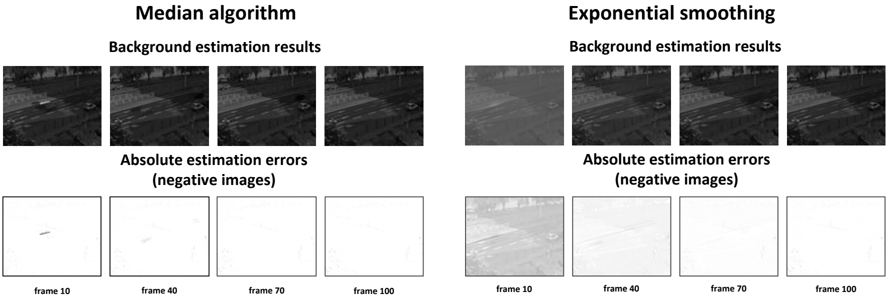


Fig. 2. The exemplary results obtained for two different background estimation algorithms (median algorithm and exponential smoothing)

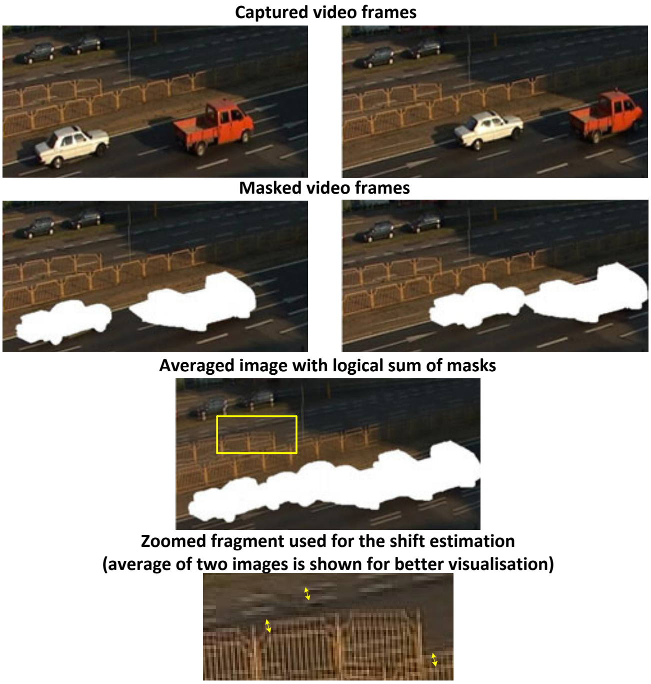


Fig. 3. The idea of masking applied for the image shift estimation

2 Methods for Estimation of the Image Shifts

Estimation of the shifts between a pair of images requires the knowledge about the area where the vehicles can move. It is essential for the elimination of the influence of the vehicles' motion, visible as changes of their locations in consecutive video frames, on the results of the shift estimation. It is necessary to determine a mask covering the representations of some physically moving objects on the image plane. This can be conducted "by hand" with necessary human assistance or automatically using a dedicated algorithm e.g. using shape classification for each video frame [9] in order to determine the possible fragments of images representing the vehicles. Without such masking the impact of the vehicles' shifts relative to the whole image can lead to significantly disrupted results of camera shift estimation. The idea of such approach is illustrated in Fig. 3.

Image matching algorithms may utilise the camera motion model. The application of some estimators, such as Benedict-Bordner, Kalman or Bayes filters [10] allows the reduction of the amount of calculations due to the prediction of the possible camera positions. Since the scene is roughly static and the motion range is limited, such solutions can be replaced by some simplified algorithms based on the image matching for a specified possible motion range.

Image stabilisation methods are based on two approaches:

– image matching according to the characteristic points,
– image matching according to the whole image plane.

The first solution is characterised by the reduced amount of calculations. It is typically applied for the image matching supervised by humans. Its main advantage is probably the robustness to the presence of some strong distortions in the image. The second approach is more sensitive to the preence of noise but the results are obtained using more informations from the larger area of the image.

One of the most typical matching criterions can be expressed as:

$$E(k_x, k_y) = \sqrt{\sum_{x,y} \left[I(x-k_x, y-k_y) - R(x,y)\right]^2} = \sqrt{\sum_{x,y} d(x,y,k_x,k_y)^2} \quad (1)$$

where E is the matching function, k_x, k_y are the horizontal and vertical shifts, I denotes the input image, R is the reference image and x, y are the pixels' coordinates. Nevertheless, this criterion takes into account only the translations without rotations and scaling (which is typically not used). The solution is:

$$E(k_x^{est}, k_y^{est}) = \arg\min_{k_x,k_y} E(k_x, k_y) \quad (2)$$

The translation causes the change of the image area used for the comparison, so the obtained result should be corrected considering the number of pixels used for the comparison:

$$E^*(k_x^{est}, k_y^{est}) = \frac{1}{N} E(k_x^{est}, k_y^{est}) \quad (3)$$

where N is the number of pixels used in the comparisons considering the thresholding and binary masks applied for both images.

The comparison of images is conducted with some errors being the result of noise and light fluctuations. For those reasons a modified criterion using the threshold T can be applied as:

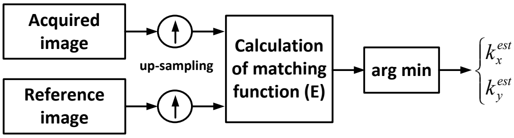


Fig. 4. The idea of image matching with up-sampling and nearest neighbour interpolation

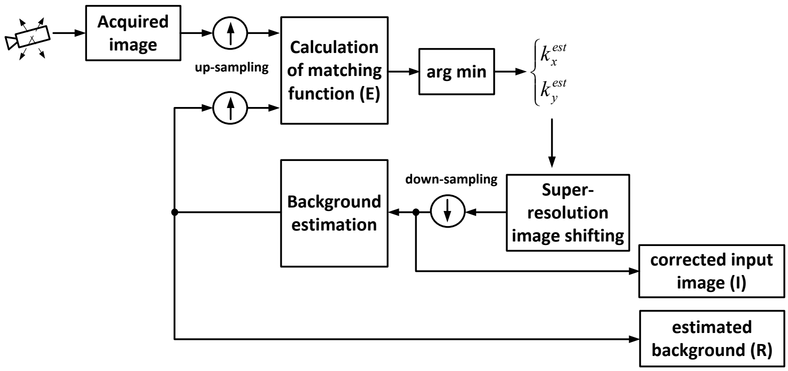


Fig. 5. The idea of the background estimation with sub-pixel image matching for camera motion estimation

$$E^*(k_x^{est}, k_y^{est}) = \frac{1}{M}\sqrt{\sum_{x,y} d(x,y,k_x,k_y)^2 \cdot [d(x,y,k_x,k_y) > T]} \qquad (4)$$

where M is the number of pixels fulfilling the thresholding condition (for which the difference between the compared pixels exceeds the threshold T) and d is defined as shown in equation (1). The purpose of this operation is the decrease of the influence of noise on the matching results.

The image matching operation is conducted typically with the assumption of integer values of the translations k_x, k_y corresponding to the coordinates of the image pixels. Nevertheless, for the images acquired from the real cameras can be characterised by sub-pixel translations. In such case the image matching can be conducted using up-sampling with nearest-neighbour interpolation as illustrated in Fig. 4. The idea of the backgorund estimation using the camera motion estimation is shown in Fig. 5.

Since the images acquired from the camera during its vibrations may be considered as the observations from slightly different locations, a negative wind phenomenon can be utilised for increasing the resolution of the acquired images. For this purpose the super-resolution algorithms [11,12] can be successfully

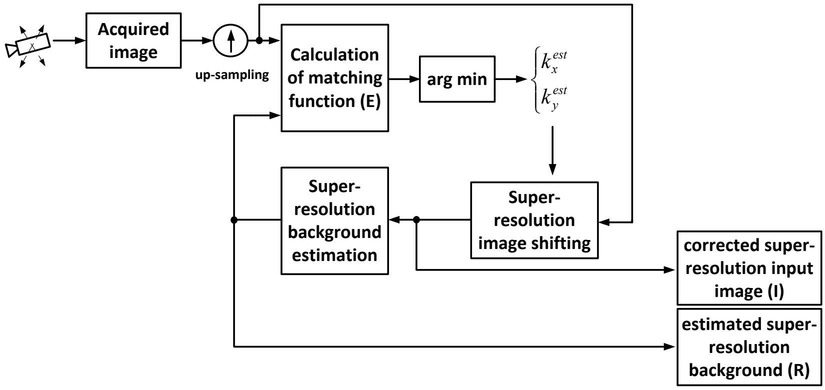


Fig. 6. The idea of the super-resolution background estimation based on the camera motion estimation

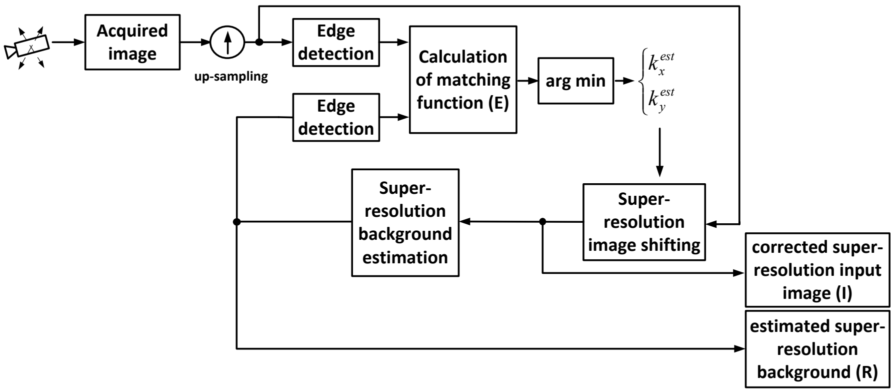


Fig. 7. The idea of the background estimation based on the camera motion estimation and edge detection (without masking)

applied leading to increase of the image resolution based on the series of shifted images. The idea is illustrated in Fig. 6.

Another solution may be based on the image matching based on some characteristics points on the image plane. Nevertheless, some lighting changes between two images can influence on the result of matching, so the method should be independent on such changing lighting conditions. For this purpose some edge detection algorithms may be applied, such as e.g. Prewitt, Sobel, Scharr or Canny filters, allowing detection of image fragments with rapid intensity changes, equivalent to the results of high-pass spatial image filtering, assuming no presence on noise. The idea of such system is illustrated in Fig. 7. Nevertheless, for simplicity, the masking operation is not included in figures.

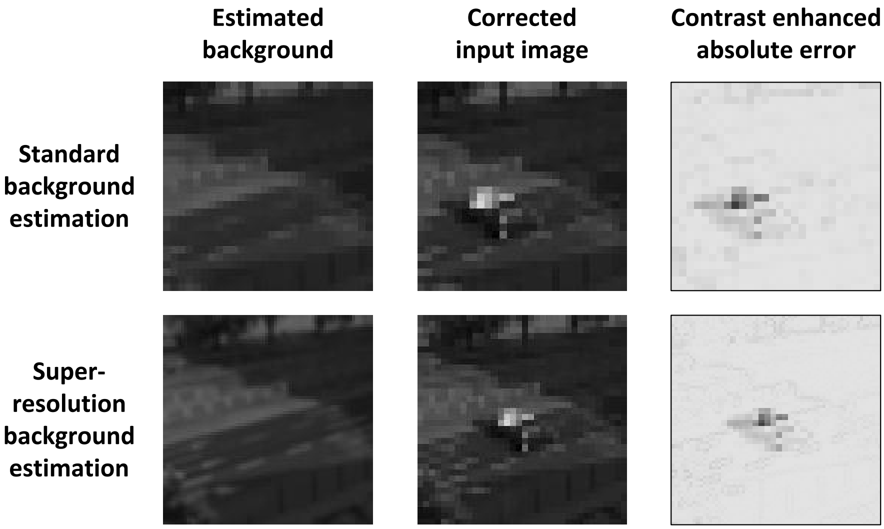


Fig. 8. Illustration of the exemplary results obtained for two different background estimation methods with the super-resolution camera motion detection

3 Results and Conclusions

In order to verify the properties of the discussed methods and their proposed modifications some experiments has been conducted using the implemented algorithms for the background estimation. Analysing the obtained results the advantages of using the methods utilising the super-resolution technology for the image shifting can be easily noticed. The greatest benefits can be observed for the super-resolution background estimation where the details of the obtained image are quite smooth.

Comparing the absolute errors presented on the right side of Fig. 8 for the frame with two vehicles visible on the road, an interesting property of the super-resolution background estimation algorithm can be observed. The largest errors resulting from the presence of vehicles are similar but the errors related to the background area are different for both methods. For the standard method some modified pixels can be observed, which occupy relatively big area of the image. Nevertheless, the corresponding fragments of the image affected by errors are much smaller so their influence on the overall result of the background estimation is also reduced. In this case only the sub-pixels are changed and for the standard method the whole pixels are changed as the smallest possible parts of the image.

The application of the super-resolution background suppression algorithm based on the super-resolution camera motion estimation can be applied as an effective pre-processing step of the video based vehicle tracking in the ITS solutions. A promising direction of further research seems to be the verification of this algorithm in the sophisticated tracking systems for small objects e.g. using Track-Before-Detect approach [13], which is especially useful for tracking distant objects in the presence of a strong noise.

Acknowledgements. This work is supported by the Polish Ministry of Science and Higher Education (Grant No. N509 399136 "Estimation of the vehicles' motion trajectories using the Bayesian analysis and digital image processing algorithms").

References

1. Klein, L.: Sensor Technologies and Data Requirements for ITS. Artech House ITS library, Norwood (2001)
2. Mazurek, P., Okarma, K.: Application of Bayesian a Priori Distributions for Vehicles' Video Tracking Systems. In: Mikulski, J. (ed.) TST 2010. CCIS, vol. 104, pp. 347–355. Springer, Heidelberg (2010)
3. Okarma, K., Mazurek, P.: Vehicle Tracking Using the High Dynamic Range Technology. In: Mikulski, J. (ed.) TST 2011. CCIS, vol. 239, pp. 172–179. Springer, Heidelberg (2011)
4. Piccardi, M.: Background subtraction techniques: a review. In: IEEE International Conference on Systems, Man and Cybernetics, vol. 4, pp. 3099–3104 (October 2004)
5. Reddy, V., Sanderson, C., Lovell, B.C.: A low-complexity algorithm for static background estimation from cluttered image sequences in surveillance contexts. EURASIP Journal on Image and Video Processing, Article ID 164956, 14 (2011)
6. Okarma, K., Mazurek, P.: Background estimation algorithm for optical car tracking applications. Machinebuilding and Electrical Engineering (7-8), 7–10 (2006)
7. Okarma, K., Mazurek, P.: Nonlinear background estimation methods for video vehicle tracking systems. Archives of Transport Systems Telematics 4(4), 42–48 (2011)
8. Okarma, K., Mazurek, P.: A modified hybrid method of nonlinear background estimation for vision based vehicle tracking systems. Logistyka (3), 1747–1752 (2012) (in Polish)
9. Okarma, K., Mazurek, P.: Application of Shape Analysis Techniques for the Classification of Vehicles. In: Mikulski, J. (ed.) TST 2010. CCIS, vol. 104, pp. 218–225. Springer, Heidelberg (2010)
10. Blackman, S., Popoli, R.: Design and Analysis of Modern Tracking Systems. Artech House (1999)
11. Irani, M., Peleg, S.: Super resolution from image sequences. In: Proceedings of the International Conference on Pattern Recognition (ICPR), Atlantic City, New Jersey, vol. 2, pp. 115–120 (June 1990)
12. Irani, M., Peleg, S.: Improving resolution by image registration. Graphical Models and Image Processing 53, 231–239 (1991)
13. Mazurek, P., Okarma, K.: Vehicle Tracking Using a Multi-scale Bayesian Algorithm for a Perspective Image of a Road. In: Mikulski, J. (ed.) TST 2011. CCIS, vol. 239, pp. 346–353. Springer, Heidelberg (2011)

Using Neural Networks for Route and Destination Prediction in Intelligent Transport Systems

Tomáš Mikluščák, Michal Gregor, and Aleš Janota

University of Žilina, Faculty of Electrical Engineering
Univerzitná 8215/1, 01026 Žilina, Slovakia
{tomas.mikluscak,michal.gregor,ales.janota}@fel.uniza.sk

Abstract. Route prediction and destination prediction based on the past routes are a missing piece in intelligent transport systems (ITS). These predictions can be useful in many areas: congestion prediction, traffic control, upcoming traffic hazards and targeting advertisements next to the roads are some of the obvious ones. Simply said, if we can estimate the future location of cars which are already on the road network, we will be able to estimate future congestions and upcoming traffic hazards. The GPS units in the new generation of smartphones provide a good data source for prediction algorithms. Google maps application already collects this data. This paper discusses several algorithms and methods which have been used in similar areas and a route prediction method based on artificial neural networks using the past routes of a vehicle.

