ISSN 1407-5806

COMPUTER MODELLING AND NEW TECHNOLOGIES

Volume 11
No 4

2007

Transporta un sakaru institūts
(Transport and Telecommunication Institute)

Computer Modelling and New Technologies

Volume 11, No.4 – 2007

ISSN 1407-5806
ISSN 1407-5814
(On-line: www.tsi.lv)

Riga – 2007

THE JOURNAL IS DESIGNED FOR PUBLISHING PAPERS CONCERNING THE FOLLOWING FIELDS OF RESEARCH:

- mathematical and computer modelling
- mathematical methods in natural and engineering sciences
- physical and technical sciences
- computer sciences and technologies
- semiconductor electronics and semiconductor technologies
- aviation and aerospace technologies
- electronics and telecommunication
- navigation and radar systems
- telematics and information technologies
- transport and logistics
- economics and management
- social sciences

Articles can be presented in journal in English. All articles are reviewed.

EDITORIAL CORRESPONDENCE
Transporta un sakaru institūts (Transport and Telecommunication Institute)
Lomonosova iela 1, LV-1019, Riga, Latvia. Phone: (+371)7100593. Fax: (+371)7100535.
E-mail: journal@tsi.lv, http:// www.tsi.lv

COMPUTER MODELLING AND NEW TECHNOLOGIES, 2007, Vol. 11, No 4
ISSN 1407-5806, **ISSN** 1407-5814 (on-line: www.tsi.lv)
Scientific and research journal of Transport and Telecommunication Institute (Riga, Latvia)
The journal is being published since 1996.

CONTENTS

Editors' Remarks

RICHES

The countless gold of a merry heart,
The rubies and pearls of a loving eye,
The indolent never can bring to the mart,
Nor the secret hoard up in his treasury.

From **'The Rosetti Manuscript'**,

William Blake (1757-1827)

This 11th volume No.4 points out an attention to some **applied statistics** problems, **computer modelling** problems, which are really topical for this day. We also continue our activities in the field of **solid state physics** problems. In particular, we present an actual paper on nanotechnology and **T**unable **E**lectronically **A**nisotropic **M**aterials on **S**emiconductor (TEAMS) structures.

This means that our journal policy is directed on the fundamental and applied sciences researches, which is the basement of a full-scale modelling in practice.

This edition is the continuation of our publishing activities. We hope our journal will be interesting for researchers' community, and we are open for collaboration both in research and publishing.

EDITORS

Yu.N. Shunin

I.V. Kabashkin

Computer Modelling and New Technologies, 2007, Vol.11, No.4, 7–14
Transport and Telecommunication Institute, Lomonosov 1, LV-1019, Riga, Latvia

RESOURCE ALLOCATION MODEL FOR CUSTOMERS WITH DIFFERENT PRIORITIES

D. Golenko-Ginzburg [1], A. Ben-Yair [2*], Y. Hadad [3], N. Badarna [4]

[1]*Department of Industrial Engineering and Management, Ben-Gurion University of the Negev,*
P.O. Box 653, Beer-Sheva, 84105, Israel, and
Department of Industrial Engineering and Management, Ariel University Centre of Samaria,
Ariel, 44837
E-mail: dimitri@bgu.ac.il

[2,3,4]*Department of Industrial Engineering and Management, Sami Shamoon College of Engineering,*
P.O. Box 950, Beer-Sheva, 84100, Israel
E-mail: [2]avnerbqnet.net@beze, [3]yossi@sce.ac.il, [4]nabeel8@sce.ac.il

A resource delivery system which supplies several customers by homogenous non-renewable resources is considered. For each customer its minimal and maximal amounts of resources to be delivered, are pre-determined, as well as its preference (importance) rate. A real-time model which at any decision-making moment reallocates optimally the total available amount of resources among the customers, is presented. The values to be optimised are the resource amounts actually delivered to each customer, while the objective to be maximized is the summarized product of the delivered resource amounts and the customers' rates. A numerical example is presented.

Keywords: *resource reallocation, homogenous consumable resources, resource delivery model, customer's preference rate, precise resource reallocation algorithm*

1. Introduction

Problems of resource reallocation in organizational systems, e.g. in Project Management, Reliability Engineering, various service systems, etc., are nowadays truly urgent. Multi-level organization systems [2] are one of the most essential components in modern technical progress and strategic management [1–2]. Such systems are usually monitored by companies with multi-level structure. At the upper level company managers transfer to lower levels various goal parameters. The latter actually determine and implement all the planning, control and scheduling procedures for the corresponding system's elements to reach their goals on time.

In order to process the elements they have to be supplied with resources. Since the total amount of the company's available resources is usually restricted, resource allocation among the system's elements has to be carried out.

It can be well-recognized from analysing various organization systems, e.g. building systems [1, 4], production systems [3, 5], that to-day proper resource reallocation techniques among the system's elements are not implemented both in on-line control or scheduling procedures. Each contractor takes all measures in order to refine his own element's parameters, independently on other company's units. Such actions, being useful for a single unit, may result in heavy financial losses for the company as a whole, since a unification of local optimums may be very far away from the global one.

The goal of our paper is to consider a resource reallocation model for a supply system with several customers of different preference (significance). The values to be optimised are the amounts of resources to be delivered to each customer, while the objective to be maximized is the summarized product of actually delivered resources to each customer and their corresponding preference rates. A numerical example is presented.

* corresponding author

2. The Delivery System's Description

The system under consideration comprises a storehouse and several supply lines which deliver consumable resources to *n* customers C_i, $1 \le i \le n$ (production units). Each unit consumes non-renewable homogenous resources (fuel, grain, steel, cement, etc.) which are similar for all units. All customers C_i place beforehand an order with the delivery system on resource supplement. An order comprises usually two values:

$V_{i\ \min}$ - the minimal amount of resources to be obtained – otherwise production process cannot be launched, and

$V_{i\ \max}$ - the maximal desirable amount of resources in order to reach the unit's goal as soon as possible. If the amount of resources exceeds $V_{i\ \max}$ it becomes redundant.

Assume that the system's resources are restricted, i.e., the storehouse contains less resources than the sum of the customers desirable amounts, but exceeds the sum of the minimal amounts. Thus, the amount of resources V_i to be delivered to each customer C_i must not exceed $V_{i\ \max}$ but has to be not less than $V_{i\ \min}$. The company attributes (usually by means of experts) to each customer C_i its preference rate η_i which designates the level of the customer's importance. The company has to take all measures to deliver as much resources as possible to customers with high preference rates. Only afterwards the company supplies other production units. The problem under consideration is to reallocate optimally the company's available resource amount among the customers. The objective is the summarized product of the delivered resource amounts and the customers' preference rates $I = \sum_i V_i \eta_i$. The optimised variables V_i^{opt}, $1 \le i \le n$, are the delivered resource amounts to each customer.

Note that in the course of planning horizon the customers' preference rates may undergo certain changes. The same goes for the company's available resource amount *V*. The first decision-making to allocate available resources has to be undertaken at $t_1 = 0$, i.e., before the units' production process starts. Later on, within the planning horizon, other decision-making moments t_s, $s = 2,3,...$, may occur. Thus, it is reasonable to substitute the terms V, V_i, $1 \le i \le n$, by V_t, V_{it} and V_{it}^{opt}.

3. Notation

Let us introduce the following terms:

C_i - the *i*-th customer, $1 \le i \le n$;

n - the number of customers;

η_i - importance rate attributed to customer C_i;

$V_{i\ \min}$ - the minimal permissible value of the resource amount to be delivered to C_i;

$V_{i\ \max}$ - the maximal possible value of the resource amount to be supplied for C_i;

V_t - the total available and non-delivered amount of resources at moment $t \ge 0$;

V_{it} - the determined amount of resources to be delivered to C_i (decision-making value to be optimised at any decision-making moment $t \ge 0$).

Assume that both relations

$$V_t \ge \sum_{i=1}^{n} V_{i\ \min}\ , \tag{1}$$

$$V_t \le \sum_{i=1}^{n} V_{i\ \max}\ , \tag{2}$$

hold. If (1) does not hold the problem under consideration has obviously no solution. In case (2) does not hold the problem obtains a trivial solution $V_{it} = V_{i\ \max}$.

4. The Problem

Given:

- decision-making moment $t \ge 0$;
- value V_t;
- values $V_{i\ \min}$ and $V_{i\ \max}$, $1 \le i \le n$;
- preference rates η_i, $1 \le i \le n$,

the problem is to determine optimal values $V_{i\,t}^{opt}$ to maximize the objective

$$I = \sum_{i=1}^{n} \eta_i V_{i\,t} \tag{3}$$

subject to (1–2) and

$$\sum_{i=1}^{n} V_{i\,t} = V_t \,. \tag{4}$$

Problem (1–4) is a resource reallocation optimisation model which refers to a broad spectrum of production planning and control problems (see, e.g. [4–5]. Problem (1–4) can be solved by means of linear programming [6]; in this study we present a simpler step-wise algorithm, which delivers the precise solution.

5. The Algorithm

Step 1. Set $V_{i\,t}^{*} = V_{i\ \min}$, $1 \le i \le n$.

Step 2. Determine the available resource reallocation volume $R_t = V_t - \sum_{i=1}^{n} V_{i\,t}^{*}$. Due to (1), value R_t cannot be negative.

Step 3. Reorder the sequence of values $\{\eta_i\}$ in decreasing order. Transform the sequence of ordinal numbers $\{i\}$ to a new one, $\{\psi_j\}$, $\{\psi_j\} \in \overline{1, n}$, where value η_{ψ_j} denotes the j-th value $\{\eta_i\}$ in the reordered sequence; thus, for any $\psi_1 \ne \psi_2$, $\psi_1 > \psi_2$ relation $\eta_{\psi_1} > \eta_{\psi_1}$ holds.

