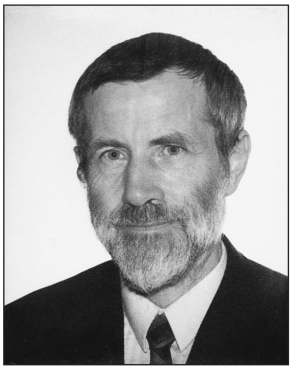
KOMUNIKÁCIE
COMMUNICATIONS



Dear reader,

You have got in your hands this volume of the university scientific letters which is again devoted mostly to informatics and its applications to a broad spectrum of scientific and professional branches.

Nowadays, various forms of informatics penetrate almost every human activity and this issue tries to reflect these phenomena. Inside this volume you will find papers written by authors not only from the Faculty of Management Science and Informatics or other faculties of the University of Žilina, but we also addressed professionals from other cooperating universities and asked them to contribute to the topic of this volume.

In the frame of this issue, there are articles dealing with optimisation techniques supported by means of informatics rather than pure problems of informatics. It is no surprise that most of the contributions are devoted to applications of informatics to transport. This issue continues with the topics concerning intelligent transportation systems and various sorts of routing and scheduling problems. Attention is also paid to economic problems and distribution logistics.

I would like to express my opinion that this volume would attract attention of professionals from relevant scientific branches and ignite their interest in some future cooperation in the area of sophisticated decision support tools as an enhancement of information systems.

Jaroslav Janáček

Peter Matis *

MANAGEMENT OF STREET ROUTING PROBLEMS USING DECISION SUPPORT SYSTEM

Servicing a large number of customers in a city zone is often a part of many logistics chains. The capacity of one delivery vehicle is limited, and at the same time, it usually serves a large number of customers. These problems are NP hard and often called a Street Routing Problem (SRP). Problems of this category are similar to the problems named as Vehicle Routing Problem (VRP). In this paper, these problems are explored and potential solution paths for their management are described. As presented in the paper, the management of street routing problems could be effectively done only by using decision support systems or other informatics systems, such as Geographical Information Systems (GIS).

1. Introduction

Street Routing Problems (SRP) as a problem of servicing a large number of customers in a city zone is often a part of many logistics chains. Into the SRP category we can include problems like newspaper delivery to the corner shops, waste removal, postal service delivery, commercial freight delivery, and meter reading.

Practical experience from the solution of the SRP in the Slovak Republic suggests that they are in most cases solved by an expert's experienced judgment. The main problem of SRP is that the number of customers is large and the number of delivery path combinations is enormous. It is not possible for a human to explore even a fragment of all these combinations. Experts usually take the first acceptable solution, which may be far from the optimal solution. As the experimental results show in the case of SRP, the error on the length of delivery routes based on an expert's judgment when compared to optimal solution is in the range of 10 % – 25 %.

The SRP is in many cases similar to the classical Vehicle Routing Problem (VRP). There are some important differences between the SRP and VRP.

- The number of customers in the SRP is large compared to the VRP. In the VRP the number of customers is in the hundreds, and in the SRP, it could be hundreds of thousand.
- The service time for one customer in the SRP is short and the route path for one vehicle contains several hundreds of these customers.
- The density of the street network in the SRP is high compared to the regular road network in the VRP.
- The distance between two close customers in the SRP is small and it is possible to get from one customer to another customer through a small number of possible network connections.
- Traffic regulations in the city zone are more complicated and more restrictive than traffic regulations in a regular road network. These regulations could greatly affect the expense for the routes.
- The access points to the customers are important and have a large effect on the total cost.

Formerly, dispatchers could communicate with the drivers of the service vehicles only using radio stations. Nowadays, they can also use other technologies, such as mobile phones, palmtops connected to the wireless internet, and GPS together with GIS map databases. In these days they could assign jobs for each vehicle, even outside of the dispatching center, as well as help navigate the driver in an unknown area.

2. Solution of SRP using VRP techniques

As shown in the previous section, there are some common properties between the SRP and VRP and some differences between them. In 2005, we explored a possible use of VRP techniques for solving SRP instances [4]. We were solving the problem of postal service in the city of Galanta. Customers in the postal deliveries usually are each house or doors in the large blocks. Customers are usually located on both sides of the street segments and could be serviced differently depending on local conditions and other regulations. In Figure 1 there are three typical cases.

As our experiments show, applying VRP techniques on these SRP cases directly, that means taking each customer in the SRP as a customer of the VRP, the results are usually not acceptable. Observe in Figure 2 that customers on one street segment are serviced by two vehicle routes.

As the result of our experiments, we can say:

- SRP solutions take more time than typical VRP solutions because of local restrictions, combined with the number of customers in the typical SRP.
- To be able to compute SRP in most cases, number of customers should be decreased by their aggregation.

* **Peter Matis**
Department of Transportation Systems, Faculty of Management Sciences and Informatics, University of Zilina, E-mal: Peter.Matis@fria.utc.sk

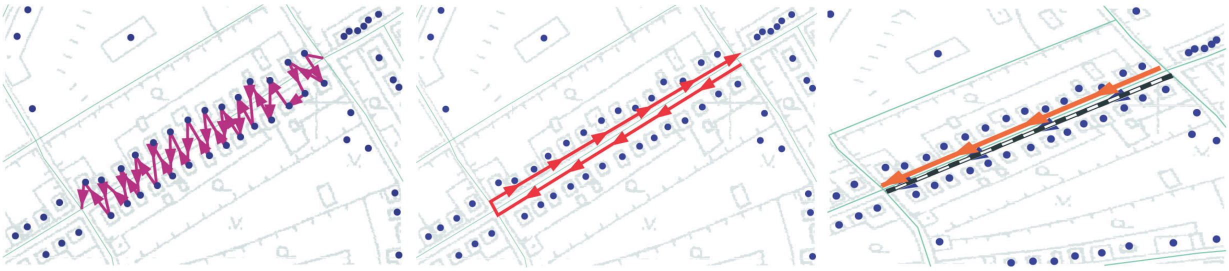


Fig. 1 Cases of service on the street segments

- Expenses to maintain customer and street network databases are high.

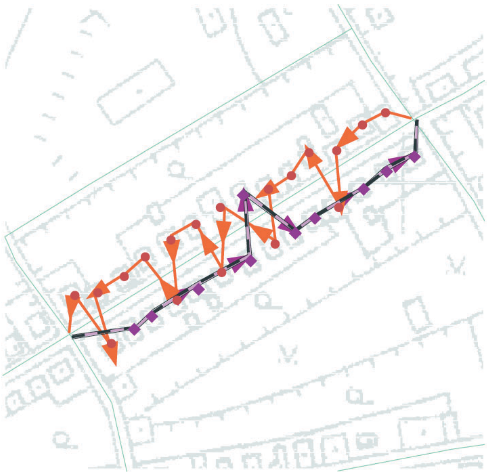


Fig. 2 Solutions of SRP using VRP strategies are not acceptable

One of the important measures for a good solution of the SRP is the parameter that tells us about the balancing of routes. It is almost always important to balance the routes, so that there are no cases when one postman has to always work overtime while another postman is using only part of his or her capacity. As presented in Tables 1 and 2, using the proper algorithm for the SRP solution, we were able to balance postal routes. By using real street networks, the average traveling speed and measured service times, we were able to decrease the number of routes in the city of Galanta by two. These results also show that no route lasts longer then 285 minutes, which was our limit. This is 48 minutes less than the longest route in the current situation. The difference in time between the shortest and the longest route is 93 minutes.

3. Customer aggregation

Practical experiments show that if we want to apply the VRP heuristics to the solution of the SRP, we need to aggregate customers to be able to solve the problem. One of the purposes for doing it is the fact that computational time expands with large numbers of customers and, by aggregation we keep this number in a range that is comparable to the typical VRP.

Current postal routes in Galanta — Tab. 1

Route	Length [m]	Total time [min]
I	8367	277
II	9881	271
III	12378	333
IV	8985	233
V	5969	182
VI	5815	233
VII	12533	257
VIII	9506	302
IX	11715	307
X	6888	326
XI	6037	268
XII	4403	172

Computed postal routes in Galanta — Tab. 2

Route	Length [m]	Total time [min]
A	7200	283
B	12450	279
C	6110	282
D	3570	284
E	6160	285
F	5020	283
G	8145	255
H	5972	266
I	9240	285
J	5380	192

- In distance aggregation, we aggregate customers based on their distance. It is possible to aggregate customers that are located on different street segment(s).
- Line aggregation allows aggregation only of customers that are located on the same side of one street segment.
- Line-segment aggregation is the special case of line aggregation where all customers on one side of one street segment create one aggregate.
- Address aggregation allows the combination of customers that have similar address.

Properties of aggregations have been explored using the following experiment. Settlements in Slovakia were chosen, and divided into 5 categories. For each settlement there was a known street network and location of customers. For the line and the distance aggregation, the maximum distance between two customers in the aggregate was set to 80 m. In all the cases, we have set customer delivery time to 10 sec and requested quantity to 2 units. We controlled the total time for the delivery inside of the aggregate area and limited it to 10 minutes and total requested units to 100. We can define the total delivery time as:

$$t_{aggregate} = \left(\sum_{i \in aggregate} t_i \right) + \max_{i,\, j \in agreggate}(c_{ij}) \quad / \; v \qquad (1)$$

Where is a service time for customer i, v is the average traveling vehicle speed (in our experiment, it was set as 4.5 km/hour) and c_{ij} is a network distance between customers i and j. For a better evaluation, an estimation could be used based on the length of the traveling salesman problem for all of the customers in the aggregate. Table 3 presents these times.

Experiments have show that the best aggregation is distance aggregation. On the other hand distance aggregation shows high deformation of aggregated positions against the original positions of customers.

4. Transformation of Arc Routing Problems to Node Routing Problems

Many SRPs could be viewed as Arc Routing Problems (ARP). There is a special case of ARP called the Capacitated Arc Routing Problem (CARP) [3], [1]. We present the definition on the graph. Let $G = (V, E)$ be an undirected graph with a set V of n nodes and a set E of m edges. A fleet of identical vehicles of capacity Q is based at a depot node S. A subset of r edges $R \subset E$, denoted as *required edges*, require service by a vehicle. Any edge can be traversed any number of times. Each edge (i, j) has a traversal cost c_{ij} and a demand q_{ij}. The goal is to determine a set of vehicle trips (routes) of minimum total cost, so that each trip starts and ends at the depot, each required edge is serviced by one single trip, and the total demand handled by any vehicle does not exceed Q.

CARP is an NP hard problem and currently there does not exist any optimal algorithm to solve cases of a realistic size. There are heuristics to solve real size CARPs, their performance is in many cases not acceptable [2]. One possible solution of this problem is to transform a CARP to a capacitated vehicle routing problem (CVRP). There are several good heuristics that perform well on realistic size problems.

Let's transform edge $(i, j) \in R$ into the nodes s_{ij}, s_{ji}, which represents edge's end points. The CVRP is then defined on the complete graph $H = (N, A)$, where the set of the nodes N is defined:

$$N = N' \cup \{S\} = \bigcup_{(i,\, j) \in R} \{s_{ij}, s_{ji}\} \cup \{S\}$$

Demand in the new nodes is defined as:

$$q(s_{ij}) = q(s_{ji}) = \frac{1}{2} q(i, j)$$

Distances in H are defined as:

$$d(s_{ij}, s_{kl}) = \begin{cases} 0 & (i, j) = (k, l) \\ c_{ij} & (i, j) = (l, k) \\ s\_path(i, k) & (i, j) \neq (k, l);\ (i, j) \neq (l, k) \end{cases}$$

$s\_path(i, j)$ represents the shortest path between the nodes i and j on the original graph G. Transformation is presented in figure 3.

Characteristics of Aggregations

Tab. 3

Cat	NCS	TNC	ANC	Agreg	NA	ATA	MTA	MeTA	SD
400–4000	75	32150	428	DIST	8167	60.2	212	56	44.7
				LINE	12096	41	183.9	35.5	32.2
				SEG	5285	130.6	600	77	158.7
4000–10000	2	2035	1017	DIST	1165	45.7	185.1	36.6	40.5
				LINE	1495	35.4	183.9	10	31.4
				SEG	579	141.5	600	65.4	188.4
30000	1	1616	1616	DIST	429	65	218	61.5	47.8
				LINE	680	41.5	147.7	35.4	31.5
				SEG	402	87	600	58.2	97.3
Average				DIST	4254	57.3	220.8	51.2	44.5
				LINE	6264	39.8	184.2	29.9	32.6
				SEG	2738	120.4	600	67.9	147.3

Cat – Category by number of habitants in cities, NCS – Number of cities in the sample, TNC – Total number of customers, ANC – Average number of customers in city, Agreg – Agregation, NA – Number of aggregates, ATA – Average time to serve aggregate, MTA – Maximum time to serve aggregate, MeTA – Median time to serve aggregate, SD – Standard deviation, DIST – Distance aggregation, LINE – Line aggregation, SEG – Line-segment aggregation

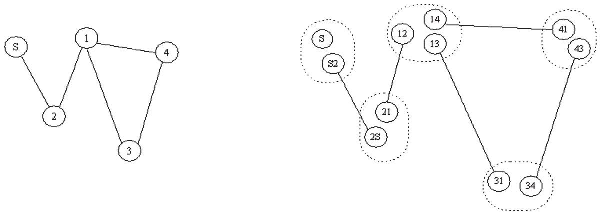


Fig. 3 Transformation of CARP to CVRP

Finally, we require that all edges $\{(s_{ij}, s_{ji}) \mid (i, j) \in R\}$ belong to the solution, i.e. we only accept CVRP solutions where s_{ij} and s_{ji} are visited in sequence, either $s_{ij} \rightarrow s_{ji}$ or $s_{ji} \rightarrow s_{ij}$. The transformed graph has $2*r+1$ nodes, and we expect that fixing these edges in the CVRP can make an algorithm perform almost as if on an instance with r nodes. [6]

5. Interactive methods and decisions support systems

Recent development in the real street routing problems shows that there is a need to make routing software a part of a larger system [7]. One of the possible solutions to this is to integrate routing software within GIS [5]. GIS can be helpful in the collection, storing and management of large geographical databases used in the routing software – i.e. transportation network databases and databases of customers. GIS can be also used for creation of all outputs from the routing software, including detailed maps for each route with the route description. One of the key properties of GIS for use with the routing software is the interactive and user-friendly environment. This is valuable for many real SRP, because as we have found out, there is always a need to work in a co-operation of expert-routing software. Routing software can find a good solution and explore the possibilities and an expert can change calculated routes to explore other possibilities based on the expert's judgment.

GIS have several useful features that could help improve routing software performance. To mention only a few here – we may pay attention to the capabilities as safe database management, flexible symbols, map management, drawing capabilities, safety and interoperability.

Decisions support systems (DSS) are important for solving real SRPs. One use of the routing software results is not only a solution of the problem, but also the ability to explore several other possibilities, or perform operations like exchanging parts of two routes or seeing how changes in the transportation network, regulations or policies could affect the routes, expenses and other parameters. Currently there exist two trends in the development of DSS for SRP.

- Independent software packages specialized for SRP with limited amount of DSS capabilities. Here we can note software like GeoRoute from the Canadian software firm GIRO. It is not an open system. It is an expensive product that costs 1,000,000 EUR for one installation. As an example of other such systems, we can use TRANSCAD created in the US, which has some similarities to GeoRoute. GeoRoute is more feasible for solving the SRP because it is specialized for the street routing. TRANSCAD is more specialized in the node routing.
- Integrated systems based on the GIS or CAD. As an example, we can put ArcView with its ArcGIS Network Analyst. It is specialized in the node routing. The user can implement extension and then make it an integral part of the whole system.

As presented here, these DSS are usually not appropriate for use with real SRP or they are expensive for use in middle size organizations that operate in cities with operations based on the routing. One can find it to be a good solution to use GIS and integrate routing software and DSS inside of GIS.

Conclusion

Paper summarizes the management of street routing problems. Management of the SRP is a new field of research and it is growing. There are expanding activities in the cities that can be presented as street-based tasks. Unfortunately, there is not a good methodology for solving the general SRP. A potential solution to the management of SRP is the integration of routing software with GIS and DSS. On of the possible solution for this problem is to aggregate the large number of customers into the clusters and then solve VRP. Paper shows characteristics of three different types of aggregations and shows possible development for future.

References

[1] AMPONSAH, S. K., SALHI, S.: *The investigation of a class of capacitated arc routing problems: the collection of garbage in developing countries*, Waste Management 24, 2004, p. 711–721.
[2] BEULLENS, P., MUYLDERMANS, L., CATTRYSSE, D., OUDHEUSDEN, D.: *A guided local search heuristic for the capacitated arc routing problem*, European Journal of Operational Research 147, 2003, p. 629–643.
[3] GOLDEN, B. WONG, R. T.: *Capacitated arc routing problems,* Networks 11, 1981, p. 305–315.
[4] JANÁČEK, J.: *Optimization on transportation networks,* EDIS-ŽU, Žilinská univerzita, 2003, p. 248.
[5] LONGLEY, P., BATTY, M.: *Spatial analysis: Modelling in a GIS environment.* John Wiley & Sons, Inc., New York, 1996, ISBN 0-470-23615-9.
[6] PEARN, W., ASSAD, A., GOLDEN, B.: *Transforming Arc Routing into Node Routing Problems,* Computers & Operations Research 14(4) 1987 p. 285–288.
[7] RUIZ, R., MAROTO, C., ALCARAZ, J.: *A decision support system for a real vehicle routing problem,* European Journal of Operational Research 153, 2004, p. 593–606.

Jan Pelikán *

A TIME LIMITED VEHICLE ROUTING PROBLEM

A vehicle routing problem is a classical problem in operation research consisting in delivery routes optimization in communications network containing depot of all routes and a given number of cities, which is necessary to include in delivery routes. The demand of all cities is given and the condition is that the sum of demands of the cities on the route should be less or equal to the capacity of a vehicle.

The paper deals with a modification of the vehicle routing problem in which the capacity of a vehicle is not limited but the times of pick-up to all cities are limited by a given value. A similar problem and model can be formulated for optimization of pick-up routes with pick-up time limited. The mathematical model proposed for the problem is based on Miller-Tucker-Zemlin formulation of the traveling salesman problem. As the time limited vehicle routing problem is NP hard, a solution to huge problems cannot be obtained in acceptable computation time. The following heuristics are proposed for the time limited traveling salesman problem: nearest neighborhood method, insert method and savings method. Use of all methods is illustrated by numerical experiment. The difference between results obtained by those methods is shown on the case study.

1. Introduction and a mathematical model

This modification of vehicle routing problem (or the traveling salesman problem) is based on the bank payment orders pick-up case study. The depot (called the centre) is the city, to which daily payment orders are transported from the bank branches. The payment orders have to be delivered to the centre in a given time, i.e. the time interval between loading payment orders at the branches and unloading them at the centre must not exceed a given value. Vehicle capacity is not the limiting factor in this problem. The transportation is executed at the public road network. The shortest distances between cities, or between the cities and the centre, are known.

Let us have n cities connected by the road network with the shortest distance matrix C. Let us suppose that the city 1 is the centre to which the payment orders are transported from other $n-1$ cities. The pick-up from each city must be finished in the time T, which forecloses the pick-up payment orders by one Hamiltonian cycle. The result forms more routes-cycles; each of them includes city 1 (as in the vehicle routing problem). The pick-up time T starts to be computed since the departure time from the first branch on the route. Thus, the traveling time from the city 1 to the first branch is not included in the time T. For the model of the problem we would need the traveling times between each pair of cities. Because of simplicity, these times are derived from the vehicle average speed and distances given by matrix C. Instead of a time limit to the pick-up time, the maximal route length to the centre 1 can be set. Let us denote it by L.

Comment. The vehicle routing problem, in which the delivery traveling time from city 1 to other $n-1$ cities[1] is limited, is similar to the above mentioned problem. The optimal routes in both problems are, in fact, the same except for their directions.

The time limited vehicle routing problem (TLVRP) is similarly as the traveling salesman problem (which is reduced on the vehicle routing problem in case of a big enough value L) NP hard.

Let us set the mathematical model of TLVRP first (for delivery case of TLVRP). The binary variable x_{ij} equals 1, if the edge (i, j) is included in the route, i.e. the vehicle goes from city i to city j; otherwise this variable equals zero.

The mathematical model of TLVRP is:

$$z = \sum_{i=1}^{n} \sum_{j=1}^{n} c_{ij} x_{ij} \rightarrow \min \qquad (1)$$

$$\sum_{j=1}^{n} x_{ij} = 1,\ i = 1, 2, ..., n, \qquad (2)$$

$$\sum_{i=1}^{n} x_{ij} = 1,\ i = 1, 2, ..., n, \qquad (3)$$

$$u_i + c_{ij} - M(1 - x_{ij}) \le u_j,$$
$$i = 1, 2, ..., n,\ j = 2, 3, ..., n,\ i \neq j, \qquad (4)$$

$$u_i \le L,\ i = 1, 2, ..., n, \qquad (5)$$

$$x_{ij} \in \{0,1\},\ i = 1, 2, ..., n,\ j = 1, 2, ..., n,\ i \neq j, \qquad (6)$$

The objective function (1) represents the total length of all the routes. The set of equations (2) assures that only one edge comes out from each city. Similarly, constraints (3) state that only one edge comes to each city. Equations (2) and (3) are not applied on city 1, because from and to city 1 there come as many edges as there are routes. The set of inequalities (4) defines variable u_i corresponding to the length of the route from city 1 to city i. The con-

* **Jan Pelikán**
Department of Econometrics, University of Economics Prague, W. Churchill sq. 4, 130 67 Praha 3, Czech Republic, E-mail: pelikan@vse.cz

[1] The return time from the last city to the city 1 is not included in that time.

ditions (5) mean that no such length exceeds the limit L. The constant M is a big enough positive number (or $L + c_{ij} - c_{1j}$).

2. Heuristic methods

As the TLVRP belongs to the class of NP hard problems, it will be useful to propose heuristic methods for a large scale problem solution. The heuristic methods used for the traveling salesman problem or the vehicle routing problem will be modified. It is assumed that the matrix C is symmetric and nonnegative, and $c_{1i} \leq L$ for $i = 2, 3, ..., n$ (otherwise, the TLVRP does not have a feasible solution).

The following notation will be used in the proposed modifications of heuristic methods. Let us denote by M the set of cities that were not included in any route. At the beginning of the method set M equals { 2, 3, ..., n}. The heuristic method ends when the set M is empty. The route will be denoted as $tr = (tr(1), tr(2), ..., tr(m))$, where $tr(1) = tr(m) = 1$. The constraints (5) will be tested for each possible change of the route (enlargement) in the form:

$$\text{P1: } \sum_{i=1}^{m-2} c_{tr(i),\, tr(i+1)} \leq L, \text{ or P2: } \sum_{i=2}^{m-1} c_{tr(i),\, tr(i+1)} \leq L.$$

If the condition P1 is satisfied, the route is feasible for delivery (in the case of pick-up the tr should be reversed). If the condition P2 is satisfied, the route is feasible for the pick-up; for delivery it is necessary to reverse tr.

2.1 The nearest neighborhood method

In this method the following steps are executed until the set M is empty:

Step 1. Let us denote the city with the shortest distance c_{1i} as k and let it form the route $tr(1) = 1$, $tr(2) = k$, $tr(3) = 1$, let it set $m = 3$. City k is deleted from the set M.

Step 2. Let us find city k from M which minimizes the distance $c_{tr(m-1),k}$ and for which the route enlargement tr when inserting this city after city $tr(m-1)$ satisfies the condition P1 or P2. If such a city k does not exist, the route is closed and the method follows by step 1 (starting the new route).

Step 3. Let us enlarge the route tr by inserting city k after the city $tr(m-1)$, let m increase by 1 and delete city k from set M. If set M is not empty, follow step 2; otherwise, the method ends.

2.2 The insert method

The following steps are taken until set M is empty.

Step 1. Let us denote city k with the greatest distance c_{1i} and put the route $tr(1) = 1$, $tr(2) = k$, $tr(3) = 1$, and $m = 3$. City k is deleted from set M. If set M is empty the method ends.