Keywords: congestion prediction, neural networks, intelligent transport systems, prediction methods, location management.

1 Introduction

The aim of this paper is to transfer experiences from next location prediction problems which have been solved in smart environments, cellular networks and other to a very similar area: route and destination prediction problem. We propose a general route and destination prediction method which can be used to solve such problems. Several algorithms and methods for route prediction are discussed as well.

Traffic congestions are a major problem in big cities in developed and developing countries. Many different more or less successful approaches have been used to reduce congestions. Research indicates that traffic congestions can be alleviated by providing timely and accurate traffic information to drivers. With such information drivers can choose another route and avoid congested roads or even change departure time. The ITS able to predict congestions could potentially help traffic control to eliminate them, or even to spread information about future congestion so that motorists may adapt their decisions appropriately. Modern ITS are trying to collect data about actual traffic using variety of sensor systems like embedded roadway loop detectors, electronic toll collection transponders, automatic vehicle location, etc. They are also trying to predict congestion according to this data. Using this sensory data, ITS can predict congestions and even estimate route duration. Unfortunately, the

J. Mikulski (Ed.): TST 2012, CCIS 329, pp. 380–387, 2012.

capability of these systems to predict congestions is not sufficient. On the other hand, the ability to predict next routes and the final destination of vehicles may play a vital role in congestion prediction. Simply said, if we can predict next route and the final destination of vehicles which are on the roads already, we can also predict possible upcoming hazards, estimate future congestions and travel times with significantly higher accuracy. As a data source we can use data which is already being collected. Every modern smart phone and every navigation device is able to provide this data and commerce navigation companies do already collect it (e.g. Google maps). Among other sources of data, there are satellite tool systems.

Even though congestion prediction is the major concern of this paper, there are many other possible applications of route and destination prediction. Some of them will also be described in the following text.

The authors of [2] propose probabilistic reasoning from observed context-aware behavior (PROCAB) system. The authors envision a future navigation device that learns driving routes from a group of experts – taxi cab drivers – who have good knowledge of alternative routes and contextual influences on driving. This device is then navigating common drivers. In fact researches shows that drivers just rarely choose the fastest possible route. According to [3] just 35% and 38% according to [2] take the fastest way even if they are using commercial navigation software. The PROCAB includes three predictions: turn prediction, route prediction and destination prediction.

One of the most innovative applications of route prediction concerns improving efficiency of hybrid vehicles. Given knowledge of future changes in elevation and speed, a hybrid control system can optimize the vehicle's charge/discharge schedule. Researchers from Nissan showed that it is possible to improve hybrid fuel economy by up to 7.8% if route is known in advance [5]. While driver can be asked for his or her route before every drive, it is suspected that most drivers would tire of this quickly. This problem has been investigated in [3], where authors use K-order Markov Model to predict routes.

2 Prediction Methods

This part contains a small review of methods which have been used for route prediction or for prediction in similar areas of research.

2.1 To What Extent It Is Possible to Predict Routes?

The assumption underlying all route prediction systems is that human behaviour and particularly routing on roads can in principle be predicted. The analysis of inhabitant's daily lifestyle will tend to show identifiable patterns: these can be learned and therefore actions can, to a certain extent, be predicted. What is more, the underlying hypothesis that human behaviour and human movement can be predicted has to date been proven by many experiments, for example [2] and [3] where success rate of predicted route of common persons is higher than 85% in both cases. The ability to predict person moving have been proved in a similar area – smart house projects [6], [7] or smart office projects [8] as well. These researches indicate that in spite of the fact, that there exist a number of possible routes, the driver will tend to have their own routine, and their own routes.

2.2 From Cellular Networks and Smart Houses to ITS

The location prediction algorithms have been mainly developed and tested in cellular networks and smart houses. Pioneering in this area is Bhattacharya and his assumption that the inhabitant's mobility is a piecewise stationary ergodic stochastic process, as hypothesized in [4]. This work concerns tracking mobile users in a wireless cellular network. It has motivated a number of researchers to explore this area and a few very good algorithms and methods have been proposed. Among methods with higher accuracy of prediction that have been be used in smart homes are Active LeZi [6], Neural networks [8, 7], Bayesian networks [9], Hidden Markov models [10] and even algorithms based on game theory [11]. Later, route prediction has come into focus and K-order Markov models [3] and Markov decision process with inverse reinforcement learning [2] have been employed to accommodate this problem.

Algorithms based on game theory have been developed with the purpose of handling multi-inhabitant environment. Human behaviour is different if there are more persons in the same environment. Route prediction does not suffer from similar issues, so we will not consider this hereinafter. The major difference between using location prediction (person movement prediction) in smart pervasive environment (SPE) and route prediction in road nets lies in the size. While one will hardly find a net with more than 100 nodes in SPEs, 10 000 node road nets are not unusual. The Active LeZi is designed to work with this amount of data and in principle, Bayesian networks, or Hidden Markov models should be able to handle work with this amount of data as well. Our aim then is to explore whether an artificial neural network would be able to handle such a large amount data (the number of input and output neurons will be rather large when using the usual representation).

3 Using Artificial Neural Networks for Prediction

There is a number of papers that investigate applications of artificial neural networks (ANN) to forecasting, or, if you like, to time series prediction. Several points are to be addressed when investigating such applications, including:

- Selecting an appropriate architecture of the ANN;
- Finding an appropriate data representation;
- Specifying a way in which the data is to be coded into inputs for the ANN.

3.1 Architecture of the Artificial Neural Network

Perhaps the first choice to be made when designing an ANN-based forecaster is whether to use feed-forward architecture of recurrent architecture. In recurrent networks, as opposed to feed-forward network, neurons are allowed to have connections to outputs of neurons from the following layers.

Recurrent networks would seem to be especially well-suited to perform the task at hand. It is quite obvious that knowledge of several previous values (states) will generally be required in order to determine the forecast. It should be noted, however, that recurrent networks tend to be significantly more difficult to train.

Also, if we are willing to provide for time delay of input signals explicitly, feed-forward networks represent a viable alternative. In fact, according to [12], of these two, this is the more widely used approach.

To summarize the approach based on the feed-forward architecture, one may say that the data is fed to the ANN in the following fashion: Let $f(k)$ be a time series we are trying to predict (and k be the discrete time step). Then, if we leave coding of the values out of consideration for the time being, the following is to hold for data fed into the ANN at point k :

$$X(k) = \{f(k-1), f(k-2),\ldots, f(k-n)\}. \tag{1}$$

where n is the number of previous values of $f(k)$ we present to the network, that is, the order of the forecaster (or, if you like, the size of its memory). When specifying the value of n , one should therefore be reasonably certain that the system producing the outputs is of order equal or smaller than n .

3.2 Data Representation and Coding

As mentioned, the issue of coding has been left aside in the previous section. Let us now return to this point and give some attention to the various options that face us.

Many works concerning ANN-based forecasting deal with continuous data such as currency exchange rates, river flow, rainfall, financial and demographic indicators, etc. For such applications, what we have said so far is actually quite accurate.

The problem at hand is not, it should be obvious, of the same nature – identifier of a crossroad, route, etc. is a discrete value. There are several approaches one can take when devising a representation for such data:

- Feed the numeric code, such as it is, into a single artificial neuron (in the input layer);
- Have a separate input (and output) neuron for every room times every time step;
- Convert the numeric code into its binary form and assign one neuron to every bit of the resulting binary code.

The first option is the same as used for forecasting of continuous numeric variables. Such approach has some significant downsides, the most important of which is the fact that the ANN will not generalize correctly [7].

If we select the second option, we have to provide the ANN with $n.r$ input neurons, where r is the number of distinct routes. Input patterns would then be generated in such way that a given route would be represented by setting the input of its corresponding neuron to 1 while setting the inputs of all other input neurons to 0. While this approach is very useful for some smaller and medium-sized tasks, an ANN for our task would be prohibitively large.

Thus, we select the third approach, which provides something of a middle ground between the other two methods.

4 Experiments

The experiments presented hereinafter are based on data sets generated using a simulation model briefly described in the following sections. The experiments themselves compare the results achieved on two such datasets: *dataset 1*, which is an easy data set with very little stochasticity, and *dataset 2*, which, in contrast, represents a very difficult problem with a considerable amount of random decisions. Dataset 1 consists of 54031 samples. Dataset 2 has 69338 samples.

4.1 Simulation Model

Advantages of simulation when compared with experiments in real environments are obvious: a simulation model saves money and time. We have developed a simulation model which includes common driving of a common employed person in a small city (1000 crossroads). According to the day and time, the simulated person drives to work, to a shopping centre, church, football stadium and to visit parents, or friends. The average distance between these points is 50 crossroads. The simulated person sometimes just randomly drives to other places in the city.

It is obvious that there some downsides to this approach: we cannot really say how stochastic a real dataset would be and – on the other hand – how pronounced would the driving patterns be. Therefore we are working with two datasets that hopefully represent the boundaries. The main motivation behind this work is to ensure that such approach is feasible, and that it is possible to do this in real time for a reasonably large number of vehicles. It is obvious, however, that to gain some more realistic insights into what accuracy to expect, real datasets would serve better.

4.2 ANN Architecture, Programming Environment, Reference Hardware

In the following experiments a feed-forward network has been used to implement the forecaster. The ANN is trained using Rprop. The experiments have been carried in the Matlab environment.

The hardware platform used to perform the experiments is composed of the following components: 8GB of RAM; 1.6GHz Intel Core i7 720QM (quadcore, each core divided into 2 virtual threads).

Only a portion of the available memory and processor cores have been utilized in the course of learning and prediction. It is to be noted that Matlab parallelizes the computation to a certain extent, when presented with all data at once.

4.3 Results – Performance of the Forecaster

As for accuracy, the following results have been achieved when varying size of the training data set (the number of samples used to train the network; the rest of the samples is used to test network's performance). The performance is a ratio of correctly predicted routes to the total number of predictions.

For all of these results, architecture with two hidden layers, each with 500 neurons and sigmoid activation functions has been used. In every case, 5 previous values are presented to the network.

Table 1. Performance with varying number of samples used for training

Dataset 1		**Dataset2**	
Training dataset size	Performance (training dataset / testing dataset)	Training dataset size	Performance (training dataset / testing dataset)
2000	98.95% / **98.74%**	2000	98.95% / **64.27%**
1500	98.53% / **97.86%**	1500	99.27% / **64.45%**
1000	99.20% / **98.25%**	1000	99.20% / **57.29%**
500	100% / **89.41%**	500	99.40% / **56.69%**
100	100% / **59.79%**	100	100% / **44.39%**

4.4 Results – Elapsed Time

In this section, the training dataset size and the number of neurons in hidden layers vary. We compare the resulting training and execution times. The number of neurons is represented in the following way: [size of the first hidden layer, size of the second hidden layer]. We are using dataset2 in this comparison.

Table 2. Training time for [500, 500]; execution time for 67337 samples

Execution time		**Training time**	
Architecture	Execution time [s]	Training dataset size	Training time [s]
[500, 500]	4.233	2000	35.898
[200, 200]	1.575	1500	26.152

4.5 Evaluation of Results and Future Research

When looking at the results concerning performance, the need for testing on real data becomes increasingly obvious. The results achieved for the highly stochastic dataset are not very good – the performance approaches 60%. This is to be expected: the theoretical maximum is in fact not much greater for the data in question. It is shown that with an increasing size of the training data set, the results become better: this can be ascribed to the fact that the ANN is able to make better guesses concerning the probabilities related to the stochastic process.

As for the elapsed time comparison, we may conclude that the approach could be used in practice. The initial training can for the most part be done offline. The execution times are reasonable even for a large number of samples. If the computation were fully parallel and re-implemented in a more low-level language such as C or C++, this would be even more apparent.

5 Conclusion

The route and the destination prediction is a novel approach, able to solve many problems of current traffic systems. In this paper we have discussed a few methods, which have been used in a similar area: location prediction (person movement prediction) in smart pervasive environments (SPEs).

The major difference between using location prediction in smart pervasive environments and route prediction in the road net is the size of the problem. We have shown in this paper that it is possible to use artificial neural networks for route prediction. From the perspective of training and execution times, this approach should be tractable. As for performance of the forecaster itself, the results vary with the degree of stochasticity and the size of the training data set. To gain better understanding of this particular problem, testing on real data (as opposed to simulation data) would be advisable.

Acknowledgments. The paper has been elaborated with the support of the Slovak scientific grant agency VEGA within the research project "VEGA-1/0453/12 Study of interactions of a motor vehicle, traffic flow and road".

References

1. Karbassi, A., Barth, M.: Vehicle route prediction and time of arrival estimation techniques for improved transportation system management. In: IEEE Intelligent Vehicles Symposium, IV 2003, pp. 511–516 (2003)
2. Ziebart, B.D., Maas, A.L., Dey, A.K., Bagnell, J.A.: Navigate Like a Cabbie: Probabilistic Reasoning from Observed Context-Aware Behavior. In: UbiComp 2008 Proceedings of the 10th International Conference on Ubiquitous Computing, pp. 322–331 (2008)
3. Letcher, J., Krumm, J., Horvitz, E.: Trip router with individualized preferences: Incorporating personalization into route planning. In: Proc. IAAI, pp. 1795–1800 (2006)
4. Froehlich, J., Krumm, J.: Route Prediction from Trip Observations. SAE Technical Paper 2008-01-0201, 13 p. (2008)
5. Deguchi, Y., et al.: Hev Charge/Discharge Control System Based on Navigation Information. In: SAE Convergence International Congress & Exposition on Transportation Electronics, Detroit, Michigan USA (2004)
6. Roy, A., Das, S.K., Basu, K.: A Predictive Framework for Location-Aware Resource Management in Smart Homes. IEEE Transactions on Mobile Computing 6(11), 1270–1283 (2007)
7. Mikluščák, T., Gregor, M.: Person movement prediction using artificial neural networks with dynamic training on a fixed-size training data set. Applied Computer Science: Management of Production Processes 7(2), 43–56 (2011)

8. Vintan, L., Gellert, A., Petzold, J., Ungerer, T.: Person Movement Prediction Using Neural Networks. Universität Augsburg, Institut für Informatik, Technical Report 2004-10, 16 p. (2004)
9. Petzold, J., Pietzowski, A., Bagci, F., Trumler, W., Ungerer, T.: Prediction of Indoor Movements Using Bayesian Networks. In: Strang, T., Linnhoff-Popien, C. (eds.) LoCA 2005. LNCS, vol. 3479, pp. 211–222. Springer, Heidelberg (2005)
10. Gellert, A., Vintan, L.: Person Movement Prediction Using Hidden Markov Models. Studies in Informatics and Control 15(1), 17–30 (2006)
11. Roy, N., Roy, A., Das, S.K.: Context-aware resource management in multi-inhabitant smart homes: a Nash H-learning based approach. In: IEEE Int'l Conference on Pervasive Computing and Communications, vol. 11, pp. 148–158 (2006)
12. Wong, W.K., Xia, M., Chu, W.C.: Adaptive neural network model for time-series forecasting. European Journal of Operational Research 207(2), 807–816 (2010)

A New Approach to Road Safety in Slovakia

Alica Kalašová[1], Ľubomír Černický[1], and Milan Hamar[2]

[1] University of Žilina, Faculty of Operation and Economics of Transport and Communications
Univerzitná 1, 01026 Žilina, Slovakia
{alica.kalasova,lubomir.cernicky}@fpedas.uniza.sk
[2] Department of Traffic Police of the Police Force Presidium in Bratislava
Račianska 45, 812 72 Bratislava, Slovakia
Milan.Hamar@minv.sk

Abstract. The road traffic accident rate represents an important indicator of road condition level (structural and technical) as well as traffic conditions (traffic organization) influencing the level of fulfilment of traffic demand of the society. Therefore, the traffic accident rate is an important criterion in planning, building, rehabilitation, and maintenance of road network. The degree of motorisation is constantly increasing. Except undoubted advantages it causes a great growth of traffic volume of road network and a constantly increasing demand on traffic and its safety. The road safety is not only an important traffic and social issue, but also an economic one. The traffic accident rate is connected to heavy material damages, permanent bodily injuries and very often irreparable casualties. Therefore, special attention is being paid to the road safety and the influence of its particular factors all over the world and of course in the European Union. By joining the European Union we are committed to achieve standards in all areas that are comparable with European developed countries. In our paper we would like to analyse the road traffic accident rate in Slovakia and to point out new approaches to road safety.