Step 4. Set counter $k = 1$.

Step 5. Calculate value

$$\Delta_k = \min\{R_t, V_{\psi_k\ \max} - V_{\psi_k\ \min}\}. \tag{5}$$

Step 6. Determine

$$V_{\psi_k t}^{opt} = V_{\psi_k\ \min} + \Delta_k \,. \tag{6}$$

Step 7. Update value R_t, $R_t - \Delta_k \Rightarrow R_t$.

Step 8. Check inequality $R_t > 0$, if $R_t > 0$ holds, apply the next *Step*. Otherwise go to *Step 11*.

Step 9. Counter k works, $k + 1 \Rightarrow k$.

Step 10. If $k > n$, apply the next *Step*. Otherwise go to *Step 5*.

Step 11. Determine optimal values $V_{\psi_k t}^{opt}$, $1 \le k \le n$, for all indices k which have changed their values at *Step 6*. For all other indices k their optimal values are set to their minimal values $V_{\psi_k \min}$ (see *Step 1*).

It can be well-recognized that the algorithm terminates either:

- in case when the available resource reallocation volume R_t is exhausted (see *Step 8*), or
- all indices k, e.g. customers C_{ψ_k}, $1 \le k \le n$, are examined.

Thus, the general idea of the algorithm centres on delivering to the most important customer as much resources as possible, on the basis of the up-dated remainder R_t. Afterwards the second, less important customer is examined, etc., until the algorithm terminates.

6. Optimality Investigation

Theorem. The algorithm provides an optimal solution of problem (1–4).

Proof. Assume that the algorithm under consideration does not provide an optimal solution, i.e. versus values $\{V_{it}\}$ obtained by (5–6), there exists another set of values $\{V_{it}^*\}$ (obtained by what we will henceforth call Algorithm 2) satisfying

$$I_1 = \sum_{i=1}^{n} \left(V_{it} \cdot \eta_i\right) \; < \; I_2 = \sum_{i=1}^{n} \left(V_{it}^* \cdot \eta_i\right). \tag{7}$$

Choose the first value $V_{\xi t}$ which, beginning from V_{1t} in descending order of values η_i, differs from its corresponding value $V_{\xi t}^*$. Because of

$$V_{\xi t} = V_{\xi \min} + \min\left\{R_t, V_{\xi \max} - V_{\xi \min}\right\}. \tag{8}$$

value $V_{\xi t}^*$ cannot exceed $V_{\xi t}$, since value R_t remains the same for both algorithms and $V_{\xi t}^* > V_{\xi t}$ results in either $V_{\xi t}^* > V_{\xi \max}$, or $V_{\xi t}^* > R_t$. Both relations do not hold because of (2) and since no solution of problem (1–4) can ever exceed the overall available amount of resources. Thus, $V_{\xi t} > V_{\xi t}^*$. Since both solutions $\{V_{it}\}$ and $\{V_{it}^*\}$ satisfy (4), an obvious relation

$$\sum_{j=1}^{n-\xi} \left[V_{\xi+j,t}^*\right] - \sum_{j=1}^{n-\xi} \left[V_{\xi+j,t}\right] = \Delta_\xi = V_{\xi t} - V_{\xi t}^* \tag{9}$$

holds. Since, due to the decreasing property of sequence $\{\eta_i\}$,

$$\eta_{\xi+1} \cdot \left\{\sum_{j=1}^{n-\xi} V_{\xi+j,t}^*\right\} > \sum_{j=1}^{n-\xi} \left(\eta_{j+\xi} \cdot V_{\xi+j,t}^*\right) \tag{10}$$

and

$$\eta_{\xi+1} \cdot \left\{\sum_{j=1}^{n-\xi} V_{\xi+j,t}\right\} > \sum_{j=1}^{n-\xi} \left(\eta_{j+\xi} \cdot V_{\xi+j,t}\right), \tag{11}$$

we obtain

$$\sum_{j=1}^{n-\xi} \eta_{j+\xi} \cdot V_{j+\xi,t}^* - \sum_{j=1}^{n-\xi} \left(\eta_{j+\xi} \cdot V_{\xi+j,t}^*\right) < \eta_{\xi+1} \cdot \left\{\sum_{j=1}^{n-\xi} V_{j+\xi,t}^*\right\} - \eta_{\xi+1} \cdot \left\{\sum_{j=1}^{n-\xi} V_{j+\xi,t}\right\} = \eta_{\xi+1} \cdot \Delta_\xi \,. \tag{12}$$

Since ξ was chosen to be the first differing number for $V_{\xi t}$ and $V_{\xi t}^*$, relation

$$\sum_{k=1}^{\xi-1} \eta_k \cdot V_{k\,t} = \sum_{k=1}^{\xi-1} \eta_k \cdot V_{k\,t}^* \tag{13}$$

holds.

Summarizing the left and the right parts of equation (13), we obtain

$$\eta_\xi \cdot V_{\xi\,t} = \eta_\xi \cdot V_{\xi\,t}^* + \eta_\xi \cdot \Delta_\xi \tag{14}$$

and thus

$$\sum_{j=1}^{n-\xi} \left(\eta_{j+\xi} \cdot V_{j+\xi,t}\right) + \eta_{\xi+1} \cdot \Delta_\xi > \sum_{j=1}^{n-\xi} \left(\eta_{j+\xi} \cdot V_{j+\xi,t}^*\right). \tag{15}$$

This, in turn, causes

$$\sum_{i=1}^{n} \left(\eta_i \cdot V_{i\,t}\right) + \eta_{\xi+1} \cdot \Delta_\xi - \eta_\xi \cdot \Delta_\xi > \sum_{j=1}^{n-\xi} \left(\eta_{j+\xi} \cdot V_{j+\xi,t}^*\right). \tag{16}$$

Since $\eta_{\xi+1} < \eta_\xi$ and $\Delta_\xi > 0$,

$$\sum_{i=1}^{n} \left(\eta_i \cdot V_{i\,t}\right) > \sum_{i=1}^{n} \left(\eta_i \cdot V_{i\,t}^*\right), \tag{17}$$

which contradicts our main assumption. Thus, the algorithm outlined in Section 5 provides indeed a precisely optimal solution.

7. Extension of the Algorithm

The developed algorithm for solving problem (1–4) may be extended to a real-time algorithm. Three different decision moments may occur:

1. Due to various economical or other influences, the customers' preference rates $\{\eta_i\}$ may undergo at moment t_2 certain changes. In this case values $\{\eta_i\}$ have to be substituted by new ones, and we must apply *Step 1* of the algorithm again. New optimal values $V_{i\,t_2}^{opt}$ will be determined, and implemented in the resource delivery process.
2. At a certain moment t_2 one of the customers is fully supplied with the amount of resources which has been determined at moment $t_1 = 0$ by means of the algorithm. Let this customer's ordinal number be γ.

<u>Corollary</u>. *Resolving the problem of resource reallocation among the remaining customers on the basis of the up-dated information, will not result in any changes for the remaining values* $V_{i\,t_2}^{opt}$, $1 \le i \le n$, $i \ne \gamma$.

<u>Proof</u>. Assume that we are mistaken, and relation

$$\sum_{\substack{i=1 \\ i\ne\gamma}}^{n} \eta_i \cdot V_{i\,t_1}^{opt} < \sum_{\substack{i=1 \\ i\ne\gamma}}^{n} \eta_i \cdot V_{i\,t_2}^{opt} \tag{18}$$

holds. Here $V_{i\,t_2}^{opt}$ stand for the optimal values obtained for the remaining $n-1$ customers and the up-dated available resources $V_{t_1} - V_{\gamma\,t_1}^{opt} \Rightarrow V_{t_2}$. Adding to both parts of (18) the same additive $\eta_\gamma \cdot V_{\gamma\,t_1}^{opt}$, we obtain

$$\sum_{i=1}^{n} \eta_i \cdot V_{i\,t_1}^{opt} < \sum_{\substack{i=1 \\ i\ne\gamma}}^{n} \eta_i \cdot V_{i\,t_2}^{opt} + \eta_\gamma \cdot V_{\gamma\,t_1}^{opt}, \tag{19}$$

where at least one value $V_{i\,t_1}^{opt}$, $i \ne \gamma$, differs from $V_{i\,t_2}^{opt}$. By setting $V_{\gamma\,t_1}^{opt} \Rightarrow V_{\gamma\,t_2}^{opt}$, we finally obtain

$$\sum_{i=1}^{n} \eta_i \cdot V_{i\,t_1}^{opt} < \sum_{i=1}^{n} \eta_i \cdot V_{i\,t_2}^{opt} , \tag{20}$$

where all n values $V_{\gamma\,t_1}^{opt}$, as well as n values $V_{i\,t_2}^{opt}$, satisfy (2–4). Relation (20) contradicts the fact that $V_{\gamma\,t_1}^{opt}$ has been obtained by using a *precise* algorithm, and objective $I = \sum_{i=1}^{n} \eta_i V_{i\,t_1}^{opt}$ cannot be further increased on the basis of the same restrictions set. Thus, our assumption is a false one.

A conclusion can be drawn that in Case 2 there is no need to resolve problem (1–4). The only amendment required boils down to changing value V_{t_2} by excluding $V_{t_2} - V_{\gamma\,t_1}^{opt}$.

3. At decision moment t_2 new ℓ customers C_j, $n+1 \le j \le n+\ell$, are fed into the system, together with corresponding parameters $V_{j\,\min}$, $V_{j\,\max}$, η_j, as well as new additional resource amount ΔV_{t_2} to cover for the extra demand, which is replenished at the storehouse.