Step 2. City k from set M being found minimizes value $d = c_{tr(i),k} + c_{k,tr(i+1)} - c_{tr(i),\, tr(i+1)}$ for all $i = 1, 2, ..., m-1$ and $k \in M$. The route tr satisfies the condition P1 or P2 if this route is enlarged by inserting city k between cities $tr(i)$ and $tr(i+1)$, where i minimizes the value d. If city k does not exist, the route is closed and the method follows by step 1 (start a new route). Enlarge the route tr by inserting city k between the city $tr(i)$ and $tr(i+1)$, increase m by 1. City k will be deleted from set M. If set M is empty the method ends; otherwise, it follows by step 2.

2.3 The savings method

The following steps are executed until set M is empty.

Step 1. If set M contains only city k, the route $tr(1) = 1$, $tr(2) = k$, $tr(3) = 1$ is formed; let $m = 3$. City k will be deleted from set M and the method ends.

Step 2. Let it find the pair of cities from M in the form (k, l), that satisfies the condition and maximizes savings. If this pair (k, l) does not exist, form direct routes $(1, k, 1)$ for all each city k from M, stop.

Step 3. Put the route $tr(1) = 1$, $tr(2) = k$, $tr(3) = l$, $tr(4) = 1$, $m = 4$.

Step 4. Find city i from set M that maximizes s_{ik}, or city j from M that maximizes s_{lj}, such that the route, being formed by inserting i into the route before city k, satisfies conditions P1 or P2, or being formed by inserting j into the route after city l, satisfies conditions P1 or P2. If either i nor j do not exist, the route is closed and the method follows by step 1.

Step 3. If (or j does not exist), city i is inserted into the route before city k and city i is deleted from set M. Let s increase by 1 and k equal i. If (or i does not exist), city j is inserted into the route after city l and city j is deleted from M. Let m increase by 1. The method follows by step 4.

3. Numerical experiments

The proposed model and heuristics were applied to a set of problems and the results are shown in table 1. It contains 20 time limited vehicle routing problems, which are solved by using heuristics proposed above, by the model (1) – (6) and LINGO8.0 solver and CPLEX9.0 solver. The value of objective function obtained using heuristics (the best value of three heuristic methods) and model are shown. Beside the objective value, the computation time and lower bound of the objective function are placed in table 1. The computation time limit was 1 hour.

Computation experiments: results. Tab. 1

			LINGO				CPLEX	
	# nodes	heuristics	model	lower bound	comp. time	optimal	model	Lower bound
1	26	439.914	508	346	1 hour	no	418	347
2	19	345.36	324.36	–	3 mins	yes	324.36	–
3	12	271.097	271.097	–	1 min	yes	271.097	–
4	29	514.598	471.071	424	1 hour	no	491	417
5	14	261.402	253.283	–	1 min	yes	253.283	–
6	20	370.395	345.079	323	1 hour	no	345.079	345.079
7	22	501.555	493.55	390	1 hour	no	473	392
8	15	286.684	286.68	–	1 min	yes	286.68	–
9	25	346.494	320.58	–	14 mins	yes	320.58	–
10	11	157.64	154.315	–	1 min	yes	154.315	–
11	26	319.074	367.6	300.4	1 hour	no	340	301
12	32	619.617	837	450	1 hour	no	650	462
13	18	272.205	272.2	–	25 mins	yes	272.2	–
14	20	356.7	322.69	287	1 hour	no	322.69	298
15	27	383.177	442	307	1 hour	no	361	307
16	53	830.402	990.8	573	1 hour	no	868	573
17	16	235.833	233.6	–	25 mins	yes	233.6	–
18	27	434.497	473.7	334	1 hour	no	424	337
19	32	619.617	837	450	1 hour	no	650	462
20	16	219.477	204.67	–	1 min	yes	204.67	–

References

[1] GUTIN, G., PUNNEN, A. P.: *The traveling salesman problem and its variations.* Kluver 2002. ISBN 1-4020-0664-0

Zsolt Tibor Kosztyán – Zoltán Perjés – Andrea Bencsik *

COST REDUCTION BY MEANS OF ALTERNATIVE SOLUTIONS

In some projects the offered remuneration is lower than the estimated total costs. There are three possibilities to handle this problem. The first two is either to resign from the implementation, or to accept it knowing that we will loose money, but can regain it later in another project. In this latter case we have to determine the optimal resource allocation with minimal total cost. In the scheduling phase we can use some cost minimizing method. After that we can determine the lower and the upper bound of the start time of the activities. The next step is to determine a feasible solution and then the optimal resource allocation. The third possibility is to accomplish the project and find alternative implementation of activities that requires lower variable costs. During the search for alternative implementation of activities the most important aspect should be the quality, the decrease of costs is only the second one. If the total cost is lower than the offered remuneration we have to find the optimal resource allocation for the problem. Sometimes we cannot find alternative implementations, e.g. when a minimal quality level is given. In this case we can decide to either refuse or accept the implementation based on the extent of possible loss of money.

1. Introduction

In the course of planning and implementation of projects it occurs frequently that after the preliminary calculation of the optimal resources allocation – in order to obtain minimal total cost – a project can not be realized at a price which is expected by the inviter of the tender.

In the following study we present a method by which, in case of a delimited maximal budget, we are able to determine the schedule, cost need and resource requirements of a project to be implemented with the highest possible quality.

2. Alternative solutions

When in case of optimal resource allocation employed for minimal total cost during the preliminary calculation the expenses are higher than the budget earmarked for the implementation of the project then we have three options: we either give up the implementation of the project or we realize the project with losses or we replace some activities by new ones in order to reduce the costs. In the course of our analysis we will not consider the first option any longer since in this case we lose this business and there is no sense in making any further optimal resource planning. The second option is sometimes undertaken when they estimate that in spite of the initial losses the deficit will return during the implementation of the subsequent projects. In this case the allocation of the resources which is optimal for the given target function involving a minimal total cost shall be determined. (For the method refer to the study: *Optimal resource allocation of active projects).*

It is the third option we are going to study in this paper, i.e. we replace some activities to be realized with other activities in order to reduce the expenses.

In case of this third option, first of all, we have to make a list of alternative implementations of individual activities.

We should take into account three criteria when setting our objective (the order of the criteria also gives a priority during selection):

1. *first of all, we have to perform the given activities with the best possible quality*
2. *next the introduction of the alternative solution shall involve the greatest reduction costs*
3. *in the third place, the lead time of the activity and the demand for resources shall change at the slightest possible extent (they preferably get reduced).*

In addition to these criteria, for each activity a minimal quality requirement can be specified which in any case shall be met under the terms of the contract.

It is practical to arrange the alternative solutions belonging to a given activity in accordance with these criteria. From the list of the possible alternative activities to be implemented we should eliminate the ones which do not meet the quality requirements. During selection we reschedule the activities by using the alternative solution which is most suitable for the purpose and try to find the optimal solution again. The search of the alternative solutions is a selection problem determined for several target functions. We can find the optimal solution by the Branch & Bound method or by means of dynamical programming.

* **Zsolt Tibor Kosztyán, Zoltán Perjés, Andrea Bencsik**
Department of Management, University of Veszprém, E-mail: kzst@vision.vein.hu

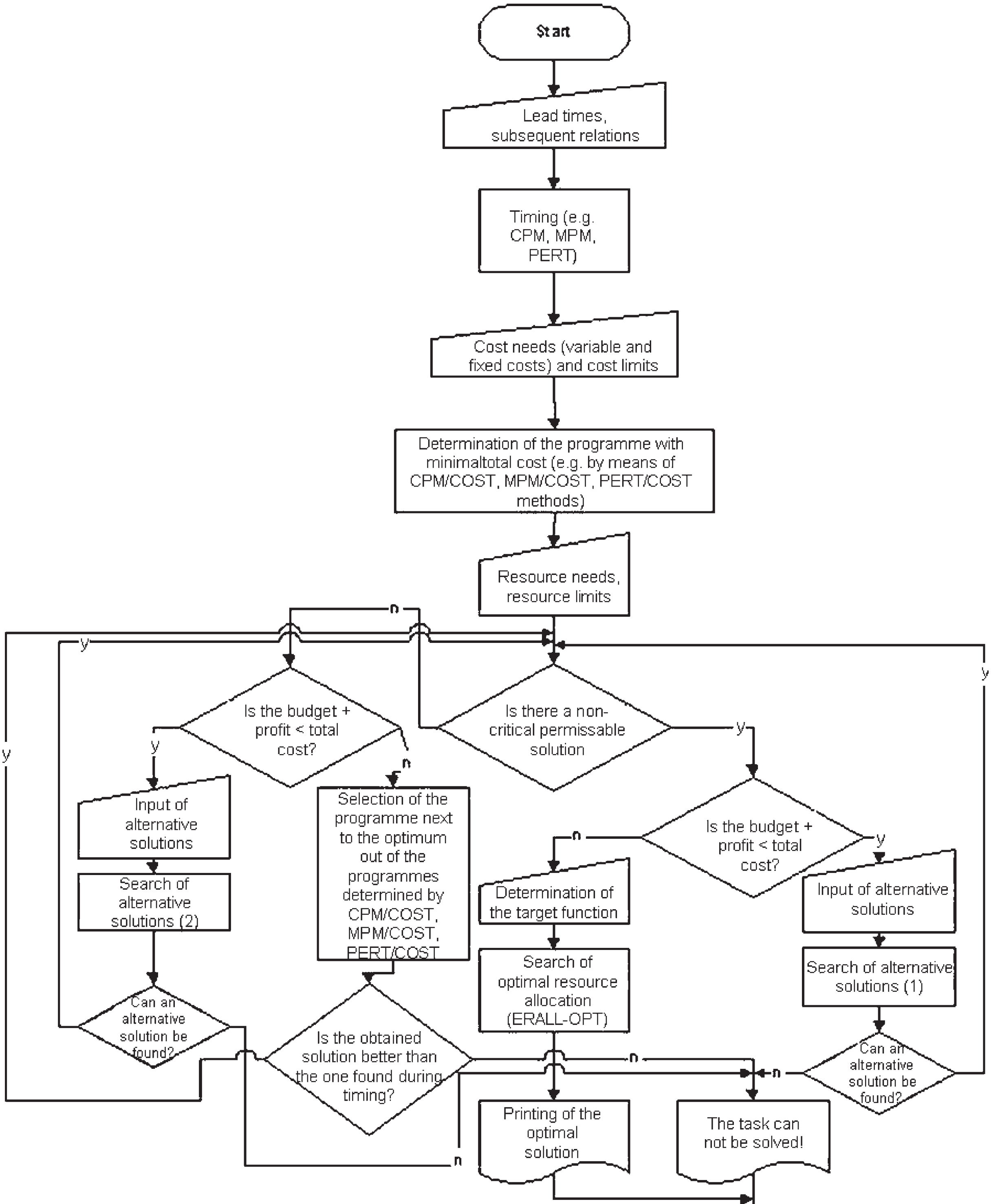


Fig. 1 Search of alternative solutions

2.1 The search of alternative solutions

The starting point of the method is the cost-optimal permissible allocation of resources or which is optimal for a given target function. We can find such a solution by determining a schedule with minimal total cost at the outset. If a non-critical permissible solution for this schedule is in existence, i.e. there is a permissible solution which does not amend the length of the critical path or does not overstep the limit of the resources then in accordance with the target function defined in Chapter 2 we should find alternative activities in a manner that during the employment of the alternative activities a non-critical solution must exist. If there is a non-critical solution then there is an optimal resources allocation as well and it can be found by a limited number of steps (see: *Optimal resource allocation of active projects).*

The method is illustrated on the following flowchart. In the first place we should make a timing to cope with the task. It requires the knowledge of the lead times and subsequent relations of the activities. We can carry out the timing by any method used in practice (CPM, MPM, Pert etc). Then we have to determine a solution with minimal total cost. This can be performed by any cost-optimising method used in practice (CPM/COST, MPM/COST, PERT/COST and so on). The utilization of the method requires the knowledge of cost needs, cost limits and the lead time – cost need functions. We estimate the lead time – cost need functions or, in case of production of small series, we determine them by means of statistical methods with a given probability. (For the details of the determination refer to the study: *Handling of uncertainty in the production management of small series).* Then we should examine whether there is a non-critical permissible solution. If there is such a solution then, according to the previous paragraph, the resource allocation optimal for the given target function can be determined by a limited number of steps. On the other hand, if there is not such a solution then we have to find an alternative solution here as well, but the 2 and 3 priority points of the target functions shown in the previous chapter are to be switched.

The objective remains that we have to carry out the given activities with the highest possible quality, but in order to find a non-critical solution we have to realize an activity of that sort in an alternative manner, which has a smaller resource need and/or shorter lead time.

If there is no other alternative solution then out of the solution list drawn up by the CPM/COST, MPM/COST and PERT/COST methods we have to pick out the solution which is next to the programme with minimal total cost need and to carry out the examination once again.

Let us take a look at the following example:

Let us examine a project with the following activities which are shown in the table below.

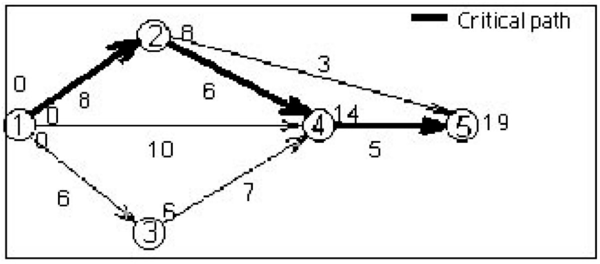


Fig. 2 CPM network diagram related to normal lead times

Let us suppose that the fix cost is 20000 thousand HUF for 19 weeks. The reduction of the lead time results in a savings of 1000 thousand HUF per week. In this case the programme with minimal total cost is the programme with minimal lead time, the total cost of the programme is 43850 thousand HUF (total variable costs) + 20000 thousand HUF (total fixed costs) – 4 × 1000 thousand HUF savings since we were able to complete the project earlier. Consequently, the total cost = 43850 + 20000 − 4000 = = **59850** *thousand HUF.*

The following chart demonstrates the possible solutions.

Table 2.1-1

Activity	Normal time (tnij)	Rush time (trij)	Normal cost (C nij) (thousand HUF)	Increase in costs/unit (ΔCij) (thousand HUF)	Normal resources demand (Rnij) (person)	Increase in resource demand/unit (ΔRij) (person)
(1.2)	8	6	4500	100	3	0.5
(1.3)	6	5	12000	200	2	1
(1.4)	10	7	10500	150	2	1
(2.4)	6	4	7500	300	4	1
(2.5)	3	3	3500	–	2	–
(3.4)	7	6	1500	500	3	1
(4.5)	5	4	2500	650	1	1
Summary			42000	–	–	–

Table 2.1-2

Number of steps	Reduced activities	Reduction	Total variable costs (thousand HUF)	Increase in costs/unit (thousand HUF)	Total increase in variable costs	Total time of the project (week)
0	–	–	42000	–	–	19
1	(1.2)	1	42100	100	100	18
2	(1.2)+(1.3)	1	42400	100+200=300	300	17
3	(4.5)	1	43050	650	650	16
4	(2.4)+(3.4)	1	43850	500+300=800	800	15

Let us examine the case where the resources limit is 10 persons. Our available cost limit is 61000 thousand HUF. In the second case the resources limit is 11 persons. Our available cost limit is 59000 thousand HUF.

The following table shows the activities which have alternative solutions.

Table 2.1-3

	(1,2)			(3,4)			(4,5)	
q	*vc*/Δ*vc*	*r*/Δ*r*	*Q*	*vc*/Δ*vc*	*r*/Δ*r*	*q*	*vc*/Δ*vc*	*r*/Δ*r*
100	4500/100	3/0.5	100	1500/500	3/1	100	2500/650	1/1
90	4000/90	2/0.5	90	1000/90	3/1	96	2500/600	1/1.5

where *q* is the quality factor. When employing this method we do not have to use figures for the quality factors by all means. It is enough to arrange the alternative solutions in order of quality. We can, however, determine the target function easier when, according to some principle, we assign a quality factor to the alternative solutions.

If the quality of the implementation of some activities needs special attention then it is practical to designate our target function used for selection in a way that we weigh the quality factors of these important functions.

Let us examine the following case where the resource limit is 10 persons. Our available cost limit is 61000 thousand HUF. The resource-loading diagram for 15 weeks is as follows:

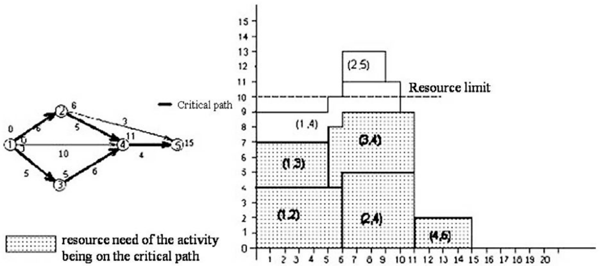


Fig. 3 Loading diagram related to the shortest lead time of the project (there is not any non-critical solution here)

On the loading diagram we can see that there is no non-critical solution. In order not to overstep the resource limit let us choose the 3rd production programme from the solutions obtained by CPM/COST.

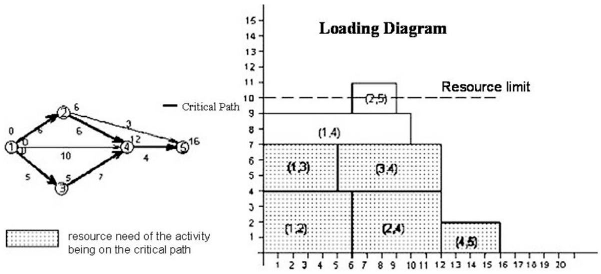


Fig. 4 Loading diagram related to a longer lead time

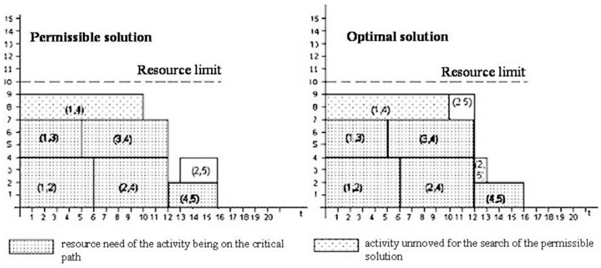


Fig. 5 Non-critical permissible and optimal resource allocation

During the determination of the optimal resource allocation the objective was to achieve a start as early as possible. At that time we did not have to reach a compromise concerning quality. The total cost: 43050 thousand HUF (total variable costs) + 20000 thousand HUF (total fixed costs) – 3000 thousand HUF (savings in fixed costs due to the shorter lead time) = 60050 thousand HUF < 61000 thousand HUF (budget).

Unfortunately, in the second case we have to carry out the activity(ies) at a lower quality level in order not to overstep the limit of the planned costs.

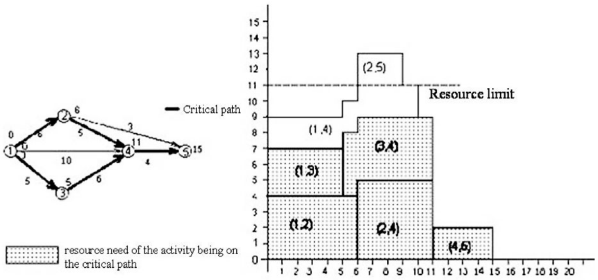


Fig. 6 Loading diagram related to the shortest lead time of the project

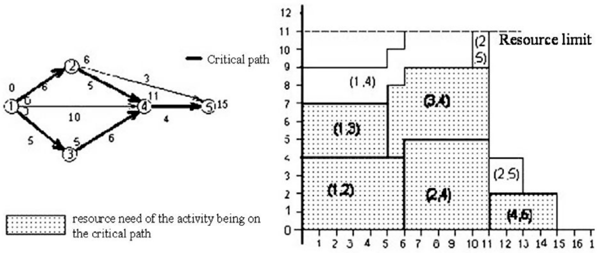


Fig. 2.1-6 Optimal resource allocation

The total cost is 59850 thousand HUF as calculated earlier which exceeds the 59000 thousand HUF budget. If the (3,4) activity is replaced by an alternative one then the resource need and the lead time remain unchanged but the cost is reduced by 900 thousand HUF. In this case the total cost will be 58850 thousand HUF which is less than the available budget.

3. Summary

The tougher the competition for the implementation of a project, the greater the chance that we have to implement the project at a low price which does not cover our expenses. Companies are often not fully aware of the costs of the implementation of the given project and frequently undertake the realization of the project and find out only later that their expenses will be much higher than their preliminary estimation.

If the contractual incomes do not cover the expenses then we either give up this business or undertake it even if we know that we can realize the project with a loss or try to replace the activities by alternative solutions. In such a case we should take into account that the activities are to be implemented at the highest possible level.

References

[1] ANDRÁSFAI, B.: *Graph theory*, Polygon, Szeged, 1997, pp 1–131
[2] KAUFMANN, A.: *Methods and Models of the Operational Research*, Technical Publishing House, Budapest, 1968, pp 11–123
[3] BENCSIK, A.: *Methodology of Organizational Techniques*, University Press, Veszprém, 1998
[4] CHRISTOPHIDES, N.: *Graph theory*, An Algorithmic Approach, Academic Press, New York, 1975
[5] DERIGS, U.: *Programming in Network and Graphs.* Lecture Notes in economics and Mathematical Systems. Vol. 300. Springer – Verlag, New York, 1988
[6] KOSZTYÁN, Z., BENCSIK, A., HOGYOR, A.: *Optimal Resource Allocation of Active Projects,* Verlag Dashöfer, 2002
[7] KOSZTYÁN, ZS., BENCSIK, A.: *Optimal Resource Allocation of the Projects with Uncertain Lead Time*, Verlag Dashöfer, 2003
[8] KOSZTYÁN, ZS., BENCSIK, A.: *Handling of Uncertainty in case of Different Logistical Tasks,* Verlag Dashöfer, 2002
[9] PAPP, O.: *Employment of Network Programming Methods in Practice*, Economic and Legal Press, 1969
[10] AHUJA, R. K., MAGNATI, TH. L., ORLIN, J. B.: Network Flows, PRENTICE HALL, *Upper Saddle River,* New Jersey 07458 1998, pp 24–46, 108–116, 135–154, 177–196, 469–173, 520–528
[11] RÓNYAI, L., IVANYOS, G., SZABÓ, R.: *Algorithms*, Typotex, Budapest, 1998, pp 110–183
[12] SZENDREI, Á.: *Discrete Mathematics*, Polygon, Szeged, 1996, pp 63–68, 295–324
[13] COHEN, TH. H., LEISERSON, CH. E., RIVEST, R. L.: *Algorithms*, (Technical Publishing House), Budapest, 1997, pp 392–543
[14] TÓTH, I.: *Mathematics for Business Administrators – Operational Research I*, Textbook Press, Budapest, 1987, pp 167–189.

Peter Czimmermann *

ON A CERTAIN TRANSPORT SCHEDULING PROBLEM FOR HETEROGENEOUS BUS FLEET

In this paper we consider a certain transport scheduling problem for heterogeneous bus fleet. We suppose that some restrictions are given for sets of vehicles and trips. We study some special cases of this problem that can be solved in polynomial time.

1. Introduction

We deal with the problem of finding an admissible transport schedule for heterogeneous bus fleet. We suppose that the place of departure and arrival is the same for every trip and the set of vehicles that can perform a given trip is restricted. This problem was studied in [2] and its special case (for two types of vehicles) in [3]. Its complexity is still an open question. It is known, that the generalisation of this problem (where the places of departure and arrival can be different) is NP-hard.