Keywords: Road safety, development of road transport, reducing traffic accidents.

1 Introduction

The traffic safety can be defined as protection of life, health and property in the road traffic. So it reflects the ability of the system to operate at an acceptable level of risk to its surrounding as well as to system itself at the acceptable level [1]. So the safe transport means that each trip to the destination was made without any accident or feeling of danger. Traffic accidents are a traditional method of the road safety measurement.

The traffic accident is defined as an unfavourable event in the road traffic, for example an accident or a collision, which happens or starts on the road and results in death or personal injury or in a damage to property. We can say that a traffic accident is an imbalance between the three components of the system: human

J. Mikulski (Ed.): TST 2012, CCIS 329, pp. 388–395, 2012.

factor–environment–vehicle.It is a random event which can be caused by many things. Let us remind that according to current statistics, 93% of road accidents are caused by a human error [3].

The development of the society is connected to the development of transport in all its branches. Especially in the road transport there is an annual growth in the number of vehicles on the roads as well as a growth of new drivers that is connected with many negative effects. Because of he lack of conditions for implementation of traffic education, low discipline, aggressive driving, breaking the basic duties, and little knowledge of law of drivers and other road users, there is a growth in the number of road accidents and their consequences.

It is impossible to completely eliminate traffic accidents, but it is possible to continue decreasing the number and severity of traffic accident by the cooperation of scientific institutes, international organizations and experts from various fields of science with political leaders of the country. The traffic accidents, as a serious social problem, require a comprehensive and effective solution with signs of coordinated and targeted progress of all stakeholders and institutions and with a broad public support. Successful implementation of regulatory, preventive, repressive and technical measures can be an instrument for a good evaluation and comparison of the social and economic development of European countries.

2 Current Status of Traffic Accidents in the Slovak Republic

Slovakia has an advantageous geographical position in Central Europe and because of this it is an important traffic intersection, where the road traffic still dominates as a mode of transport. With increasing intensity and capacity of road traffic, its accompanying negative features are still more evident and the problems have been associated with the establishment and functioning of the transport system. The mentioned problems include mainly:

- negative impact on the environment,
- increasing number of traffic accidents, threats to human health and life,
- traffic congestion, breakdowns and other negative impacts.

Despite the fact that the accidents frequency in Slovakia has been decreasing over the last years, the consequences are alarming.

A significant reduction in accidents has been a desire of all stakeholders for many years. There were several documents, which were dedicated to achieving the desire. The White Paper, European transport policy for 2010: time to decide, set a policy target of reducing by half the number of people killed on European roads by 2010 [8].

In 2010, the national plan for increasing the road safety in Slovakia for the period 2010-2012 was prepared. The plan is based not only on the programs, but also on the principles contained in the First Global Ministerial Conference on Road Safety on November 20th, 2009, on the UN General Assembly resolution proclaiming a "Decade of Action" for Road Safety 2011-2020 (A/64/255), and also on the Communication from the Commission (2010) 389: policy orientations on road safety 2011-2020 [6]. The Slovak

Republic tried to achieve the objectives by decisive attributes, which mean the creation of legislative conditions to eliminate irresponsible behaviour of road users to each other and to enforce the obligations of road users are respected in practice, which were the decisive attributes [4]. In 2009 the Slovak Republic nearly met the obligation to the EU to reduce by half the number of people killed in road accidents by 2010. But the obligation was not met despite enormous effort to prepare legislative and technical conditions for achieving the goal of halving the traffic accidents consequences by 2010, which Slovakia had adopted by joining the European Union, on May 5th, 2004 [7]. See Fig. 1

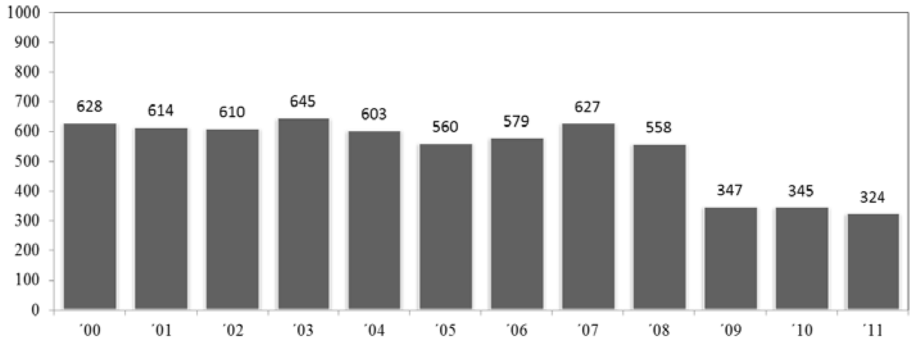


Fig. 1. The number of people killed in the Slovak Republic for the period 2000-2011

You can see in Fig. 2 that on the 31st December 2010 the required target of 305 road accident fatalities was not achieved in the Slovak Republic (305 means half of fatalities in comparison to 2002). This result is caused by several circumstances, but especially because the legislative-technical and educational field was not sufficiently prepared. Therefore we can objectively say that EU Member States, which adopted the commitment in 2001, were 3 years in advance in these 3 areas. It is clear that Slovakia had time for 7 years to meet the goal, while the "old EU Member States" had time for 10 years to meet the same goal. Although in 2011 there were 21 fewer people killed in road accidents as compared to 2010, Slovakia still does not meet the EU commitment. But we can notice that the target is met by only very few countries such as France and Portugal, which have met the target by intensive controlling or by increasing sanctions [5].

The reduction in the number of fatal accidents is supposed to continue till 2020 with the goal to reduce their number by 75% in comparison to 2001. For the Slovak Republic it means the reduction in fatalities to 172 people. From the road safety point of view the Slovak Republic should pledge to adopt and implement such legislative, technical and educational conditions for drivers of motor vehicles, so that in 2020 Slovakia can contribute to the EU target of reducing the number of fatalities.

On the basis of current trend we can say that without radical measures it will be impossible to meet the commitments. The only way to meet the commitments is a systematic implementation of ITS.

In 2011 Slovakia managed to decrease the number of road accidents to its historical minimum. There were registered 14,991 accidents, thatmeans 31% decrease in comparison to 2010 (Fig. 2).

Thanks to this in the history of the Slovak Republic there is the lowest:

- Number of people killed in road accidents; 324 (annual decrease of 6%).
- Number of people seriously injured in road accidents; 1140 (annual decrease of 6%).
- Number of people slightly injured in road accidents; 5905 (annual decrease of 15%).

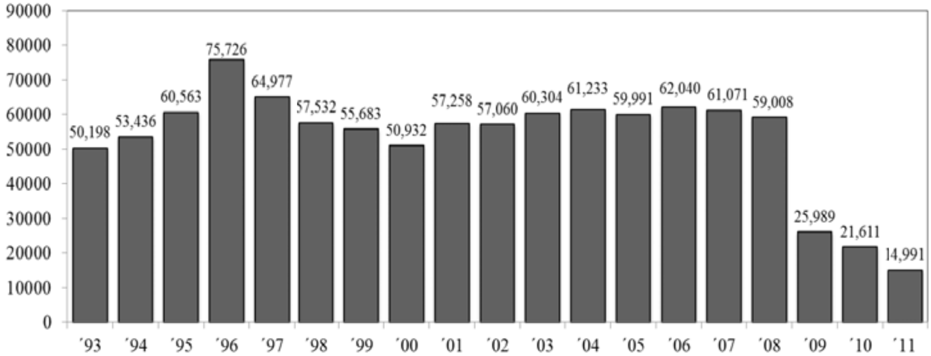


Fig. 2. The numbers of traffic accidents between 1993 and 2011

Because of constant development of motoring, every year new drivers and new vehicles became road users and it results in increasing the road traffic volume. This fact affects and determines to a large degree the overall evolution in the road safety situation. The mobility of the population is growing as well. Table 1 gives us an overview of the traffic volume in 2010 and 2011 by months.

Table 1. Overview of the traffic volume

Month\Year	2010	2011
January	2 772 260	3 133 386
February	2 987 539	3 318 760
March	3 813 967	4 029 021
April	3 554 637	3 867 415
May	3 750 848	4 152 496
June	4 011 888	4 086 690
July	3 998 524	3 777 902
August	3 952 798	3 962 443
September	4 081 055	4 057 408
October	4 211 267	4 043 414
November	3 990 317	3 871 065
December	3 241 558	3 191 628
Sum	44 366 658	45 491 628

If we take into account the traffic volume, which in 2011 raised by 2.5% while traffic accidents decreased by 31% (see Fig. 2), we can say that there is a very positive development in the Slovak Republic. Fig. 3 is the evaluation of accidents in Slovakia in the context of Central European countries, which shows us that among these states the Slovak Republic has the best results in reducing the traffic accidents.

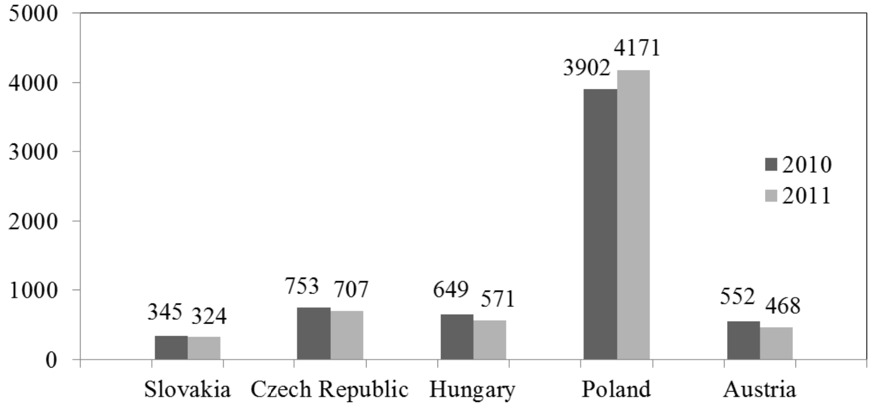


Fig. 3. People killed in road accidents in Central European countries

3 The Measures for Reducing Traffic Accidents

The road traffic safety is a measure of society-wide effort expressed in the quality of road infrastructure, skills of drivers, and in the level of IT technical equipment. Police have a role in protecting participants from low discipline and thoughtfulness. The reduction of traffic accidents was conditioned by [2]:

1. The improvement in the preventive part, where one of the most important measures is the implementation of information technologies into transport systems. The traffic safety solutions based on information systems should increase in the future (with respect to repressive component). An implementation of a good information system results in reducing traffic accidents and also in decreasing the number of seriously and slightly injured people. Nowadays, there are various portals where drivers can find information about their trips. On the website of the Ministry of Interior, in the Traffic Police section, there is information about accidents in the Slovak Republic, most frequent black spots, as well as about the capacity of bridges and underpasses, which were added just this year to better inform the public. There are some other websites that inform the public about the current state of road traffic safety by various locations, namely websites of the Ministry of Transport, Slovak Road Administration, National Highway Company etc.

2. The implementation of repressive measures, which include the implementation of strict responsibility for breaches of selected provisions, related to road traffic rules. The task of strict responsibility was to increase the road safety and to ensure discipline and thoughtfulness of road users through enforcing the respect particularly of those provisions of the Act No. 8/2009 ("Law on Road Traffic"), the result of which are fatal traffic accidents or accidents where people are injured.

The road safety can be affected by 12 basic effects plus 2 new ones that can be seen in Fig. 4. These effects can be divided into technical-organizational areas (in the left part of the figure) and legislative-informative (in the right part of the figure). Among them there is a supervision of respecting the rules of road traffic safety and continuity, but it is only one twelfth from the positive effect on the traffic safety point of view. This fact is often generalized by the public as a major area of impact on the traffic safety and therefore the responsibility for the number and consequences of road traffic accidents is often wrongly given to the Police Force.

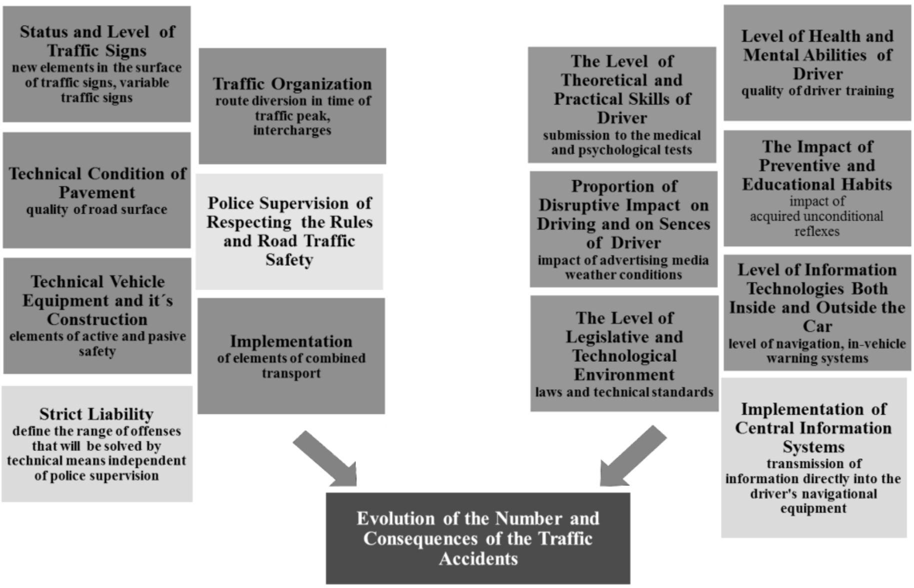


Fig. 4. Current and future state of the main impacts on traffic safety

4 Conclusion

The correct education in the traffic discipline, education of a responsible road user is a long process that starts at an early age. As soon as the child is over his first steps, the responsible parents teach him/her the first knowledge of transport. The education systematically continues in schools and it can be said that the education process of a responsible road user becomes lifelong. The Slovak Republic, as a full EU member, respects the European Commission´s recommendations on road traffic safety and tries to fulfil them. This fact is guaranteed by the Government Programme of the Slovak Republic from 2010, which among other things is declared in: "The Government will support the development of quality, accessible and integrated transport infrastructure, competitive transport services, and user-friendly, environmental and energy-efficient safe transport." The road traffic safety in Slovakia is related not only to the national safe transport, but also to a safe transport on European roads in the context of the Slovak Republic activities in the European Union and to the European area. The direction of road traffic safety in the Slovak Republic is based on the Strategy of the Slovak Republic and on the European transport policy. The White Paper, Roadmap to a Single European Transport Area – Towards a competitive and resource efficient transport system, gives a significant attention to the matters of road safety. It is not only a harmonization of new intelligent transport systems development, but also it is a development of comprehensive strategic measures in the area of traffic accidents, emergency services, definitions for classifying injuries and deaths as a preparation for the adoption of a target to reduce the number of injuries; to focus on special preparation and education of users, to support using of safety equipment and to give special attention to vulnerable users such as pedestrians, cyclists and motorcyclist; and this all through a safer infrastructure and technology of vehicles as well. The essence of these objectives is the prospect of establishing a "vision of zero accidents" in the area of traffic safety. In this context, towards a European space of road traffic safety, the European Commission advocates as a priority for the years 2011 - 2020 the reduction in the number of road traffic accident victims by half compared to 2010 [5].

Acknowledgments. This contribution is the result of the project implementation: Centre of excellence for systems and services of intelligent transport II., ITMS 26220120050 supported by the Research & Development Operational Programme funded by the ERDF.