In Case 3 an inspection has to be undertaken to determine the resource amounts which *all the customers, beginning from* $t_1 = 0$, *have actually obtained towards moment* t_2. For all those customers which are under way, parameters $V_{i\,\min}$ and $V_{i\,\max}$ have to be updated, and later on the updated initial data for the delivery system has to be determined. The available resource amount at moment t_2 has to be updated: $V_{t_1} + \Delta V_{t_2} - \Delta V_{t_1,t_2} \Rightarrow V_{t_2}$, where $\Delta V_{t_1,t_2}$ designates the resource amount being already delivered. Later on we apply *Step 1* of the algorithm to establish new optimal values $V_{i\,t_2}^{opt}$. The on-line algorithm terminates when all production units are fully supplied with resources.

8. Numerical Example

To demonstrate implementation of the suggested procedures, let us consider the extended on-line version of the allocation algorithm. A company delivers fuel in pipe-lines to ten different customers C_i, $i = \overline{1,10}$, with appropriate preference rates η_i. The system's initial data is presented in Table 1.

Assume that all the pipe-lines are similar, i.e., provide equal delivery rates.

Table 1. The system's initial data at $t_1 = 0$

C_i	$V_{i\,\min}$	$V_{i\,\max}$	η_i
1	15	25	0.65
2	20	45	0.80
3	15	25	0.76
4	40	60	0.73
5	60	80	0.67
6	75	100	0.72
7	50	75	0.70
8	60	90	0.68
9	25	40	0.78
10	30	50	0.71

At the beginning of the fuel supply, at $t_1 = 0$, the total amount equals $V_{t_1} = 480$. By implementing the allocation algorithm at $t_1 = 0$, we obtain values V_{it}^{opt} as follows:

$V_{1t} = 15$, $V_{2t} = 45$, $V_{3t} = 25$, $V_{4t} = 60$,
$V_{5t} = 60$, $V_{6t} = 95$, $V_{7t} = 50$, $V_{8t} = 60$,
$V_{9t} = 40$, $V_{10t} = 30$,

and the fuel delivery process starts.

At moment $t_2 = 3$ customer C_1 is fully supplied with fuel and leaves the system. The fuel delivery process proceeds with $V_{t_2} = 465$ until the moment $t_3 = 5$ when customer C_3 is fully supplied with fuel as well. It can be well-recognized that at moment t_3 the total resource amount equals $V_{t_3} = 440$, and the process proceeds until moment $t_4 = 6$, when three new customers C_{11} , C_{12} and C_{13} enter the system; simultaneously, the overall resource amount to be distributed is up-dated by an additional total of $\Delta V_{t_4} = 140$.

Assume that the considered company implements on-line control on the following principles (see Section 7):

1. Due to the Corollary outlined above, when at a certain moment t one or several customers finish to be supplied with fuel, for other customers still being under supply their delivery volumes are not re-calculated.
2. If at a certain moment t one or several *new* customers are fed into the delivery system, an inspection has to be undertaken, to determine the actual amount of fuel which all the remaining customers have already received to that moment. For the customers under way their appropriate parameters $V_{i\ \min}$ and $V_{i\ \max}$ are diminished by the supplied amounts of fuel (if the difference becomes negative it is set equal to zero), and optimisation problem (1–4) is resolved for the updated system data.

It can be well-recognized that in Case 2, whenever newly considered customers have higher preference rates than the remaining "older" ones, the reallocation algorithm will redistribute the available fuel amount by taking care of more important customers first. In some cases the "older" customers may even "loose" some of the supply amounts which have been previously assigned to them, for the benefit of the "new" ones. In the considered numerical example, an inspection is therefore undertaken at $t_4 = 6$, to determine that each customer under way (this does not count for those who have already left the system) has received $V_{i\,t_4} = 30$ units of fuel amount.

Thus, the system's management has to update the system's initial data in order to re-apply the allocation algorithm. The updated system's data at moment $t_4 = 6$ is presented in Table 2:

Table 2. The system's updated data at $t_4 = 6$

C_i	$V_{i\ \min}$	$V_{i\ \max}$	η_i
2	0	15	0.80
4	10	30	0.73
5	30	50	0.67
6	45	70	0.72
7	20	45	0.70
8	30	60	0.68
9	0	10	0.78
10	0	20	0.71
11	20	40	0.80
12	15	60	0.69
13	50	100	0.78

To enable decision-making for customers with the same preference rate, assume that the company management, when applying the reallocation algorithm at *Step 3*, acts as follows:

- when at least two customers have equal preference rates, they are re-ordered with preference given to the customer who entered the system before;
- if several competing customers entered the system simultaneously, preference is granted to those who have a smaller "tolerance gap" in terms of $V_{i\ \max} - V_{i\ \min}$;
- in case both the remainder $V_{i\ \max} - V_{i\ \min}$ stays the same, re-ordering customers is carried out by means of implementing random Monte-Carlo simulation.

The updated available amount of resources at $t_4 = 6$ may be estimated therefore as

$$V_{t_4} = 480 - 15 - 25 - 8 \times 30 + 140 = 340 .$$

Applying the resource reallocation algorithm, we finally obtain:

$$V_{2t_4} = 15, \quad V_{4t_4} = 30, \quad V_{5t_4} = 30, \quad V_{6t_4} = 50,$$
$$V_{7t_4} = 20, \quad V_{8t_4} = 30, \quad V_{9t_4} = 10, \quad V_{10t_4} = 0,$$
$$V_{11t_4} = 40, \quad V_{12t_4} = 15, \quad V_{13t_4} = 100 .$$

It may be noted that customer C_{10} ceases to be supplied ($V_{10t_4} = 0$), since he already obtained his minimal permissible fuel amount according to his preference rate. The fuel delivery process proceeds until the next decision making moment, etc.

Conclusions

The following conclusions can be drawn from the study:

a) The developed delivery model for homogenous consumable resources can cover a broad spectrum of organization systems, e.g. building enterprises, various production systems, service systems, etc.
b) The developed model incorporates a resource reallocation algorithm, which delivers the precise solution.
c) The developed algorithm can be incorporated in real on-line control models. The algorithm is simple in use and can be easily programmed on PC.

Acknowledgement

This research has been partially supported by the Paul Ivanier Center of Robotics and Production Management, Ben-Gurion University of the Negev.

References

1. Ben-Yair, A. *Harmonization Models in Strategic Management and Safety Engineering: Ph. D. Thesis.* Beer-Sheva, Israel: Ben-Gurion University of the Negev, 2004.
2. Burkov, V. V. and Irikov, V. A. *Models and Methods for Controlling Organization Systems*. Moscow: Nauka, 1994. (In Russian)
3. Chase, R. and Aquilano, N. *Production and Operations Management*. Illinois: Irwin, 1989.
4. Gonik, A. *Planning and Controlling Multi-Level Stochastic Projects: Ph. D. Thesis.* Beer-Sheva, Israel: Ben-Gurion University of the Negev, 1995.
5. Kats, V. *A Hierarchical Production Control Model with Variable Speeds and Inspection Points: Ph. D. Thesis.* Beer-Sheva, Israel: Ben-Gurion University of the Negev, 1997.
6. Taha, H. A. *Operations Research: An Introduction*. New York: MacMillan, 1987.

Received on the 20th of October 2007

Computer Modelling and New Technologies, 2007, Vol.11, No.4, 15–24
Transport and Telecommunication Institute, Lomonosov 1, LV-1019, Riga, Latvia

MANAGING LONG-TERM CONSTRUCTION PROJECTS WITH ALTERNATIVE OUTCOMES

***D. Golenko-Ginzburg* [1]*, Z. Laslo* [2]*, A. Ben-Yair* [3]**

[1]*Department of Industrial Engineering and Management (Professor Emeritus)*
Ben-Gurion University of the Negev
Beer-Sheva, P.O. Box 653, 84105, Israel
E-mail: dimitri@bgumail.bgu.ac.il
and
Department of Industrial Engineering and Management, Academic College of Judea and Samaria
Ariel, 44837

[2]*Department of Industrial Engineering and Management, Sami Shamoon College of Engineering*
Beer-Sheva, P.O. Box 950, 84100, Israel
E-mail: zohar@sce.ac.il

[3]*Department of Industrial Engineering and Management, Sami Shamoon College of Engineering*
Beer-Sheva, P.O. Box 950, 84100, Israel
E-mail: avnerb@sce.ac.il

Some levels of alternativity in long-term network projects by examining both fully divisible and more generalized controlled alternative activity networks will be considered. We will demonstrate that those alternative (multi-variant) projects may be controlled in the course of their actual realization. Moreover, the level of indeterminacy of such projects may influence the process of the project's capital investment, i.e., the procedure of contracting has to comprise consecutive step-wise stages.

Keywords: *project management (PM); alternative network project; deterministic and stochastic decision-making; joint variant; outcome tree; long-term construction project; consecutive step-wise contracting*

1. Introduction

Our main philosophy in stochastic project management with indeterminacy, as well as the basic Western conceptions of planning and controlling with uncertainty [3–14], is as follows. What is the essence of the philosophy when controlling a project system with uncertainty and being, at the outset, something which is basically indeterminate? Many examples from high performance practice in the Western World, as well as in Japan, show that under such circumstances, the control system should not work to a predetermined plan, but should be inherently adaptable, seeking at each decision node to assess the *best route forward*, reconfiguring the ultimate goals, if appropriate.