Let $V = \{v_1, ..., v_m\}$ be the set of vehicles and $S = \{s_1, ..., s_n\}$ be the set of trips. Every trip $s_i = (t_i^D, t_i^A, L_i)$, $0 \leq t_i^D < t_i^A$ is determined by the time of departure t_i^D, time of arrival t_i^A and by the list of vehicles $L_i \subseteq V$ that can perform the trip s_i. We suppose that

1. the place of departure and arrival is the same for every trip,
2. $t_i^X \neq t_j^Y$ for all $i, j \in \{1, ..., n\}$, $i \neq j$ and $X, Y \in \{D, A\}$
3. $\forall s_i, s_j \in S$: $(i < j \Leftrightarrow \mathrm{t}_\mathrm{i}^\mathrm{D} < t_j^D)$,
4. the trip s_i precedes the trip s_j ($s_i \prec s_j$) if $i < j$ and $t_i^A < t_j^D$.

From the condition (4), it is easy to show that $(S, \prec)$ is a partially ordered set with the relation "$\prec$". More about partially ordered sets (posets) can be found in [8].

We say that the trips $\forall s_i, s_j \in S$ are of the same type if $L_i = L_j$. An admissible running board of the vehicle v_i is the chain $T_i = S_{i_1} \prec S_{i_2} \prec ... \prec S_{i_k}$ where vi $\in L_{i_1} \cap L_{i_2} \cap ... \cap L_{i_k}$.

A schedule $R = (T_1, ..., T_m)$ is admissible if T_i is admissible running board of the vehicle v_i (for $i = 1, ..., m$), $T_1 \cup T_2 \cup ... \cup T_m = S$ and $T_i \cap T_j = \emptyset$ $(i,j \in \{\forall 1, 2, ..., m\}$. A set of all admissible schedules will be denoted by $\mathfrak{R}$. We will deal with the decision problem, whether $\mathfrak{R} = \emptyset$ or $\mathfrak{R} \neq \emptyset$.

2. Mathematical models

In this section we present two models of the mentioned decision problem. The first is *the model based on a bivalent programming formulation*.

By ordering increasingly the times of departures and arrivals, we obtain the vector

$$\vec{t} = (t_1^D, ..., t_1^A, ..., t_i^D, ..., t_i^A) = (t_1, t_2, ... t_{2n})$$

where $t_k < t_l \Leftrightarrow k < l$. Let $K = \{1, 2, ..., 2n\}$ be the set of the indices of the times in the vector $\vec{t}$. Let $I = \{1, 2, ..., m\}$, be the index set of the vehicles and J = {1, 2, . . . , n} be the indices of the trips.

Let x_{ij} be a decision variable with the value $x_{ij} = 1$ if s_j is performed by v_i and $x_{ij} = 0$ otherwise. We need to decide if there exists a matrix $X = (x_{ij})m \times n$ for which the following constraints hold:

$$\sum_{i \in I} a_{ij} x_{ij} = 1 \quad j \in J, \tag{1}$$

$$\sum_{j \in J} a_{ij} b_{kj} x_{ij} \leq 1 \quad i \in I, k \in K \tag{2}$$

$$x_{ij} \in \{0, 1\} \quad i \in I, j \in J \tag{3}$$

where

$$a_{ij} = \begin{cases} 1 & if\, v_i \in L_j \\ 0 & if\, v_i \notin L_j \end{cases} \tag{4}$$

$$b_{kj} = \begin{cases} 1 & if\, t_j^D \leq t_k \leq t_j^A \\ 0 & if\ otherwise \end{cases} \tag{5}$$

* **Peter Czimmermann**
Faculty of Management and Informatics, University of Žilina, Slovak Republic, E-mail: petocimo@frcatel.fri.utc.sk

Conditions (1) ensure that every trip s_j is assigned to exactly one vehicle (for which $v_i \in L_j$).

Conditions (2) ensure that the vehicle can perform at most one trip in every moment. The existence of some polynomial method for finding the matrix X is still an open question.

The model based on graph theory.

Definition 1. (see [6]) Let $G = (V, E)$ be a graph and $C \neq \emptyset$ be a set of colors.

A (vertex) coloring of G is a function $c : V \rightarrow C$ such that subgraphs induced by each color class have no edges.

Definition 2. (see [5]) Let the sets $Lv \subseteq C$ be assigned to each vertex $v \in V$. We call the vertex coloring $c : V \rightarrow C$ of G with property $\forall v \in V \;\; c(v) \in L_v$ a coloring from the lists (or list coloring).

Definition 3. (see [1]) A conditional coloring of G with respect to a graphical property $\mathfrak{P}$ is an assignment of colors to its vertices so that subgraphs induced by each color class satisfy the property $\mathfrak{P}$.

Since $(S, \prec)$ is the poset, we can define a transitive digraph $G_S = (V_S, E_S)$ where $V_S = S$ and $\forall s_i, s_j \in S \; s_j \; (s_i, s_j) \in E_S \Leftrightarrow s_i \prec s_j$. The problem of finding the admissible schedule can be solved as the problem of finding some conditional coloring of G_S from the lists which satisfies the property P_{op}: vertices of the same color form an oriented path (or a transitive tournament) in G_S. Every color corresponds to one vehicle. For every vertex (trip), the list of admissible colors (the set of vehicles that can perform this trip) is specified. Vertices of the same color form a chain in $(S, \prec)$ which corresponds to some admissible running board.

The existence of polynomial algorithm for conditional coloring from the lists with respect to the property that vertices of the same color form an oriented path in G_S is an unsolved problem.

3. Cases that can be solved in polynomial time

In this section we study some special cases that can be solved in polynomial time. We prefer a graph theoretical approach.

1. Let $|L_1| = \cdots = |L_n| = 1$. This case is trivial, every vertex $v_i \in G_S$ is colored by unique color from the list L_i. There is an admissible schedule if and only if the vertices of the same color form an oriented path in G_S. We are able to verify this in polynomial time.
2. Let $|L_1| = \cdots = |L_n| = 2$. Since G_S is transitive, vertices of every oriented path form also a transitive tournament in G_S. Let $G = (V, E)$ be a graph for which $V = V_S$ and the edge $\{s_i, s_j\} \in E$ if and only if the trips s_i, s_j are incomparable by relation $\prec$. It is easy to show that there is a coloring of G from the lists $L_1 = L_2 = \cdots = L_n$ if and only if there exists a conditional coloring of G_S from the lists with respect to the property P_{op}. It is known (see also [7]) that the list coloring of the graph is solvable in polynomial time if every list contains exactly two elements (colors).
3. If $L_1 = L_2 = \cdots = L_n \subseteq V$ then we are able to transform this case to the problem of covering a poset by disjoint union of chains – the solution of this problem is based on the famous Dilwort's Chain Decomposition Theorem [4] and we are able to find it in polynomial time.
4. Let $L_i \cap L_j = \emptyset$ or $L_n \subseteq V$ for all $i, j \, \{1, ..., n\}$. We can decompose $(S, \prec)$ into subposets $(S_1, \prec)$, ..., $(S_k, \prec)$ (see [8]) where $S_i \cap S_j = \emptyset$, $S_1 \cup \cdots \cup S_k = S$, and two trips s_i, s_j belong to the same subposet if $L_i = L_j$. We can solve this case for every poset $(S_i, \prec)$ separately as in (3).

Acknowledgement: The research was supported by the Slovak Scientific Agency, Grant No.1/0490/03.

References

[1] CHARTRAND, G., GELLER, D. P., HEDETNIEMI, S. T.: *A Generalization of the Chromatic Number,* Proc. Camb. Philos. Soc. 64:265–271, 1968
[2] CZIMMERMANN, P.: *A Note on Bus Schedules with Several Types of Vehicles, Journal of Information,* Control and Management Systems, vol. 2, No. 2/2003
[3] CZIMMERMANN, P., PEŠKO, S.: *The Bus Schedules with Two Types of Vehicles,* 3rd International Conference APLIMAT, 2004
[4] DILWORTH, R. P.: *A Decomposition Theorem for Partially Ordered Sets,* Ann. of Math. 51, 161-166, 1950
[5] DIESTEL, M.: *Graph Theory,* Springer-Verlag, New York, 1997
[6] HARARY, F.: *Graph Theory,* Reading(Mass.), Addison-Wesley, 1969
[7] PEŠKO, S.: *Optimization of NP-hard transport schedules (in Slovak),* Habilitation Thesis, University of Zilina, 2002
[8] Schröder, B. S. W.: *Ordered Sets,* Birkhäuser, Boston, 2003.

KOMUNIKÁCIE COMMUNICATIONS

Jaroslav Janáček – Lýdia Gábrišová *

LAGRANGEAN RELAXATION BASED APPROXIMATE APPROACH TO THE CAPACITATED LOCATION PROBLEM

When a distribution system is to be designed, limits on terminal capability often must be taken into account. These capacity constraints in this and other facility location problems constitute severe obstacles in exact solving processes. Within this paper, we focused on study of approximate methods based on Lagrangean relaxation of the capacity constraints, which has several advantageous properties. The first of them is that the relaxed problem, known as the uncapacitated location problem, can be solved exactly even for real sized instances [6], [4]. The second useful property of the Lagrangean relaxation is that the objective function value of the optimal solution of the relaxed problem provides lower bound of the optimal solution of the original problem. We present two methods for obtaining suitable values of Lagrangean multipliers. The classical one is based on a sub-gradient method applied on capacity constraints after their special adjustment. The second method is designed as an adaptive method with random experiments for determination of candidates for move from the current solution to the next one. These two methods were tested, compared and the associated results are reported in the concluding part of this paper.

1. Introduction

Cost optimal design of the most of distribution and servicing systems consists in decisions on number and locations of facilities, from which customer's demands are satisfied. When such public or private servicing system is designed, some limits are put on abilities of the particular facilities. These constraints follow from the limited ability of the facilities to satisfy demands of customers. In spite of the fact that capacity of a facility is a very vague value, the capacities of the located facilities bring serious complications concerning a solving technique, which may be used to solve a real-sized facility location problem. In contrast to an uncapacitated facility location problem, which can be solved exactly for real-sized instances containing hundreds of possible locations and thousands of customers, the capacitated location problem resists to all attempts to solve it exactly in reasonable time. To avoid this complication and to obtain a good solution of the capacitated facility location problem in sensible time, we employ approximate methods based on Lagrangean relaxation of the capacity constraints, which has several advantageous properties. One of these properties is that the relaxed problem known as the uncapacitated location problem can be solved exactly by an algorithm based on Erlenkotter's approach [2]. Real sized instances of the uncapacitated problem were broadly tested [6], [4] and it was proved that it is possible to obtain their optimal solution in a sensible time. The next useful property of the Lagrangean relaxation, which can be exploited with the previous one, is that the objective function value of optimal solution of the relaxed problem provides lower bound of the optimal solution of the original problem. In this way, the obtained approximate solution can be compared with the estimated value of the optimal one. The third property is that any optimal solution of the Lagrangean relaxation preserves so called compactness of particular customer clusters, which arise from a customer assignment to the individual facility locations. This property is not directly connected with the facility location, but it has great impact on applications following up with a distribution system design. This property enables good service of the customers by vehicle routes starting and terminating at the facility location.

We present two methods for obtaining suitable values of Lagrangean multipliers. The classical one is based on a sub-gradient method applied on capacity constraints after their special adjustment. The second method is designed as an adaptive method with random experiments for determination of candidates for move from a current solution to the next one. Both approaches make use of the strengthening constraint, which considerably improves quality of the obtained lower bound. Both methods were tested, compared and the associated results are reported in the section "Numerical experiments" of this paper and furthermore some explanation of their various behaviors is suggested.

2. Mathematical model of the solved problem

Mathematical programming approach to the capacitated facility location problem originates from the assumption that goods distribution is performed from a primary source via warehouses or other goods transshipment places to the particular customers. The number and positions of the warehouses, which will be called facilities in this paper, should be determined so that the total yearly cost for both located facilities and customer demand satisfaction are minimal.

* **Jaroslav Janáček, Lýdia Gábrišová**
Department of Transportation Networks, Faculty of Management and Informatics, University of Žilina, Slovak Republic, 010 26, E-mail: jardo@frdsa.fri.utc.sk
Department of Mathematical Methods, Faculty of Management and Informatics, University of Žilina, Slovak Republic, 010 26, e-mail: gabriska@frcatel.fri.utc.sk

The problem is described by finite set I of possible facility locations and finite set J of the customers, whose demands should be satisfied. Associated costs and charges are connected to particular elements of these sets and to their pairs.

A fixed charge for location of a facility at possible location $i \in I$ is denoted by f_i. This fixed charge includes all costs connected with keeping this facility at the location for one year. This charge does not include items, size of which depends on amount of demands, which are satisfied via this location.

The cost of *j-th* customer yearly demand satisfaction via facility located at place i is denoted by coefficient c_{ij}. This coefficient includes all transportation costs for goods transport from the primary source to the facility location, from this location to the customer and manipulating cost in the facility.

It is presumed that a facility may be placed only at some place of the above-introduced finite set I of possible locations. To model the decision on placing or not placing a facility at location i, variable $y_i \in \{0, 1\}$ is introduced for each location i from set I.

Let us assume that each customer $j \in J$ should be supplied by a yearly amount of goods b_j. To be able to express that a customer is assigned to a given facility location and that he is supplied via this location, another set of zero-one variables is established. Variable $z_{ij} \in \{0, 1\}$ models a decision on assigning or not assigning customer j to facility location i. Each solution of the following integer programming problem will be described by zero – one variables y_i and z_{ij}. This zero – one range of the decision variables is assumed in each model within this paper and that is why this associated obligatory constraint is not resumed at the following models.

Let us consider that a_i denotes the capacity of a facility located at i then the complete model of the cost minimal capacitated facility location problem can be formulated as follows:

$$\text{Minimize} \quad \sum_{i \in I} f_i y_i + \sum_{i \in I} \sum_{j \in J} c_{ij} z_{ij} \qquad (1)$$

$$\text{Subject to} \quad \sum_{i \in I} z_{ij} = 1 \qquad \text{for } j \in J \qquad (2)$$

$$y_i - z_{ij} \geq 0 \qquad \text{for } i \in I, j \in J \qquad (3)$$

$$\sum_{j \in J} b_j z_{ij} \leq a_i \qquad \text{for } i \in I \qquad (4)$$

In this integer programming model, constraints (2) ensure that each customer demand must be satisfied from exactly one facility location and constraints (3) force out the placement of a facility at location i whenever a customer is assigned to this facility location. Constraints (4) ensure that the total demand satisfied via facility location i doesn't exceed given capacity a_i. Having omitted or relaxed constraints (4), the problem (1)–(3) is known as the uncapacitated facility location problem and it can be effectively solved making use of implementation of the branch and bound method with Erlenkotter's lower bounding [1]. Computation behaviour of the technique was broadly examined in [4], [6], [7], [8] and it was shown that this approach is able to manage large size problems of practice. The capacitated problem (1)–(4) loses integrality property of variables z_{ij} due to the capacity constraints and so it constitutes a very hard problem for exact solving when a real-sized instance of the problem is considered.

Model (1)–(4) can be replaced by following formulation, which is equivalent from the point of set of feasible (zero – one) solutions:

$$\text{Minimize} \quad \sum_{i \in I} f_i y_i + \sum_{i \in I} \sum_{j \in J} c_{ij} z_{ij} \qquad (5)$$

$$\text{Subject to} \quad \sum_{i \in I} z_{ij} = 1 \qquad \text{for } j \in J \qquad (6)$$

$$z_{ij} \leq y_i \qquad \text{for } i \in I, j \in J \qquad (7)$$

$$\sum_{j \in J} b_j z_{ij} \leq a_i y_i \qquad \text{for } i \in I \qquad (8)$$

In this formulation, constraints (7) become surplus, but we let them in the model for a better numerical convergence of the branch and bound method solving the relaxed problem. The performed reformulation of constraints (8) has proved to be a very efficient form for subsequent Lagrangean relaxation.

The above – mentioned feasible solution equality of the models (1)–(4) and (5)–(8) follows from the following proofs: Each zero – one solution described by variables y_i and z_{ij}, which satisfies (3) and (4), must satisfy (8). For $y_i = 1$ constraint (8) takes the same form as (4) and for $y_i = 0$ is $z_{ij} = 0$ for each $j \in J$ and it follows that left – hand – side of (8) is equal to zero as well as the right – hand – side. Let us assume that (7) and (8) hold. If $y_i = 1$, then (8) takes the same form as (4). If $y_i = 0$, then $z_{ij} = 0$ for each $j \in J$ and (4) holds also.

3. Lagrangean relaxation and strengthening constraint

The studied Lagrangean relaxation of problem (5)–(8) consists in removing constraints (8) from the model and in embedding a measure of their dissatisfaction into the objective function. The measure of the capacity constraint dissatisfaction is enumerated as the difference between the capacity and the demand, which is modeled by the expression on the left-hand-side of the constraint. This overload of a facility is weighed by a nonnegative weight called Lagrangean multiplier. This multiplier represents a penalty paid for breaking the constraints.

Having denoted u_i the Lagrangean multiplier of the capacity constraint $i \in I$, the Lagrangean relaxation takes the form:

$$\text{Minimize} \quad \sum_{i \in I} f_i y_i + \sum_{i \in I} \sum_{j \in J} c_{ij} z_{ij} + \sum_{i \in I} u_i \left(\sum_{j \in J} b_j z_{ij} - a_i y_i\right) \qquad (9)$$

$$\text{Subject to} \quad \sum_{i \in I} z_{ij} = 1 \qquad \text{for } j \in J \qquad (10)$$

$$z_{ij} \leq y_i \qquad \text{for } i \in I, j \in J \qquad (11)$$

For fixed nonnegative values of all Lagrangean multipliers, the objective function of the Lagrangean relaxation has the form:

$$\sum_{i \in I} (f_i - a_i u_i)\, y_i + \sum_{i \in I} \sum_{j \in J} (c_{ij} + b_j u_i)\, z_{ij} \qquad (12)$$

which is the objective function of an uncapacitated location problem. Due zero – one variables y_i and z_{ij}, it is obvious that an optimal solution exists for each setting of the Lagrangean multiplier values.

The preliminary experiments with this model showed that the set of feasible solutions enlarges too much in this way of relaxation. This consequence can be minimized to some extent by addition of a simpler constraint, which is redundant with respect to the relaxed ones, but which diminishes the set of feasible solutions of the relaxed problem. The strengthening constraint is derived from the capacities and is laid upon the number of located facilities. It is obvious that at least such a number of facilities must be located to cover all customer demands. If a denotes $max\ \{a_i : i \in I\}$ and B denotes sum of all customer demands, then minimal number r of the necessary located facilities is $r = B / a$ and the strengthening constraint can be constructed in this way:

$$\sum_{i \in I} y_i \geq r \qquad (13)$$

To keep the form of the uncapacitated location problem, this constraint can be relaxed as well as the other capacity constraints using nonnegative Lagrangean multiplier v. After this reformulation we get

$$\textit{Minimize} \ \sum_{i \in I} (f_i - a_i u_i - v) y_i + \sum_{i \in I} \sum_{j \in J} (c_{ij} + b_j u_i) z_{ij} + rv \qquad (14)$$

$$\textit{Subject to} \sum_{i \in I} z_{ij} = 1 \qquad \text{for } j \in J \qquad (15)$$

$$z_{ij} \leq y_i \qquad \text{for } i \in I, j \in J \qquad (16)$$

To determine value v, for which the optimal solution of (14)–(16) satisfies complementary constraint (17), a simple dichotomy algorithm can be used.

$$\Big(\sum_{i \in I} y_i - r\Big) v = 0 \qquad (17)$$

It is true that the dichotomy algorithm does not ensure satisfaction of constraint (17) in general, but we did not meet any real instance of the solved problem, in which the algorithm failed.

4. Sub-gradient method

The sub-gradient method used in our approach was designed with the goal to maximize a lower bound of the objective function value of an optimal solution of the original capacitated facility location problem [1]. The method improves the lower bound by particular steps, in which it adjusts the values of the Lagrangean multipliers u_i. During this iterative process, the overloads of facilities with dissatisfied capacity constraints are penalized by higher values of the associated Lagrangean multipliers and hence forced to minimize size of dissatisfaction. This process may lead to a complete capacity constraint satisfaction or, which is a more frequent case, to such solution, which is only slightly infeasible.

The sub-gradient method maximizes function $F(u)$, the value of which is equal to the minimal value of expression (14) subject to (15), (16,) (17) and integer nonnegative values of variables y_i and z_{ij} for the given values of $u = \langle u_1, u_2, ... \rangle$.

Value $F(u)$ for the given u is obtained using the above mentioned dichotomy algorithm applied on the branch and bound method developed for the uncapacitated location problem.

The algorithm of the sub-gradient method starts from an initial point from nonnegative part of $|I|$ dimensional space of the Lagrangean multipliers. This point is usually set to the origin of the coordinates system i.e. its coordinates are set to zeros.

A move from current point u^k to next point u^{k+1} is made along a direction, in which the function increases in the neighborhood of point u^k. The steepest increase of $F(u)$ in the neighborhood of point uk can be found along the gradient of this function at point u^k. The coordinates of the gradient can be enumerated as values of partial derivatives of (17) or (14) or (9) by the individual multipliers u_i with consequent substitution of values u_i^k. In this way, the i-th component takes the value of the associated located facility overload.

To follow the direction of the gradient, the move should be performed in accordance with equality

$$u^{k+1} = u^k + \alpha\, grad / \| grad \|,$$

$$\text{where } grad_i = \sum_{j \in J} b_j z_{ij} - a_i y_i \qquad \text{for } i \in I.$$

Unfortunately, keeping this formula, the requirement of nonnegative multipliers should be broken. That is why the following formula is used in the algorithm

$$u^{k+1} = max\ \{0, u^k + \alpha \cdot grad / \| grad \| \}.$$

That is why the move is not performed exactly along the gradient, but along the direction of sub-gradient. The length of the step is given by parameter α, the value of which is chosen from interval $\langle \alpha_{min}, \alpha_{max} \rangle$. After each step, values of $F(u^k)$ and $F(u^{k+1})$ are compared, and if $F(u^k) \geq F(u^{k+1})$ holds, then the return to u^k is done and the move is repeated with a lower value of α. The process terminates if parameter α reaches value of α_{min} or if the resulting improvement of the last step is less than the given value ϵ.

5. Adaptive method SUPRA

The method SUPRA was designed to overcome the weak side of the sub-gradient method, which emerges when the minimization sequence of $\{u^k\}$ approaches a point, in which partial derivation of $F(u)$ is not defined. In such situation the sequence $\{u^k\}$ converges prematurely in the direction given by the sub-gradient even if it can continue to some better solutions in other direction.

The method also improves the lower bound by particular steps in which it adjusts the values of the Lagrangean multipliers u_i. Similarly to the sub-gradient method, the overloads of facilities with dissatisfied capacity constraints are penalized by higher values of the associated Lagrangean multipliers and hence forced to minimize the size of dissatisfaction.

Maximization of function $F(u)$ is performed where the value of this function is also equal to the minimal value of expression (14) subject to (17) and integer nonnegative values of variables y_i and z_{ij} for the given values of $u = <u_1, u_2, ... >$. As in the previous approach, the dichotomy algorithm applied on the branch and bound method is employed here to obtain value $F(u)$ for the given u.

The algorithm of method SUPRA also starts from an initial point, which can be set to the origin of the coordinates system, i.e. its coordinates are set to zeros but a move from the current point u^k to the next point u^{k+1} is made in a completely different way, which does not utilize the property of the gradient computed from the facility overloads.

The move from the current point u^k to point u^{k+1} is the best found move obtained as a result of two phases, which explore neighborhood of the current point uk.

The first phase consists of a random generation of changes r^j for $j = 1, ..., s$, from which s trial points u^{kj} are derived in accordance with the formula $u^{kj} = P(u^k + r^j)$, where P denoted a projection of a general point to the nonnegative part of $|I|$ dimensional space of the Lagrangean multipliers. The values $F(u^{kj})$ are compared and the best found solution is updated. Furthermore, the statistical gradient r is enumerated in accordance with the equation:

$$r = \sum_{j=1}^{s}(f(u^{kj}) - F(u^k))r^j.$$

The random generation of the changes is based on an adaptive principle, in which a "memory" is modeled by vector w of memory coefficients, which are connected with the particular coordinates of the space of the Lagrangean multipliers. The generation of r^j is done using equation $r^j = w + x$, where x is obtained as a realization of $|I|$-th dimensional random variable with a uniform probability distribution of their components over interval $<-2A, 2A>$.