References

1. Kalašová, A., Hamar, M.: Telematické aplikácie – dôležitý nástroj zvyšovania bezpečnosti. In: Nehody s hromadnýmpostihnutímosôb. 4. Medzinárodný kongres spojený s výstavou, November 24-25, hotel HolidyInn, Žilina, Slovakia. – Euroedu, s.r.o. Košice, pp. 50–56 (2011) ISBN 978-80-969219-8-0 EAN 978809692180
2. Kalašová, A., Hamar, M.: Možnosti znižovania dopravnej nehodovosti v SR pomocou telematických aplikácií. In: Dopravná Infraštruktúra v Mestách, 7. Medzinárodná konferencia: Žilina, zborník: Žilina: Žilinská univerzita, October 20-21, pp. 149–154 (2010) ISBN 978-80-554-0254-3

3. Krchová, Z.: Využitie teórie chaosu v doprave. In: CMDTUR 2009, zborník z 5. Medzinárodnej Konferencie, Žilina, November 4-5, str. I – 118-122, Vydala Žilinská univerzita v Žiline v EDIS – vydavateľstve ŽU (2009) ISBN 978-80-554-0132-4, vyšloajna CD – ISBN 978-80-554-0124-9
4. Kupčuljaková, J.: Preference městské hromadné dopravy. In: Mezinárodní Masarykova Konference pro Doktorandy a Mladé vědecké Pracovníky Hradec Králové. Sborník Příspévků [elektronickýzdroj]: CD. Magnanimitas Hradec Králové, December 6-10, pp. 1303–1307 (2010) ISBN 978-80-86703-41-1
5. Národný plan nazvýšenie bezpečnosti cestnej premávky na obdobierokov 2011 – 2020 In Internet [elektronickýzdroj] (2010), dostupnéna, `http://www.becep.sk/images/download/N%C3%A1vrh%20NP%20BCP%202011%20az%202020.pdf` (November 10, 2011)
6. Program podpory rozvoja inteligentných dopravných systémov Národný systém dopravných informácií. In Internet [elektronickýzdroj] (2008), dostupnéna, `http://www.telecom.gov.sk/pk/08118656z55968/vlastnymat.pdf` (May 22, 2011)
7. Prezídium policajného zboru: Štatistický pre hľadnehôd v cestnej premávke, SR 2008, Bratislava (Apríl 2009)
8. PAPER European transport policyfor 2010: time to decide. Brusel : Commission of the European Communities, 2001. COM(2001) 370 final.3.Foster, I., Kesselman, C.: The Grid: Blueprint for a New Computing Infrastructure. Morgan Kaufmann, San Francisco (1999)

Safety and Functionality Assessment of Railway Applications in Terms of Software

Peter Lüley[1], Mária Franeková[2], and Marek Hudák[1]

[1] EVPÚ a.s., Trenčianska 19, 018 51 Nová Dubnica, Slovakia
{luley,hudak}@evpu.sk
[2] University of Žilina, Faculty of Electrical Engineering,
Univerzitná 1, 010 26 Žilina, Slovakia
maria.franekova@fel.uniza.sk

Abstract. The safety and functionality assessment is an integral part of development of every device. For majority of products particular European standards already exist (EN standards), but the safety assessment of the railway applications in terms of software is described without sufficient details. The issue of software for railway applications is described in the standard EN 50128:2001, which is focused only on the phases of software life-cycle, on the inputs and outputs of different phases of the life-cycle and on the software documentation with regard to the software safety integrity level SIL 1 to 4. In this paper authors deal with the safety of railway applications with regard to the functionality of software. The main part of this article is devoted to the description of methodology for revealing potential dangerous situations in the equipment, which can be subsequently subjected to testing and to a practical example of the described methodology usage.

Keywords: software, safety, testing, methodology, railway application, functionality, assessment**.**

1 Introduction

Standards for almost every kind of equipment exist, which describe the safety tests and put requirements on the equipment functionality. These standards simplify the manufacturer's decisions, whether the device is safe enough, since after testing according to standards requirements there are legal proofs on standardized safety. The problem arises for the equipment for which there are no safety requirements in standards but only requirements for development. According to the Council directive 85/374/EEC [1], the manufacturer is responsible for every product and therefore all manufacturers are trying to make sure that their devices are safe. Today, when nearly every device contains a programmable component used for the device control there are software-related requirements in standards only for a small part of devices. The manufacturers have to decide themselves how to make sure that their software is working properly. The basic requirements to carry out software checks for any equipment can be summarized as follows:

J. Mikulski (Ed.): TST 2012, CCIS 329, pp. 396–405, 2012.

- manually and automatically generated data shall not permit invalid data to cause an error in the program execution,
- software shall be designed to avoid the occurrence of a deadlock in the program flow,
- the execution of the program shall be monitored,
- memory shall be monitored,
- every hazardous situation shall lead to a fail-safe state.

Although the basic software requirements are met, it certainly does not guarantee the program functionality and safety, because the software alone cannot endanger the safety of persons, animals or property, which is a requirement of Article 2 of the Directive 2006/95/EC of the European Parliament and of the Council [2]. Every device is potentially dangerous as a whole that means the software in combination with the hardware, where we have to reduce the level of risk to the lowest possible level. Therefore, while assessing the software safety, it is necessary to subject to tests the whole device and to expose it to conditions that may occur during a normal operation as well as during malfunctioning. In most cases the device consists of the hardware, software and communication media. It may contain elements which read the data, store the data in the memory or transmit them further. What part of overall safety depends on the software depends on the particular application.

The control system of the E/E/PE (electrical/electronic/programmable electronic safety-related systems) has to react to hazardous failure modes within the safety-relevant process (from the physical and external environment) and according to IEC 61508-1 [3] it has to perform safety functions with the required level of safety. A quantitative evidence of safety based on defined parameters is needed, in which the intensity of a dangerous failure of the device λ_{dev} is most important, which after the assumption simplification is the sum of partial intensities of dangerous failures of particular system components (hardware, software, communication), as shown in equation (1).

$$\lambda_{dev} = \lambda_{HW} + \lambda_{SW} + \lambda_{com}. \tag{1}$$

Qualitative evaluation methods of software safety are now enforced for the area of control of safety critical processes, for which the procedure for railway applications is described in the EN 50128 standard [4].

Also modeling is important when specifying the system requirements, during designing the device structure and also during the validation and verification. In some cases it may mean also the design optimization or setting-up parameters of the existing device so that the requirements for the safety integrity level would be fulfilled as requested by customers or as shown in the risk analysis. The meeting of these challenges requires a combination of appropriate modeling methods and modeling tools. In these cases the device is represented by an abstract mode, which graphically or mathematically describes the desired properties of the device.

In many cases the input data for qualitative evaluation methods are not available and often only a qualitative evaluation is used. The issue of safety evaluation of railway application software can be found in [5] and [6].

2 Risk Management

The life cycle of the overall software safety of the E/E/PE according to the EN 61508-1 standard [3] requires to perform a risk analysis to identify the hazard and the hazardous situations of the equipment which results in the overall safety requirements. The term risk analysis mean a systematic use of available information on the hazard identification and risk assessment [7]. However, the risk analysis is only one part of risk management. If we want to achieve the equipment on the highest level of safety we should perform:

- Risk analysis
- Risk assessment
- Risk reduction

The entire risk management process is shown in Fig. 1.

2.1 Risk Analysis

The concept of risk has two components. It is the likelihood of the risk and the consequence. The risk analysis should be prepared by the device designer, which means a person which determines how the equipment will work based on its intended use, on the components used and on modules supplied by subcontractors, bearing in mind the two components of risk. It is clear that no designer can accurately predict how the modules from subcontractors will behave, therefore the risk analysis must cover cases when the module stops performing its function partially or completely or starts providing invalid data about its function to the master control system. The majority of standards consider the fault condition only when there is only one failure. Generally, two independent failures at the same time are not expected, therefore the risk analyses do not have to consider such a situation. However, the case where one failure leads to another failure is a common situation and it is necessary to analyze such case. The risk analysis should consider:

- every potentially hazardous event and steps leading to this event,
- the circumstances and the likelihood of the sequence which is associated to each hazardous event,
- measures to reduce or eliminate the hazard or risk.

Preconditions established during the risk analysis, including the number of challenges and failures, any recognized functional limitation or human intervention, must be described in detail. A record of the risk analysis should be part of the risk management documentation.

2.2 Problems of Risk Evaluation

During the risk evaluation it is important to realize that when a hazardous event has arisen as a result of failure, the probability of failure and of damage is not the same, because the result of a failure is not always a hazardous situation and the result of a

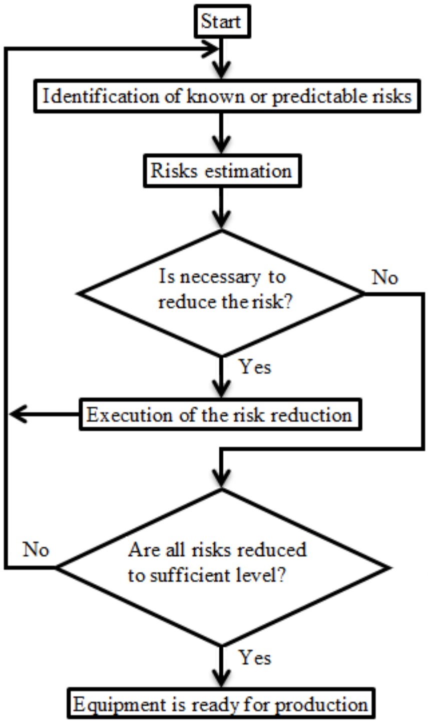


Fig. 1. Overview of risk management activities

hazardous situation is not always a damage. Because of that, the following should be examined during the risk evaluation:

- what can cause the damage,
- the probability of damage or injury,
- the type of damage or injury.

In order to make the output of risk evaluation clear the level of risk should be expressed in discrete values. To estimate the risk it is necessary to use all the available information including the information from older versions of the device, the information from simulations and from components reliability estimations. For risks, where it is not possible to perform the estimation based on any available information it is necessary to make an estimation based only on the type of potential damage or injury. A record of the risk evaluation should be part of the risk management documentation.

2.3 Risk Reduction

In the case, where some risks have been estimated as high it is necessary to reduce the risk to an acceptable level. The method of risk reduction is the designer choice, where the following options are available:

- change of the design to remove a specific risk,
- change of the design to minimize the sensitivity of the damage or injury,
- change of the design to automatically switch into a fail-safe state in the case of a risk,
- change of the design to initiate an automatic warning in the case of risk.

As shown in Figure 1 after each risk reduction it is necessary to assess the changes impact on the overall safety by the re-execution of risk analysis and of risk assessment. During this reanalysis of risks new risks arising from changes in the design can be identified or the level of some other risk might be changed.

3 Mathematical Apparatus Used in the Quantitative Safety Evaluation

Random failures in HW and SW can be expressed using mathematical tools for random a variable using the probability theory and mathematical statistics. Indicators of safety for safety-relevant devices are based on well-known reliability indicators, but it is necessary to apply them to a dangerous (critical) failure. They can be divided into two groups [8]:

- Total characteristics

They express a full description of the probability distribution of a random variable. The most common characteristics are:

- probability of dangerous failure $F(t_1,t_2)$ of SW in a defined period of time,
- discrete distribution function of failure $f(t)$,
- probability that the SW carries out requested functions in given conditions and in a given time $R(t)$,
- intensity of dangerous failures of SW λ(t).

The above mentioned safety indicators are theoretically equivalent expressions of a full description of the probability distribution of the random variables and clear relationships exist among them, (2) to (5).

$$F(t_1,t_2) = \int_{t_1}^{t_2} f(t)dt, \tag{2}$$

$$f(t) = \frac{dF(t)}{dt}, \tag{3}$$

$$R(t) = 1 - F(t), \tag{4}$$

$$\lambda(t) = \frac{f(t)}{1 - F(t)}. \tag{5}$$

Graphical relations of probability of dangerous failure $F(t)$ and probability of non – failure operation $R(t)$ in chosen time period are illustrated in Fig. 2.

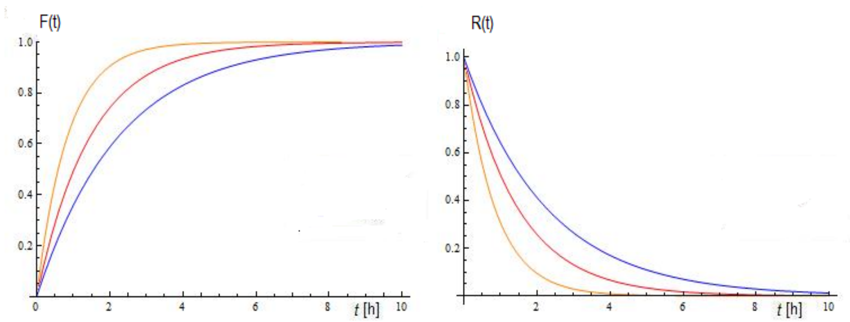


Fig. 2. Graphical relations of total characteristics $F(t)$ and $R(t)$

- Local characteristics

They are the numerical characteristics of random variables. During the software safety evaluation we can used the following characteristics:

- mean time to failure $MTTF_N(t)$,
- the probability of not detecting the storage data by the safety coder p_{ned},
- the average error probability of the cryptographic decoder.

Equations for these characteristics are (6) to (8).

$$MTTF_N(t) = \int_0^{\infty} R_N(t)dt, \tag{6}$$

$$p_{ned} = \sum_{\left\| i=\frac{d_{\min}+1}{2} \right\|}^{n} A_i . p_b{}^i (1-p_b)^{n-i}, \tag{7}$$

$$\overline{p}_{CW} \approx \left(1-2^{-n}\right)^{-1}\left(1-2^{-k}\right) n P_b . \tag{8}$$

Where:

n	- the length of storage data,
d_{min}	- minimum Hamming distance of the safety code used,
A_i	- number of code words of the safety code with the weight i,
p_b	- bit error rate of one element in the transmission communication channel,
k	- size of block of the cryptographic word.

A detailed description of safety indicators for specific applications of safety analysis can be found in the literature, e.g. [9, 10, 11].

Criteria for the selection of techniques and measures used in the software testing are recommended in the EN 50128 standard [4]. For each technique or measure illustrated in Table 1 the requirement for each software safety integrity level (SIL) is indicated.

Table 1. Techniques and measurements recommended for SW verification and testing

Technique/Measure	SIL 0	SIL 1	SIL 2	SIL 3	SIL 4
Formal Proof	-	R	R	HR	HR
Static Analysis	-	HR	HR	HR	HR
Dynamic Analysis and Testing	-	HR	HR	HR	HR
Metrics	-	R	R	R	R
Traceability	R	HR	HR	M	M
Software Error Effect Analysis	-	R	R	HR	HR
Test Coverage for code	R	HR	HR	HR	HR
Functional/ Black-box Testing	HR	HR	HR	M	M
Performance Testing	-	HR	HR	HR	HR
Interface Testing	HR	HR	HR	HR	HR

The meaning of symbols used in the Table 1:

'M' this symbol means that the technique or measure is mandatory for this safety integrity level.

'HR' this symbol means that the technique or measure is highly recommended for this safety integrity level.

'R' this symbol means that the technique or measure is recommended for this safety integrity level.

'-' this symbol means that the technique or measure has no recommendation for or against being used,

'NR' this symbol means that the technique or measure is positively not recommended for this safety integrity level.

4 Example of Software Assessment in Railway Applications

One of the last processes before placing the device on the market is to prove the safety and functionality of the device by testing. When the test schedule is well prepared the entire device can be examined that means the hardware and software parts. It may be dangerous to test all types of railway applications in real conditions especially during the tests of failure modes, therefore simulated conditions in testing laboratories are often used. Along with properties and performance tests it is also necessary to examine conditions specified as potentially dangerous during the risk evaluation.

During testing of railway applications manufactured by EVPÚ a.s. Nová Dubnica, Slovakia the analysis of hardware failure modes is performed together with tests of correct functioning of the software.

4.1 Example of Risk Management Use to Assess the Safety and Functionality

The development of risk analysis is the first step during the risk assessment. It is necessary to identify every possible hazard and to examine the risk for each hazardous situation. Hazards associated with the software can be:

R_1 – Damaged integrity of the program.
R_2 – Power supply cut-off of the control element.
R_3 – Input data out of range.
R_4 – Deadlock in the program flow.
R_5 – Incorrect data transmission.

Quantitative measures of probability are preferred during the risk estimation. However, this approach needs adequate data. In the case we do not have the required extent of data we choose a quantitative description. A good quantitative description is more appropriate than an inaccurate quantitative description. A typical example of qualitative analysis is to use a N×M matrix to describe the sensitivity and probability of risk. We consider the 5×5 size sufficient, but the final choice is always to be made by the manufacturer. A small size of the matrix cannot be detailed enough, while a large size can be hard to elaborate and very time consuming. An example of the sensitivity and probability levels can be found in Tables 2 and 3.