Note that the sub-problem of determining the *best route* may be very difficult and complicated, especially for systems with a high level of indeterminacy. Solving this sub-problem usually results in solving the general control problem.

Further, our philosophy in project planning and control with indeterminacy centres on avoiding predetermining the initial network model; moreover, in certain cases the structure of such a model may be indeterminate. At the initial stage of the project's realization, the network may be restricted to a source node and several alternative sink nodes (goals) together with some milestones (a decision-tree model). Such a restricted project is called an aggregated project. Various activities are usually of random duration. Such a stochastic alternative network is renewed permanently over time, including changes in the ultimate goals. At each decision node, our techniques enable us to choose the optimal outcome. Decision-making is repeatedly introduced for the renewed network at every sequentially reached decision node.

We will examine henceforth a network model with a very high level of uncertainty – a branching network to control a project with two kinds of alternative events: stochastic (uncontrolled) branching of the development of a project, as well as deterministic branching where the outcome direction is chosen by the project's decision maker.

Note that while the literature on PERT and CPM network techniques is quite vast, the number of publications on *alternative networks* remains very scanty. Various authors, e.g. Eisner [6], Elmaghraby [7], Pritsker [18], Whitehouse [20], etc., introduced the concept of a research and development (R&D) project as a complex of actions and problems towards achieving a definite goal. Several adequate network models for such projects have been considered. The first significant development in that area was the pioneering work of Eisner [6] in which a "decision box" with both random and deterministic alternative outcomes was introduced. Elmaghraby [7] introduced additional logic and algebra in network techniques, while Pritsker, Happ and Whitehouse [16–17] developed the GERT techniques for alternative network models with stochastic outcomes in key nodes. Xespos and Strassman [21] introduced the concept of the stochastic decision tree, while Crowston and Thompson [3] and later on Hastings and Mello [14] suggested the concept of multiple choices at such alternative nodes, when decision-making is of deterministic nature (Decision-CPM models). Lee, Moeller and Digman [15] developed the VERT model that enables the analyst to simulate various decisions with alternative technology choices within the stochastic decision tree network.

Golenko-Ginzburg [10, 11, 13] developed a unified controlled alternative activity network (CAAN model) for projects with both random and deterministic alternative outcomes at key nodes. At each routine decision-making node the developed algorithm based on lexicographic scanning, singles out all the sub-networks (the so-called joint variants) that correspond to all possible outcomes from that node. The joint variants of the CAAN model are enumerated by introducing a lexicographic order to the set of maximal paths in the CAAN graph. The corresponding look over algorithm is very simple in usage. Decision-making results in determining the optimal joint variant and following the optimal direction up to the next decision-making node.

Thus, the previously developed alternative network models GERT, Decision-CPM and VERT are particular cases of the CAAN model. The CAAN model does not include non-fully-divisible networks. It can be well-recognized therefore that the model is not relevant to those projects, which are structured from non-divisible sub-networks.

A more generalized alternative activity network model [12] is not fully divisible. It comprises all types of the previously considered alternative network models, including Eisner's model and the CAAN model. Note that the outlined above alternative network models enter the so-called SATM model which comprises various additional logical restrictions [19]. However, the computational algorithm to determine the joint variants has not been developed as yet. Later on we will use the CAAN alternative network model which suits mostly to long-term construction projects (LTCP).

2. Formal Description of the Alternative Stochastic Model

The alternative CAAN network model is generally a finite compendent directed graph, $G(U, Y)$, with the following properties:

(1) Graph G has one initial vertex, y_0 (the network entry), for which $\Gamma^{-1} y_0 = \acute{\varnothing}$ and $\Gamma y_0 \neq \acute{\varnothing}$.

(2) Graph G contains a set Y' of vertices y' (called terminal vertices, or network exits), where $\Gamma y' = \acute{\varnothing}$, $\Gamma^{-1} y' \neq \acute{\varnothing}$ and $|Y'| \geq 2$.

(3) The set of vertices Y of graph G is not uniform and consists of vertices of type $\tilde{\chi} \in \tilde{X}$ (classical PERT model) and of more complex logical types, $\tilde{\alpha} \in \tilde{A}$, $\tilde{\beta} \in \tilde{B}$, and $\tilde{\gamma} \in \tilde{\Gamma}$, being represented in the below Table 1:

Table 1. Logical possibilities of alternative network model events

Designation of an event in the model	Logical relations at the event's receiver	Logical relations at the event's emitter
$\tilde{\chi}$	and	and
$\tilde{\alpha}$	and	exclusive "or"
$\tilde{\beta}$	exclusive "or"	and
$\tilde{\gamma}$	exclusive "or"	exclusive "or"

(4) The set of arcs U of graph G is split into a subset U' of arcs corresponding to the actual functioning of the alternative network, and subset U'' of arcs representing the logical interconnections between actual and imaginary functions.

(5) Vector W_{kl} of values characterizing actual work is constructed preliminary for every arc, $U_{kl} \in U'$, representing an actual activity. Among such values are the time of the activity duration t_{kl}; the required cost C_{kl}; and other components of this vector. The vector's components $\omega_{kl}^{(\rho)}$ ($\rho = 1 - k$, k being the vector' dimension) can be represented, depending on the degree of indeterminacy, either by determined estimations or by random values with a given distribution function, $f\left(\omega_{kl}^{(\rho)}\right)$, on the interval $\left[\alpha\left(\omega_{kl}^{(\rho)}\right),\ \beta\left(\omega_{kl}^{(\rho)}\right)\right]$, where $\alpha\left(\omega_{kl}^{(\rho)}\right)$ and $\beta\left(\omega_{kl}^{(\rho)}\right)$ are boundary estimations of the ρ-th component of vector W_{kl}.

(6) For the stochastic alternative model of a combined type, the set of alternative vertices, $\widetilde{A} \cup \widetilde{\Gamma}$, is split into subsets $\overline{A}$ – alternative vertices that show the branching of determined variants, and $\overset{=}{A}$ – alternative vertices that represent the situations of branching stochastic variants, where $\widetilde{A} \cup \widetilde{\Gamma} = \overline{A} \cup \overset{=}{A}$.

(7) When the network vertex is of alternative nature, it is assigned a set of estimations of corresponding local variant probabilities. In other words, a nonnegative number, $\overset{=}{p}_{ij} \le 1$, such that $\sum_{j=1}^{n_i} \overset{=}{p}_{ij} = 1$ (where $\overset{=}{p}_{ij}$ is the *a priori* probability of transferring from i to j and n_i stands for the number of local variants appearing in event i), is related to each alternative path starting from vertex i of type $\widetilde{\alpha} \in \overset{=}{A}$ or $\widetilde{\gamma} \in \overset{=}{A}$ and leading to outcome j.

(8) If event i is related to an alternative vertex of class $\widetilde{A}$, the corresponding conditional transfer probability, $\overset{=}{p}_{ij}$, is usually assumed to be equal 1. This means that the process of choosing the direction in which the system has to move towards its target is of a determined character; it is the prerogative of the system's controlling device.

Problems of alternative network model analysis and synthesis are solved by applying the principle of network enlarging and obtaining a special graph – the outcome tree [8–10], which is usually designated as $D(A,V)$ and represents a graph that can be constructed by modifying the original model, $G(Y,U)$, as follows:

(a) The set, which consists of the initial vertex, finite vertices, and vertices that are branching points of alternative paths of graph G, is taken as the set of vertices of graph D. The initial vertex, $\alpha_0 = y_0$, is called a hanging vertex.

(b) The set of arcs $V = \{v_{ij}\}$ of graph D is obtained thorough an equivalent transformation of a set of sub-graphs, $\{G_{ij}\}$, extracted from network G according to the following procedure:

- any vertex α_i, except for the finite ones, α', can be the initial vertex of sub-graph $G_{ij} = (L_{ij}, U_{ij})$, where $\alpha' \in y_{ij}$ and $\Gamma^{-1}\alpha_i \cap Y_{ij} = \acute{\emptyset}$;
- $Y_{ij} \subset \widetilde{\Gamma}\alpha_i$, where $\widetilde{\Gamma}\alpha_i$ stands for the transitive closure of mapping α_i;
- only an α-vertex of graph G, except for the initial vertex, $\alpha_0 = y_0$, can be a finite vertex of sub-graph G_{ij}, and

- no $(\alpha_i, \ldots, \alpha_j)$-type paths that connect the initial vertex, α_i, with sub-graph finite vertex α_j in G_{ij}, contain other α -vertices of graph G.

(c) every arc, v_{ij}, of outcome tree D is obtained by reducing fragment G_{ij} of network $G(Y, U)$ to one arc beginning at α_i and ending at α_j. In addition, realization probability p_{ij}, fulfilment time t_{ij}, and other parameters equivalent to the corresponding characteristic values for initial fragment G_{ij} are brought into correspondence with the enlarged arc v_{ij}.

If different fragments, G_{ij}, of the model do not intersect, the alternative network is called entirely divisible; all vertices of the corresponding outcome tree prove to be γ -type vertices. We will require a supplementary definition. A <u>partial variant</u> is a variant of the network model's realization; it corresponds to a definite direction of its development at an individual stage, characterizes one of the possible ways of reaching the intermediate target, and does not contain alternative situations. The variant of realization of the whole project, which does not contain alternative branching and is formed by a sequence of partial variants, is called a <u>full variant</u>. On the outcome tree, $D(A, V)$, a certain arc, v_{ij}, corresponds to the partial variant, while some path connecting root vertex α_0 with one of the hanging vertices, corresponds to the full variant.