The vector of the memory coefficients is initialized to zero values at the beginning of the optimization process, and whenever new point u^{kj} is constituted and its objective function value $F(u^{kj})$ is found, the vector of memory coefficients is updated in accordance with the rule $w_i = Bw_i + C.(F(u^{kj}) - F(u^k))(u_i^{kj} - u_i^k)$, where B is the parameter of oblivion and C is the parameter of learning intensity. Both parameters were chosen from interval (0, 1).

The second phase utilizes the statistical gradient enumerated in the first phase and explores the neighborhood of the current point u^k in the associated direction. The statistical gradient is normalized to the unit length and the first experiment is undertaken with point $u = u^k + 4.A.r/\|r\|$. Then the step length is diminished and the new point is tested. This searching process is repeated until the number of failed of attempts to improve solution reaches the prescribed maximal value.

The best-found solution in both phases represents the resulting move to u^{k+1} subject to condition that the new point has a better function value than the current one.

The whole searching process of the method SUPRA is repeated until one of the following conditions is met. The first condition tests the equality of given number N and the total number of move designs and the second one tests the equality of given number Nb and the number of move designs performed from the last improvement of the best-found solution.

6. Numerical experiments

To test and compare both approaches, the associated algorithms were implemented using Delphi 7 programming environment. To perform the numerical experiments, Pentium 4, 3.0 GHz, 1 GB was used. The approach to the capacitated facility location problem was tested with data originating at the Slovak road network with 2907 dwelling places, where each of them represents one customer. Seventy-one centres of the former districts formed the set of the possible facility locations. The testing problems are referred as LAS*071xy2911.txt, where * denotes D, B, A, F, $x = 3, ..., 9$ and $y = 2, ..., 5$. Then the number of the test problems is 28. The set of the test problems is partitioned into four groups in accordance with different values of y, which corresponds to the ratio cost coefficients used in enumeration of c_{ij}.

In the performed experiments the following attributes of the resulting solutions are observed and evaluated in Tables 1 – 4. The first attribute of the solution quality is resulting lower bound (14) of the objective function value of an optimal solution of the original problem (LB), but it must be taken into consideration together with some measure of capacity constraint violation, because this way of handling with capacity constraint admits some violation. Two parameters were chosen to evaluate this possible capacity constraint violation and they are considered to be the second and third quality characteristics of a solution. The first parameter (Max Over.) is ratio of the maximum value of positive differences between the workloads of facilities and their capacities and the associated capacity. The second parameter (Sum Over.) is ratio of the sum of positive differences between the workloads of facilities and their capacities and the sum of all located facility capacities. The computational time in seconds is reported in column (Time).

7. Conclusions

As concerns lower bound of the original problem, the above-described approaches to the capacitated facility location problem

The group location problems for $y = 2$ Table 1

Problem	Method	LB	Max Over.	Sum Over.	Time [s]
LAS*071322911	SUPRA	30961178	0.54	0.54	55
	sub-gradient	23195365	0.89	1.31	5
LAS*071422911	SUPRA	28961178	0.54	0.54	48
	sub-gradient	26667913	0.37	0.75	7
LAS*071522911	SUPRA	27161178	0.54	0.54	51
	sub-gradient	19401236	0.89	1.31	4
LAS*071622911	SUPRA	25161178	0.54	0.54	49
	sub-gradient	23818899	0.84	1.50	7
LAS*071722911	SUPRA	24161178	0.54	0.54	48
	sub-gradient	18062008	0.61	1.96	5
LAS*071822911	SUPRA	23761178	0.54	0.54	47
	sub-gradient	17646692	0.61	1.96	5
LAS*071922911	SUPRA	22534075	1.97	3.29	52
	sub-gradient	15954497	0.89	1.31	4

The group location problems for $y = 3$ Table 2

Problem	Method	LB	Max Over.	Sum Over.	Time [s]
LAS*071332911	SUPRA	37621850	0.50	0.50	63
	sub-gradient	31505261	0.97	1.67	8
LAS*071432911	SUPRA	36359338	0.50	0.50	71
	sub-gradient	30243514	0.97	1.67	8
LAS*071532911	SUPRA	31309354	0.50	0.50	71
	sub-gradient	25196505	0.97	1.67	6
LAS*071632911	SUPRA	26259338	0.50	0.50	70
	sub-gradient	20149498	0.97	1.67	4
LAS*071732911	SUPRA	25959338	0.50	0.50	68
	sub-gradient	19833076	0.97	1.67	5
LAS*071832911	SUPRA	21196252	1.01	4.39	51
	sub-gradient	19518233	0.97	1.67	3
LAS*071932911	SUPRA	20527067	1.38	3.61	82
	sub-gradient	19202592	0.97	1.67	5

promise a good solution, especially if constraints (8) and the strengthening constraint (13) are employed.

Comparing the lower bounds provided by both approaches on the groups of tested instances has shown that there is no unambiguous winner.

When applying the approaches on the first two groups of instances, the SUPRA wins in each case. As concerns the third group, the SUPRA wins only in half of the instances and the SUPRA completely fails in the fourth group.

This effect can be caused by bias setting of the parameters A, B, C of the adaptive process of SUPRA. The excellent performance of the SUPRA considering facility overload minimization is worth noticing. In most of the solved problems, SUPRA achieves half of the maximal overload produced by the sub-gradient method.

The negligible computational time of the sub-gradient method in comparison with the computational time of SUPRA evokes an idea for the future research, where composed heuristic should be designed and studied. This heuristic could combine both above – mentioned approaches so that it could use the sub – gradient method in the first phase and then utilize the SUPRA to find a new improving direction.

Acknowledgement: This work has been supported by the grant VEGA 1/3775/06.

The group location problems for $y = 4$ Table 3

Problem	Method	LB	Max Over.	Sum Over.	Time [s]
LAS*071342911	SUPRA	43287380	0.48	0.48	67
	sub-gradient	44916936	1.21	2.01	9
LAS*071442911	SUPRA	40087380	0.48	0.48	58
	sub-gradient	39943420	0.68	1.79	8
LAS*071542911	SUPRA	33687380	0.48	0.48	61
	sub-gradient	34965704	0.73	1.27	6
LAS*071642911	SUPRA	30487380	0.48	0.48	54
	sub-gradient	31543841	0.73	1.53	7
LAS*071742911	SUPRA	27287380	0.48	0.48	57
	sub-gradient	22276893	1.61	2.48	2
LAS*071842911	SUPRA	27287380	0.48	0.48	58
	sub-gradient	26964051	0.79	1.85	6
LAS*071942911	SUPRA	26887380	0.48	0.48	64
	sub-gradient	24999437	0.65	2.35	5

The group location problems for $y = 5$ Table 4

Problem	Method	LB	Max Over.	Sum Over.	Time [s]
LAS*071352911	SUPRA	48402614	0.38	1.00	537
	sub-gradient	52338055	0.65	1.86	11
LAS*071452911	SUPRA	44580789	0.13	0.13	225
	sub-gradient	46237898	0.83	1.22	5
LAS*071552911	SUPRA	34593026	0.95	2.06	188
	sub-gradient	46634256	0.79	2.25	8
LAS*071652911	SUPRA	29159703	0.94	1.53	370
	sub-gradient	32047752	0.89	1.75	4
LAS*071752911	SUPRA	29263127	0.45	0.45	151
	sub-gradient	29674754	0.89	2.71	7
LAS*071852911	SUPRA	27318213	0.29	0.77	93
	sub-gradient	28906129	0.89	1.75	5
LAS*071952911	SUPRA	25274108	1.14	1.56	401
	sub-gradient	38436538	1.16	1.34	7

References

[1] BEASLEY, J. E. (1993): *Lagrangean relaxation.* In: Modern Heuristic Techniques for Combinatorial Problems (edited by Reeves C.R.), Blackwell Scientific Publications, London, pp. 243–303

[2] ERLENKOTTER , D. (1978): *A Dual-Based Procedure for Uncapacitated Facility Location.* Operations Research, Vol. 26, No 6, pp. 992–1009

[3] DREZNER, Z. (2004): *Facility Location Application and Theory.* Springer Verlag, Berlin, Heidlberg, New York, 457 p.

[4] JANÁČEK, J., BUZNA, Ľ. (2004): *A Comparison Continuous Approximation with Mathematical Programming Approach to Location Problems.* Central European Journal of Operations Research, Vol. 12, No 3, pp. 295–305

[5] JANÁČEK, J., GÁBRIŠOVÁ, L. (2005): *Fuzzy Set Operations for Capacity Constraint Relaxation.* Journal of Information, Control and Management Systems, ŽU Žilina, No 2, pp. 81–89

[6] JANÁČEK, J., KOVÁČIKOVÁ, J. (1997): *Exact Solution Techniques for Large Location Problems. I*n: Proceedings of the Math. Methods in Economics, Ostrava, pp. 80–84

[7] JANÁČKOVÁ, M., SZENDREYOVÁ, A. (2005): *An Impact of Transportation Network Topology on Time Consumption of an Exact Algorithm for Distribution System Design.* Zeszyty naukove, No 1691, November 2005, Wydawnictwo politechniky Šlaskiej, pp. 175–179

[8] KÖRKEL, M. (1989): *On the exact solution of large – scale simple plant location problem.* European Journal of Operational Research 39, pp. 157–173.

Petr Fiala *

COMBINATORIAL AUCTIONS FOR SELLING NETWORK CAPACITY

The network economy is a term for today's global relationship among economic subjects characterized by massive connectivity. Today's network systems provide the infrastructure and foundation for the functioning of societies and economies. They come in many forms and include physical networks such as transportation and logistical networks, communication networks, energy networks, as well as more abstract networks as economic, financial, social, and knowledge networks. The paper presents modeling of auctions in network economy. Auctions are important market mechanisms for the allocation of goods and services. Combinatorial auctions are those auctions in which bidders can place bids on combinations of items. Combinatorial auction is an appropriate instrument for selling network capacity. The winner determination problem in a combinatorial auction for selling network capacity is formulated. The model can be formulated as a multicommodity flow problem and efficient algorithms can be used.

Keywords: Network economy, combinatorial auction, winner determination problem, multicommodity network flows

1. Introduction

The paper presents modeling of combinatorial auctions in network economy. It is an integration of two characteristic elements of today's economic reality, the network economy and auction mechanisms. The network economy is a term for today's global relationship among economic subjects characterized by massive connectivity. Network industries play a crucial role in modern life. Today's network systems provide the infrastructure and foundation for the functioning of societies and economies. Auctions are important market mechanisms for the allocation of goods and services. Auctions are preferred often to other common processes because they are open, quite fair, easy to understand by participants, and lead to economically efficient outcomes. Their popularity is also caused by expansion of e-commerce. Design of auctions (see [7]) is a multidisciplinary effort made of contributions from economics, operations research, computer science, and other disciplines.

There is a possible classification of auctions by different aspects:
1. Traded items (indivisible, divisible, pure commodities, structured commodities).
2. Participants' roles in auctions (one-sided, multilateral auctions).
3. Objectives of auctions (optimization, allocation rules, pricing rules).
4. Complexity of bids (simply, related bids).
5. Organization of auctions (single-round, multi-round, sequential, parallel, price schemes).

For auctions for selling network capacity it is useful to use so called combinatorial auctions. Combinatorial auctions refer to auctions in which participants are allowed to bid on combinations of items. A classical problem of combinatorial auctions is the winner determination problem. The problem can be formulated as an integer programming problem and is well-known to be NP-hard. In the paper the winner determination problem for selling network capacities is formulated. The model is based on combinatorial auctions with a network structure of items.

2. Network economy

The network economy (see [4], [5], [9]) is a term for today's global relationship among economic subjects characterized by massive connectivity. The central act of the new era is to connect everything to everything in deep web networks at many levels of mutually interdependent relations, where resources and activities are shared, markets are enlarged and costs of risk are reduced. Connections are enabled by an explosive development of information and communication technologies. Network connections enable tighter relations between firms and stakeholders. New technologies provide a permanent feedback that enables activity modifications and quick responses and therefore fundamentally change business models. Network industries play a crucial role in modern life. Today's network systems provide the infrastructure and foundation for the functioning of societies and economies. They come in many forms and include physical networks such as: transportation and logistical networks, communication networks, energy networks, as well as more abstract networks comprising: economic and financial networks, environmental networks, social, and knowledge networks. Many important non-network industries share many essential economic features with network industries. These non-network industries are characterized by strong complementary relations.

* **Petr Fiala**
Department of Econometrics, Faculty of Informatics and Statistics, University of Economics, W. Churchill Sq. 4, 130 67 Prague 3,
E-mail: pfiala@vse.cz

The reality of today's networks includes features:
- large-scale nature and complexity,
- increasing congestion,
- complementarity,
- externalities,
- switching costs,
- alternative behaviors of users of the networks,
- interactions between the networks themselves.

Many of today's networks are characterized by both a large-scale nature and complexity of the network topology. Congestion is playing an increasing role in not only transportation networks but also in telecommunication networks. The crucial relationship in networks is the complementarity between the pieces of the network. Complementarity turns out to be a crucial factor in the markets for information goods. Networks exhibit positive externalities. The value of a unit increases with the expected number of units to be sold. Costs of switching to a different service or adopting a new technology are significant. The decisions made by the users of the networks, in turn, affect not only the users themselves but others, as well, in terms of profits and costs, timeliness of deliveries, the quality of the environment, etc. The behavior of the users of the networks themselves may be non-cooperative. An example is behavior of users of transport or telecommunication networks, where optimization from single users' perspective may not be optimal from a system one. This situation is illustrated by the famous Braess's paradox (see [2]), where an addition of a new link with identical demand faces the increasing costs for all users.

Example 1

Braess's paradox is illustrated by a simple example. The original network (Fig. 1) consists of four nodes 1, 2, 3, 4 and four edges h_1, h_2, h_3, h_4, the origin and the destination of the network are presented by nodes 1 and 4.

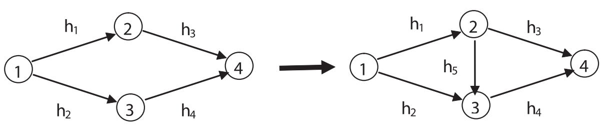


Fig. 1 Braess's paradox

There are two paths from the origin to the destination of the network $C_1 = \{h_1, h_3\}$ a $C_2 = \{h_2, h_4\}$. Let we suppose the costs on the edges depending on the flow quantities x_1, ..., x_4 are

$$n_1(x_1) = 10\,x_1,\ n_2(x_2) = x_2 + 50,\ n_3(x_3) = x_3 + 50,$$
$$n_4(x_4) = 10\,x_4$$

and the total required network flow $X = 6$.

In the case of user-optimization, the equilibrium solution is given by the situation, where all paths connecting the origin-destination pair have equal and minimal costs and therefore no user has any incentive to switch this path. The equilibrium solution is given by flows on edges

$$x_1^* = 3,\ x_2^* = 3,\ x_3^* = 3,\ x_4^* = 3$$

and by associated path costs

$$n(C_1) = 83,\ n(C_2) = 83.$$

A new edge h_5 joining node 2 to node 3 with user cost $n_5(x_5) = x_5 + 10$ is added. The change creates new path $C_3 =$ $= \{h_1, h_5, h_4\}$. The original solution is no longer in equilibrium. The new equilibrium solution has the flow of amount 2 for all three paths. Edge flows are

$$x_1^* = 4,\ x_2^* = 2,\ x_3^* = 2,\ x_4^* = 4,\ x_5^* = 2$$

and the associated path costs are

$$n(C_1) = 92,\ n(C_2) = 92,\ n(C_3) = 92.$$

Costs grow up for every user of the network from the value 83 to the value 92. This cost increase is caused by the fact that edges h_1 and h_4 are shared by two paths and the flows and cost are increasing on these edges. The addition of a path connecting an origin-destination pair that shares no links with the original connection will never result in Braess's paradox.

3 Combinatorial auctions

An auction mechanism is denoted as a combinatorial auction, if combinations of items are traded and not single items only (see [3], [8]). Combinatorial auctions are increasingly considered as an alternative to simultaneous single-item auctions. The advantage of combinatorial auctions is the more precise expression of bidder's preferences. This advantage is primarily important in the case of complementary items. The items are complementary, if the utility of set of items is greater than a sum of utilities of single items. Two items A and B are complementary, if it holds

$$u(\{A, B\}) > u(\{A\}) + u(\{B\}).$$

From different types of combinatorial auctions we present an auction of indivisible items with one seller and several buyers. Let us suppose that one seller offers a set G of m items, $j = 1, 2, ..., m$, to n potential buyers. Items are available in single units. A bid made by buyer i, $i = 1, 2, ..., n$, is defined as

$$N_i = \{S, p_{i,S}\},$$

where
S □ G, is a combination of items,
$p_{i,S}$ is the offered price by buyer i for the combination of items S.

The objective is to maximize the revenue of the seller given the bids made by buyers. Constraints establish that no single item is allocated to more than one buyer and that no buyer obtains more than one combination. The winner determination problem belongs to NP-hard problems.

Notations and variables are introduced for model formulation:
$\delta_{j,S} = 1$, if item $j \square S$, $\delta_{j,S} = 0$, if item $j \square S$,
$x_{i,S}$ is a bivalent variable specifying if the combination S is assigned to buyer i ($x_{i,S} = 1$).

The winner determination problem can be formulated as follows

$$\sum_{i=1}^{n} \sum_{S \subseteq G} p_{i,S}\, x_{i,S} \rightarrow \max$$

subject to

$$\sum_{i=1}^{n} \sum_{S \subseteq G} \delta_{j,S}\, x_{i,S} \square 1, \;\; \forall j \square G,$$

$$\sum_{S \subseteq G} x_{i,S} \square 1, \; \forall i, \; i = 1, 2, ..., n,$$

$$x_{i,S} \square \{0, 1\}, \; \forall S \square G, \; \forall i, \; i = 1, 2, ..., n.$$

For general solving of the winner determination problem, dynamic programming (see [6]) is proposed. The authors also consider several restrictions on allowable bids that make the problem computationally manageable.

4. Auctions on networks

The traded commodities can be network capacities, which enable various types of flows in network industries. A classical example is capacity of telecommunication networks where the capacities of certain links are supplied and demanded. The objective is the optimal combination of segments in required paths. Principles of combinatorial auctions are useful for selling network capacities. The utility of the path capacity is greater than a sum of utilities of edge capacities.

For illustration we present a basic formulation of the winner determination problem in a combinatorial auction for selling network capacities. Let $G = (U, H)$ be a network, where U is a set of nodes and H a set of edges. To each edge $h_j \square H, j = 1, 2, ..., m$ a capacity $k(h_j)$ is associated. Capacities of the network are owned by a single seller and there are n buyers, potentially interested in path capacities. The combinatorial aspect of the problem ensues from the fact that buyers desire to obtain path capacities (combinations of edges) rather than on individual capacities of edges. We suppose that buyer i, $i = 1, 2, ..., n$, submits a single bid specified by the following specifications

$$N_i = \{Z_i, K_i, G_i, k_i, p_i\},$$

where
Z_i, K_i is an origin-destination pair of nodes specifying the path for buyer i,
G_i is a subgraph of the graph G specifying edges for possible paths,
k_i is the required capacity for the path,
p_i is the offered price of buyer i for the combination.

Other notations and variables are introduced for model formulation:
C_i is the set of all paths between Z_i and K_i in the subgraph G_i,
c is a path from the set C_i,
$\delta(h_j, c) = 1$, if $h_j \square c$, $\delta(h_j, c) = 0$, if $h_j \square c$,
y_c is a variable specifying the capacity on the path $c \square C_i$,
x_i is a bivalent variable specifying if bid N_i is winning ($x_i = 1$).

The winner determination problem in a combinatorial auction for selling network capacities can be formulated as follows

$$\sum_{i=1}^{n} p_i x_i \rightarrow \max$$

subject to

$$\sum_{c \in C_i} y_c = k_i x_i , \; i = 1, 2, ..., n,$$

$$\sum_{i=1}^{n} \sum_{c \in C_i} \delta(h_j, c)\, y_c \square k(h_j) , \; h_j \square H,$$

$$x_i \square \{0, 1\}, \; i = 1, 2, ..., n,$$

$$y_c \square 0, \; c \square C_i, \; i = 1, 2, ..., n.$$

When paths are specified by buyers and single units of capacity are available on edges as requested by buyer, the model is equivalent to the model of the winner determination problem, in which items are edges and combinations are paths. The particularity of the model lies in the fact that buyers do not need to indicate a specific path along which the capacity should be allocated. It is an auctioneer's task of routing the requested capacities in order to determine the winning bids. The model could be solved by commercial mixed integer programming software. However, the model can be formulated as a multicommodity flow problem. Multicommodity network flow problems involve several flow types or commodities, which simultaneously use the network and are coupled through edges with limited capacity. Each commodity has an associated demand and an origin-destination pair of nodes. Important examples of such problems arise in communication, transportation, and manufacturing networks. Effective methods can be used for solving multicommodity network problems (see [1]). The solution methods for solving the multicommodity flow problem generally attempt to exploit the network structure of the individual commodity flow problems. The approaches are based on price-directive decomposition, resource-directive decomposition, or partitioning methods.

Example 2

This simple example is an illustration of terms and notations. Fig. 2 presents a network with following capacities

$k(h_1) = 3$, $k(h_2) = 3$, $k(h_3) = 3$, $k(h_4) = 3$, $k(h_5) = 3$,
$k(h_6) = 3$, $k(h_7) = 3$, $k(h_8) = 3$, $k(h_9) = 3$.

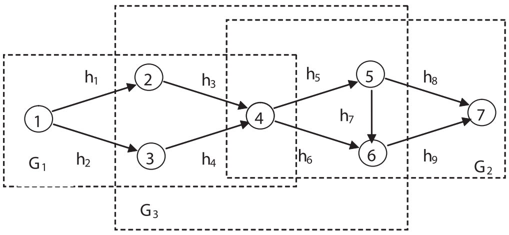


Fig. 2 Selling network capacities

Let us suppose three bidders with their bids

$N_1 = \{Z_1 = 1, K_1 = 4, G_1, k_1 = 3, p_1 = 10\}$,

$N_2 = \{Z_2 = 4, K_2 = 7, G_2, k_2 = 3, p_1 = 10\}$,

$N_3 = \{Z_3 = 3, K_3 = 6, G_3, k_3 = 3, p_3 = 10\}$.

Subgraphs G_1, G_2, G_3 specifying edges for possible paths are shown in Fig. 2. However, for the bid N_3 the edge h_3 will be never used.

There is the optimal solution of the example

$$x_1 = 1, x_2 = 1, x_3 = 1, \sum_{i=1}^{n} p_i x_i = 30.$$

The paths are identified for particular bids

$C_1 = \{h_1, h_3\}$,

$C_2 = \{h_6, h_9\}$,

$C_3 = \{h_4, h_5, h_7\}$

or alternatively

$C_1 = \{h_1, h_3\}$,

$C_2 = \{h_5, h_8\}$,

$C_3 = \{h_4, h_6\}$.

5. Conclusions

Analysis and optimization of network economy functioning are challenges for application of modeling approaches. Auctions are important market mechanisms for the allocation of goods and services. Recently research and applications of combinatorial auctions are significantly increasing. In the paper a basic model for selling network capacity is formulated. The model is based on combinatorial auctions. The advantage for solving this model is a possibility of utilization algorithms for multicommodity network problems. This basic model can be extended for other types of auctions on network structures. Integration of findings from economics, operations research and computer science is promising for interesting results.

The research project was supported by Grant No. 402/05/0148 from the Grant Agency of the Czech Republic "Network economy – modeling and analysis."