Table 2. Example of qualitative sensitivity levels

Sensitivity level	Description
Critical	Leads to a damage to human health or death
Significant	Leads to complete damage of the device
Serious	Leads to a damage which requires service actions
Small	Leads to a damage which does not affect functionality or safety
Negligible	Minor temporary malfunction with unimportant impact on safety do not causing a damage

Table 3. Example of semiquantitative probability level

Probability level	Description
Large	$\geq 10^{-3}$
Likely	$\Re\ 10^{-3} \geq 10\text{-}^{4}$
Occasional	$\Re\ 10^{-4} \geq 10^{-5}$
Very small	$\Re\ 10^{-5} \geq 10^{-6}$
Unlikely	$\Re\ 10^{-6}$

After determining the levels of sensitivity and probability we can bring risks to the risks matrix. During the evaluation it is necessary to consider what event would the risk cause and the probability of such situation.

The risks matrix clearly shows all the risks and after comparing with the boundaries of acceptable risks, the risks that need to be tested in detail are highlighted. In the case of unacceptable risks an ideal solution consists in reducing them by one of methods described in 2.3, even if the testing proves this risk as without problem.

Table 4. Example of quantitative risks matrix

Prob./Sens.	Negligible	Small	Serious	Significant	Critical
Large					
Likely		R_5			
Occasional	R_3				
Very small					R_2
Unlikely					R_1, R_4

Not acceptable risk

Acceptable risk

5 Conclusion

The evaluation of the railway application software is not possible without a joint evaluation of hardware because together they make one device. The process of risk management, which we understand as consisting of the risk analysis preparation and documentation, of risk assessment and of risk reduction, may reveal weaknesses of the device that the manufacturer can subject to tests and evaluate as not acceptable risks and therefore actions to remove them must be taken. The process of risk management is a method that is easily applicable to other types of risks than those arising from the software. It can be easily used to detect electrical, mechanical and other risks depending on what the manufacturer considers as important to examine.

The described methods are used during the testing of railway applications manufactured by EVPÚ a.s. Nová Dubnica, Slovakia.

Acknowledgement. This work has been particularly supported by the Educational Grant Agency of the Slovak Republic (KEGA) Number: 024ŽU-4/2012: Modernization of technology and education methods orientated to area of cryptography for safety critical applications and by the scientific grant agency VEGA, grant No. VEGA-1/0388/12 “Quantitative safety integrity level evaluation of control systems in railway applications”.

References

1. Council directive 85/374/EEC. 25. 07. 1985. OJ L 210, 7.8.1985, pp. 29–33 (1985)
2. Directive 2006/95/EC of the European Parliament. 12. 12. 2006. OJ L 374, 27.12.2006, pp. 10–19 (2006)
3. EN 61508-1: Functional safety of electrical/electronic/programmable electronic safety-related systems – Part 1: General requirements. CENELC (2010)
4. EN 50128: Railway applications - Communication, signalling and processing systems - Software for railway control and protection systems. CENELEC (2011)

5. Mazzeo, A., Mazzocca, N., Nardone, R., D'Acierno, L., Montella, B., Punzo, V., Quaglietta, E., Lamberti, I., Marmo, P.: An Integrated Approach for Availability and QoS Evaluation in Railway Systems. In: Flammini, F., Bologna, S., Vittorini, V. (eds.) SAFECOMP 2011. LNCS, vol. 6894, pp. 171–184. Springer, Heidelberg (2011)
6. Huhn, M., Hungar, H.: 8 UML for SW safety and certification. In: Model Based Development of Safety Critical Software – Intensive Systems, pp. 201–237. Springer, Berlin (2010)
7. Zahradník, J., Rástočný, K., Kunhard, M.: Safety of Railway Interlocking Systems. EDIS-ŽU Žilina (2004) (in Slovak) ISBN 80-8070-296-9
8. Mykiska, A.: Safety and reliability of technical systems. ČVUT, Czech (2006)
9. Franeková, M.: Mathematical Apparatus for Safety Evaluation of Cryptography and Safety Codes Used in Safety – Related Communication System. In: Mikulski, J. (ed.) TST 2011. CCIS, vol. 239, pp. 126–135. Springer, Heidelberg (2011)
10. Rástočný, K.: Risk Analysis of a Railway Interlocking System. Journal: Advances in Electrical and Electronic Engineering 2(3-4), 24–29 (2003)
11. Andonov, A., Hubenova, Z.: Principles and methods for creation of models in automated information management systems. Vasil Levski NMU – Artillery, AAD and CIS faculty 2, 257–263 (2010)

Monitoring System of Vibration Propagation in Vehicles and Method of Analysing Vibration Modes

Rafał Burdzik

Silesian University of Technology, Faculty of Transport
Krasińskiego 8, 40-019 Katowice, Poland
rafal.burdzik@polsl.pl

Abstract. Article provides a presentation of a prototype monitoring system for the vibrations occurring in vehicles. The mathematical algorithm prepared and programmed enables the analysis and separation of the crucial properties of vibration signals. The application of an automatic algorithm developed and proposed by the author for the recognition of stationary and non-stationary states makes it possible to use the appropriate tools assuming the form of frequency as well as time and frequency transforms. The paper also provides a discussion concerning the concept of a neural classification system which may be applied as a decision-making module in a control system. Preliminary numerical tests have confirmed the system's functional correctness. The monitoring system is based on the assumption that vibrations can be recorded at any chosen structural point of a vehicle which enables the vibration propagation to be analysed as well as their time, frequency or time and frequency modes to be assessed.

Keywords: monitoring system, vibration propagation in vehicles.

1 Introduction

Vibrations are one of the most important negative phenomena generated by systems and means of transport. They are purely undesirable phenomena, which exert a significant impact on the safety level as well as on the degradation of technical and natural environment. Therefore, there are many different assemblies in vehicles with the sole task to minimise or eliminate vibrations and their propagation, not only the material issue [1, 2, 3]. The article provides a presentation of a prototype monitoring system for the vibrations occurring in vehicles.

2 Data Registration and Acquisition System

A Programmable Unit for Diagnostics PUD (Fig. 1) was proposed for the data registration and acquisition. The unit is composed of separate analogue and programmable modules, which are tightly-coupled with the ARM extension board. The data acquired from external sensors are digitalized by AD converters,

J. Mikulski (Ed.): TST 2012, CCIS 329, pp. 406–413, 2012.

pre-processed by the FPGA module and stored in a local SDRAM memory. A number of hardware modules (e.g. a linear decimation unit) were implemented in the FPGA, which makes the system perfectly suited for non-stationary signals analysis. Furthermore, the presented system is equipped with a user-friendly interface implemented on the ARM extension board. The ARM processor supports a well-known and widely used Linux operating system, which provides a user with a whole collection of ready-to-install software bundles such as Octave, Emacs etc [4].

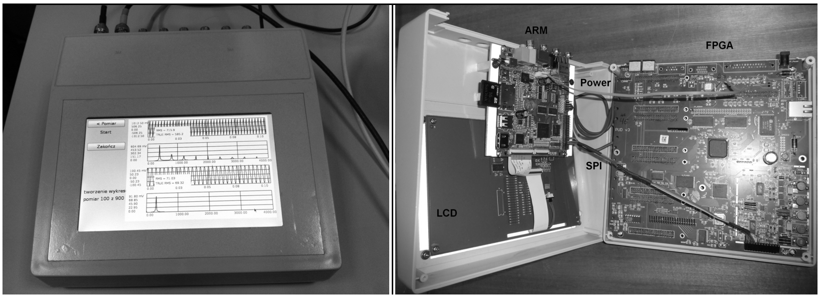


Fig. 1. PUD for the data registration and acquisition

3 Algorithm for Recognition of Stationary and Non-stationary States of the Signal

Vehicles vibrations are random processes resulting from a wide range of possible forces and nonlinear characteristics of suspension elements of the vehicles construction. The stationary and non-stationary signals require the use of different transformations for the signal processing. The mathematical algorithm prepared and programmed enables the analysis and separation of the vibration signals in accordance with the stationary and non-stationary states. It is based on the core of the frequency comparison to localize the signal with a constant frequency as the main condition for a stationary signal. The changes of mean values of the signal in the processed window are simultaneously compared. Those operations allow dividing the signal into windows with stationary and non-stationary part of the recorded signal. The markers of frequency comparison based on STFT transformation were used. A short execution time was the main reason of choosing this transformation. The active experiment was conducted with a passenger car forced to vibration on a harmonic test stand. The results of the algorithm are presented in Fig. 2. It can be observed that the developed programmed algorithm allows identifying the run up and coasting of forced machine and vibration signal [5,7].

The STFT transformation enables the signal observation in time-frequency domains. It is possible to choose a constant frequency band and to observe the signal realization in time. The algorithm for frequency comparison can be identified based on a time function. The elaborated algorithm is based on the comparison of the next value of analysed frequency band ("analysis of edge") around set parameters. The programmable algorithm has been depicted in Fig. 3.

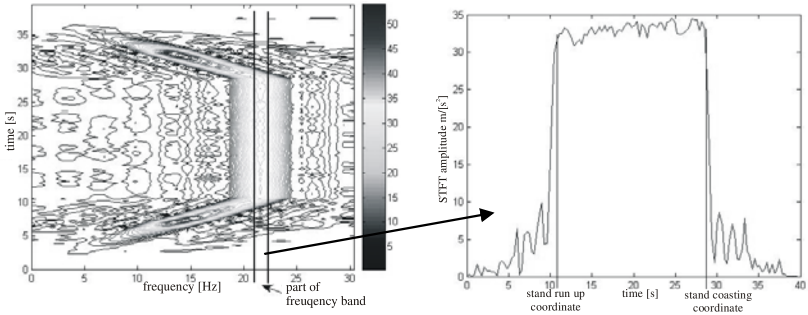


Fig. 2. Execution of the algorithm of frequency comparison

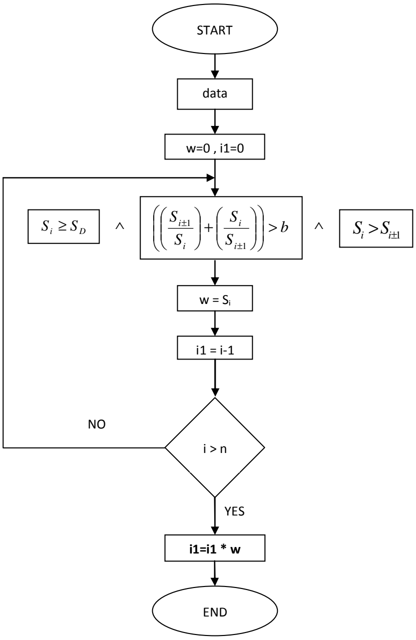


Fig. 3. The programmable “analysis of edge” algorithm

The application of automatic developed algorithm for the recognition of stationary and non-stationary states makes it possible to use the appropriate tools assuming the form of frequency as well as time and frequency transforms.

4 Vibroacoustic Signals Analyser

The modules of data acquisition and recognition of stationary and non-stationary states of the signal allow collecting the signals and separating the stationary and non-stationary parts of the signal nearly in real-time. It enables to conduct the dedicate signal processing for vibration resonances and propagation analysis. For the stationary signal analysis the FFT (Fast Fourier Transformation) was used. The non-stationary signal was analysed by WT (Wavelet Transformation). A wavelet transformation of signal x(t) may be defined as follows:

$$WT_x(a,b) = \frac{1}{\sqrt{a}} \int_{-\infty}^{\infty} x(t)\Psi\left(\frac{t-b}{a}\right)dt \tag{1}$$

where:

$\Psi(t)$ – wavelet family,

a – scaling parameter effecting the frequency change $a \in R^+ / 0 \wedge a \sim \frac{1}{f}$,

f – frequency,

$\frac{1}{\sqrt{a}}$ - wavelet normalising constant,

b – shift parameter locating the wavelet position against the time axis $b \in R$.

The wavelet transform represents the correlation between the signal analysed and function $\Psi(t)$ having been appropriately scaled. The very idea of wavelet transformation is the decomposition of signal $x(t)$ into wavelet coefficients $WTx(a,b)$ by the application of the base function. As a result of such a transformation, coefficients being functions of scale and time are obtained. By altering the parameters of scale a and time shift b, the time-frequency distribution may obtained.

For the application of the monitoring system of vibration propagation in vehicles the interface for the Vibroacoustic Signals Analyzer (VSA) was programed. The communication with the user is based on the next dialogue windows programed in Graphical User Interface Tools in MATLAB environment. A graphical user interface (GUI) is a graphical display in one or more windows containing controls, called components, that enable a user to perform interactive tasks. GUIs created using MATLAB tools can also perform any type of computation, read and write data files, communicate with other GUIs, and display the data as tables or as plots. The article presents a VSA module dedicated for the monitoring system of vibration propagation (AMOR).

The next windows of the interface of the vibration signal processing system were presented in Fig. 4.

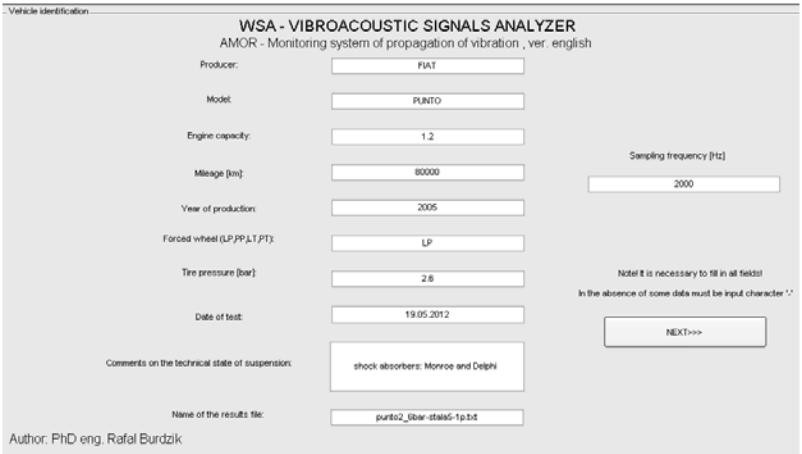


Fig. 4. Interface for the vehicle identification data and the source and sampling of recorded signal

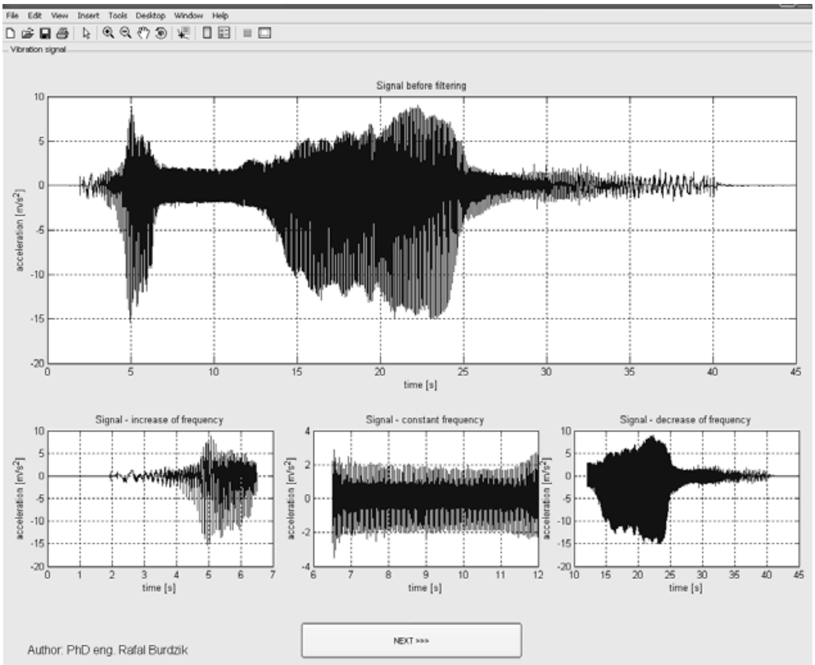


Fig. 5. Interface of the results of algorithm for recognition of stationary and non-stationary states of the recorded signal

The monitoring system allows correcting manually the divided signal if the results are not satisfactory.