The combined outcome tree, $D(A, V)$, can be regarded as a union of purely stochastic outcome trees that reflects some homogenous alternative stochastic network models. The latter are obtained by choosing different directions in the controlled devices. Such stochastic outcome trees, which are all part of the combined outcome tree, $D(A, V)$, are called <u>joint variants</u> of realizing the stochastic network model.

The joint variant can be extracted from the original graph, $D(A, V)$, by "fixing" certain directions in interconnected vertices of type $\overline{\alpha}$ and excluding unfixed directions. In other words, every joint variant can be regarded as a realization variant of the network model. Such a variant has a determined topology, but it contains probability situations and has certain possible stochastic finite states.

Let us examine an outcome tree of a CAAN type alternative project presented on *Figure 1*. Here $\{\overline{\alpha}\}$ denote decision-making nodes of deterministic nature, where the outcome direction is fully governed by the project's manager. Nodes $\{\overline{\overline{\alpha}}\}$ are of stochastic nature and, as such, are not controlled. Each $\overline{\overline{\alpha}}$ -type node comprises several outcome probabilities which form a full group of events.

Stage a) on Figure 2 presents six joint variants which can be singled out by analysing the outcome tree. Note that none of the joint variants comprise alternative deterministic nodes and are determined by choosing non-contradictive directions in nodes $\overline{\alpha}$. Thus, all joint variants are either purely alternative stochastic non-controllable networks of α -type, or non-alternative fragments. Stage *b)* on Figure 2 demonstrates simplified joint variants of both types.

3. Optimal Joint Variant

The managing a controlled alternative activity network with two types of branching events means choosing an optimal joint variant, which optimises the project's goal function. For the case of an LTCP we consider several optimality criteria:

A. Since a joint variant can be regarded as a purely stochastic alternative project, we may calculate the entropy level as a measure of indeterminacy for each joint variant. The joint variant with the least entropy level has to be chosen.

B. For the case of an averse risk manager the strategy is as follows. Calculate (for each joint variant) the goal function for the <u>worst possible probability outcome</u> (i.e., the worst possible full variant). In other words, we determine the worst goal function value which may be actually (i.e., with probability exceeding zero) achieved in the course of realizing the joint variant. Call such a goal function value for the i -th joint variant $G_{\min}(J_i)$. The joint variant which delivers an optimal value from all $G_{\min}(J_i)$, $1 \le i \le n$, has to be chosen as the optimal one.

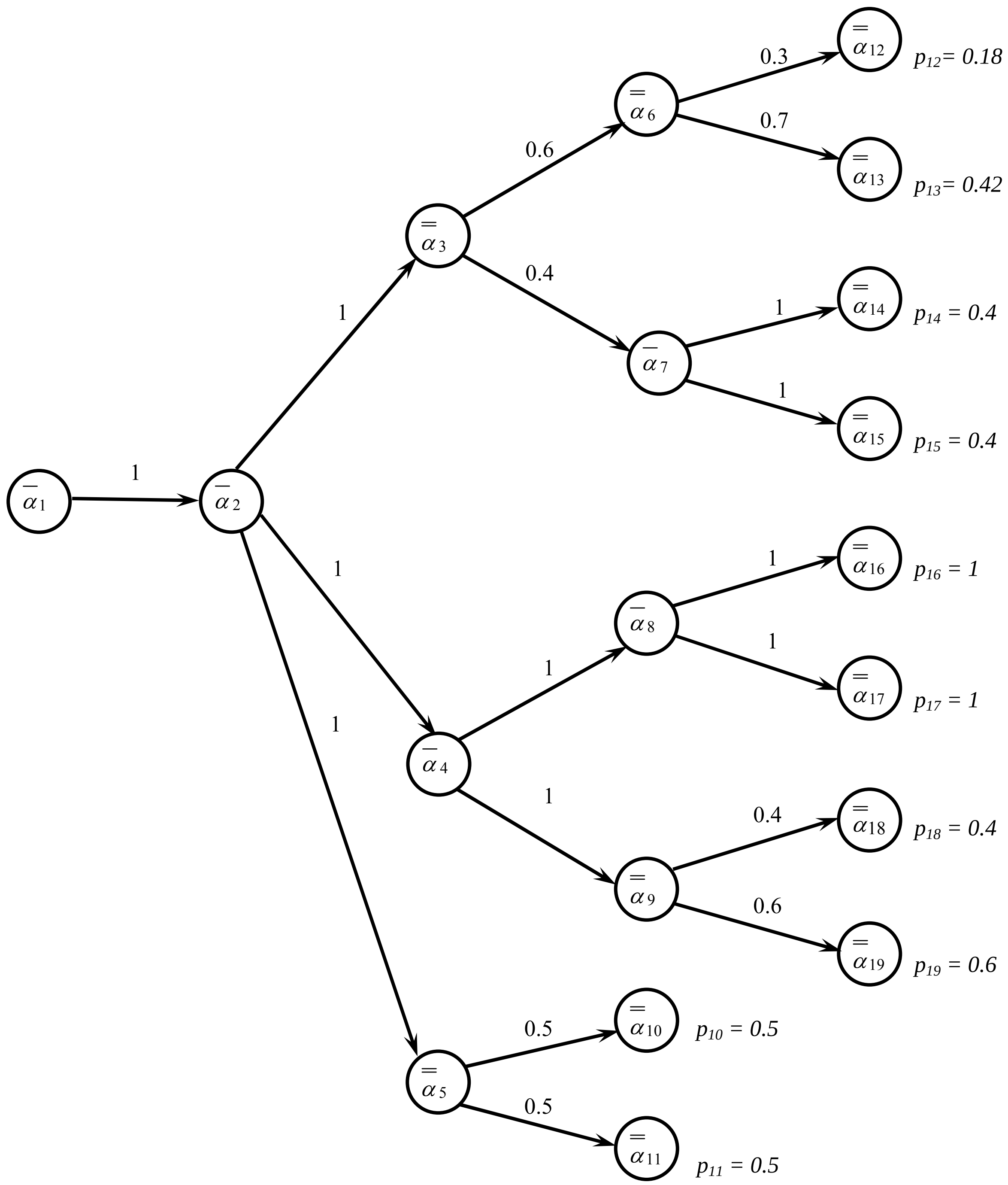


Figure 1. Controlled alternative network project

C. Calculate for each i-th joint variant J_i, $1 \le i \le n$, the average value of the goal function, i.e., the mathematical expectation given in the form

$$\overline{G}(J_i) = \sum_{j=1}^{n_i} p_j G(F_j), \tag{1}$$

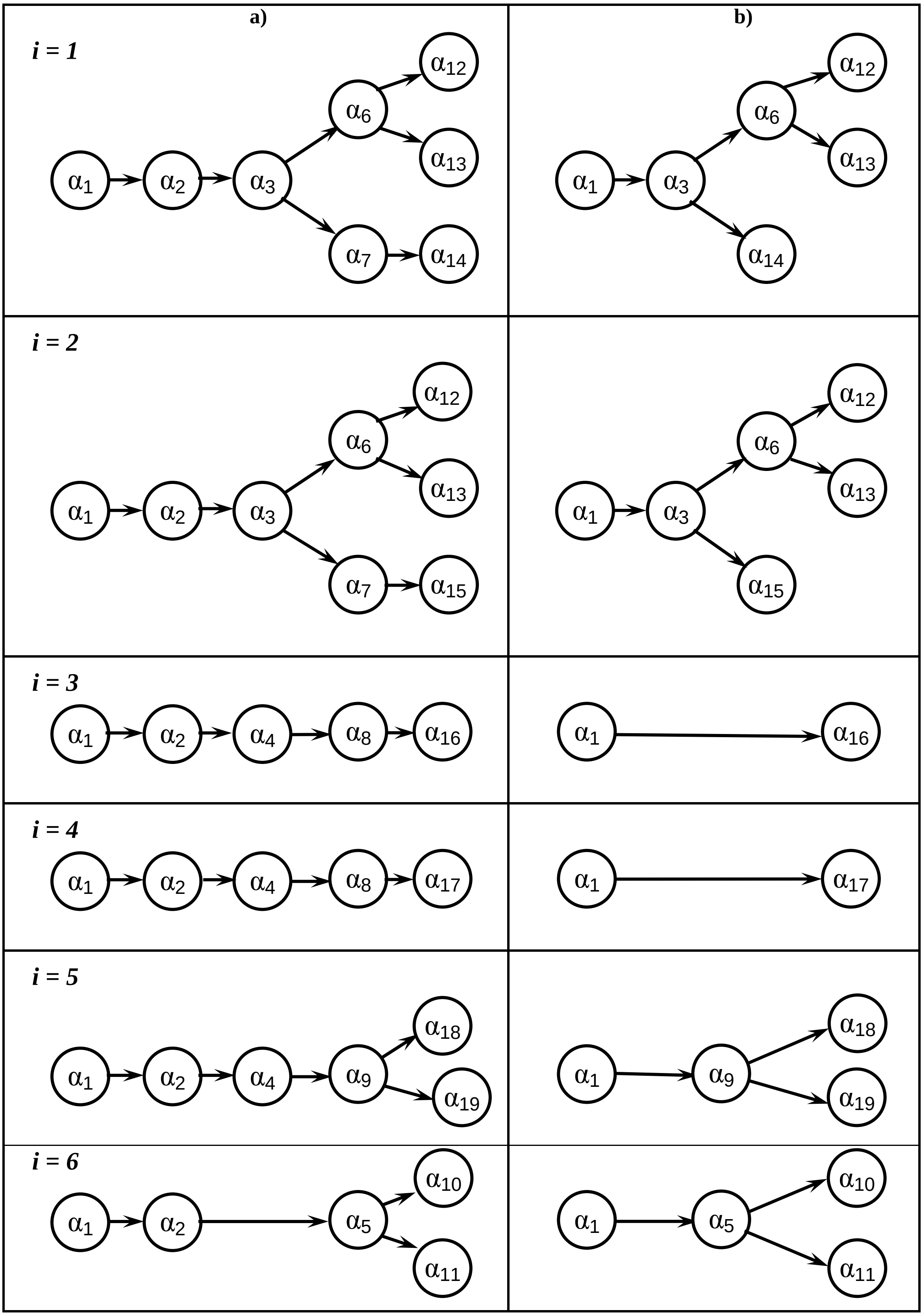


Figure 2. The project's joint variants

where p_j denotes the probability of realizing the full variant F_j, and $G(F_j)$ stands for the goal function of that full variant. Joint variant J_ξ satisfying

$$\overline{G}(J_\xi) = \max_{1 \le i \le n} \left\{ \sum_{j=1}^{n_i} p_j G(F_j) \right\}, \tag{2}$$

has to be preferred as the optimal one.