References

[1] AHUJA, R. K., MAGNANTI, T. L., ORLIN, J. B. (1993): *Network flows: Theory, algorithms, and applications.* Prentice Hall, Englewood Cliffs.
[2] BRAESS, D. (1968): *Über ein Paradoxon aus der Verkehrsplanung.* Unternehmensforschung 12, 258–268.
[3] CRAMTON, P., SHOHAM, Y., STEINBERG, R. (eds.) (2006): *Combinatorial Auctions. MIT Press,* Cambridge, Massachusetts.
[4] ECONOMIDES, N. (1996): *The Economics of Networks.* International Journal of Industrial Organization 14, no. 6, 673–699.
[5] FIALA P. (2005): *Network's economics.* Automa 11, no. 7, 13–15.
[6] PEKEČ, A., ROTHKOPF, M. H. (2003): *Combinatorial auction design.* Management Science 49, no. 11, 1485–1503.
[7] ROTHKOPF, M. H., PARK, S. (2001): *An elementary introduction to auctions.* Interfaces 31, no. 6, 83–97.
[8] ROTHKOPF, M. H., PEKEČ, A., HARSTAD, R. M. (1998): *Computationally manageable combinatorial auctions.* Management Science 44, 1131–1147.
[9] SHAPIRO, C., VARIAN, H. (1999). Information Rules: A Strategic Guide to the Network Economy. Harvard Business School Press, Boston.

KOMUNIKÁCIE COMMUNICATIONS

Michal Kaukič *

MATHEMATICS IN APPLIED INFORMATICS EDUCATION – NEW CHOICES AND CHALLENGES

In this paper, new ways of teaching Mathematics, which have been opened by the widespread of computers and notebooks and also by adequate software tools, are discussed. Based on our experience, we advocate the use of Open Source software as the mainstream tools in Mathematics and Informatics education. We give the more concrete examples of tools, used for teaching of Numerical Analysis topics, which can be successfully explored also for majority of other subjects in Applied Informatics and Mathematics. Some recommendations for the choices of suitable software and for further coordination in this field are given.

1. The traditional view of teaching Mathematics in Engineering environment

In Engineering education, there was traditionally a big difference between the way of teaching Mathematics and other, specialised subjects e.g. from mechanical, electrical, civil engineering or informatics.

The mathematicians like to expose the essence of Mathematics in the strictly logical form of reasoning using axioms, definitions an theorems, eventually with rigorous (or informal, but more intuitive) proofs whenever it can contribute to better understanding of explained concepts.

The engineers are more interested in the solution of real-world problems, using selected mathematical apparatus for formulation of (simplified) computationally tractable model and for solution of that model problem with the aid of some of available (or suitably modified) algorithms, usually only with certain (prescribed or estimated) accuracy.

Standing on this viewpoint, the practical engineers perceive many of traditional topics of algebra or calculus as more or less unnecessary or simply useless for engineering practice. On the other hand, mathematicians must follow the logical continuity and dependencies between introduced concepts. There is no way to fully understand the concept of Fourier transform without the basic knowledge concerning the definite integrals. In linear algebra, the basic concept is that of matrix. Without proper theoretical grounding (linear independence, determinants, inverse matrix, eigenvalues and eigenvectors, etc.) the students cannot fully understand and explore many of practical algorithms for solution of linear systems, optimisation algorithms such as simplex method, spectral analysis of linear processes, etc.

Two principal questions in this field are about teaching material – (what to teach?) and teaching methods – (how to teach?). In our opinion, the latter question is far more significant. In present, the concrete information we give to students is not so important as the message it cares about ideas, leading to some intellectual or practical achievements. *We should communicate ideas not the plain information.* Our students should be able to make individual progress in the process of complementing their existing skills and knowledge with new tools and information as demanded by needs of their professional career.

Thus, the common point for both mathematics and engineering educators is to prepare students to *independent, logical, clear and continuous thinking.* It is not sufficient to give the students many (maybe, highly practical) isolated pieces of information in the cookbook style. We should bring to the students the feeling of beautiful integral view of *The Forest of Science,* not to walk with them in the darkness of labyrinth of individual trees. If such walk is necessary, then it should be not too long and lead to a peak with far, clear views.

We (the communities of mathematicians and engineers) should choose the adequate tools to be in state of resonance and to support each another, not to make the useless remarks about (supposed) weak points of either "engineering" or "mathematical" mode of thinking. In this paper, we will report our experience with computer assisted teaching of some topics in Numerical analysis and some negative habits observed in students' behaviour.

2. Dangers coming from mindless use of computers

With the wide accessibility of computers, there is the growing principal danger of overestimating their role in education and in the engineering practice. In the naive students' perception the computers with highly intelligent software can be used as substitute for the (painful) process of creative thinking. But also on the teacher's side there are some overly optimistic expectations from using tools like LMS (Learning Management Systems, e-learning).

* **Michal Kaukič**
Department of Mathematical Methods, Faculty of Management Science and Informatics, University of Zilina, Slovakia
E-mail: mike@frcatel.fri.utc.sk

The common sense, of course, tells us that nothing can replace the individual creative thinking and that the personality of teacher will always play the key role in the process of knowledge, intuition and skill transfer to the students.

Let us bring the quotation from SEFI (European Society for Engineering Education) document [1]:

There are some signs that the involvement of computers in the teaching of undergraduate engineering mathematics is beginning to gain momentum. However, the experience of the last thirty years warns that care must be taken. When the pocket calculator arrived we were told that there would be two advantages: first, the students would be relieved of hours of tedious calculation, leaving them free to concentrate on concepts and understanding; second, it was more likely that the calculations would be performed correctly. The reality is somewhat different. The most trivial of calculations is often subcontracted to the machine, the students have little or no feel for what they are actually doing or to what precision they should quote their answers. What they can do is to obtain obviously unrealistic results more quickly and to more significant figures. There has also been a tendency to replace analytical reasoning by trial-and-error methods.

This is the clear analysis of dangers, caused by unqualified use of computers. We should bear this in mind and to create the environment where computers are helpful but where critical, logical thinking is all the time at the first plane. In the next section we will show an example of such environment, we now use for the teaching of Numerical analysis.

3. Pylab – the environment for experimenting and visualization of Numerical analysis concepts

Effective teaching of Numerical Analysis is closely related to the problem of using computers as the laboratory, experimental equipment, in the sense which is well known from other natural sciences. Numerical analysis must take into account the real world imperfections, which can be happily ignored by theoreticians. Thus, in the field of Numerical analysis there is a big room for experiments. We believe that numerous computer-based experiments are essential for successful teaching of Numerical Analysis.

The choice of appropriate software tools is of the fundamental importance. This choice, once made, determines the effectiveness and long-term maintainability of educational and research activities for many future years.

On the first look, we will probably find three well-known commercial software systems for mathematical teaching and research: MATLAB, Mathematica, Maple.[1)]

Besides of certain advantages, commercial software systems have also many drawbacks (for more detailed discussion, see [2]). Our search for alternative software leads us to several Open Source systems with MATLAB-like capabilities. The most MATLAB-compatible among them was Octave [3]. In the course of using MATLAB and later Octave we constantly perceived the apparent deficiencies of underlying simple programming language. Moreover, programming in MATLAB can be done in very bad style by inexperienced students not encouraging them to learn new ways of thinking about problems. The MATLAB language simply is not general enough to be useful outside of its primary domain - matrix and vector manipulations.

Although there is no universal system suitable for all areas of mathematical education, we can try to minimise the number of different tools used. The choice should be made so that students can reuse our tools in possibly widest area of their future professional career (having in mind preferably the students of Informatics).

This brings us to the idea of using some of general purpose programming languages. The language of our choice should allow to express mathematical ideas with minimum programming effort and in "nearly mathematical" notation.

In present, we use the Pylab environment, comprised of interpreted language Python [4], complemented by user-friendly interactive environment (IPython shell, [5]), modules for linear algebra and numerical analysis (see [6], [7]) and also module Matplotlib [8] for excellent quality two-dimensional graphics. We will show the typical use of this environment on the following example.

Example 3.1 *Find positive solutions (i.e. with positive components) of the system of nonlinear equations*

$$\sin(xy^2) - \cos(x^2 - y) + 0.2 = 0$$

$$x^3 + y^3 - 3xy = 0$$

First, we can give the nice graphical interpretation for this example. The solutions will be exactly the common points of zero contour levels of surfaces $z_1(x, y)$, $z_2(x, y)$, given by left sides of equations, i.e.

$$z_1(x, y) = \sin(xy^2) - \cos(x^2 - y) + 0.2,$$

$$z_2(x, y) = x^3 + y^3 - 3xy$$

The corresponding plot will be generated by the commands (exactly the same, as it would be in MATLAB):

```
xx=linspace (-2, 2, 120);   yy=linspace (-2, 2, 120)
X, Y=meshgrid (xx, yy)
```

[1)] Maple, MATLAB, Mathematica are registered trademarks of their respective owners (Waterloo Maple Inc., The MathWorks, Inc., Wolfram Research, Inc.). Other trademarks mentioned in this article are registered trademarks of their respective owners too.

```
Z1=sin(X*Y*Y)-cos(X*X-Y)+0.2;     Z2=X**3+Y**3-3*X*Y
contour (x, Y, Z1, [0.0])
contour (x, Y, Z2, [0.0],colors='g')
```

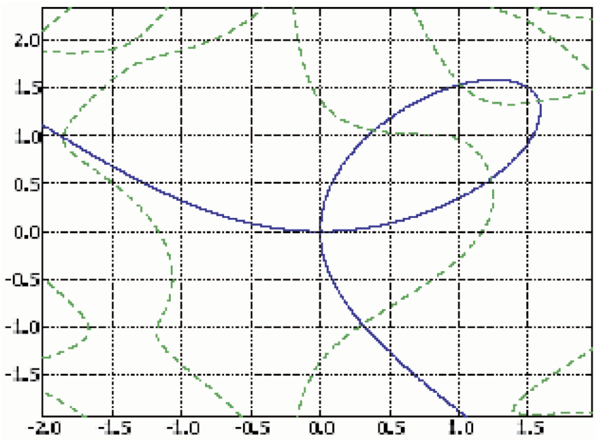


Fig. 1 Solution of system nonlinear equations

From this figure we can see that there are exactly four positive solutions. We can zoom the figure window and to make the good initial guesses for solutions:

```
(0.38, 1.06), (1.23, 0.55), (1.02, 1.54), (1.58, 1.36)
```

For the numerical solution of nonlinear systems we can use the function `fsolve`. The first argument of this function is the vector function, describing the given nonlinear system – the input is in our case the tuple (x, y) and the function returns the tuple of values (z_1(x, y), z_2(x, y)). The code for this function is as follows (we can save it to the file `nlfct.py`):

```
def nlfct(xy):
x,y = xy # xy has two components, separate them
z1=sin(x*y*y)-cos(x*x-y)+0.2
z2=x**3+y**3-3*x*y
return (z1, z2)
```

The second argument for fsolve is the initial approximation of solution. Thus, we can get the first solution in interactive IPython environment by commands:

```
run nlfct
from scipy.optimize import fsolve
r=fsolve(nlfct, (1.23, 0.55))
# The answer is: r=(1.2328735, 0.5521787)
```

As always, we cannot blindly trust the results, we get from computer. But we can easily verify that the command `nlfct(r)`, which computes the values of functions z_1(r), z_2(r) returns very small values (approximately $-4.552 \cdot 10^{-15}$, $-3.997 \cdot 10^{-14}$. The same is true for remaining three solutions, but we will not repeat the calculations here.

We can see that Pylab is, indeed, the very high level language – no type declarations of variables and functions, no troubles with memory management. There are many ready-made useful functions for numerical analysis (data interpolation and approximation, linear algebra functions – eigenvectors and eigenvalues, determinant, inverse matrix, solution of linear equations, also in least-squares sense, numerical integration, optimisation with bound constraints, solution of differential and differential-algebraic equations and many others). Pylab has also many functions for statistics and probability theory. There is also a signal processing toolbox analogical to that of MATLAB.

Using additional Python modules, we can easily add e.g. the SQL DBMS interface for working with commercial (Oracle) or Open Source (PostgreSQL, mysql, SqLite) SQL databases, there are other modules for discrete and continuous simulations, linear and, more general, convex programming.

Now, we can summarise some experience we gained in teaching Numerical Analysis with Pylab. The very positive thing is the possibility of rapid prototyping and interactive debugging of little scripts. We have prepared the material about basic programming in Pylab (see [9]), which was sufficient for students to start working in the interactive environment. Unfortunately, many students have strange and non-productive habits of programming.

IPython environment has excellent command completion, but students prefer the (erroneous) lengthy typing. Or, there is no need to position the cursor on the end of line before entering the command, but they have constantly made this action. IPython saves command line history, so any previous commands can be conveniently re-entered or edited, but students like to type the same thing many times, instead of using this feature.

The biggest misbehaviour from students' side is the programming without thinking. They start typing something, having little understanding of the given task and the methods they will use for solution. They put emphasis on programming routine, but not on "The Art of Computer Programming" as seen by the classical works of D. Knuth [10]. We have the strong opinion that for beginners the clear and simple programming language should be used. Neither Pascal, nor C/C++ or Java can fulfil these basic requirements. We propose to try (at University level) to unify the programming tools for basic courses on algorithmization and programming using Python-based environment.

The considerable amount of work should be done for analysis of existing tools and environments (especially non-commercial, OpenSource alternatives), the choice of basic software for applied informatics education and the customisation of chosen software for the specific needs of faculty and university. We think there should be first a broad discussion at inter-faculty level and afterwards the work group responsible for the implementation of decisions coming from this discussion should be formed. The process of implementation will, of course, take several years to be reasonably complete.

4. Conclusion

We hope that the reader got some feeling of what can be done with newly available Open Source software tools in the field of mathematical education at University level. The cultivation of mind in Mathematics and in Computer Science can be done in close cooperation. This can lead to the synergy effect of better understanding of the common points and the role of clear, logical, independent thinking in all applied informatics activities. This is especially important for making right key decisions, which will be crucial for the future of the fields of Applied informatics and Mathematics at our University and in our country.

We saw the example of successful use of user-friendly environment, based on Open Source tools in teaching of Numerical Analysis. This environment, in our opinion, has the great potential of becoming the tool of choice for nearly all educational topics in traditional and modern branches of Mathematics (algebra, calculus, probability and statistics, graph theory, mathematical programming and optimisation, cryptography, numerical computations, computer graphics, computational geometry, etc.) and also for many important concepts of Applied Informatics (the basic principles of algorithmization, data structures, discrete and continuous simulation, stochastic processes, signal processing, parallel computations, etc.).

For optimal results in applied informatics education (bachelor, undergraduate, graduate and doctoral study) we cannot use the multitude of software tools without any coordination. There should be clearly defined "central path" software, which will be used continually during all the time of study. We are convinced that the right choice for this durable software tool cannot be based on traditional, low level languages like Pascal, C/C++. The good candidate should be very high level language with clear, simple syntax and high expressive power – at present, the language of our choice is Python with suitable modules. Using it has also the additional benefit of learning a simple, useful general-purpose language, which students can use later in their career (important not only for students of Computer Science but also for general engineering students and mathematicians, too).

References

[1] SEFI Math. working group: *Mathematics for the European Engineer,* SEFI HQ, 2002
[2] KAUKIČ, M.: *Open Source software resources for numerical analysis teaching,* International Journal for Mathematics Teaching and Learning, CIMT, Univ. of Exeter, Exeter, Oct. 2005, ISSN 1473-0111, electronical form, 9 pages
[3] EATON, J. W.: *Octave Manual, available on http://www.octave.org/docs.html*
[4] VAN ROSSUM, G.: *Python Documentation, on http://www.python.org/doc*
[5] PEREZ, F.: *IPython – An enhanced Interactive Python,* available from http://ipython.scipy.org
[6] ASHER, D., DUBOIS, P. F., HINSEN, K., HUGUNIN, J., OLIPHANT, T.: *Numerical Python,* available on *http://numpy.sourceforge.net*
[7] SciPy Devolopers: *SciPy,* available on *http://www.scipy.org*
[8] HUNTER, J. D.: *Matplotlib*, available on *http://matplotlib.sourceforge.net*
[9] KAUKIČ, M.: *Essentials of Pylab Programming, (in Slovak),* draft of textbook in preparation, can be downloaded from *http://frcatel.fri.utc.sk/mp2.pdf*
[10] KNUTH, D.: *The Art of Computer Programming,* vol. 1, Addison-Wesley, 1969.

KOMUNIKÁCIE COMMUNICATIONS

Matúš Jurečka *

FEATURE EXTRACTION USING PULSE-COUPLED NEURAL NETWORK IN ISOLATED SPEECH RECOGNITION

This article presents achieved results concerning an feature extraction in isolated speech recognition problem using the Pulse-Coupled Neural Network (PCNN) approach. PCNN based feature extraction is analyzed for a direct Pulse Coded Modulation (PCM) input and a Fast Fourier Transform (FFT) coefficients input.

1. Introduction

The speech recognition problem may be interpreted as a speech-to-text conversion problem. A speaker wants a voice to be transcribed into text by a computer. Automatic speech recognition has been an active research topic for more than four decades. With the advent of digital computing and signal processing, the problem of speech recognition was clearly posed and thoroughly studied. These developments were complemented with an increased awareness of the advantages of conversational systems. The range of the possible applications is wide and includes: voice-controlled appliances, fully featured speech-to-text software, automation of operator-assisted services, and voice recognition aids for the handicapped persons. Different approaches in speech the recognition have been adopted. They can be divided mainly into two trends – hidden Markov model (HMM) and artificial neural network (ANN).

2. Speech signal processing

Speech acquisition begins with a person speaking into a microphone. Generally, a speech signal is converted onto a digital form using the pulse coded modulation (PCM). This means of speech signal representation is not so suitable for a pattern recognition. However, it can be represented by a limited set of features. There are several methods available for features extraction and dimension reduction. The dimension reduction is a transformation of an input signal space into a feature space with a lower dimension. The goal of the dimension reduction is to obtain significant features for a unique pattern representation. Classical methods of dimension reduction include Karhunen – Lo_ve transform [11], singular value decomposition (SVD), etc.[5] Dimension reduction methods based on ANN are for example Kohonen Self-Organized maps [6] or principal component analysis (PCA neural networks) [10, 9].

Classical methods of features extraction in digital signal processing for speech recognition include coefficients of discrete Fourier transform, linear predictive coefficients (LPC), filter bank, mel scale frequency cepstral coefficients (MFCC) etc. [7,8] The method scoring growing interest for a dimension reduction & feature extraction in a field of image processing is the Pulse Coupled Neural Network (PCNN). My work focused on the feature extraction in isolated speech recognition process using the PCNN.

3. A PCNN structure

The structure of a standard PCNN comes out from the structure of an input pattern which will be processed. Let us consider that the input pattern is a matrix of values for input of an isolated word. The PCNN is a single layered, two-dimensional, laterally connected neural network of pulse coupled neurons connected with values of an input matrix. Each input matrix value is associated with a pulse coupled neuron of a specific structure. The PCNN neuron consists of an input part, linking part and a pulse generator. The neuron receives the input signals from feeding and linking inputs. The feeding input is a primary input from the neuron's receptive area. The neuron receptive area consists of neighboring values of the corresponding value in the input matrix. The linking input is a secondary input of lateral connections with neighboring neurons. The difference between these inputs is that the feeding connections have a slower characteristic response time constant than the linking connections. The standard PCNN model is described as iteration by the following equations:

$$F_{ij}(n) = S_{ij} + F_{ij}(n-1) \cdot e^{-\alpha F} + V_F \cdot (M * Y(n-1))_{ij} \quad (1)$$

$$L_{ij}(n) = L_{ij}(n-1) \cdot e^{-\alpha L} + V_L \cdot (W * Y(n-1))_{ij} \quad (2)$$

$$U_{ij}(n) = F_{ij}(n) \cdot (1 + \beta \cdot L_{ij}(n)) \quad (3)$$

$$\Theta_{ij}(n) = \Theta_{ij}(n-1) \cdot e^{-\alpha\Theta} + V_\Theta \cdot Y(n-1) \quad (4)$$

$$Y_{ij}(n) = \begin{cases} 1 \text{ if } U_{ij} > \Theta_{ij} \\ 0 \ otherwise \end{cases} \quad (5)$$

* **Matúš Jurečka**
Department of Technical Cybernetics, Faculty of Management Science and Informatics, University of Žilina, Veľký diel, 010 26 Žilina, Slovakia, jurecka@frtk.fri.utc.sk

Where F_{ij} is the feeding input, L_{ij} is the linking input, n is an iteration step, S_{ij} is a value at i,j coordinates in the input matrix. W and M are the weight matrices, is the convolution operator, Y_{ij} is the output of the neuron at i,j coordinates, V_L and V_F are potentials, α_L and α_F are decayed constants.

Single signals of the linking input are biased and then multiplied together. Next, the input values F_{ij}, L_{ij} are modulated in the linking part of a neuron. We also obtain internal activity of the neuron U_{ij} in the specific iteration step. If internal activity is greater than dynamic threshold Θ_{ij}, then the neuron generates output pulse. Otherwise, the output equals to zero. The neuron output Y_{ij} does not necessarily need to be binary. It is possible to use a sigmoid pulse generator where the neuron takes the analogue value from 0 to 1. The input matrix is transformed through the PCNN into a sequence of temporary binary matrixes. Each of these binary matrixes has the same dimension as the input matrix. The sum of all activities in a specific iteration step gives one value representing one feature for the classification. If we have N iteration steps, we obtain N features. The one-dimensional time signal generated from the values of the output matrix $Y_{ij}(n)$ in every iteration step n can be defined as follows:

$$G(n) = \sum_{i,j} Y_{ij}(n) \qquad (7)$$

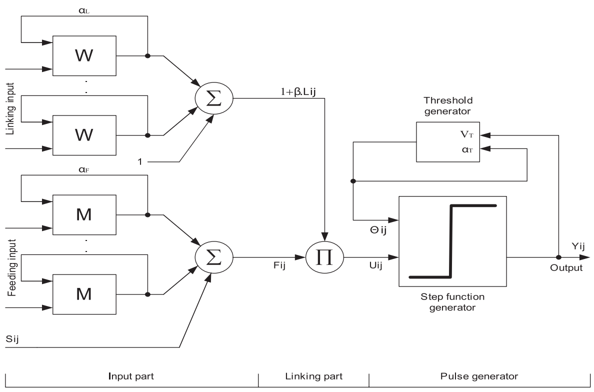


Fig. 1: PCNN neuron structure – taken from [4]

Significant advantage of the PCNN, which is useful mainly in image recognition, is the invariance of a generated time signal to rotation, dilatation or translation of images [4]. Therefore, the PCNN is advisable for the feature generation and pattern recognition in the classification tasks using conventional neural networks or other methods. Thanks to translation invariance of generated features, the PCNN method used in speech recognition does not relay on an outstanding endpoint word detection. It is evident that the PCNN is not the neural network in the term of classification. It is only a means of feature extraction for a pattern classification using conventional neural network models, like that of multi-layer perceptron. Several models of the PCNN have been developed. The most used PCNN models are, for example, a PCNN with modified feeding input [1], fast-linking PCNN [3] or feedback PCNN [2].

4. Experiments and Results

The following experiments with feature extraction using the PCNN were made in my testing database consisting of 36 isolated Slovak words uttered once by 23 different speakers:

The abovementioned PCNN approach was applied directly to a sequence of PCM of an isolated word. I used 16-bit PCM with f_s = 8kHz. The PCNN with 200 iteration steps produced 200 features for every isolated word. The 200x1 feature vectors $G_{ab}(n)$ where a is the word index ($1 \le a \le 36$) and b is the speaker index ($1 \le b \le 23$), were than divided by their maximum values for certain normalization reasons. Figure 2 shows the mean courses of $G_a(n)$ functions for all 36 input words, the mean course of $G_a(n)$ function for the input word a was computed as follows:

$$G_a(n) = \frac{1}{23} \sum_{b=1}^{23} G_{a,b}(n) \qquad (8)$$

In my next experiment I used the Fourier transform for the PCNN input coefficients. The sequence of PCM values was partitioned into the sequence of consecutive frames. The frame length was chosen as 128 samples with 64 samples overlap. After applying the Hamming window function which prevents some spectral leakage, the fast Fourier transform (FFT) was computed in these frames. The input matrix was formed by the first 64 FFT coefficients (as the product of FFT is symmetrical) from every time frame. The PCNN with 200 iteration steps produced 200 features for every isolated word. The feature vectors were then divided by their maximum values for certain normalization reasons. Figure 3 shows the mean courses of $G_a(n)$ functions for all 36 input words computed similarly as in my first experiment.