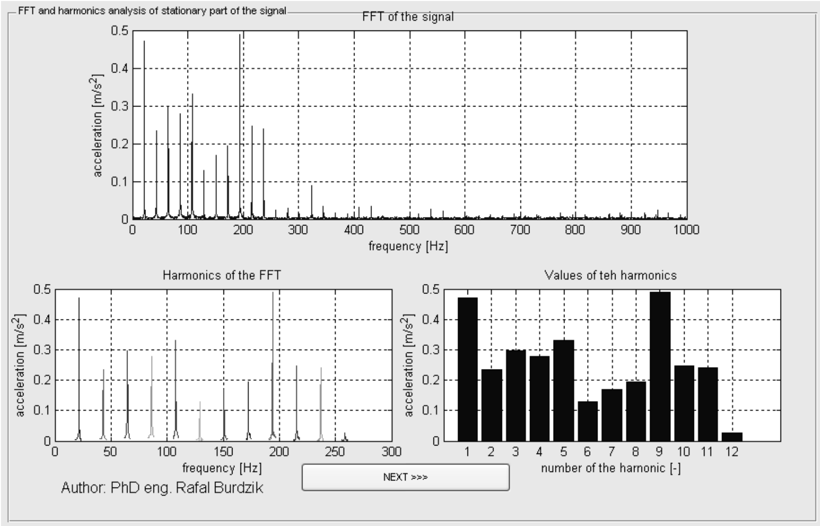


Fig. 6. Interface of the FFT and harmonics analysis of the stationary signal

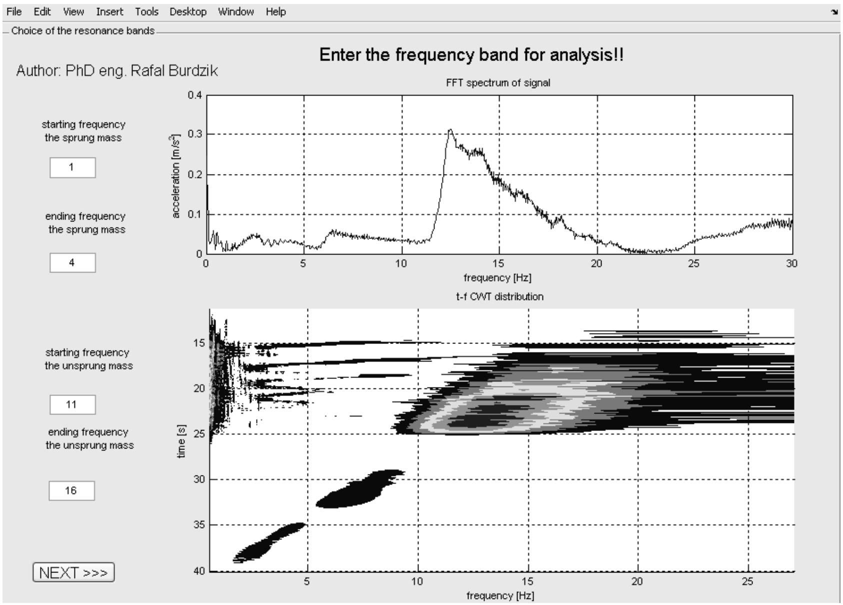


Fig. 7. Interface for the manual choosing of the resonance bands of the signal

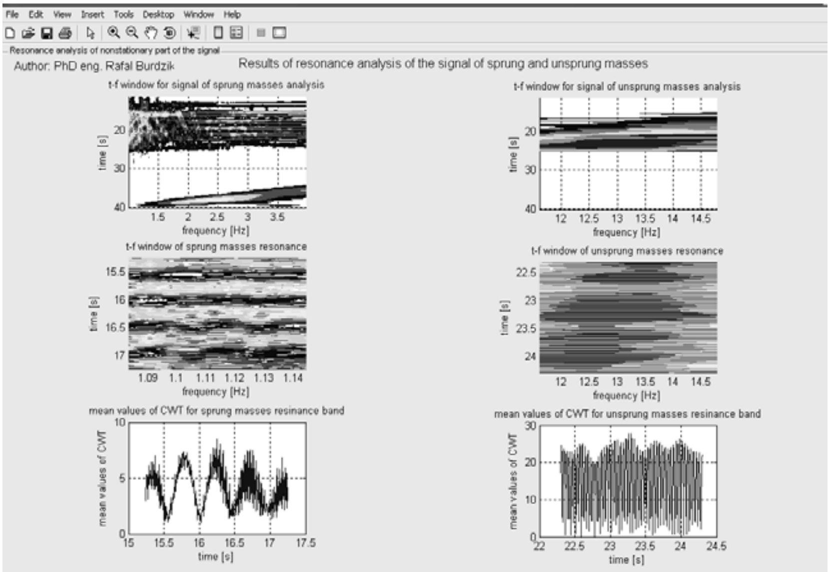


Fig. 8. Interface of the resonance analysis of the non-stationary signal and propagation from unsprung to sprung masses of the vehicles

5 Classification System

The obtaining of vibration signal vector estimators is the result of Vibration System Analyser. Classification modules were based on the neural network theory. That choice was influenced by their attributes such as: generalisation of knowledge, approximation of functions and recognition of patterns. The possibility of constant development of the basic knowledge by numerous examples and a simple adaptation of a monitoring system is very significant as well. The neural classification system can be applied as a decision-making module in a control system.

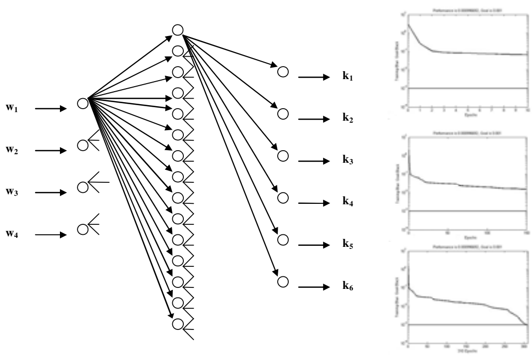


Fig. 9. Neural network classification module with training process graphs

The simulation tests of neural network application to detect an abnormal tyre pressure and a shock absorber defect based on vibration signals were conducted. The architectures of the network were designed and the transfer functions in the layers were programmed. The obtained results were satisfactory [6]. The architectures of the example feed-forward network and training process has been depicted in Fig. 9.

6 Conclusion

The monitoring system is based on the assumption that vibrations can be recorded at any chosen structural point of a vehicle which enables the vibration propagation to be analysed as well as their time, frequency or time and frequency modes to be assessed. The monitoring system of vibration propagation in vehicles includes methods and algorithms for signals acquisition, pre-processing, transmission and data processing and storage. Consequently, the whole system consists of the following modules: data acquisition, stationary and non-stationary signal localisation, VSA and neural classification system.

References

1. Mikulski, J.: Using Telematics in Transport. In: Mikulski, J. (ed.) TST 2010. CCIS, vol. 104, pp. 175–182. Springer, Heidelberg (2010)
2. Burdzik, R.: The influence of the rotational speed of engine on vibrations transferred on vehicle construction. Silesian University of Technology Scientific Papers, s. Transport 72, pp. 13-23 (2011)
3. Łabaj, J., Oleksiak, B., Siwiec, G.: Analysis of the options of copper removal from liquid iron by evaporation. Metalurgija 50(3), 173–175 (2011)
4. Jamro, E., Wielgosz, M., Bieniasz, S., Cioch, W.: FPGA – ARM heterogeneous system for high speed signal analysis. Solid State Phenomena 180, 207–213 (2012)
5. Burdzik, R., Gardulski, J.: Frequency analysis decimation vibration signals of passenger car's suspensions. Journal of Transport Problems, 23–29 (2007)
6. Burdzik, R., Gardulski, J., Konieczny, Ł.: Neuronowy klasyfikator falkowy uszkodzeń amortyzatorów samochodowych. Przegląd Mechaniczny (3), 32–36 (2007) (in Polish)
7. Burdzik, R., Stanik, Z., Warczek, J.: Method of assessing the impact of material properties on the propagation of vibrations excited with a single force impulse. Archives of Materials and Metallurgy 57(2), 409–416 (2012)

Identification of Leakages in the Inlet System of an Internal Combustion Engine with the Use of Wigner-Ville Transform and RBF Neural Networks

Piotr Czech

Silesian University of Technology, Faculty of Transport
Krasinskiego 8 Street, 40-019 Katowice, Poland
piotr.czech@polsl.pl

Abstract. In recent years the number of vehicles using the roads all over the world has been constantly increasing. A continuous progress in the field of new technologies causes that they become cheaper and cheaper to manufacture and at the same time more accessible. In order to balance the positive and negative effects of motorisation development, world organisations introduce stricter and stricter regulations concerning the safety and the natural environment protection. Each car produced now is equipped with an on-board diagnostic system OBD. Such systems serve to assess the efficiency of the basic elements in a vehicle. The article attempts to determine the lacks of tightness in the inlet system of the internal combustion engine on the basis of vibration effects emitted by the engine. On the basis of a model prepared with the use of Wigner-Ville transform a series of experiments has been conducted to achieve a design of a properly functioning neural classifier which may extend the operation of an OBD system.

Keywords: internal combustion engines, artificial neural networks, diagnostics, On-Board Diagnostics OBD.

1 Introduction

During the last ten years, the number of vehicles moving on the roads has almost doubled and only last year the number of engine vehicles registered on the territory of Poland was equal to 24.562.530. Such increase in the number of vehicles reflects in negative results of civilisation development such as a bigger number of road accidents and the rise in the level of the natural environment pollution. The level of road and traffic safety is the sum of the levels of safety of each element of the person-vehicle-surrounding system. In order to keep the right level of road and traffic safety the vehicles should be kept in the proper technical condition and to do that the earliest possible detection of all types of damages is essential. The right diagnostic tools should be available both for the people diagnosing the vehicles in diagnostic stations as well as for mechanics working in the service points. The right equipment in such places would improve the efficiency of their work and, in a broader sense, it would cause the increase in the road traffic safety. The literature simultaneously presents a

J. Mikulski (Ed.): TST 2012, CCIS 329, pp. 414–422, 2012.

huge interest of scientists in the problems connected with the improvement in the reliability and safety of transport means [1, 2, 4, 6, 7, 8, 9, 12].

The danger to the environment caused by motorisation is the most noticeable modern inconvenience resulting from the civilisation development. The tests show that the early identification of damages proved by a bigger emission, which is currently applied by diagnostic systems, is not always possible. Here, we observe the lack of efficient diagnostics methods which give clear results and which enable to assess the technical condition of the engine as well as to identify the possible damages at the earliest possible stage of their development. It is really important due to a relatively large number of defects connected with the inefficiency of engine systems [3, 5, 6, 7, 12, 13]. Diagnosing such types of defects at the early stage of development and their fast elimination would definitely cause a decrease of emission of harmful substances released to the atmosphere and at the same time would improve the quality of our lives, especially in big conurbations.

The results of tests presented in this article serve as the basis to consider how to expand the OBD system applied in motor vehicles [7] with the use of a vibroacoustic diagnostic system based on artificial neural networks which would help to detect various types of defects occurring in the engines of motor vehicles.

2 Description of Experiment

The experiment was aimed at detecting the leakages in the inlet system of an internal combustion engine with the use of vibrations which accompany them.

The provision of the most identical cylinder filling is the basic task of the inlet system of an internal combustion engine. In the case of power regulation in an engine connected with a change of coefficient of filling it is required that the system regulates precisely the air mass in the cylinders. The system should meet the requirements of maximum filling particularly at high loads. The system is also required to swirl correctly the load in a cylinder and to measure precisely the air mass which is supplied to cylinders [3, 5, 12, 13].

The following aspects affect the value of maximum coefficient of filling:

- degree of compression;
- type of the combustion chamber;
- intensity of cooling;
- geometry of the crankshaft-piston system.

Recently, there is a tendency to automate various functions in internal combustion engines. Moreover, the functions of engine fuelling and the choice of air mass to required power have been automated as well. It was undoubtedly influenced by the aspects connected with the protection of environment.

During the engine start-up, gear switch, engine braking or idle running, the control system of filling by setting the air mass entering the cylinders:

- maintains the right rotational speed;
- balances the changes of engine load;

- prevents the engine stoppage as a result of overload;
- minimises the fuel consumption;
- reduces the emission of harmful substances to the atmosphere;
- compensates for the changes caused by ageing and exploitation wear;
- improves the ride performance of vehicles.

An automatic control applied in the inlet system:

- simplifies the procedures of engine start-up;
- reduces the rotational speed of idle run;
- optimises the unstable conditions;
- regulates the inner re-circulation of the exhaust gas;
- broadens the methods of switching off cylinders;
- improves the efficient engine braking;
- allows the execution of complicated thermodynamic cycles;
- allows an easier adjustment of turbo-charging of load.

An important parameter in the cylinder filling is the value of pressure in the inlet collector. Usually, the filling of cylinders is proportional to its value. The location of the throttle decides mainly about the value of pressure. By full load the value of pressure in the inlet collector is close to the ambient pressure but at a full closure of the throttle the pressure reaches the value of about 20kPa. Charging of the engine allows achieving the value above the ambient pressure [12, 13].

The object of tests was a 4-cylinder internal combustion engine ZI (installed in a vehicle) with capacity of 2.2 dm^3 fuelled with petrol.
Tests were conducted on a Bosch FLA 203 engine test bench, for two speeds and at two loads.

During the tests the signal of vibrations acceleration was registered in a direction parallel and perpendicular to the cylinder axis as well as the signal of crankshaft rotation angle with the marker of upper dead centre of the piston position.

Vibration signals were registered with the use of piezoelectric sensors placed on the body using a multi-channel measuring device, which made synchronous sampling with high frequency possible. The device worked with an application created in the LabView environment.

A defect of the exchange system of load, which would disturb the work of the engine, was simulated by unsealing of the inlet collector behind the throttle. A leakage of the inlet collector was realised with the use of a precise needle valve with a micrometre screw. Various stages of engine run were simulated by four different valve positions.

Registered vibration signals served as the basis to build descriptors of damages for artificial neural networks. In order to perform it, the time-frequency distributions were created with the use of Wigner-Ville transform (WVD). The WVD transform belongs to a group of methods allowing a simultaneous analysis in time and frequency domains. It is used for variable signals in terms of time in the amplitude sense, that is for time-variant signals [1, 6]. It is defined as follows:

$$WVD(t,f)=\int_{-\infty}^{+\infty} w(\tau)\cdot x\left(t+\frac{\tau}{2}\right)\cdot x^{*}\left(t-\frac{\tau}{2}\right)\cdot e^{-j2\pi f\tau}d\tau \qquad (1)$$

where:

$x^{*}(t)$ – time complex signal, conjugated with *x(t)*,

$w(\tau)$ – symmetrical gravimetric function.

Example arrangements of Wigner-Ville are shown in Fig. 1.

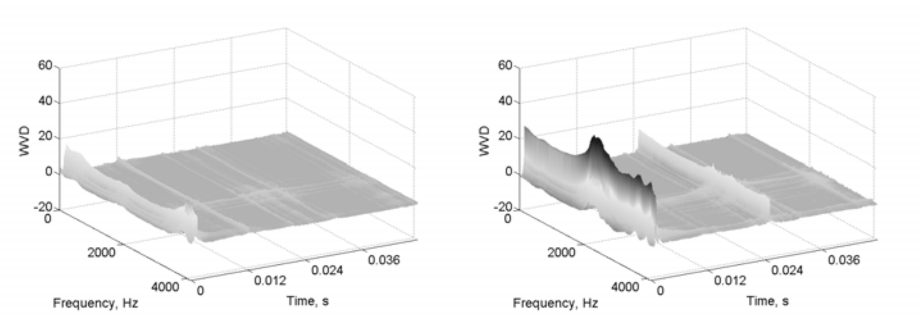


Fig. 1. Wigner-Ville distribution in the case of lack and in the case of presence of leakages in the inlet system of an internal combustion engine

Fig. 2 presents schematically the method of proceeding by the models creation. To describe the character of changes in the WVD distribution, depending on the degree of damage, a two-stage method of proceeding was assumed.