D. *Criterion D* is contrary to *Criterion B*. We have to calculate for each joint variant the goal function corresponding to the best goal function outcome which may be actually obtained in the course of realizing the joint variant. Call it $G_{\max}(J_i)$, $1 \le i \le n$. Joint variant J_η satisfying

$$G_{\max}(J_\eta) = \max_{1 \le i \le n} \left[G_{\max}(J_i) \right] \tag{3}$$

has to be chosen as the optimal one.

The choice of the optimal strategy depends on the nature of LTCP. If the value of the quality of he project's product is extremely important, *Strategy D* has to be preferred. Note that both *Strategies B* and *D* are, in fact, game strategies. In case if we are interested in a less nervous progress of the regarded projects, *Strategy C* seems to us to be a better choice.

4. Numerical Example

Let us present a numerical example for the outcome tree appearing in *Figures 1* and *2*. The alternative model of CAAN type comprises 6 joint variants $J_1, ..., J_6$ and 10 full variants F_{11}, F_{12}, F_{13}, F_{21}, F_{22}, F_{23}, F_{31}, F_{41}, F_{51}, F_{52}, F_{61}, F_{62} (note that full variants F_{11} and F_{12} coincide with full variants F_{21} and F_{22}). Let the goal function be the project's cost (to be minimized) and preset the local activities' costs as follows: $C_{12} = 10$, $C_{23} = 6$, $C_{24} = 15$, $C_{25} = 12$, $C_{36} = 14$, $C_{37} = 9$, $C_{6,12} = 10$, $C_{6,13} = 16$, $C_{7,14} = 15$, $C_{7,15} = 20$, $C_{48} = 11$, $C_{49} = 13$, $C_{8,16} = 15$, $C_{8,17} = 8$, $C_{9,18} = 10$, $C_{9,19} = 18$, $C_{5,10} = 12$, $C_{5,11} = 36$. Assume, further, that the alternative graph under consideration refers to the LTCP class of projects.

It can be well-recognized from examining *Figures 1* and *2* that implementing *Strategy A* results in comparing two alternative joint variants J_3 and J_4, both with zero level of entropy. Since

$$C(J_3) = C_{12} + C_{24} + C_{48} + C_{8,16} = 51$$

exceeds

$$C(J_4) = C_{12} + C_{24} + C_{48} + C_{8,17} = 44,$$

joint variant J_4 has to be determined as the optimal one.

Implementing *Strategy B*, i.e., risk-averse decision-making, boils down to calculating the following values:

$$C_{\max}(J_1) = \max \begin{bmatrix} C_{12} + C_{23} + C_{36} + C_{6,12} = 40;\ C_{12} + C_{23} + C_{36} + C_{6,13} = 46; \\ C_{12} + C_{23} + C_{37} + C_{7,14} = 40 \end{bmatrix} = 46;$$

$$C_{\max}(J_2) = \max[40;\ 46;\ 45] = 46;$$

$C_{\max}(J_3) = C_{12} + C_{24} + C_{48} + C_{8,16} = 51;$

$C_{\max}(J_4) = C_{12} + C_{24} + C_{48} + C_{8,17} = 44;$

$C_{\max}(J_5) = \max\left[C_{12} + C_{24} + C_{49} + C_{9,18} = 48;\ C_{12} + C_{24} + C_{49} + C_{9,19} = 56\right] = 56;$

$C_{\max}(J_6) = \max\left[C_{12} + C_{25} + C_{5,10} = 34;\ C_{12} + C_{25} + C_{5,11} = 58\right] = 58.$

Thus, joint variant J_4 which delivers the extreme (the minimal) goal function value if the worst comes to the worst for all joint variants, has to be chosen as the optimal one. Using the "opposite" *Strategy D*, we may calculate

$C_{\min}(J_1) = \min[40;\ 46;\ 40] = 40;$

$C_{\min}(J_2) = \min[40;\ 46;\ 45] = 40;$

$C_{\min}(J_3) = 51;$

$C_{\min}(J_4) = 44;$

$C_{\min}(J_5) = 48;$

$C_{\min}(J_6) = 34.$

Thus, when implementing risky decision-makings, the result of the procedure is different, namely: joint variant J_6 has to be determined as the optimal one. When adopting *Strategy C*, the mathematical expectations of the cost to realize the considered joint variants may be calculated as follows (refer again to *Figures 1* and *2*):

$$\overline{C}(J_1) = (10+6+14+10)\cdot 0.18 + (10+6+14+16)\cdot 0.42 + (10+6+9+15)\cdot 0.4 = $$
$$= 7.20 + 19.32 + 16 = 42.52;$$

$$\overline{C}(J_2) = 40\cdot 0.18 + 46\cdot 0.42 + 45\cdot 0.4 = 7.20 + 19.32 + 18 = 44.52;$$

$$\overline{C}(J_3) = 51;$$

$$\overline{C}(J_4) = 44;$$

$$\overline{C}(J_5) = (10+15+13+10)\cdot 0.4 + (10+15+13+10)\cdot 0.6 = 19.2 + 28.8 = 48;$$

$$\overline{C}(J_6) = (10+12+12)\cdot 0.5 + (10+12+36)\cdot 0.5 = 46 .$$

Since J_1 results in the minimal mean cost expenses required, it has to be chosen as the optimal one. Thus, adopting different optimality concepts may result in corresponding changing of the joint variant determined as optimal.

5. Capital Investment in Long-Term Alternative Projects under Random Disturbances

It can be well-recognized that in recent years undertaking capital investments and contracting long-term projects which are carried out under random disturbances, has been the subject of lengthy debate and a very sharp criticism (see, e.g., [1]). This is because nowadays it is extremely difficult to implement into commercial agreements both the projects' durations and especially the required volume of the corresponding capital investments (see, e.g., Ananjin [1]). This refers mostly to long-term construction projects based on future geological surveys with a high level of indeterminacy, projects involving implementation of new unique technology, etc. It goes without saying that for LTCP comprising both deterministic and stochastic alternative variants, the challenge of determining with a more or less accuracy the future project's parameters (like cost, duration, reliability attributes, etc.) becomes practically impossible. However, something has to be decided and has to be done immediately, otherwise the losses originating from failure to compete with accelerating technical and technological progress, may prove to be tremendous.

To meet the challenge, we suggest a new step-wise procedure in order to manage long-term construction projects with alternatives of both deterministic and stochastic nature. The main stages of the procedure are as follows:

Stage I. If possible, determine an alternative graph of the future LTCP. The graph has to be similar to that outlined on *Figure 1*.

Stage II. Determine all the joint variants entering the graph. The corresponding algorithm is outlined in [8–11], and is based on lexicographical simulation.

Stage III. Determine the strategy for recognizing the optimal joint variant. We remind that different conceptual strategies may result in different principles of optimality and indeterminacy and, thus, result in variety of the optimal joint variant identity, as it was demonstrated in the previous *Section*. In our opinion, the majority of LTCP projects may use the average criterion value in order to determine the optimal joint variant, i.e., *Strategy C*.

Stage IV. After determining the optimal joint variant, one may start the contracting process. We suggest to undertake this process sequentially. On the first step the capital investments have to cover the progress of the project from the very beginning until the first branching node of stochastic type. If, for example, we have chosen J_1 presented on *Figure 2*, as the optimal one, the signed agreements have to cover expenses starting from event $\bar{\alpha}_1$ until the next alternative (branching) node $\bar{\bar{\alpha}}_3$, i.e., the primary capital investments have to cover the realization of fragment $\bar{\alpha}_1 \Rightarrow \bar{\alpha}_2 \Rightarrow \bar{\bar{\alpha}}_3$. Thus, the corresponding contract has to cover expenses estimated as $C_{12} + C_{23} = 16$.

Stage V. After reaching event $\bar{\bar{\alpha}}_3$ the contract has to be rewritten anew, depending on the realization of the uncontrolled direction ($\bar{\bar{\alpha}}_3 \Rightarrow \bar{\bar{\alpha}}_4$ or $\bar{\bar{\alpha}}_3 \Rightarrow \bar{\bar{\alpha}}_5$) of the progress of the project.

Stage VI. In the course of the project's realization the joint variant we have chosen before, besides being updated, may undergo other changes as well, both in the structure of the graph itself and in the values of the probability outcomes. Thus the consecutive progress of the project results in consecutive updating the contract's agreement. We do not see another managerial principle applicable to multi-variant alternative projects under consideration. Note that such a form of monitoring enables both on-line and financial control procedures.