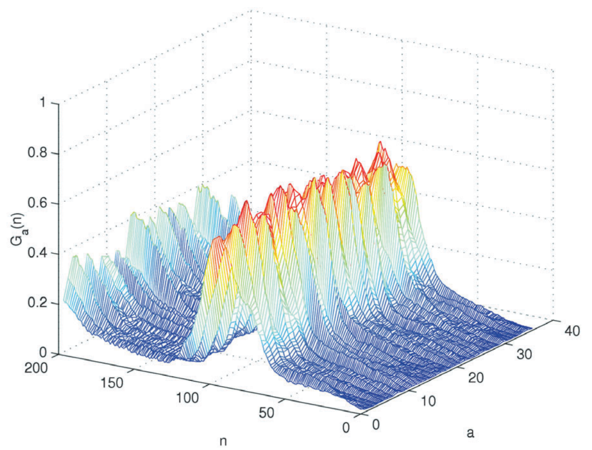


Fig.2 Mean courses of $G_a(n)$ functions for PCM input

Figure 4 shows the variance of the mean courses of $G_a(n)$ functions for input words No. 4, 5, 6, 7 with their 95% reliability level for FFT coefficients input. The experiments described were carried out in the MATLAB environment.

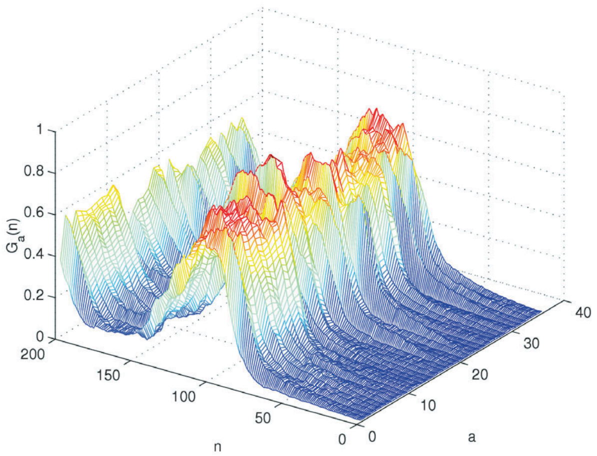


Fig. 2 Mean courses of $G_a(n)$ functions for PCM input

5. Conclusion

It is clear that for recognition it is very important for the feature vectors $G_{ab}(n)$ to be as similar as possible to the same word uttered by different speakers and at the same time these vectors should be as different as possible in different words. Difference measure can be understood, for example, as the Euclidean distance between the feature vectors. As it can be seen from the above mentioned results, the feature vectors for different words do not differ too much. On the other hand, the variance of these vectors for the given input word is much bigger than the difference measure of distinct words. It clearly shows that the speech recognition system which will rely on the introduced PCNN based feature extraction will fail. The PCNN approach has been very successful in the field of image recognition, but there is still hope that some other methods of speech signal preprocessing will be helpful in dealing with speech recognition.

Acknowledgement

This work was supported by the institutional grant FRI 2/2006.

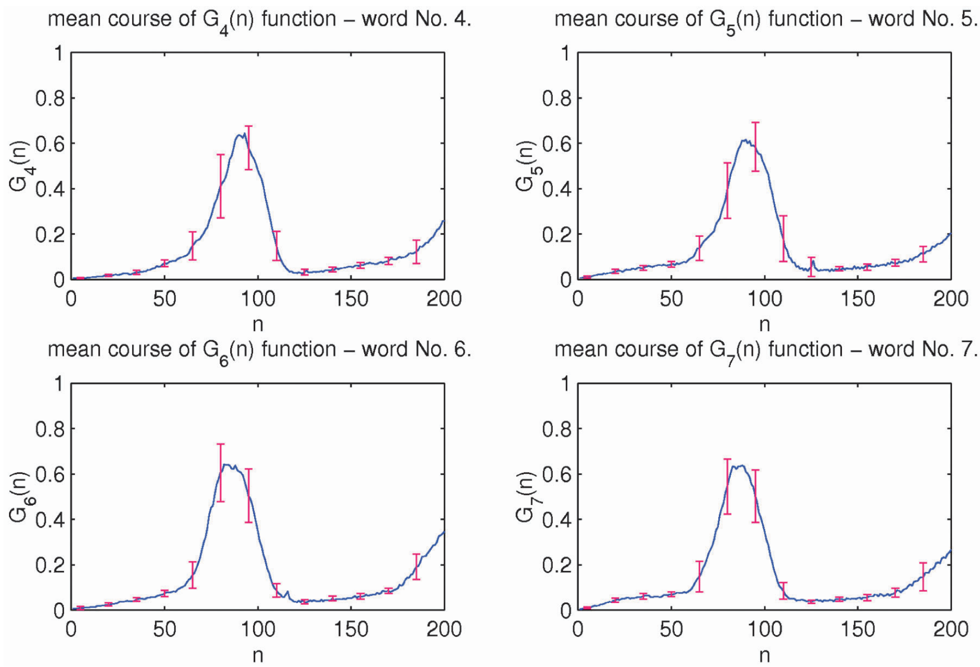


Fig.4 Mean courses of $G_a(n)$ functions for input words 4, 5, 6, 7 with their 95 % reliability level

References

[1] RANGANATH, H. S., KUNTIMAD, G.: *Iterative Segmentation using Pulse Coupled Neural Networks,* SPIE, Vol. 2760, pp. 543–554

[2] JOHNSON, J. L., PADGETT, M. L.: *PCNN Models and Applications,* IEEE Transaction on Neural Networks, Vol. 2, No.3, 1999, pp. 480–498

[3] KISER, J. M., JOHNSON, J. L.: *Implementation of Pulse-Coupled Neural Networks in the CNAPS Environment,* IEEE Transactions on Neural Networks, Vol. 10, No. 3, 1999, pp. 584–590

[4] FORGÁČ, R., MOKRIŠ, I.: *Artificial neural networks for feature space dimension reduction and classification (in Slovak),* Banská Bystrica: Univerzita Mateja Bela, 2002, ISBN 80-8055-743-8
[5] ORAVEC, M., POLEC, J., MARCHEVSKÝ, S.: *Neural networks for digital signal processing (in Slovak),* Bratislava: FABER, 1998
[6] WASSERMAN, D. P.: *Neural Computing Theory and Practice,* New York: VNR, 1989
[7] PSUTKA, J.: *Communication with Computer by Speech (in Czech),* Praha: Academia, 1995
[8] PROAKIS, J. G., MANOLAKIS, D. G.: *Digital Signal Processing,* New York: MPC, 1992
[9] KHANNA, T.: *Foundations of Neural Networks,* New York: Addison-Wesley Publishing Company, 1990. ISBN 0-201-50036-1
[10] HAYKIN, S.: *Neural Networks – A Comprehensive Foundation,* Macmillan College Publishing Company, New York, 1994
[11] ROSENFELD, A., KAK, A. C.: *Digital Image Processing,* John Wiley & Sons, New York 1982.

KOMUNIKÁCIE COMMUNICATIONS

Ivana Olivková *

CHOOSING THE RIGHT TYPE AND LOCATION OF THE CITY PUBLIC TRANSPORT STOPS

This paper is concerned to tram stop safety in the Czech Republic. Recently some experimental solutions of tram stops (with alternative successes) have been realized in Prague and Brno. These experimental solutions of tram stops, which very often emerged as trial and error, were based on elevated carriageways. Due to the fact that there is lack of qualified information and projection recommendations, I would like to present the selection of information based on real operation experience. The paper has been compiled using results of the grant project No. 103/04/0476 "Proposal of financing traffic attendance methodology", which is being realised thanks to the financial support from the public Grant Agency of the Czech Republic [1].

1. Introduction

Passengers' safety in the city public transport is a sense of personal security being felt by passengers. This feeling emerges both from factually installed safety precautions and also from activities ensuring passengers' awareness about these precautions.

It is inevitable to consider traffic safety as a matter of the highest importance. It is also unacceptable to improve other quality criteria, such as velocity, observance of timetable, comfort, etc. to the exclusion of traffic safety. Carriers are obligated to transport passengers safely without jeopardizing their health, lives and properties.

Fear of being a victim of some criminal attack is one of the most significant factor discouraging people from public transport usage. It represents a problem in many cities and results from various type of crime (such as physical attack, robberies in vehicles and stops, vandalism etc.). Those carriers who are able to cope with personal security issues and provide physical help and information enjoy respect among their users.

Fig. 1 Access must be free of any dangerous situations and places

2. Location of stops, technical requirements

Passengers encounter serious obstacles when trying to reach their stop in many cities. Pedestrian access to the stop often comprises complicated crossings through busy roads or unpleasant and dangerous zones.

Location of a public transport stop is more or less a technical matter. However, surveys presented in various EU-projects indicates necessity to consider complex of social, psychological and safety aspects during the process of municipal planning, such as:

Fig. 2 Clear transparent axes among various components

* **Ivana Olivková**
Institut of Transport, Technical University of Ostrava, tř. 17. listopadu 15, 708 33 Ostrava-Poruba, Tel.: 59-732 3122, E-mail: ivana.olivkova@vsb.cz

Interior of a public transport stop should create awareness of security among passengers. This could be achieved by eliminating dark corners and restricting number of pillars to minimum.

Additionally, there should be clear transparent axes between entrances, exits, station halls and platforms in order to enlarge space for orientation and also passengers' sense of security.

Fig. 3 Satisfactory length of a refuge with regard to trains dimensions

Dimensions of vehicles and trains should be taken into account with regard to space needed for motion and manoeuvring within the stop areas. Design of stops should respect easy cleaning and maintenance.

Fig. 4 Daylight usage

Daylight and artificial light should be maximally preferred; for instance installation of glass surfaces in large numbers making public transport travelling more pleasant and increasing sense of comfort and passengers' safety.

Refuges and corridors should be wide enough so as the passengers could feel safely. Travel distance between changes should be as short as possible, with comprehensible marking that gives required information to the passengers and assure them about the right direction of walking.

An easily localizable lifts and escalators should be on grade separation places at the passengers' disposal.

Ticket issue machines should be located near entrances and in vestibules. Harmonized timetables for all means of transport should be also appropriately located.

Fig. 5 Refuges and corridors must be designed with regard to carrying capacity and safety

Traffic measures for the disabled are very needful. Moreover, they affect much bigger group of passengers than only disabled. Every measure (such as low-floor vehicles, accessibility of metropolitan railway stops, elevated platforms...) helps also other passengers, especially those with big amount of luggage, children etc.

Intelligent transport system – ITS might be helpful when improving connection among individual transport activities and services, accessibility and safety of stops for passengers. Therefore, it is important to locate stops in a way allowing installation and accessibility of ITS equipment; for instance creating complex informational systems independent on carriers or means of transport. These systems should be multilingual, adapted for demands of disabled people, be located both in stops areas and outside and, last-but-not-least, combine static (i.e. timetables) and dynamic data about public transport services.

Fig. 6 Low-floor vehicles allow access to public transport also for disabled and elderly people

3. Choosing the right type of a public transport stop

There is a compact tramway transportation network with a long tradition and advanced traffic-engineering solution (including tram stops) in city of Vienna, Austria. A new and at the same time very promising type of a tram stop – so-called stop-with-elevated-carriageway (wrongly also as "passing-traffic-island") – has been designed and tried for the very first time in Vienna, 1992. Efforts for modernizing and improving attractiveness of the tram stops are universal.

Main issues in modernizing public transport stops (actually, public transport preferences) are as follow:

- Passengers exchange development (i.e. more comfortable getting in and out);
- Decreasing difference in elevation between the ground and a vehicle, which is very important for the disabled;
- Withdrawal from stops having exit directly to the carriageways (they are both the least comfortable for passengers and the most difficult problem for the disabled) in favour of stops with either elevated carriageways or bay stops.

During the design stage of a new stop it is essential to respect a lot of evaluating criteria, especially functional requirements (user's requirements) and site available in the street area.

That's why *there isn't any general variant preferred* or considered as "panacea". Nevertheless, there are some distinctive tendencies towards understanding appropriate order of some stop forms:

- *Bay stops* – they offer the highest possible protection and passengers may use bigger waiting areas (the carriageway is substituted for pavement within the stop area and motor traffic is led on the track).
- *Elevated pavement stops* – they are preferred due to their safety merits and low area requirements. While traffic island stops very often mean potential threat to passengers and a lot of accidents of passengers catching trams in the last possible moment are registered, elevated pavement stops accident review are much more favourable.
- *Traffic island stops* – the most widespread type of stops despite their area requirements. Sometimes it's not possible to find enough space for an island and carriageway axe shift if there are some competing functional requirements in the side areas and at the same time to guarantee side areas wide enough (pavements, cycle tracks and/or residency areas).

As a very important criterion for choosing the right type of a stop form is *the way of motor traffic conducting* before the stop. In the case of conducting motor traffic predominantly on the track before stop, because there's a parking zone, this should continue even in the stop area. These are the cases of using bay stops.

Building stops with elevated pavement is rather problematical in the case of conducting motor traffic before stops on the tracks. This is due to the fact that as a result of such solution the axe shift of motor traffic carriageway will occur and a lot of drivers would continue to drive on the tracks even in the stop area. They usually do not understand the system of traffic there – see experience from Prague. Additionally, the axe shift of motor traffic also hides significant perils. When used, it should be rather slow in order to allow motor vehicles leaving the tracks in due course.

The stops with elevated pavements were recognised and accepted by the public as effective stop form. Traffic experience with stops having elevated pavements is excellent when keeping some basic conditions. Due to the fact that the stop area is passable even for drivers, they are accepted by them more easily in comparison to the bay stops. Conducting the motor traffic on the tracks, which is the case of the bay stops, often leads to discussion about the waiting of vehicles while a tram is stopping etc.

Fig. 7 Tram stop with elevated pavement in the city of Vienna

4. Quality standards of public service from the safety point of view

A quality of carrier service in public transport is being determined by quality *standards* set for providing public transport. During the design stage of carrier quality standards in public transport it is essential to set such inviolable qualitative and quantitative limits so as to stay competitive with private car transport. These data might be used as referential for evaluating current conditions of public transport system. The intention of the quality standards

CRITERION: TECHNICAL CAPABILITY OF THE EQUIPMENT

	PUBLIC STOP	PUBLIC TRANSPORT VEHICLE
DEFINITION OF A STANDARD	*Distance post* safely fixed, having good technical conditions, with a good surface finish. *Reflex panel* (or plastic "beacon") safely attached and in good technical conditions. *Platform* in a good condition, flat surface, with gradient and functional draining. *Shelter* safely anchored, in a good technical condition, undamaged. *Street furniture – bench, rubbish bin* in a good technical condition, properly fixed.	Technical condition of a vehicle must be adequate, ground surface without any projections and failures, seats and handles must be properly attached. Device for opening and closing doors, ventilating system, heating system, radio equipment must work properly. Information cavettos, device for attaching the course number and for attaching orientation directional boards must be in good technical conditions. No scratched, splashy or doodled glasses are allowed.
UNACCEPTABLE LIMITS	*Distance post* impends to fall, corroded, heavily damaged surface finish. *Reflex panel* (or plastic "beacon") is missing, badly attached, heavily damaged. *Platform* has uneven surface, dangerous for passengers, deep puddles occur. *Shelter* is incorrectly fixed and damaged. *Street furniture – bench, rubbish bin* damaged, incorrectly fixed.	Technical devices are clearly out of order. Damage of windows or bodyworks by vandals is noti ceable.
DEMAND LEVEL	95 % of conditions must be out of unacceptable limits.	95 % of conditions must be out of unacceptable limits.
MEASUREMENT METHODS	Visual check. This is carried out by 1 worker from the transport operation check group with 1 worker from the operation group. Each stop is being inspected 2x per year, in the case of backbone exposed stops 12x per year.	Visual check. This is carried out by workers from the measurement transport operation group. The inspection comprises of 100 samples per month, reliability performance inspection of fleet acoustics devices is carried out 1x per year.

CRITERION: PASSENGER SAFETY

	VISIBLE MONITORING	STAFF/ POLICE ATTENDANCE
DEFINITION OF A STANDARD	Newly built terminals must be monitored. The CCD-cameras are in operation and must be situated in a way that all stations/stops are clearly visible. Transmission of a signal to the transport operation centre is ensured. The CCD-cameras are in operation 24/7 year-round (workdays, weekends and bank holidays included).	Security service/Police must be present in newly built terminals. They should ensure security in the terminals by regular checks every 2 hours.
UNACCEPTABLE LIMITS	The CCD-cameras are inconveniently situated (stations/stops are not clearly visible). There's no transmission of a signal to the transport operation centre. Monitoring is off.	Security service/Police is not present in terminal. Security is not ensured.
DEMAND LEVEL	95 % of conditions must be out of unacceptable limits.	95 % of conditions must be out of unacceptable limits.
MEASUREMENT METHODS	Random check 3x per year.	Random check 3x per year.

is presentation and service-monitoring reasons of carriers. The main objective is to create conditions for ensuring an eminent share of transport services by improving their quality.

Carrier quality service from the safety point of view refers to the following criteria:

- Technical capability of the equipment (stops, public transport vehicles);
- Private safety (visible monitoring, staff attendance, first-aid station marking).

Standards for each criterion of the quality of service have been set. They comprise:

- *Definition of a standard* (being characterized by the service provided or – in the case of technical equipment – determines its operation capability);
- *Unacceptable limit* (being characterized by absolutely unacceptable situations considered as service failure, which bring about immediate reactions for restoring the service;
- *Demand level* (being determined by the share of service provided in the demanded and suitable quality to the total number of monitored samples);
- *Measurement method* (presents instruction for choosing appropriate measurement methods)

The following table presents quality standards of individual criteria with regard to the object.

5. Conclusion

The paper deals with passengers' security issues in the means of public transport and stops in municipalities, where all participants of road transport might be jeopardised. An objection is to create conditions for increasing public transport share during division of transport labour by improving quality in safety matters. Offering quality in passenger transport is a basis for decreasing volume of cars in cities. As for the quantity point of view, passenger transport is being ensured rather easily, but user's demands rising especially in quality area in terms of safety matters.

References

[1] SUROVEC, P., OLIVKOVÁ, I., KŘIVDA, V.: *Proposal of financing traffic attendance methodology.* VŠB TU Ostrava, Institut dopravy, 2004. Grant Project No.103/04/0476 realized under financial support of the Grant Agency of the Czech Republic.

[2] ANGENENDT, W., BRÄUER, D., KLÖCKNER, D.: *Mittellage-Haltestellen mit Fahrbahnanhebung,* Berichte der Bundesanstalt für Straßenwesen, Bergisch Gladbach, 2002.

[3] *Quality of interchanges,* Conquest Research for London Transport Marketing (1997).

[4] RUNKEL, M: *Stations in public transport,* UITP Revue, Vol 43, N^{o} 5. (pp.32–36) (1994).

[5] *Public transport in stations: location, functional design and management and operations of public transport stations.* UITP (1993). Brussels.

Andrea Bencsik – Judit Fehér *

QUALITY OF FRONT-LINE WORK = HUMAN RESOURCE MANAGEMENT?

"Good quality" human resource management is a source of competitive advantage for companies in the field of personal services, especially in customer service.

Customers' opinion of overall service quality is very much influenced by the impression, when they encounter front-line staff. HR management can be regarded as "good quality" if it is able to ensure that staff members working on the front line are professionally and emotionally well-prepared, and committed. It can only perform its function if it is able to adopt a new approach. As an internal service provider it strives to support workers in both meeting corporate requirements and achieving their individual goals. In the case of service companies with networks change inevitably requires giving local management a broader scope of authority in the field of human resource management, as well as preparing local management for carrying out its increasing range of functions.

1. Introduction

"Good quality" human resource management is a source of *competitive advantage* for companies in the field of personal services, especially in customer service. Ensuring "good quality", that is efficacy and efficiency, presents a significant challenge to companies in the industry due to the special features of services (heterogeneity, intangibility, perishability, inseparability. Companies that have national or regional branch networks, such as banks, insurance companies, travel agencies or postal service providers, find it even more difficult to meet this challenge. They run a large number of independently functioning premises, branches, or offices (we will collectively refer to them as 'network units' in this paper), where they have to offer well-balanced and high quality services to their customers' satisfaction. It is well-known that customers' opinion of the quality of services is greatly influenced by the impression that they have of front-line personnel as they meet them. To make sure that these companies have satisfied customers, they have to have well-trained, motivated, satisfied, what is more, committed staff.

Network units occupy a pivotal role in the value chain of the above-mentioned companies and they affect the value creating ability of companies in a significant way. However, the fact that a service provider has a large number of network units does not "automatically" generate a competitive advantage for it.

Research has shown that those corporate abilities and traits can turn into competitive advantages that are *valuable, rare, inimitable*, and *non-substitutable*.

The above outlined requirements should be met to ensure that the network is not just one important element of the value chain but an actual source of competitive advantage. In this respect, the network is significant as a "tangible" asset, that is as physical *infrastructure*, but we believe that the role of human resources is far more important in gaining competitive advantage. Therefore, a network can only offer a real competitive advantage if the knowledge of the front-line staff and the organizational knowledge based on this represent a value that is unique and difficult for other companies to imitate. This requires a strong *corporate culture*, an appropriate *selection* and *training system*, efficient *incentive* and *communication systems*, and a management team that strives to lead and not just to manage the staff and the company. HR holds a crucial role in creating all these, but shaping corporate culture, designing the above systems, or management development can only take place by involving other corporate functions as well, and in accordance with their needs. HR can perform its function successfully if it defines itself as an internal service provider and adopts this approach in its operations.

Barber and Strack also stress the vital role of HR management when they introduce a new classification of companies and fields of business. They distinguish "people-based", "people-oriented" and "capital-intensive" businesses. "People-based" businesses have significant staff costs, the ratio of which to capital costs is relatively high, and have relatively low expenses related to future revenue generation. They include fields such as postal and express postal services, brokerage and financial services, hospital and hotel management. Barber and Strack claim that in "people-based businesses" HR management is not a simple supporting function but – due to its direct effect on the company's financial results – an essential, "core" business process. [1]

Experts in HR and in related fields of study have developed several comprehensive concepts on what tools, systems and practices can contribute to the "good quality" management of human resources. Companies that understood the need for improvements have made serious efforts to implement them. However, changes

* **Andrea Bencsik[1], Judit Fehér[2],**
[1]University of Veszprém, Egyetem u. 10., 8200 Veszprém, Hungary, E-mail: bencsik@gtk.vein.hu
[2]Hungarian Post Office Ltd, Hungary, E-mail: kolkazsu@axelero.hu

in the environment and in the needs of stakeholders – primarily of customers and employees – make it necessary to rethink the efficacy and efficiency of models and real-life practice, as well as the direction of potential improvements. That is why we believe that it is important to present a summary of what typical problems may arise in the management of front-line staff, and what kind of new expectations are presented to companies by changes in the environment.

2. Human resource management on the front line – factors contributing to HR-problems

A common classification of HR management tasks distinguishes administrative tasks from "substantive" ones. Administrative tasks include ones related to employees joining or leaving the organisation, transfers, supplying statistical data for payroll processing or for other purposes etc. Substantive activities include the non-administrative tasks of selection, labour retention, development and motivation. It is useful to separate the two fields from the point of view of quality management. Administrative tasks lend themselves well to standardisation and centralisation, whereas substantive HR activities include several tasks that may not be carried out in a centralised way at all, or only with less success.

Human resource management is an extremely complex field, in which specific activities – such as sourcing and retaining the right staff, motivating employees, effective staff development and communication – operate in close interaction, and, due to the intricate interrelationships between them, it is difficult to identify their individual impact and results. For reasons of clarity, however, it is necessary to separate the sub-systems on a theoretical level.