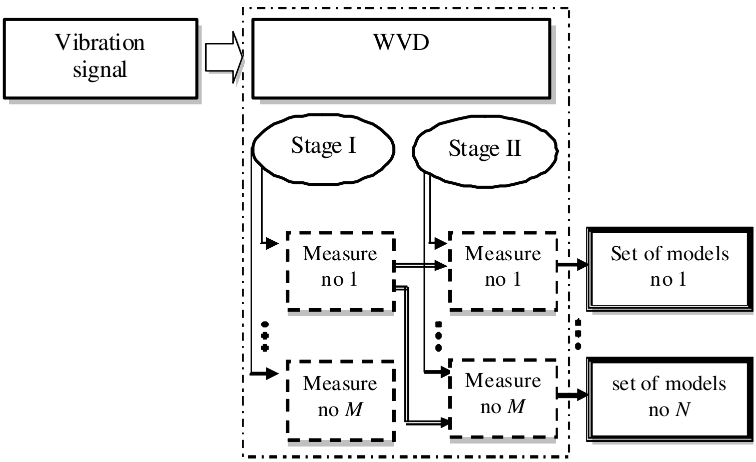


Fig. 2. Method of building sets of models

In stage I a statistical measure is marked for each following frequency.

The usefulness of 36 estimates was checked (all of which are broadly described in the literature – coefficients of changeability, peak, clearance, shape, switching mode and asymmetry, quarter and average aberration, arithmetic average, geometric and harmonic mean, quartile, central moments, cumulants, energy of signal, root-mean-square values, peak-to-peak values, maximum and minimum values, variance, positional variation coefficient) [1]. Such procedure was aimed at describing the character of changes in the time domain for next frequencies marked according to the assumed sampling frequency.

In stage II, from characteristics of frequencies achieved in such a way, ranges 500 or 1000 Hz long were separated. The aim of the division into ranges with different options of length was to check their size influence on the classification result. In each separated part of spectra (achieved in such a way) the changeability character of distribution was described with the use of 36 measures. A vector composed of marked measure in each of the parts of the spectrum served as the input data for the neural classifier.

RBF neural networks were used in conducted tests, where a function with a circular symmetry is used as the neuron activation function. Radial neurons represent a hyper-sphere, which performs circular separations around the central point. RBF network has a layered structure with a gas input layer, a hidden layer and an output layer. The hidden layer is built from radial neurons which most often have a Gauss function as the activation function. Neurons in the output layer have a linear or logistic activation function. The role of neurons in this layer is mainly to sum up the weight of signals from the hidden layer. The basic problem in designing such types of networks is the proper choice of the number and types of basic functions (neurons in the hidden layer). A too small number causes that the quality of adjusting to models is poor and a too big number causes the loss of generalisation of knowledge by the network. The values of central points, which place the basic functions inside the space stretched on the input data should be determined together with the balance values of connection weights between neurons in the hidden layer and the output layer. The tests were focused mainly on the choice of the right character of the basic function optimally separating the data space. By changing the coefficient γ the shape of the function was determined together with the size of receptive field for which the function has a non-zero value. This coefficient should be chosen in a way which would allow all receptive fields of all basic functions to cover the whole area of the input data with only a slight overlap of two neighbouring receptive fields permissible [10, 11]. Tests were conducted for 86 values of coefficient γ.

In the first experiment an attempt was made to classify two conditions of the inlet system of an internal combustion engine:

- inlet system in good condition;
- occurrence of leakages.

For all sets of standards achieved from the two-stage way of building descriptors – 36 measures x 36 measures, by two ways of spectrum division into ranges – a window sized 500Hz or 1000Hz, the influence of the coefficient γ value on the achieved classification level was checked.

An example influence of the coefficient γ value on the classification error is shown in Fig. 3.

The best results achieved for vibration signals registered in a direction parallel to the cylinder axis are shown in Fig. 4 whereas Fig. 5 presents the best achieved results for signals registered in a direction perpendicular to the cylinder axis.

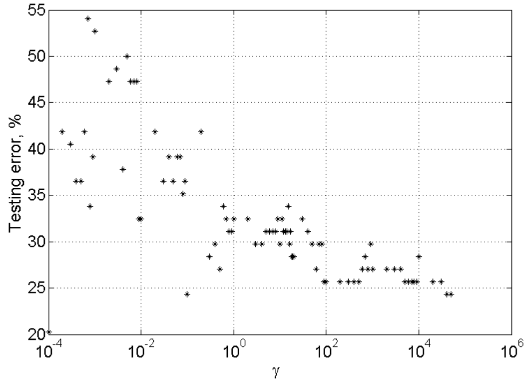


Fig. 3. Influence of parameter γ values on the damage classification error

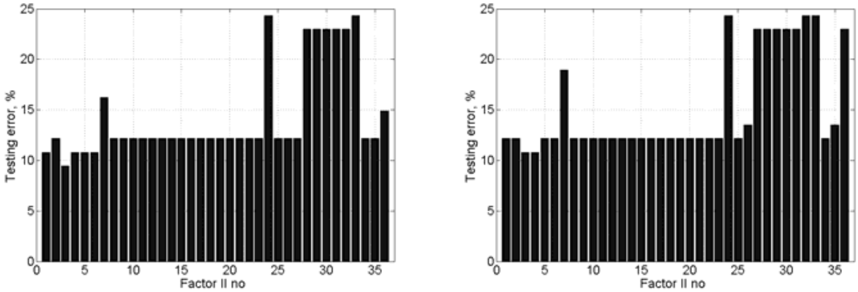


Fig. 4. Influence of measure choice in stage II on the damage classification error for registered vibration signal a_x, Δf=500 Hz (a), Δf=1000 Hz (b) – the best results achieved

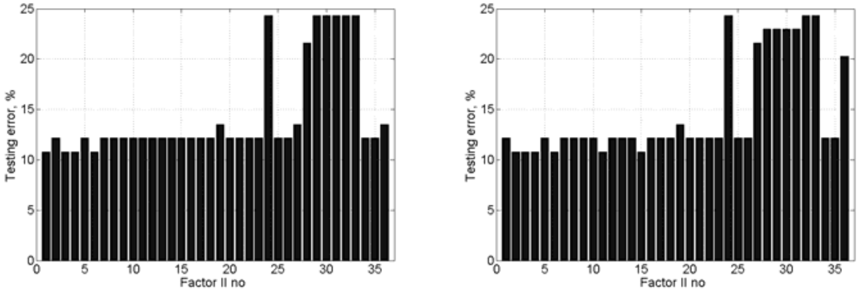


Fig. 5. Influence of measure choice in stage II on the damage classification error for registered vibration signal a_y, Δf=500 Hz (a), Δf=1000 Hz (b) – the best results achieved

Independently of the choice of the direction of vibration signal registration and of the way the spectrum is divided into ranges of different width, for most cases of applied measures in stage II a similar level of correctness of damage classification was achieved. The best achieved results were on the level of 10% of classification error.

Due to the fact that the experiment ended with a success an attempt was made to classify the damage degree of the inlet system. RBF type neural networks were taught on the same models as in the previous experiment. The output data achieved from the neural network consisted in this case of information concerning qualification of leakages of the inlet system of the engine in one of four stages.

The best achieved results for vibration signals registered in directions parallel and perpendicular to the cylinder axis are shown in Fig. 6 and 7.

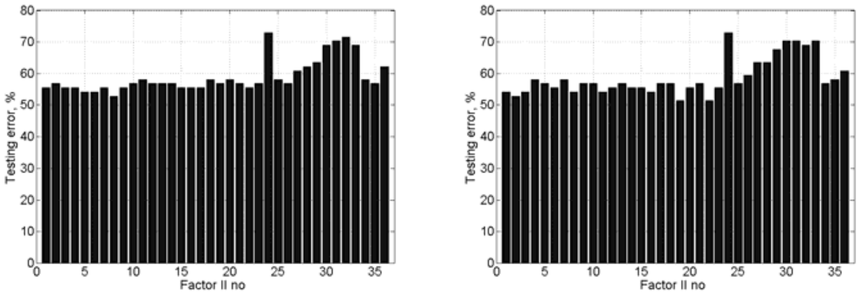


Fig. 6. Influence of measure choice in stage II on the damage classification error for registered vibration signal a_x, Δf=500 Hz (a), Δf=1000 Hz (b) – the best results achieved

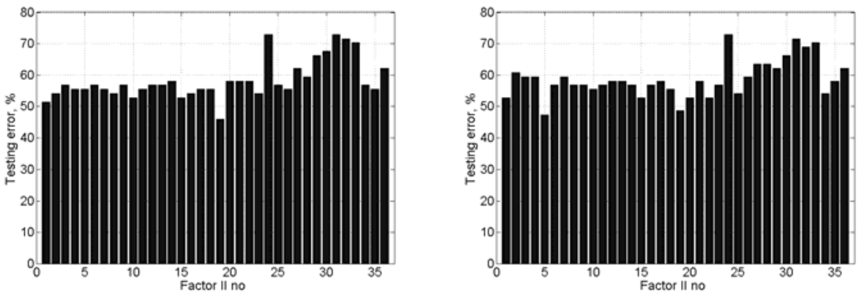


Fig. 7. Influence of measure choice in stage II on the damage classification error for registered vibration signal a_y, Δf=500 Hz (a), Δf=1000 Hz (b) – the best results achieved

Results of experiment did not bring expected positive results. Irrespective of the way of constructing models of leakages of the inlet system in an internal combustion engine, RBF type neural networks were unable to learn to identify them correctly.

That is why further experiments need to be carried out with the use of other types of neural networks or work out a different way of constructing descriptors meant for teaching networks so as to definitely determine the assessment possibility of condition of leakage in the inlet system of an internal combustion engine.

3 Conclusion

The operation conditions affecting the vehicle during driving cause constant changes of driving motor load. All manoeuvres which are performed on the road require an adjustment of the crankshaft rotational speed to them as well as the right filling of the cylinders. Everything takes place very fast and the control system of the motor need to deal with it. The basic aim of the cylinder filling control is to provide the right air mass. Thanks to it, the right conditions of engine work can be achieved at a decreased level of used fuel and decreased emission of harmful substances to the atmosphere.

The article presents the results of tests which aimed at the determination of leakages in the inlet system of an internal combustion engine on the basis of generated vibration effects. RBF type neural networks were used to perform experiments. Descriptors were built with the use of time-frequency distributions achieved from the Wigner-Ville transform.

As a result of performed experiments the method of building models was worked out and on their basis the way to teach a neural network how to classify a good condition and occurrence of leakages in the inlet system of engine was found. The experiments which were aimed at the identification of degree of occurring leakage did not work out. They require carrying out further research.

References

1. Czech, P., Łazarz, B., Wojnar, G.: Detection of local defects of gear teeth using artificial neural networks and genetic algorithms. ITE, Radom (2007)
2. Grega, R., Krajňák, J., Kaššay, P.: Impact of pneumatic flexible coupling on effective vibration value within the mechanical systems. Transactions of the Universities of Košice 3 (2009)
3. Heywood, J.B.: Internal combustion engines fundamentals. McGraw Hill Inc. (1988)
4. Homišin, J.: Tuning methods of mechanical systems by means of torsional oscillation tuner application. Pneumatyka 61(6) (2006)
5. Kneba, Z., Makowski, S.: Power supply and steering of the engines. WKiŁ, Warsaw (2004)
6. Madej, H.: Diagnosing the mechanical defects in internal combustion engines masked by electronic steering devices. ITeE, Radom (2009)
7. Merkisz, J., Mazurek, S.: On-board diagnostics OBD in an automotive context. WKŁ, Warsaw (2007)
8. Mikulski, J.: The Possibility of Using Telematics in Urban Transportation. In: Mikulski, J. (ed.) TST 2011. CCIS, vol. 239, pp. 54–69. Springer, Heidelberg (2011)
9. Młyńczak, J.: Analysis of Intelligent Transport Systems (ITS) in Public Transport of Upper Silesia. In: Mikulski, J. (ed.) TST 2011. CCIS, vol. 239, pp. 164–171. Springer, Heidelberg (2011)

10. Osowski, S.: Neural networks for information processing. House of University of Technology in Warsaw (2000)
11. Tadeusiewicz, R., Lula, P.: Introduction to neural networks. StatSoft, Cracow (2001)
12. Wajand, J.A.: Piston internal combustion engines of average rotational speed and high-speed. WNT, Warsaw (2005)
13. Wendeker, M.: Steering of ignition in car engine. On-board diagnostic systems of car vehicles. Lublin (1999)

Implementation and Managing of Innovation in the Conditions of Legal and Economic Constraints on the Based of Rail Transport

Adam Jabłoński and Marek Jabłoński

Southern Railway Cluster Association
Ligocka 103, bud. Nr 7, 40-568 Katowice, Poland
{adam.jablonski,marek.jablonski}@ottima-plus.com.pl

Abstract. The authors of the article present the key problems in implementing innovation in Polish railway companies in the context of legal and economic constraints. Legal requirements resulting from the processes of authorizing structural subsystems and interoperability constituents determine application of specific approaches to the creation of innovation in rail transport. A process of developing and implementing innovation in rail transport requires a number of test, experimenting research and pilot analysis before a product has a chance of developing its capacity for its practical application. The author presents own approach to defining innovation in rail transport and the problems during its implementation and identifies barriers in technology transfer in the context of the requirements for safety of railway traffic.

Keywords: innovation, a railway company, interoperability, rail transport.

1 Introduction

Nowadays the Polish railway requires a new approach to development. On the rail market there is an extensive network of companies focused on production, supply of products and provision of services related to ensuring continuity of transport processes. This applies both to activities related to transport of goods or expert services, such as certification of products, services, research aimed to certify products fit for their use, laboratory tests, expert evaluations, design, etc. All these activities are done in the conditions of legal and economic constraints, which significantly influence development of economic activity in rail transport. Due to the fact that in the last 20 years investment into development of rail transport in Poland was inadequate, an interest in the industry declined significantly, which contributed greatly to a low level of innovation in this sector in the Polish conditions, contrary to European Western countries like Germany, France, or Spain, where the rail industry is developing rapidly. In Poland, interest in rail transport is also increasing, which is indicated by an increase in rail transport performance, according to a number of industry reports. The European Union authorities put an emphasis on the need for

J. Mikulski (Ed.): TST 2012, CCIS 329, pp. 423–432, 2012.

developing rail transport, particularly for growth of its innovation[1], in many documents of strategic importance for its development. In Poland in 2011 compared to 2010 263.8 million passengers used railway services. While carrying them work was done at the level of 18.1 billion- passenger- kilometers. In comparison with 2010, a number of passengers carried increased by 1.5 million (0.6%). The volume of transport performance increased by 181 million-passenger-kilometers (1%). In 2011 licensed carriers of goods transported 249.3 million tonnes of goods, by 13.8 million more than in 2010. An increase in weight transported was 5.9%. Carriers' transport performance equaled 54 195 million, increased by 5 353 million tonne-kilometers (10.9%) [1]. This is not a significant increase, but there is consistent progress in this respect. The White Paper, the Plan to establish a single European transport area, striving to achieve a competitive and resource-efficient transport system in the European strategy for research, innovation and implementation in the field of transport, says, inter alia, that technological innovation can ensure that a more efficient transport system in Europe will be achieved more quickly and easily. Moreover, it will be in conformity with sustainable development, affecting three main factors: efficiency of vehicles thanks to new engines, materials and construction, using more green energy thanks to new fuels and propelling systems, a better use of networks and safer and more reliable operation owing to information and communication systems. Making research on technologies more effective requires a systemic approach, which will include infrastructure and regulation requirements, coordination of various entities and big demonstration projects encouraging their launch on the market. This will also involve introducing intelligent mobility systems developed under the EU-funded research, such as the Future Air Traffic Management System (SESAR), the European Rail Traffic Management System (ERTMS), and railway information systems, Maritime Surveillance Systems (SafeSeaNet), River Information Services (RIS), Intelligent Transport Systems (ITS) and combined interoperable solutions for information and management systems for next-generation multi-modal transport. Innovation and implementation must be accompanied by regulatory standards. Moreover, in the area of using more efficient means of transport, the EU's objective is to finish a rapid European rail network by 2050, a threefold increase in the existing speed rail network by 2030, and retaining dense railway network in all member states. By 2050, bigger parts of passenger traffic on average distances should be by rail. This overriding document, strategic in its nature, provides a platform for implementing an innovation strategy at the national level, and at the level of specific companies. The emphasis on innovation should be the base for

---

[1] On 28.03.2011 the European Commission adopted a document of the White Paper– Transport until 2050. The strategy intends to transfer from road transport to environmentally-friendly transport branches, and rail transport in particular and to reduce emission of harmful substances into environment. It is essential to use modern and innovative traffic and information management technologies (e.g. ERTMS, optimizing logistic and market chains of the European transport area, and striving to achieve a competitive and resource-efficient transport system Bruksela 28.03.2011.

building competitiveness of Polish railway companies and effectiveness of their business models.