Conclusions

The following conclusions can be drawn from the study:

1. Long-Term Construction Projects may deal with a high level of indeterminacy, as well as with various types of branching nodes in key events. Those nodes may be the result of unpredictable outcomes of future pioneering hi-tech experiments, geological surveys with possible alternative outcomes, etc.
2. We have described and presented a numerical example of the class of stochastic alternative networks which comprise both decision-making nodes with deterministic branching and un-controllable alternative nodes with probabilistic outcomes.
3. We have demonstrated the possibility of singling out an optimal joint variant from the previously given stochastic alternative network graph. The structure of the optimal joint variant depends on the concept of optimality, as it has been presented by means of the numerical example. A joint variant does not comprise controllable branching events and is, in fact, a purely homogenous alternative stochastic network.

4. We have suggested a new procedure of contracting capital investments for the considered class of stochastic alternative models. On our opinion, the suggested mechanism may be effectively used in the course of drawing out financial contracts and other agreements in order to supply complicated long-term projects of alternative structure.

Acknowledgement

This research has been partially supported by the Paul Ivanier Center on Robotics and Production Research, Ben-Gurion University of the Negev.

References

1. Ananjin, V. Indeterminacy in Contracts. In: *Proceedings of the International Symposium on Project Management, Nizhni Novgorod, February 14–18, 2007, Russia*. Nizhni Novgorod, Russia, 2007.
2. Ben-Yair, A. *Harmonization Models in Strategic Management and Safety Engineering: Ph. D. Thesis.* Beer-Sheva, Israel: Ben-Gurion University of the Negev, 2004.
3. Crowston, W. B., Thompson, G. L. Decision CPM: A Method for Simultaneous Planning, Scheduling and Control of Projects, *Operations Research*, Vol. 15, 1967, pp. 407–426.
4. Crowston, W. B. Decision CPM: Network Reduction and Solution, *Operations Research Quarterly,* Vol. 21(4), 1970, pp. 435–445.
5. Crowston, W. B. Models for Project Management, *Sloan Management Review,* Vol. 12(3), 1971, pp. 25–42.
6. Eisner, H. A Generalized Network Approach to the Planning and Scheduling of a Research Project, *Operations Research*, Vol. 10(1), 1962, pp. 115–125.
7. Elmaghraby, S. *Activity Networks: Project Planning and Control by Network Models*. New York: Wiley, 1977.
8. Golenko, D. *Statistical Methods in Network Planning and Control*. Moscow: Nauka, 1968. (In Russian)
9. Golenko, D. *Statistical Models in Production Control*. Moscow: Statistika, 1973. (In Russian)
10. Golenko, D., Livshitz, S., Kesler, S. *Statistical Modelling in Technical-Economical Systems (Project Management)*. Leningrad: Leningrad University Press, 1977. (In Russian)
11. Golenko-Ginzburg D. (1988) Controlled Alternative Activity Networks in Project Management, *European Journal of Operations Research* 7, pp. 336–346.
12. Golenko-Ginzburg, D., Blokh, D. A Generalized Activity Network Model, *Journal of Operations Research Society,* Vol. 48, 1997, pp. 391–400.
13. Golenko-Ginzburg, D., Gonik, A. Project Planning and Controlling by Stochastic Network Models. In: *Managing and Modelling Complex Projects/T. Williams (ed.) NATO ASI Series*. Kluwer Academic Publishers, 1997, pp. 21–43.
14. Hastings, M. A., Mello, J. M. *Decision Networks*. New York: Wiley, 1979.
15. Lee, S. M., Moeller, G. L., Digman,, L. D. *Network Analysis for Management Decisions.* Boston: Kluwer-Nijhoff, 1987.
16. Pritsker, A. A. B. and Happ, W. W. GERT: Graphical Evaluation and Review Technique, Part I (Fundamental), *Journal of Industrial Engineering,* Vol. 17(5), 1966, pp. 267–274.
17. Pritsker, A. A. B., Whitehouse, G. E. GERT: Graphical Evaluation and Review Technique, Part II (Probabilistic and Industrial Engineering Applications), *Journal of Industrial Engineering*, Vol. 17(6), 1996, pp. 293–301.
18. Pritsker, A. A. B. *Modelling and Analysis Using Q-GERT Networks*. New York: Wiley, 1977.
19. Voropajev, V., Ljubkin, S., Titarenko, B., Golenko-Ginzburg, D. Structural Classification of Network Models, *International Journal of Project Management*, Vol. 18, 2000, pp. 361–368.
20. Whitehouse, G. E. *Systems Analysis and Design Using Network Techniques*. New Jersey: Prentice-Hall, Englewood Cliffs, 1973.
21. Xespos, R. F., Strassman, P. A. Stochastic Decision Trees for the Analysis of Investment Decisions, *Management Science B*, Vol. 11, 1965, pp. 244–259.

Received on the 1st of November 2007

Computer Modelling and New Technologies, 2007, Vol.11, No.4, 25–35
Transport and Telecommunication Institute, Lomonosov 1, LV-1019, Riga, Latvia

MEASURING EFFICIENCY OF RESTAURANTS USING THE DATA ENVELOPMENT ANALYSIS METHODOLOGY

***Y. Hadad* [1], *L. Friedman* [2], *M. Z. Hanani* [3]**

[1]*Department of Industrial Engineering and Management, Sami Shamoon College of Engineering*
Bialik / Basel Sts. Bialik / Basel Sts., P.O. Box 950, Beer-Sheva 84100 Israel
Phone: +972 86475645. Fax:+972 86475643. E-mail: yossi@sce.ac.il

[2]*Department of Industrial Engineering and Management, Ben-Gurion University of the Negev*
P.O. Box 653, Beer-Sheva 84105, Israel
Phone: +972 86472224. Fax: +972 86472958. E-mail: leaf@bgumail.bgu.ac.il

[3]*Department of Industrial Engineering and Management, Sami Shamoon College of Engineering*
Bialik / Basel Sts. Bialik / Basel Sts., P.O. Box 950, Beer-Sheva 84100 Israel
Phone: +972 86475644. Fax: +972 86475643. E-mail: micha@sce.ac.il

The Data Envelopment Analysis (DEA) is a methodology for evaluating the relative efficiency of units based on multiple inputs and multiple outputs. The efficiency score is defined as the ratio between weighted outputs to the weighted inputs. The weights for the inputs and outputs are chosen in order to maximize this efficiency score. These optimal weights differ from unit to unit, and can have a very large range. Therefore, one cannot rank the units according to these weights. Another reason for not ranking the units based on the DEA is that, according to these optimal weights, one can only distinguish between an efficient unit and an inefficient one. Especially when the number of units is small, relative to the number of inputs and outputs, most of the units will be efficient. In this paper, we refer to several ranking methods in the DEA context: Super Efficiency, Cross Efficiency (CE) Canonical Correlation Analysis (CCA), Discriminant Analysis of Ratio (DR / DEA) and the Global Efficiency (GE). We illustrate all the ranking methods based on data from 30 restaurants. Each restaurant has 2 outputs and 4 inputs.

Keywords: *The Data Envelopment Analysis (DEA), Common Weights (CW), Super Efficiency, Multi-criteria Decision Analysis (MCDA), Cross Efficiency (CE) Canonical Correlation Analysis (CCA), Discriminant Analysis of Ratio (DR / DEA), Global Efficiency (GE), Ranking*

1. Introduction

The Data Envelopment Analysis (DEA) was first introduced by Charnes, Cooper and Rhodes (CCR) in 1978. The DEA is a non-parametric methodology for evaluating the relative efficiency of Decision-Making Units (DMU) based on multiple inputs and multiple outputs. The efficiency score is measured as a ratio between weighted outputs and weighted inputs, even if the production function is unknown. The weights are chosen so as to find the best advantage for each unit to maximize its relative efficiency, under the restriction that this score is bound by 100% efficiency. If a unit with its optimal weights receives the efficiency score of 100%, it is efficient, while a score of less than 100% is considered inefficient. The DEA provides only a dichotomy classification into two groups; efficient and inefficient. This classification into only two groups may cause several problems when there is a need to rank all the units on one scale. Especially when the number of inputs and outputs is large in comparison to the number of units, most of the units may be efficient and it will be difficult to distinguish among them. In order to reduce the number of efficient units, there are several solutions: to put constraints on the weights (Sueyoshi (1999)), or through assurance regions (Thompson (1992)) or by a cone-ratio (hyperplane) models (Charnes et al. (1989)). Another way to refine the ranking of the efficient units is by the super efficiency method that was introduced by Anderson and Peterson (1993). Silkman et al. (1986) suggested ranking all the units utilizing the Cross Efficiency Matrix. Cooper and Tone (1997) argue that the DEA only classifies the units into two dichotomic sets: efficient and inefficient. They do not rank the efficient unit since they claim that they are all on the efficient frontier. Moreover, they do not accept the DEA score as a ranking score for the inefficient units, since their weights vary from unit to unit. However, they suggest another ranking method based on the slack variables of the dual problem (the improvements of each variable).

Practical examples are given by a number of authors, among them: Doyle and Green (1994) who ranked 20 universities in the UK; Sinuany-Stern et al. (1994) ranked 22 academic departments in a university; Friedman and Sinuany-Stern (1998) ranked industrial branches in Israel; Sinuany-Stern and Friedman (1998b) ranked towns in Israel; Sueyoshi and Kirihara (1998) use linear discriminant analysis for classification of Japanese Banking Institutes, and they extended the method in 2001; Sueyoshi (1999), ranked Japanese agriculture cooperatives. Sharman et al. (1998, 1999) investigated fish farms in Nepal (1998) and China (1999), and Sueyoshi (1992), measured the industrial performance of Chinese cities by Data Envelopment Analysis.