2.1. Workforce and quality

One of the key problems of organising customer service – due to its special features – is to manage fluctuations in demand and to determine optimum *staffing levels* correspondingly.

It is clear that competition requires companies to operate cost-effectively. Wages and staff costs are dominant in the cost structure of service companies, so it is easy to understand that companies wish to reduce them. However, too low staffing levels may prove a serious obstacle to providing customer services of the right quality. If the company tries to minimise its staffing levels purely on the basis of *economic considerations* and not according to the *expected service standard*, there will certainly be a decrease in the quality of services.

If the labour shortage is only temporary, peaks may be managed by involving back-office staff or by changing the division of labour temporarily. If, however, it becomes permanent, the company will likely face a decrease in quality. Staff may become frustrated, and, after a while, uninterested by endless queues and grumbling customers. Since only indirect methods are available for measuring quality, the company only learns about the quality problems from the decreasing number of customers.

Providing the *right quality staff* is even more important than having the right number of staff. This may be affected by the pool that is available to a given network unit, that is the number and quality of applicants available when the local management wants to recruit new staff. Front-line work is basically a routine activity in many cases, but it may involve great responsibility in financial or other terms. If this stress factor is not recognised by means of cash or non-cash rewards (in other words, employees are underpaid), front-line jobs will not necessarily be very attractive to jobseekers. The company may be forced to employ candidates who are not sufficiently prepared, which may result in growing staff turnover, and if it wishes to prevent the decrease in quality, it will have to increase the budget for staff development.

The weaknesses of the *system of selection* may also present a threat to maintaining the desirable quality in customer service. If there are no effective recruitment processes in place and there is a lack of relevant expertise, employees whose personality or moral standing is not appropriate, might be given front-line jobs. On the one hand, this may incur extra costs as the company has to reinforce its monitoring activities. On the other hand, it will definitely damage the image of the company as mistakes made by customer service staff members are difficult to rectify afterwards.

2.2. The measure of value – performance measurement and motivation

It is obvious that an essential requirement of "fair" *performance evaluation* is the establishment of the *measurability* of performance. The performance of the front line is basically the same as the quality of the interaction between the front-line staff and the customer. It is, however, difficult to present with exact indicators. For want of a better method, companies measure the outcome of the interaction, that is the output of the customer service process (customer numbers, sales figures etc.) and try to obtain information about the performance of the front line in this indirect way.

Undoubtedly, the advantage of using *indicators* of *customer numbers* or *sales revenues* (sales revenue per employee, number of customers, number of repeat customers etc.) is that they are relatively objective. However, when introducing and using such indicators, service companies with networks have to bear in mind the following:

- Target figures, expectations, etc., that is the target values forming the basis of comparison in performance evaluation, have to be properly differentiated (e.g. to allow for different sales potentials).
- All too often companies regard the number of services sold and the revenue from them as the sole measure of performance. If they assess performance and give incentives relying only on such indicators, employees will primarily concentrate on meeting these performance criteria, too. This may very well be favourable in terms of short-term financial indicators, but will not necessar-

ily contribute to improving quality in the long run. It is a fact, too, that customers have a negative attitude to an overly strong sales orientation.

Indicators related to *complaints* are often used when appraising the performance of front-line staff members. It was Kotler who pointed out the adverse effects of this practice. The utilisation of the number of complaints in this respect does not assist in achieving the long-term objectives, namely, the elimination of the sources of deficiencies in quality. (If this figure becomes an element of the employees' performance appraisal, they will go to all lengths to prevent customers from making even justified complaints. Consequently, companies will be unable to receive valuable comments from customers.)

The applicability of *customer satisfaction* as an indicator is also limited. One reason for this is that customers' value judgement of the overall service process, and their opinion of customer service, which is one part of this process, are difficult to separate. A case in point is the survey that has proved that if trains run on schedule, passengers tend to believe that carriages are cleaner, queues in front of railway ticket offices are shorter etc. [2]

Also, customer satisfaction surveys – especially when conducted by service companies operating in mass markets – are usually representative, therefore they are rarely useful in measuring individual performance exactly.

The *weight* that is given to specific *performance indicators* depends on the actual company's culture and its strategic goals, as well as on some practical conditions, such as the availability of IT infrastructure that collects and classifies data.

Incentive schemes can only serve their function if they are *effective, flexible* and able to *differentiate*. Ensuring flexibility presents a serious challenge mainly to large companies due to the complex nature of the operations of these organisations. No matter what change the company is planning to make, it has to realise that "everything is interrelated", thus a survey of aspects to be affected by the change should be carried out beforehand. Barber and Strack argue the importance of establishing pay differentials at "people-based companies" as the performance of individual employees or teams may differ significantly. At such companies, unlike at capital-based ones, the performance of lower ranking employees can have a direct effect on financial results, therefore performance-based differential pay should be used down to the lowest level of the organisation. [1]

Financial reward schemes may be used to recognise individual, team or company performance, or a combination of these. *Individual* reward systems may be applied successfully at companies where employees have an actual effect on their job outcomes. What may actually lead to conflicts in the case of front-line staff members is that they feel that factors like product quality, brand or work organisation influence performance to a much greater extent than their own aptitude and efforts, and the size of available commissions are not in line with employee input.

Team-based incentives can offer benefits in several respects. They may have a positive effect on the efficient division of labour between front-line and back-office workers, as well as on the co-operation between front-line staff members, if performance-based motivation is linked to e.g. team sales performance or other team-level turnover indicators.

The third type of *performance-based* incentive schemes is linked to *corporate performance*. This solution is often used by service companies, too, as there are exact measures for assessing the performance of the company. In this case the difficulty may lie in determining the extent of the contribution of a particular individual or certain groups to corporate-level performance.

If we examine the issue of *motivation* in broader terms, and look at job satisfaction and internal motivation, too, we can see that the development of these may be hindered by factors related to the job and to the "image" of the job. Front-line staff can pace or plan their work to a relatively small degree, since the company can offset the fluctuations in demand only to a certain extent. There are a lot of routine elements in their tasks, but, at the same time, they have to be prepared to be able to handle customer reactions. Yet, in the informal hierarchy of the company they often rank rather low, which is also reflected in their non-financial rewards.

2.3. Short- and/or long-term staff development

Service companies have to place great emphasis on developing *emotional competencies* when training front-line staff members. Developing such competencies effectively may require totally different training methods than developing professional competencies. The search for solutions has even led to the emergence of extreme trends like bringing religion into education.

In front-line staff development companies have to strike a balance between the requirements of long-term human resource management and short-term cost-efficiency. Although the importance of learning has increased, companies still try to cut staff training budgets on the basis of financial considerations. At the same time, *effective problem resolution*, what is more, "solution sales" are expected from employees by service providers. To be able to do this, it is not enough for employees to acquire professional knowledge related to e.g. a particular range of products. Employees need to understand and embrace the company's business policy, and they have to be able to make corresponding decisions in the course of transactions with customers. One- or two-day training courses that focus on preparing staff for a specific task do not create the basis for meeting this requirement, yet companies often settle for this solution for financial reasons.

Product innovation in the field of services does not provide a lasting competitive advantage. Services, especially those with low asset needs, may be copied fast, and the advantages that can be gained when launching the product dissolve quickly. Therefore, it is crucial for service companies with networks to introduce their new products in the entire network in the shortest time possible. To be able to do so, they have to set up a system of knowledge transfer that can provide *a large number of staff members* with the necessary knowledge *in a short period of time*.

4. Communication between "higher-ups" and "lower-downs"

Information flows in two directions in *internal communication:* from network units to corporate head office and its management, and from headquarters and management towards network units. Whatever direction communication follows, the two prerequisites of a proper flow of information are a corporate culture encouraging dialogue, and an information system supporting this culture. Role culture, which is often characteristic of large organisations, upholds one-way and mostly formalised communication, where the official way of doing things is adhered to and enforced. Communication between corporate headquarters and network units is rendered more difficult by geographical distances, since head office staff can have access to a lot of information indirectly, too.

When interpreting *information*, the *subculture* of a specific field is an important factor. In the case of one-way and formal flows of information, management does not get any feedback on how the receivers of the information transferred by management actually interpret it. This fact should be recognised especially by those companies that – due to their size – are inevitably "forced" to transfer information in writing, and frequently in formalised ways.

5. The expectations of external and internal customers

HR specialists have developed several new models and practices that chart the course towards "substantive" human resource management, and typically aim at supporting leaders or key personnel *individually*. Examples include selection processes comprising several rounds, personalised incentive schemes designed to truly meet individual needs, the continuous development of competencies, e.g. with the help of a personal coach, etc. The key role of human resources in customer service and the specific issues outlined above indicate that large service companies *should move towards substantive human resource management* in this particular employee segment, too. The need for this shift is reinforced by the following noteworthy trends:

- Research findings from McKinsey show that there is an increasing number of jobs in which most of the tasks involve complex problem resolution, and success hinges on employees' co-operation with co-workers and/or customers. Researchers have created three categories on the basis of what kind of interactions are predominant in particular jobs. Jobs where complex interactions are typical belong to the category of "tacit" jobs. "Transactional" jobs are characterised by routine interactions. The third, "transformational" category includes jobs in which employees are involved in industrial activities, such as the extraction or conversion of raw materials. According to the findings of the McKinsey survey there will be an increase in the proportion of "tacit" jobs in the future. That is why it is vital that companies make the necessary arrangements to develop employees holding down such jobs and to increase the efficacy of these employees' work. [3]
- It is a well-known fact that subjectivity cannot be eliminated from the quality assessment of services. Recent surveys have also indicated what an important role emotions play in value judgements. Research by Fleming et al. shows that "emotionally committed" customers feel more attached to and tend to remain loyal to a particular company, and use its services more often than those who are simply "generally satisfied". [4] This means that the emotional abilities of customer service employees become more valuable. On the other hand, it clearly confirms the fact that the work attitude of front-line staff greatly affects quality.
- Looking at the issue from the point of view of the labour market, it can be assumed that the increase in services, and within services primarily in personal services, will increase the demand for labour force, thus improving the bargaining power of service workers. Consequently, companies will have to make more efforts to retain "good quality" workforce.
- Not only external customers but internal customers, that is employees, have become more conscious, too. Employee expectations related to work and the workplace are changing. Though financial incentives are still dominant in most cases, employee satisfaction is affected by multiple complex criteria. Some researchers distinguish between satisfaction and commitment, and argue that it is not enough if employees are satisfied, they should be committed. A survey of ten companies by Fleming et al. showed that the emotional commitment of customers and that of employees reinforce each other at the local level. The financial results of those network units which have above median values – calculated from the entire sample – for customer and employee emotional commitment are three or four times higher than those produced by outlets which have below median values for both indicators. [4]

Companies have to respond to constantly changing employee needs. Drucker points out that „[i]ncreasingly, 'employees' have to be managed as 'partners' – and it is the definition of a partnership that all partners are equal". He argues that partners cannot be given orders, they have to be persuaded. Therefore, managing people is turning into a marketing activity. Marketing is about finding out what the other party wants, what their values and goals are and what they consider to be a result. [5] The success of adopting this approach in practice depends on how well the company can handle individual needs, in other words, what kind of customised solutions it can offer to employees. Satisfying employees' individual needs in the fields of training, motivation or job design is far more difficult and complex at companies that are particularly cost-sensitive and have a large number of employees than at small, knowledge-based companies rendering project-type services. Large service companies cannot avoid embracing the marketing orientation in their human resource management either, and have to try to introduce systems, tools and processes that support this developmental direction.

6. Marketing and economies of scale in HR

As Deming puts it, *quality*, which may also be regarded as the ability of creating value, is primarily a function of the *activities of corporate management*. Management sets the goals, develops the strategy, decides on allocating resources, establishes the system of operation etc. Deming differentiates two groups of factors affecting quality: general and special causes. It is the task of the management to handle general causes, to eliminate factors causing system failures, whereas employees may only be held accountable for factors under their direct supervision. [6]

If the above principles are applied to service companies with networks and their human resource management, it can be concluded that head office management, and within that HR, may cause system failures in two ways. Firstly, if corporate level systems – related to jobs, competencies, performance evaluation, motivation etc. – do not match strategy, or do not meet the needs of internal customers, e.g. the needs of specific arms of the business. Secondly, if management does not create the conditions that are needed to operate HR systems in line with objectives. The first condition to be mentioned is that a properly prepared local management team possessing the right professional and emotional competencies should be put in place. The reason for this is that the effectiveness of HR systems is very much dependent on how they are implemented in specific network units. *The management/managers of network units play the role of a bridge between the management of the company and the front line.* Corporate systems developed by HR specialists reach employees partly directly, partly through the management of network units. Local managers may "share the contents of a box" that has arrived from corporate management in various ways, in other words, they can transmit the information to subordinates in different ways. Additionally, they can also contribute to the success of HR systems by shaping the local subculture. [7]

Naturally, it is true for human resource management, too, that the feasibility of "customisation" is largely affected by the costs it incurs. Bearing this fact in mind, making progress may be possible by way of providing a "building blocs" service. What marketeers mean by building blocs is that the company does not render "ready-made" services, but develops service modules to satisfy the needs of well-identifiable customer groups. These modules can be mixed and matched with the active participation of the front line so that customers perceive the service rendered to them as unique. To apply this to human resource management: corporate-level, centralised systems should inherently contain "mass customisation", and local management, playing the role of the front line in this context, should be given the chance and the conditions to convert them into individual solutions.

Certain methods that can work cost-effectively due to the fact that they affect a large number of employees doing the same or similar jobs may be adopted when developing the above mentioned corporate-level systems.

Employee satisfaction surveys similar to customer satisfaction measurement could form the starting point for improvements. Anonymity encourages respondents to give honest opinions, and a sufficiently large sample could ensure the representative nature of the survey. The findings of such a survey could provide information for the differentiation of the offers to be made to employees, too.

One of the most important preconditions of providing customised solutions is extensive differentiation. Employees may be segmented according to their jobs, demographic features or, e.g. their motives. There have been reports on large-scale applications, such as the one Tesco, one of the world's leading international retailers, introduced. Several thousand employees were interviewed, or surveyed in others ways to map the company's workforce. Then employees were classified according to their aspirations, work methods, and even lifestyles, and "customised" incentive packages were offered to them accordingly. [8] Naturally, whether this method actually helps to satisfy individual needs depends on how successful the selection of relevant criteria for differentiation was. Cost-efficiency may be reached by differentiating between employee segments in terms of the degree to which the company fine-tunes solutions for them. Marketing considers establishing customer value as the basis of customer care. Customer value means the degree to which a particular customer or customer segment can contribute to the successful performance of the company in the long run. The value assigned to specific jobs, though, can be calculated according to factors such as the replacement costs of one employee, which should include the actual administrative and training costs as well as elements such as a loss of profits, which may arise from customer defection, or from a fall in revenues during the training period.

Large companies tend to or may actually be forced to apply standardised solutions and more centralisation due to their size, or, possibly, their technology and strategy. However, the selection of employees, the evaluation of their performance, as well as motivating them are based on a lot of information or even impressions that emerge locally and are very difficult to specify, therefore they may only be utilised in decisions made at the local level. Consequently, HR management can make its activities more effective if it trusts the *local management* and provides it with the right degree of *independence*.

It is essential, of course, to *prepare local management* for this task, and to develop tools that support execution. Supporting tools may include, e.g. *software for wages management, methodology for selection*, or, as a matter of fact, *a network of HR specialists* the local management can turn to for support when in need. Although developing or creating them may incur considerable costs in some cases, due to the large number of potential users *economies of scale* may be reached once they are used.

The success of human resource management thus largely depends on whether or not the *local management* has the authority to make the necessary *decisions*, and whether it is properly prepared to do so. A broader *scope of authority* makes it possible to use special information that is not quite available for "higher-ups", it reduces reaction time, and increases the sense of *responsibility* for decisions made, too. A prerequisite for providing a broader scope of authority is that those who are assigned such authority should have the right *skills, abilities*, in sum, the *professional, intellectual* and *emotional competencies*.

References

[1] BARBER, F., STRACK, R.: *The Surprising Economics of a „People Business"*, Harvard Business Review 2005, June.
[2] NEELY, A., ADAMS, C., KENNERLEY, M.: T*he Performance Prism – The Scorecard for Measuring and Managing Business Success*, Pearson Education Limited/2002
[3] BEARDSLEY, S., JOHNSON B., MANYIKA, J.: *Competitive advantage from better interactions*, The McKinsey Quarterly/2006/2
[4] FLEMING, J. H., COFFMANN, C., HATER, J. K.: *Manage Your Human Sigma,* Harvard Business Review 2005, July-Aug.
[5] DRUCKER, P. F.: *Management Challenges for the 21st Century,* Harper/Collins Publisher, NY, 1999.
[6] TENNER, A. D., DETORO, I. J: *TQM: three steps to continuous improvement,* Addison-Wesley Publishing Company, Inc, 1992.
[7] HUSELID, M. A., BEATTY, R. W., BECKER, B. E.: *A players or positions,* Harvard Business Review /2005 Dec.
[8] KANTOR, R.: *Human Resources in the 21st Century*, Wiley, 2003.

KOMUNIKÁCIE
COMMUNICATIONS

Milan Dj. Blagojevich *

TIME-TEMPERATURE CURVE DEFINITION ACCORDING TO FUEL TYPE

An idealized temperature curve of compartment fire has three, distinct phases: growth phase - development phase from ignition to flashover, steady-burning phase, and decay phase. The experimental results performed in recent years show that the standard temperature curve proposed by the International Standards Organization (ISO) as well as ASTM E119 time-temperature curve selected for a standard fire resistance test do not satisfactorily describe fire phases for different fuel types. In this paper we propose the definition of the time-temperature curve by means of a unique function in which the fuel type is defined with parameters.

Key words: temperature curve, adaptive control, software

1. Introduction and background

Generalization of the unique temperature-time fire curve is very complicated because many parameters as ignitability and combustibility of fire material, rate of heat release, flame spread, etc. are included. Further, shape of the temperature-time curve depends on parameters as open doors or windows, vent flows and others.

No single time-temperature curve can represent all fires because there are many ways in which the fire can develop after ignition. However, the most fire curves are similar in their growth phases, independently of fuel type. In Fig. 1 fire curves for different kinds of fuels and for different quantity of the same fuel (fuel: rubber) are shown.

An idealized temperature curve of compartment fire has three, distinct phases: growth phase – development phase from ignition to flashover, steady-burning (fully developed) phase, and decay phase. In most cases the front edge of a fire signal may be considered to be a reliable confirmation about fire growth. The growth phase has primary importance for obtaining a realistic prediction of detector and sprinkler activation, time to start evacuation, and time to initial exposure of occupants. Moreover, with suitable analytical function, which describes the growth phase, it is possible very fast – immediately after ignition, to predict the temperature development, its maximal value, and the time interval needed to achieve this value.

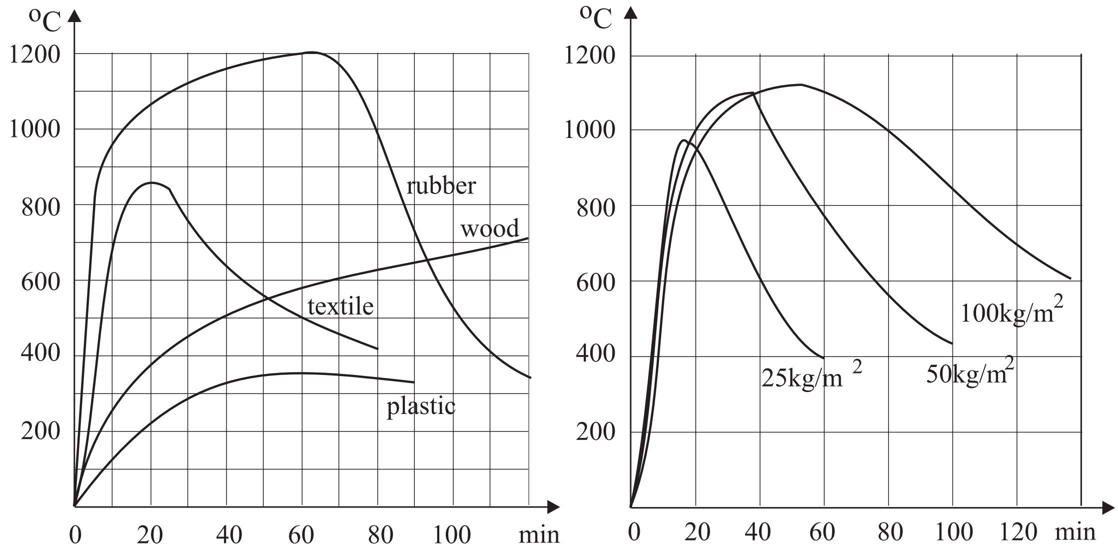


Fig. 1 Fire curves for different types of fuel (on the left), and fire curves for different quantities of the same fuel (on the right)

* **Milan Dj. Blagojevich**
Faculty of Occupational Safety, Carnojevica 10a, 18000 Nish, Yugoslavia, E-mail: milan@znrfak.znrfak.ni.ac.yu

In order to design an appropriate time-temperature curve it is necessary to know some facts about the growth, steady burning and decay phases of fire.

From the previous figures it is obvious that the growth phase of fire may be treated like pulse phenomena. This means that it is possible to approximate the fire growth phase with Heaviside function, with a linear combination of exponential functions or with some suitable function that satisfactory describes the pulse phenomena. The steady burning phase – the phase of fully developed fire may be at a sub-flashover or post-flashover level. This phase may originate with or without flashover. At the post-flashover level the steady burning phase is fuel or ventilation controlled. From this reason, the shape of the time-temperature curve will have characteristic peak before the decay phase – in a point of flashover, or slow decrease before the decay phase – in absence of flashover.

The decay phase is characterized with a fast decrease because a burning rate declines as the fuel is exhausted. In engineering applications and in absence of experimental data, this phase may be considered as the inverse of the growth phase. It is often assumed that the decay phase begins when 20% of fuel is left. With this assumption, the inverse shape of the decay phase from the growth phase is technically reasonable. The previous consideration is illustrated in Fig. 2.

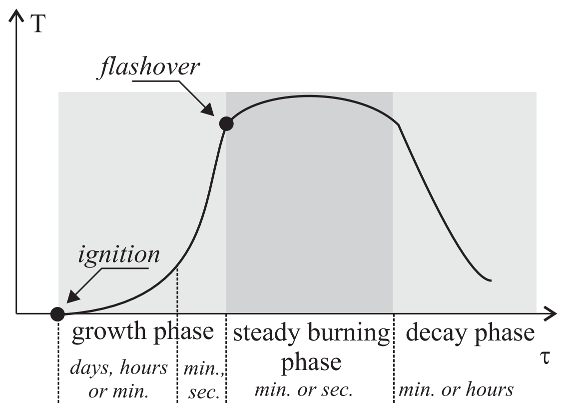


Fig. 2 Schematic presentation of time-temperature phases (an idealized fire curve)

2. Time-temperature curves in engineering applications

Physical modeling does not always involve major reductions in a physical scale. Physical models, in general, simplify complex phenomena in a manageable and understandable way. An example of this type of physical modeling is the well known ASTM E119 fire-resistance test. The ASTM E119 time-temperature curve shown in Fig. 4 was selected as a representative time-temperature curve for testing structural assemblies. The furnace environment is intended to model the fire environment experienced by structural members in a fire.

On the other hand, in many engineering applications the standard time-temperature relationship according to ISO 834 can be used:

$$T = \begin{cases} T_0 + 504\tau^{0.141} & \tau < 10 \text{ min} \\ T = T_0 + 345\log(8\tau + 1) & \tau \geq 10 \text{ min} \end{cases} \quad (1)$$

where: T_0 – temperature at the start of the fire, [°C]
T – temperature in the compartment, [°C], and
τ – time [min]

These representative time-temperature curves are shown in Fig. 3. and regardless many limitations observed over the past decades, these models are still widely used.