2 Innovation in Business Models of Polish Railway Companies

Innovation should be a driving force of designing modern business models. Innovation may occur at the company level, and may refer to development of products, processes and work organization. Various definitions of innovation are shown in Table 1.

Table 1. Font sizes of headings. Table captions should always be positioned *above* the tables.

No.	Innovation definitions
1	Innovation is deliberate changes people make or cybernetic systems they design, which replace the existing state of affairs with the others, positively evaluated in the light of certain criteria and making up progress.
2	Innovation is a product (idea, method, thing) perceived by an adoption entity as norms, about which information is passed on through certain means of communication. The entity can be both individual and a social group.
3	Innovation means a new, yet unknown way of meeting new needs.
4	Innovation is changes in the production process connected with different quality of inputs and production factors.
5	Innovation refers to any good, service or idea that somebody perceives as new. The idea may exist for a long time, but it is an innovation for the person who sees it as new.

Source: Own study based on [2].

The definitions of innovation presented in table1 imply interdisciplinary nature of the concept. Using innovation and applying it to improve efficiency of railway company management seem to be a fundamental assumption.

The next phase of innovation is its use in company development, in some cases even leading to changing its business model, especially while changing configuration of the supervision chain.

The creator of the innovative theory of a company was J. Schumpeter [3] and this theory is built from innovation. Institutional and organizational boundaries between the company, market and household economies are formed by maximizing benefits of introducing innovation. A company is an organization and institution of economy, which by its activity implements inventions (that is, it makes innovation) in the following forms: new products, new technologies. new markets, new organisation of products and sales. In [3] you obtain an entrepreneur status only when you are a creator of something new: new products, new technologies, new markets, new methods of organizing production and new sources of raw materials. According to the author, not all businesses can survive in the economy. Those that survive are the most enterprising, they decide about healthy growth and economic development, and development must be by nature imbalanced. Another major theory of an innovation-based company is the theory of J. Schmookler. The elements of this theory are processes of creating inventions (ideas). Institutional and organizational boundaries between the company, market and households are formed by minimizing costs of creating inventions. A company is an organization that deals with both manufacturing and service activity, but its most important distinctive feature is that it uses its major resources to search for new technological solutions (products, technology, knowledge, and thus intellectual capital) [4].

The above mentioned assumptions of the theory of innovation may be a starting point for a development strategy of Polish railway companies, which have a lot to do in the area of innovation. It refers mainly to typical railway companies, namely licensed Railway Undertakings, and authorized Infrastructure Managers. Leading manufacturing plants of rolling stock have relatively well-developed research and development processes.

3 Barriers to Development of Innovation of Polish Railway Companies

As it has been written above, according to one of the leading theories of a company by J. Schumpeter, an objective of company existence is to launch innovative solutions that determine its development.

In general, solutions used in the Polish railways, are the solutions of old type, which in the coming years should be replaced with modern ones, in conformity with the objectives of the EU transport policy to ensure interoperability of rail transport.

Interoperability of a railway system – it is an ability of the railway system to ensure safe and uninterrupted traffic of trains, fulfilling technical, motor, operating and legal requirements, whose application ensures adherence to essential requirements for railway system interoperability and enables efficient traffic of the trans-European railway network [5]. On the one hand interoperability requirements demand something from the railway systems in the area of their production and operation, specifying certain limits of technical compliance with other systems. On the other hand they allow unlimited implementation of innovative solutions in the area of pre-defined criteria. This process enables the creation and implementation of innovative solutions, which preceded by a number of text studies and analyses can be successfully put into practice of the railway industry. [6] highlight the fact that innovation can be categorized as strategic, operational, or related to changing business models.

It seems that in Polish conditions in the railway sector, the majority of innovations will be connected with processes. As a result, they will be categorized as operational. Innovations implemented at the level of a railway company itself are used to improve its efficiency, reduce operating costs of the transport system and improve its safety.

With reference to the railway company interfaces with environment, new business models that can change the entire industry may be established. For example, in the framework of intermodal transport [7] define business models for intermodal transport, whose popularity is rising now and therefore new assumptions of the value chain are suggested. They highlight the following proposals for business models, namely: a model of an intermodal operator, a railway company acting as an operator and provider of logistic services, acting as operator, and their combinations in the context of specific transport corridors.

Innovations will therefore play a leading role in the functioning and development process of railway companies. Fig. 1 presents levels of innovation for railway undertakings.

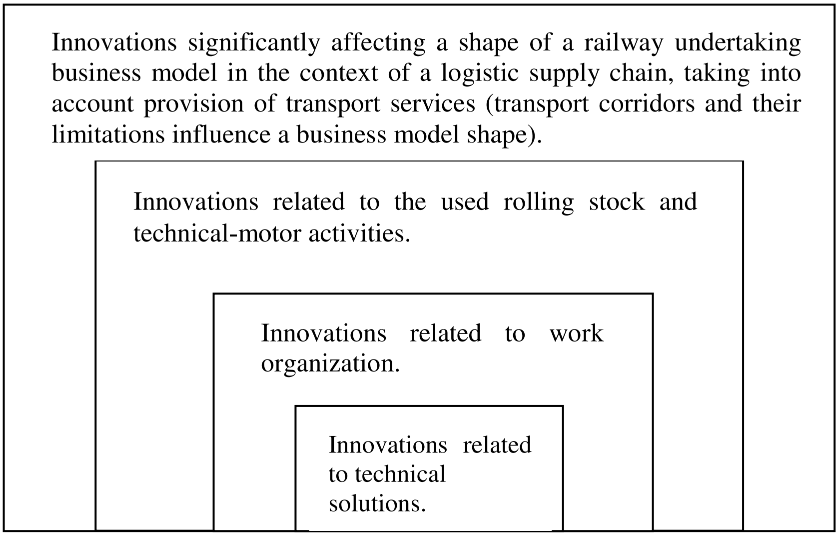


Fig. 1. The proposed levels of innovation on the example of a railway undertaking
Source: Own study

The presented model of innovation levels shows that legal requirements, for example in the area of interoperability will be classified on the lowest level of technical applications, while transport logistics, in the context of geographical constraints (the sites of transport services) may create innovations connected to a proposal of value for customers and related logic of earning income by a company. It should be noted that Polish railway companies now suffer so called technological gap. This is due to the difference in the level of scientific and technical thinking in relation to those countries in Europe and the world, where the most advanced countries may allocate a higher percentage of national income to scientific research.

Although in Poland there are a lot of research centers dealing with rail transport, their financial strength does not allow them to carry out projects that would bring revolutionary solutions, including patents, which would be widely adopted by the European or global market. Of course you should not generalize because now there are numerous examples of innovation, especially in production of rolling stock. They are not, however, groundbreaking, as they often copy European solutions but at the lowest costs. According to [8] international transfer of technical innovations is a necessity in the conditions of deepening specialization in the field of scientific and technical research. The essence of import is a possibility to obtain not only the material means, but often such a resource of technical knowledge (skills) which will be exploited [8]. Figure 2 presents the key factors influencing the process of implementing innovation in rail transport.

Idea, product/service concepts	
Concept implementation	**Concept copyright**
Idea/market analysis Techncial analysis e.g. in TSI (Technical Specification for Interoperability) aspect. Designing construction Technical analysis, including assessment of a product life cycle, e.g. using RAMS [9] Building a prototype Research, tests, analyses etc. Designing technology Production pre-series Tests, analyses Certifying the type fit for use Production	Identifying legal aspects of a product Recognizing economic aspects of product copyright in the country, Europe and in the world Patenting, trademark etc.

Fig. 2. A general process of implementing innovation in rail transport
Source: Own study

4 Managing Change and Innovation in Rail Transport

By principle, rail systems are designed and used bearing a safety criterion in mind. Any changes in the technical railway system should be made in accordance with appropriate procedures ensuring an acceptable level of safety. Introduction of innovations to the practice of the railway sector should be preceded by a procedure of common safety assessment in terms of risk valuation and assessment [10].

CSM (Common Safety Method) related to valuation and risk assessment is applicable to any changes in the railway system, which are considered to be significant. Such changes may be technical, operational or organizational.

If a proposed change has an impact on safety, a railway company, guided by professional judgment, decides about its importance, on the basis of the following criteria:

1) effects of system failures and a reliable worst-case scenario in the event of failure of the assessed system, taking into account the existence of safety barriers outside the system;

2) innovation used in introducing a change: this criterion includes innovations in both the railway sector, and only in the organization implementing the change

3) the complexity of the change;

4) monitoring: reliability of monitoring the implemented change during the entire life cycle of the system and undertaking appropriate intervention;

5) reversibility of the change: an inability to return to the system prior to the change;

6) additionality: change assessment, taking into account all the recent changes in the evaluated system, which were related to safety and were not assessed as significant.

Innovations, as defined in section 2, may affect even the entire railway sector. The changes may result from implementing technical, operational and organizational innovation. In this case, the innovation should be implemented taking into account a safety criterion. The safety criterion is leading, as far as barriers in rail transport are concerned, which creators of innovation must overcome. Knowledge of the TSI requirements (Technical Specification for Interoperability) and a number of additional criteria for accepting this regulation are the basis for implementing innovation in this sector. Therefore, the difficulty in implementing solutions for the railway results from the complexity of the railway system and its mutual interfaces inside and outside the system, but it reasonably relates to the safety criterion.

An example of labour intensity and high complexity of implementing innovative solutions in the railway sector is a process of implementing the Rail Traffic Management System (ERTMS). According to the idea of interoperability of European railways, written in the European Union directives such as 91/440/EEC, 96/48/EC, 2001/16/EC, 2004/49/EC, 2004/50/EC, 2008/57/EC, in order to improve international rail transport, it is necessary to integrate and harmonize the activities of national management systems. Examining the layered division of traffic control devices, you may notice that the strongest need to implement interoperability in the control of train movement and train communication systems concerns the lowest (basic) layer. This is a consequence of the fact that uncoordinated development in the past led to creating technically heterogeneous, different train control systems[2]. Rüdiger K. Hase highlights the problems in implementing this system, in terms of the complexity of

[2] See: Commission Regulation (EC) No 352/2009 of 24. April 2009, on the adoption of common safety methods in the valuation and risk assessment.

principles of accepting various methods of this system in the context of practical implementation [11].

In this example a notified certification body in the process of rail traffic control device approval plays a key role in launching a product.

Another example of difficulties in implementing innovations in rail transport is implementing ECM (Entity in Charge of Maintenance) models. Like in most areas of the railways, a criterion of access to the sector in many cases depends on the certification process. The certification process carried out in Polish conditions by the Railway Transport Office requires a proof that requirements of the MMS (Management System Maintenance) have been fulfilled, but it is also necessary to present a model of an organization and its responsibility structure, requiring a new innovative business model of an entity responsible for maintaining freight wagons. Figure 3 shows the business models proposed by the European Railway Agency ERA [12].

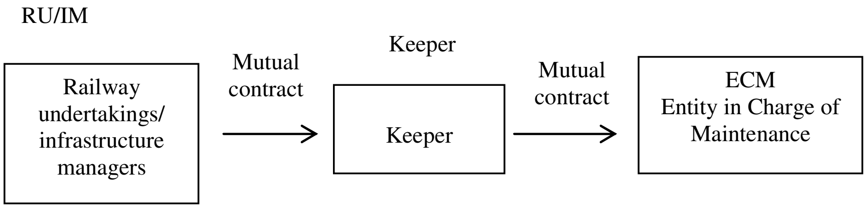


Fig. 3. A general model of the relationship between the ECM, keeper and RU. Source [12].

In addition to the standard model, there are other solutions – Fig. 4.

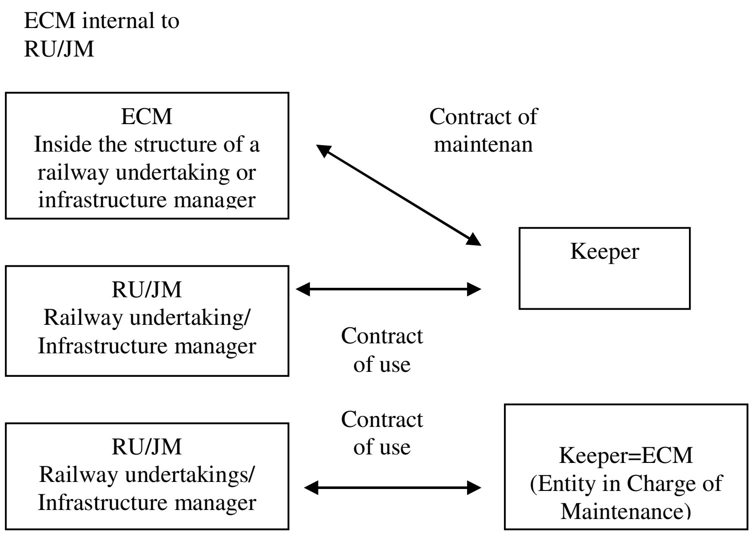


Fig. 4. Other possible relationship between ECM, Keeper and RU. Source [12].

The ERA (European Railway Agency) guides, as well as legal requirements in rail transport create the boundaries of innovation. They determine their barriers and opportunities for creating innovations, including the needs for safety, reliability, usability, cost-efficiency, energy efficiency, environmental impact, quality, and other characteristics of a product or service.

Knowledge of product certification processes, motivation, accreditation and authorization of certifying bodies and laboratories is essential to create innovation in rail transport. It is also often a barrier in implementing modern solutions. It significantly affects the duration of the procedure of introducing innovations onto the market and time to achieve economic growth.

New legal requirements connected with the legislation of EU law on interoperability and a procedure of authorizing the use of the subsystem have complicated the process significantly.

At the request of a subsystem manufacturer or his authorized representative, a manager, railway undertaking, keeper, importer, investor, or a contracting entity, a notified certification body verifies (EC) on the basis of the TSI if the subsystem fulfills the essential requirements for interoperability of the rail system. A above-described mode is uniform throughout the EU, therefore it should not be treated as a significant barrier in implementing innovation. Nevertheless, at present in Poland it is necessary to disseminate knowledge on the subject to focus primarily on creating innovation and not only on issues related to its practical implementation in the railway sector.

References

1. Raport Urzędu Transportu Kolejowego na temat przewozów koleją (Report of the Railway Transport Office on rail transport), UTK (February 2, 2012)
2. Penc, J.: Innowacje i zmiany w firmie, Transformacje i sterowanie rozwojem przedsiębiorstw (Innovations and changes in the company, Transformations and controlling development of businesses), pp. 141–143. Placet Publishing Agency (1999)
3. Schumpeter, J.A.: Teoria rozwoju gospodarczego (Theoretical problems of economic growth), p. 104 and further. PWN (1960)
4. Noga, A.: Teorie przedsiębiorstw (Theories of companies), pp. 150–152. PWE, Warszawa (2009)
5. Railway Transport Act of 28 March 2003, as amended, Art. 4. par. 29
6. Prahalad, C.K., Krishnan, M.S.: Nowa Era Innowacji (The New Age of Innovation), p. 41. PWN Publishing House, Warszawa (2010)
7. Lehtinen, I., Bask, A.H.: Analysis of business models for potential 3 Mode transport corridor. Journal of Transport Georgaphy 22, 49 (2012)
8. Janasz, W. (ed.): Innowacje w modelach działalności przedsiębiorstw (Innovations in models of company activity), p. 62. University of Szczecin, Szczecin (2003)
9. RAMS (Reliability, Availability, Maintainability, Safety)
10. König, M., Völker, R.: Innovations – Management in der Industrie, pp. 168–169. Carl Hanser Verlag, München Wien (2002)

11. Białoń, A., Gradowski, P., Toruń, A.: Nowoczesny system zarządzania ruchem kolejowym (Modern Rail Traffic Management System) (ERMTS). Problemy Kolejnictwa (Problems of the Railway) (148), 138–139 (2009)
12. Rüdiger Hase, K.: Applying 'Open Proofs' to European Train Control. European Railway Review 18(3) (2012)
13. Guide for the application of the Commission Regulation No 445/2011 on a system of certification of entities in chargé of maintenance for freight wagons, Version 0.5, p. 22 (December 17, 2011)

Author Index