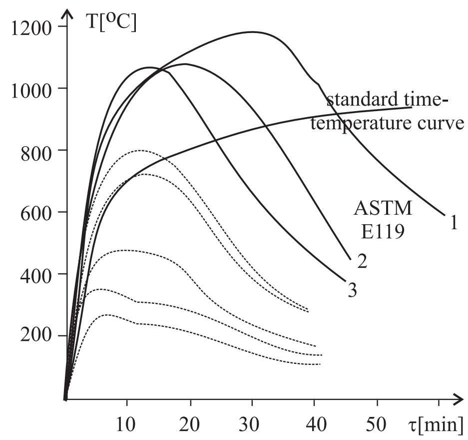


Fig. 3 ASTM E119 curves and standard time-temperature curve

The basic idea presented in this paper is to define the family of curves that satisfactorily and accurately describe the phases of fire. With this approach, each of the curves differs only in one parameter - according to fuel type.

3. The new function for time-temperature relationship

For approximation of the fire phases we suggest, as very simple and accurate, the following function:

$$T = T_{max}\left[\frac{\tau}{\tau_{max}} e^{(1-\frac{\tau}{\tau_{max}})}\right]^c \quad (2)$$

where: T – temperature [°C],
τ – time [s],
T_{max} – maximal temperature in steady burning phase [°C],
τ_{max} – time instance for maximal temperature [°C],
e – parameter that depends on fuel type.

In Fig. 4. the curves in accordance to the function defined with (2) for values $c > 1$ and $c < 1$ are shown. The parameter c that defines fuel type is, in fact, the rate of rise of the given function.

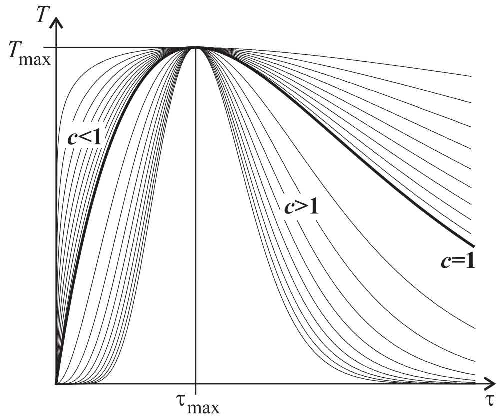


Fig. 4 Time-temperature curves

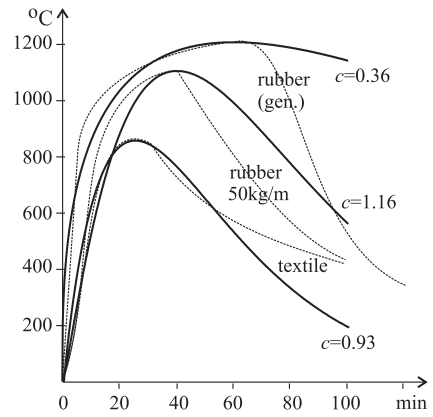


Fig. 5 Time-temperature and approximation curves

In order to prove the proper choice of function (2) for time-temperature, we must know the value of maximal temperature raised during the steady burning phase and the time instance for maximal temperature. In this way, it is necessary to find only a parameter *c*. On the other hand, it is obvious that the shape of curves shown in Fig. 4. can be used to describe a time-temperature process for the most fires, especially flaming ones (with characteristic peak). For smouldering fires, this function can be used only for the growth phase of fire – between ignition and flashover.

4. Using the new function

The function given with (2) can be used in the following way: if we know the maximal temperature of fire and the corresponding time coordinate, the parameter *c* can be calculated by adjusting the function to the accurate value in any point (τ_1, T_1) in the growth phase.),

The adjusting of function (2) in point (τ_1, T_1) gives:

$$c = \frac{ln\frac{T_1}{T_{max}}}{ln\frac{\tau_1}{\tau_{max}} + 1 - \frac{\tau_1}{\tau_{max}}} \qquad (3)$$

For numerical experiments and for illustration of this approach, some curves from Fig. 1. and Fig. 3. were chosen. The aim was to compare shapes of the experimental curves with curves obtained by approximation.

The results of numerical experiments are shown in Fig. 5. and in table 1.

The new function can also be applied on ASTM E119 curves, as shown in Fig. 6 and table 2.

Results of approximation with function (2) for different types of fuel (rubber and textile) — Table 1

Approximations for curves from Fig. 1.			
τ[s]	T [°C] rubber (gen.)	T [°C] rubber (50kg/m)	T [°C] (textile)
600.00	850.16	633.14	545.59
1200.00	1027.36	831.88	889.11
1800.00	1119.48	836.12	1055.26
2400.00	1169.27	753.09	1100.00
3000.00	1193.26	638.73	1067.92
3600.00	1200.00	521.54	991.17
4200.00	1194.61	414.82	891.82
	$c = 0.36$	$c = 0.93$	$c = 1.10$

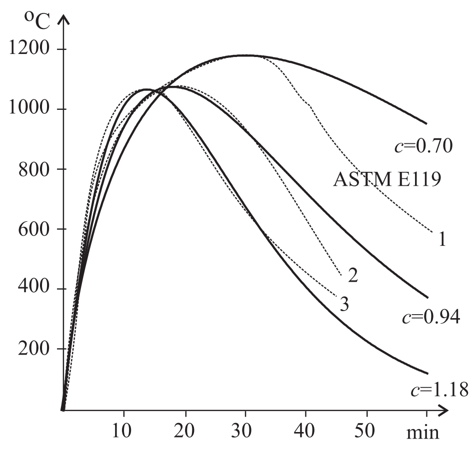


Fig. 6 Time-temperature and approximation curves

Results of approximation with function (2) for ASTM curves — Table 2

Approximations for ASTM curves from Fig. 6.			
τ[s]	T [°C] – 1	T [°C] – 2	T [°C] – 3
150.00	391.61	380.69	370.24
300.00	601.35	639.87	678.60
450.00	754.26	821.41	886.41
600.00	870.85	944.28	1007.93
750.00	960.83	1021.84	1062.24
900.00	1030.11	1064.28	1066.94
1050.00	1082.72	1079.60	1036.69
1200.00	1121.63	1074.20	983.15
1350.00	1149.14	1053.16	915.24
1500.00	1167.08	1020.60	839.65
1650.00	1176.95	979.77	761.24
1800.00	1180.00	933.28	683.45
1950.00	1177.27	883.22	608.59
2100.00	1169.66	831.21	538.15
2250.00	1157.92	778.52	473.01
	c = 0.70	c = 0.94	c = 1.18

5. Conclusion

Time-temperature curves having the shape similar to pulse phenomena are very convenient for approximation. For smouldering fires (without characteristic peak) this function can be used for approximation of the development phase of fire. In any way, this function provides significantly better results related to the already known standard time-temperature curves. The type of fuel can be defined by only one parameter or by couple of values of this parameter.

The method described in the above text is complex and, therefore, can be useful in reconstruction of fire. It is necessary to know most of the data before approximation and, consequently, this method may not be used in real time, during the fire data acquisition. Unfortunately, when all of the data are known, the fire is fully developed, and no prediction has any sense.

References

[1] ABRAMOWITZ, M., STENGUN, I.: *Handbook of mathematical functions,* Dover publications, New York, 1964.
[2] BLAGOJEVICH, M., PETKOVICH, D.: *Detecting fire in early stage – a new approach,* Facta Universitatis, Facta Universitatis, Series: Working and Living Environmental Protection, Vol. 2, No 1, 2001, pp. 19–26.
[3] JANSSENS, L. M.: *Heat release rate (HRR)*, FORUM Workshop of Measurement Needs for Fire Safety, NIST, USA, 2000.
[4] JONES, W.W, FORNEY, P. G.: *Comparison of CFAST prediction to real-scale-fire tests*, Proceedings of Fire Safety Conference on Performance Based Concepts, Switzerland, 1996.
[5] PITTS, M.W.: *Temperature measurements in fire,* Proceedings of FORUM Workshop on Measurement Needs for Fire Safety, NISTIR 6527, NIST, USA, 2000.

Jerzy M. Wolanin *

DOMAINS OF SAFETY

It is considered that providing safety in universal dimension is exclusively responsibility of a public authority and the authority is fully responsible for efficient implementation of systems concerned with it. In reality, there is no prospect of not having incidents and therefore, there is no absolute safety. If such ideas exist anywhere, they are only theoretical concepts. Safety is a condition of the surroundings and natural environment of a local community. The condition level is defined by the degree of existing risk in that community. The risk is ubiquitous and inevitable as it exists in nature and civilian space. So the local governing body must decide what level of risk it is able to accept.

Safety includes a wide range of undertakings be realized in order to provide security in all areas of concern in the state. Therefore, safety is the result of activities of different institutions, services, organizations, and also individual human being, its closest surroundings, local and central authorities.

Every day we are notify of tragic events which means that our life is filled with incidents dangerous for us. Terrorist attacks, natural disasters, technical damages, plane crashes, or road accidents constitute a background of our lives. Breaking one's leg on a slippery surface, thefts, attacks, drowning, or spine injuries after jumping to water in a wrong place, accidents at work or on holidays are just examples of those unwilling events that can happen for everyone. Flooding, damages caused by gales, terrorist attacks are undoubted too. The numerous above examples clearly show that unwilling events vary greatly in their causes, character, results, social, political and economic context as well as individual realm. But no matter how varied they are, there is one common feature that can be described as:

All unwilling events disturb development and existence of individuals, local communities, or even the whole societies and in extreme examples they even make their survival impossible.

This universal feature has crucial influence on perception of safety by most people. It is thought that safety has universal character and can be provided unconditionally, with no importance of what is being secured. Safety of the state is perceived as an "umbrella" that can cover everyone and everything but divided on smaller (local community) and the smallest (individual) parts. Thus, the concept of safety is understood as a lack of unwilling events and, in case of their appearance, causing unimportant results (safety equals lack of threats). But there is no such situation. In reality, there is no prospect of not having incidents and therefore, there is no absolute security. If such ideas exist anywhere, they are only theoretical concepts. Safety (an individual or a state), as well as all connected actions, are often faced with identical requirements, no matter who they concern and what the character of the unwilling event is. It is believed that providing safety in universal dimension is exclusively a responsibility of a public authority and the authority is fully responsible for efficient implementation of response systems.

The problem is worth closer consideration. Let's analyze a road accident where a pedestrian is a victim. He or she crossed a street when the red light was on, straight in front of a correctly driven car. In this case, the pedestrian is found guilty. The red light did not protect him against the accident. The question is if the authorities responsible for safety on roads had done everything in order to provide it. The answer is yes as the red light is a sufficient means of providing safety when crossing a street. Even though, the accident happened. Decision to cross the street was made by the pedestrian himself or herself with no respect to the principles. This is an important moment in our considerations. The pedestrian took on himself or herself the risk of collision with no regard to safety measures employed, like traffic lights, which in this case had a supportive character. Thus, people responsible for safety on roads can only support it and their responsibility does not have and cannot have absolute character. Here, full responsibility is on the pedestrian. That is how it is. Such individual decisions, connected with our own safety, are numerous and examples can be many. On the one hand, there are individual decisions that may cause threat and protection against it and on the other, those decisions can be supported but support is a key word here. There are areas, or better to say, an individual safety domains to which external access is denied. Generalizing the above considerations a following rule can be formulated:

An individual area of a man's actions that may cause a threat or protect against it and that is dependent only on his/her free decision constitutes an individual safety domain (ISD).

An important feature of an individual safety domain is lack of outside influence including external safety systems which play only supportive role in decision making process. Support, in case of the individual safety domain, brings the best effects by promoting awareness of threats, i.e. by education. The education can be included into activities connected with safety system functioning. Long term

* **Jerzy M. Wolanin**
The Main School of Fire Service, Slowackiego 52/54, 01-629Warsaw, Poland, E-mail: Jerzy.Wolanin@sgsp.edu.pl

investment in education on safety is the most effective form of supporting safety in the case of ISD.

It is worth considering whether there are other domains of safety. It is known that exposure to threats is not often connected with individual decisions only. In case of sickness or other misfortune we receive help from our family. Building up a surrounding with our neighbour we constitute an element of safety system, together with other functions that the surrounding plays. One can rely on their help. Many decisions, connected with the possibility of threat occurrence, are made within this micro – community (family or neighbourhood). Those decisions, as in a previously mentioned example, are of supportive character and do not provide absolute safety. In this case the following definition can be formulated:

An area of unconstrained activity within a given micro – community capable of creating a threat or protecting against it constitutes a micro – community safety domain (MSD).

This domain has also limited access for external safety systems. Although, in some cases we can have direct influence on it. The possibility of a house being robbed is one of them. In order to prevent it security is provided. The possibility of hiring security teams is an element of safety system that exists outside of MSD. On the other hand, police force is responsible for fighting against this type of threats, acting within safety system organized by a state. Members of micro – community have a choice then, either to relay on police protection or provide individual protection for a house by hiring security. In the second case, the risk of being robbed is much lower. Here, a phenomenon of micro – community safety domain expansion into a state safety domain of a (SSD) occurs. From the above considered example results a fundamental feature of all domains: safety domains do not have strict boundaries. We can change them, i.e. expand or shrink them according to efficiency of the other domains. Hiring security teams has a number of advantages. It increases safety of micro – community members at the same time decreasing state budget expenses. For instance, police forces do not have to dispatch patrols into protected areas too often or even at all. In this case, transformation expenses of micro – community safety domain (MSD) into state safety domain (SSD) are covered by a micro – community.

A domain of safety that covers a wider concept than the domain of micro – community safety is the local safety domain, which can be defined as follows:

An area of unconstrained activity within a given local community capable of creating a threat or preventing it constitutes a domain of local safety (LSD).

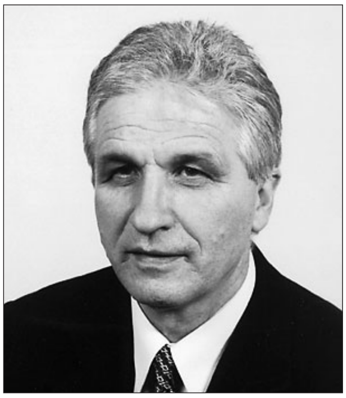
It is a highly interesting domain (LSD) that influence all other... within this one. Different concepts i.e. theoretical, political, social, organizational, etc. collide here with the problem of other domains influence on LSD. Thus, some of its constant and interesting features shall be described here. Historical experience of democratic countries shows that this domain is the basic one influencing in real safety in a state as well as citizens' feeling of safety. The stronger external connections of a local community are the greatest expansion of LSD on a state safety domain and the more effective protection against local threats. It is connected with better self-organization of this community, which is similar to previously mentioned self – organization of a micro – community. It influences its greater activity concerning risk management on a given area and consequently decreases in number of claims. This in turn leads to lowering public money expenditure. An important role in self – organization of a local community and building expansive LSD is education of society. Replacing claims with an active attitude allows not only for preventing unwilling events but it contributes to support state services during a catastrophe. On a territory of local communities with strong domains of local safety there are no spectators observing rescuers from outside. Also, inhabitants do not expect rescuers to wash their potatoes of mud as a part of flood results removal activities. These simple examples show the importance of citizens' awareness concerning threats in effective rescue actions.

There is no doubt that facing some threats, their scale and results they may cause, even the most developed local safety domain is not sufficient. Then, a state safety domain (SSD) marks its influence. This domain is generally based on central safety management system. Its role is to support other domains. As it was shown previously, some domains are "impermeable" for a state safety domain, especially of an individual safety domain.

An area of ruler decisions of state authority organs in shaping state's safety policy and their activities providing safety constitutes essence of the domain of a state safety.

As it was mentioned above, providing safety to citizens has no absolute character. According to a domain kind, the influence of the state safety domain is varied. Starting with lack of direct influence on safety area in case of an individual safety domain to complete state of providing safety in case of external and internal threats.

Safety then is not a universal and homogenous being. Its character is rather of varied and grained structure whose basic elements are domains. Domains differ from one another by the scale of interaction but also a level of autonomy in respect to one another. Subsequent domains, starting with the individual one, are more advanced complexity. That, in turn, requires different tools used in safety management. Domains are natural reflection of safety nature and all protecting systems should be connected with it and based on it. The best results are achieved by building systems according to "upwards" rule. However, every higher level should support the lower one. With regard to possible conflicts of legal and authoritative character, those elements of safety system that define boundaries of every domain should be precisely defined. Taking into consideration autonomy of domains as well as freedom to start activities within each of them, providing complete safety is impossible in its nature.



We are saddened by the death of our friend and colleague Ján Čorej, Professor of the University of Žilina, who died on 25th June 2006 at the age of 63 years.

Prof. Čorej was born on 27th August, 1942 in Remeniny near Prešov. In the years 1956 – 1960 he studied at a secondary school of civil engineering in Prešov. He entered the College of Railway Transport in Žilina in 1960 and graduated with a civil engineering degree in 1965. For a short period he worked as civil engineer in road construction companies in Košice and Banská Bystrica. Back at the university in 1969 he occupied various posts until he was appointed university professor in the field of Theory and Construction of Engineering Buildings. He went through various posts from head of the ground communication section, deputy of the department, chair of the Department of Road Civil Engineering, vice dean of the Faculty of Civil Engineering of the University of Žilina and vice Rector for science and research in the years 1996 – 2002.

He spent some time studying abroad, for example, at universities in Dresden, Köln, Delft, Amsterdam, Moscow, Kiev, Vienna, and Hanover.

In his research and scientific activities, Prof. Ing. Ján Čorej, CSc. focused on the area of road buildings, particularly on pavement structures and on their water and temperature regimes. He actively participated in solution of a wide range of projects both at national and international levels. He was the coordinator of an international project of Regional Development and Logistic Park within the frame of the EU INTERREG III C program. Professor Čorej published results of his scientific research in many domestic and foreign journals, in his monograph The Water and Temperature Regime of the Pavement and Sleeper Subgrade as well as in the textbook on Mechanics of the Pavement. He is author of numerous publications and textbooks dealing with road construction, mechanics of the pavement and urban roads. He significantly contributed to the building of a road laboratory of the Department, helped both to create a new field of study devoted to road construction and to set up the Highway Engineering Department.

Prof. Ing. Ján Čorej, CSc. headed the expert board responsible for road designing in the Slovak Road Association, was member of two editorial boards, member of various university and faculty scientific boards.

We have lost a dedicated teacher, researcher, colleague who considerably contributed to the University community.

Ján Čelko

COMMUNICATIONS – Scientific Letters of the University of Žilina Writer's Guidelines

1. Submissions for publication must be unpublished and not be a multiple submission.
2. Manuscripts written **in English language** must include abstract also written in English. The submission should not exceed 7 pages (format A4, Times Roman size 12). The **abstract** should not exceed 10 lines.
3. Submissions should be sent: **by e-mail** (as attachment in system Microsoft WORD) to one of the following addresses: *holesa@nic.utc.sk* or *vrablova@nic.utc.sk* or *polednak@fsi.utc.sk* **with a hard copy** (to be assessed by the editorial board) **or on a 3.5″ diskette** with a hard copy to the following address: Zilinska univerzita, OVaV, Univerzitná 1, SK-10 26 Zilina, Slovakia.
4. Abbreviations, which are not common, must be used in full when mentioned for the first time.
5. Figures, graphs and diagrams, if not processed by Microsoft WORD, must be sent in electronic form (as GIF, JPG, TIFF, BMP files) or drawn in contrast on white paper, one copy enclosed. Photographs for publication must be either contrastive or on a slide.
6. References are to be marked either in the text or as footnotes numbered respectively. Numbers must be in square brackets. The list of references should follow the paper (according to **ISO 690**).
7. The author's exact **mailing address of the organisation where the author works, full names, e-mail address or fax or telephone number,** must be enclosed.
8. The editorial board will assess the submission in its following session. In the case that the article is accepted for future volumes, the board submits the manuscript to the editors for review and language correction. After reviewing and incorporating the editor's remarks, the final draft (before printing) will be sent to authors for final review and adjustment.
9. The deadlines for submissions are as follows: September 30, December 31, March 31 and June 30.
10. Topics for the next issues: 4/2006 – Fatigue and endurance of constructional materials, 1/2007 – Mechatronics in electrotechnical systems.

POKYNY PRE AUTOROV PRÍSPEVKOV DO ČASOPISU KOMUNIKÁCIE – vedecké listy Žilinskej univerzity

1. Redakcia prijíma iba príspevky doteraz nepublikované alebo inde nezaslané na uverejnenie.
2. Rukopis musí byť **v jazyku anglickom**. Príspevok by nemal prekročiť 7 strán (formát A4, písmo Times Roman 12 bodové). K článku dodá autor **resumé** v rozsahu maximálne 10 riadkov (v anglickom jazyku).
3. Príspevok prosíme poslať: **e-mailom**, ako prílohu spracovanú v aplikácii Microsoft WORD, na adresu: *holesa@nic.utc.sk* alebo *polednak@fsi.utc.sk* príp. *vrablova@nic.utc.sk* (alebo doručiť na diskete 3,5″) **a jeden výtlačok** článku na adresu Žilinská univerzita, OVaV, Univerzitná1, 010 26 Žilina.
4. Skratky, ktoré nie sú bežné, je nutné pri ich prvom použití rozpísať v plnom znení.
5. Obrázky, grafy a schémy, pokiaľ nie sú spracované v Microsoft WORD, je potrebné doručiť buď v digitálnej forme (ako GIF, JPG, TIFF, BMP súbory), prípadne nakresliť kontrastne na bielom papieri a predložiť v jednom exemplári. Pri požiadavke na uverejnenie fotografie priložiť ako podklad kontrastnú fotografiu alebo diapozitív.
6. Odvolania na literatúru sa označujú v texte alebo v poznámkach pod čiarou príslušným poradovým číslom v hranatej zátvorke. **Zoznam použitej literatúry** je uvedený za príspevkom. Citovanie literatúry musí byť **podľa STN 01 0197 (ISO 690)** „Bibliografické odkazy".
7. K rukopisu treba pripojiť **plné meno a priezvisko autora a adresu inštitúcie v ktorej pracuje, e-mail adresu** alebo číslo telefónu event. faxu.
8. Príspevok posúdi redakčná rada na svojom najbližšom zasadnutí a v prípade jeho zaradenia do niektorého z budúcich čísiel podrobí rukopis recenziám a jazykovej korektúre. Pred tlačou bude poslaný autorovi na definitívnu kontrolu.
9. Termíny na dodanie príspevkov do čísiel v roku sú: 30. september, 31. december, 31. marec a 30. jún.
10. Nosné témy ďalších čísiel: 4/2006 – Únava a životnosť konštrukčných materiálov, 1/2007 – Mechatronika v elektrotechnických systémoch.

KOMUNIKÁCIE
COMMUNICATIONS

VEDECKÉ LISTY ŽILINSKEJ UNIVERZITY
SCIENTIFIC LETTERS OF THE UNIVERSITY OF ŽILINA
7. ROČNÍK – VOLUME 7

Šéfredaktor – Editor-in-chief:
Prof. Ing. Pavel Poledňák, PhD.

Adresa redakcie:
Address of the editorial office:
Žilinská univerzita
Oddelenie pre vedu a výskum
Office for Science and Research
Univerzitná 1, Slovakia
SK 010 26 Žilina
E-mail: komunikacie@nic.utc.sk, polednak@fsi.utc.sk,

Každý článok bol oponovaný dvoma oponentmi.
Each paper was reviewed by two reviewers.

Časopis je excerpovaný v Compendexe
Journal is excerpted in Compendex

Vydáva Žilinská univerzita
v EDIS – vydavateľstve ŽU
J. M. Hurbana 15, 010 26 Žilina
pod registračným číslom 1989/98
ISSN 1335-4205

It is published by the University of Žilina in
EDIS - Publishing Institution of Žilina University
Registered No: 1989/98
ISSN 1335-4205

Objednávky na predplatné prijíma redakcia
Vychádza štvrťročne
Ročné predplatné na rok 2006 je 500,– Sk

Order forms should be returned to the editorial office
Published quarterly
The subscription rate for year 2006 is 500 SKK

Jednotlivé čísla časopisu sú uverejnené tiež na:
http://www.utc.sk/komunikacie
Single issues of the journal can be found on:
http://www.utc.sk/komunikacie