The DEA methodology provides, for each unit, its own optimal set of weights. Because of this variability of weights, one cannot rank all the units on one scale. In order to rank all the units on the same scale, the use of common weights for calculating each unit's efficiency score will solve this problem. There are several criteria to find these common weights. Friedman and Sinuany-Stern (1997) utilize the CCA / DEA (the Canonical Correlation Analysis) to calculate the common weights that maximize the correlation between the weighted inputs and weighted outputs. Another method was the DR / DEA investigated by Sinuany-Stern and Friedman (1998a) to find the common weights that will maximize the ratio of the sum of squares between the averages of the two groups (efficient and inefficient) and the sum of squares within variance of the groups. Ganely and Cubin (1992) found these common weights by maximizing the sum of efficiency scores of all the units. Trout (1997) maximized the minimum efficiency score.

Not much research has been done in the application of the DEA in the service industry. One exception is Anderson et al. (1999), which was done in an area close to the restaurant sector of the service industry. The significance of applying the DEA model to the service industry is expressed in the article, as this industry is characterized by extensive competition and the importance of performance in providing the service. Shiuh-Nan and Te-Yi (2001), and Israeli et al. (2002) are among, the few other published research efforts in the hotel industry.

In this paper, we introduce several ranking methods in the DEA context; Super Efficiency, Canonical Correlation Analysis (CCA), Discriminant Analysis of Ratio (DR / DEA) and the Global Efficiency (GE). We illustrate these ranking methods using data collected in a case study done on 30 restaurants. Each restaurant has 2 outputs and 4 inputs.

The paper is organized as follows: Section 2 describes the Data Envelopment Analysis and Ranking Methods in the DEA Context and Section 3 presents the case study. Section 4 deals with the results. Section 5 presents a summary and conclusions.

2. Data Envelopment Analysis and Ranking Methods in the DEA Context

2.1. Data Envelopment Analysis

DEA is a procedure designed to measure the relative efficiency in situations when there are multiple inputs and multiple outputs and no obvious method how to aggregate both inputs and outputs into a meaningful index of productive efficiency. DEA was developed by Charnes, Cooper and Rhodes (CCR) (1978). The DEA provides a mechanism for measuring a Decision-Making Unit (DMU) Pareto efficiency compared with other DMUs. The mechanism is extensively employed in diverse industries and environments (an extensive review of DEA applications is provided by Seiford (1996)). In the service sector, applications of DEA include education (Sexton et al. (1994)), recreation and health care management (Sherman (1984)) to name just a few.

The efficiency in DEA is termed Technical and Scale Efficiency (TSE) and the relative efficiency of a DMU is defined as the ratio of its total weighted output to its total weighted input. The question is how to select the weights if no standard unit of measure can be assigned to the inputs and outputs? Here lies the seed of DEA procedure. DEA permits each DMU to select any desirable weight for each input and output, provided that they satisfy certain reasonable conditions: first those weights can not be negative, and second that the weights must be universal, which means that the resulting ratio should not exceed 1. The BCC model, named after Banker, Charnes and Cooper (1984) allows the production function to exhibit non-constant return to scale (Banker and Chang (1995)) while the CCR model imposes the additional assumption of constant returns to scale on the production function.

The Technical and Scale Efficiency (TSE) with constant return to scale is computed according to the CCR model (Charnes, Cooper&Rhodes (1978)). Consider *n* DMUs, where each DMU *j* ($j = 1,\ldots, n$) uses *m* inputs $\vec{X}_j = (X_{1j}, X_{2j}, \ldots\ldots, X_{mj})^T > 0$ for producing *S* outputs $\vec{Y}_j = (Y_{1j}, Y_{2j}, \ldots\ldots, Y_{sj})^T > 0$. The CCR model is: For each unit k we find the best weights U_r^k $(r = 1, 2, .., S)$ and V_i^k $(i = 1, 2, .., m)$ that maximize the ratio between the weighted output and weighted input:

$$
\begin{aligned}
& h_k = Max \sum_{r=1}^{s} U_r^k Y_{rk}, \\
& s.t \\
& \sum_{i=1}^{m} V_i^k X_{ik} = 1 \quad , \\
& \sum_{r=1}^{s} U_r^k Y_{rj} - \sum_{i=1}^{m} V_i^k X_{ij} \leq 0 \; ; \; j = 1,\ldots, n \\
& U_r^k \geq 0 \; ; \; r = 1,2,.., s; \\
& V_i^k \geq 0 \; ; \; i = 1,2,.., m.
\end{aligned}
\tag{1}
$$

The weights are all positive and the ratios are bounded by 1 (100%). Each unit k is assigned the highest possible efficiency score by choosing the most optimal weights. If a unit reaches the maximum possible value of 100% it is efficient, otherwise it is inefficient.

Obviously, the values of the weights would differ from unit to unit, and they sometimes have great variability. Therefore, we cannot perform a full rank of all the units based on the DEA scores.

The Technical Efficiency (TE) with Decreasing Return to Scale is computed according to the BCC model (Banker Charnes and Cooper (1984)). The BCC model is as follows:

$$
\begin{aligned}
& h_k = Max \sum_{r=1}^{s} U_r^k Y_{rk} - W_k \\
& s.t \\
& \sum_{i=1}^{m} V_i^k X_{ik} = 1 \; ; \\
& \sum_{r=1}^{s} U_r^k Y_{rj} - \sum_{i=1}^{m} V_i^k X_{ij} - W_k \leq 0; \; j = 1,\ldots, n; \\
& U_r^k \geq 0; \; r = 1,2,.., s \; ; \; V_i^k \geq 0; i = 1,2,.., m; \\
& W_k \; free \; .
\end{aligned}
\tag{2}
$$

There are different methods for ranking units in the DEA context. (See the Review of Ranking Methods by Adler et al. (2000)). Ranking is a well established approach in social science (see Young and Hammer (1987)), historically much more established than the dichotomic classification of DEA for efficient and inefficient organizational units (see Adler et al. (2002)). Also, economics applied the classical measurement of efficiency which rank-scales economic units. Rank scaling in the DEA context has become well established in the last decade. Sexton (1986) was the first to introduce full rank scaling of organizational units in the DEA context, by utilizing the Cross-Efficiency Matrix. Anderson and Peterson (1993) developed the super efficiency approach for rank-scaling that was followed by other researchers. The ranking in relation to rank-scaling has the advantage that it can be tested statistically by a nonparametric analysis (see, for example, Friedman and Sinuany-Stern (1997) and (1998), Sinuany-Stern and Friedman (1998a) and Sueyoshi and Aoki (2001)).

2.2. The Super Efficiency Method

Anderson and Peterson (A&P) (1993) view the DEA score for the inefficient units as their rank scale. In order to rank scale the efficient units, they suggest allowing the efficient units to receive a score greater than 1 by dropping the constraint that bounds the score of the evaluated unit k; namely the primal problem of A&P of unit k will be formulated as follows:

$$
\begin{aligned}
& h_k = Max \sum_{r=1}^{s} U_r^k Y_{rk} \\
& s.t. \\
& \sum_{r=1}^{s} U_r^k Y_{rj} - \sum_{i=1}^{m} V_i^k X_{ij} \le 0 \,; \quad \text{for } j = 1,2,.., n, \; j \ne k \;; \\
& \sum_{i=1}^{m} V_i^k X_{ik} = 1; \\
& U_r^k \ge \varepsilon > 0 \,; \qquad r = 1,2,.., s; \\
& \mathrm{V}_{\mathrm{i}}^{\mathrm{k}} \ge \varepsilon > 0 \qquad i = 1,2,..., m,
\end{aligned}
\tag{3}
$$

where $\varepsilon > 0$ is a non-Archimedean infinitesimal.

2.3. The Cross Efficiency Method (CE)

The Cross Efficiency (CE) rating was first introduced by Sexton et al. (1986). The results of all the CCR ratios can be summarized in matrix which is called the Cross Efficiency Matrix. Its elements are:

$$
h_{kj} = \frac{\sum_{r=1}^{s} U_r^k Y_{rj}}{\sum_{i=}^{m} V_i^k X_{ij}} \qquad k = 1,2,...,n \quad j = 1,2,..,n \,; \tag{4}
$$

thus, h_{kj} represents the ratio given to unit j in the CCR run of unit k. This score evaluates the efficiency of unit j by the optimal weights of unit k. The elements on the diagonal h_{kk} are the CCR efficiency score. Let us define $\overline{h}_k = \frac{\sum_{j=1}^{n} h_{kj}}{n}$ as the average cross- efficiency score given to unit k. The maximum value of $\overline{h}_k$ is 1, which occurs if unit k is efficient in all the runs i.e. all the units evaluate unit k as efficient. In order to rank the units, we can assign the unit with the highest score a rank of one and the unit with the lowest score a rank of n. For more details see Friedman and Sinuany-Stern (1998).

2.4. Ranking Methods Based on Common Weights

There are other methods for ranking units in the DEA contexts based on common weights for all the units depending on different criteria. This is in contrast to the standard approach of the DEA, which generates different sets of weights for each unit.

2.4.1. The Canonical Correlation Analysis (CCA/DEA) for Ranking

The canonical correlation analysis (CCA) is an extension of regression analysis. While the regression model explains a single output using multiple inputs, canonical correlation analyzes multiple inputs and multiple outputs. CCA searches for a single vector weight for the inputs and outputs, common to all the units. CCA constructs a composite input variable Z_j, as a linear combination of the m inputs and W_j as a linear combination of the s outputs.

$$
\begin{aligned}
Z_j &= V_1 X_{1j} + V_2 {}_X {}_{2j} + ... + V_m X_{mj} \,, \\
W_j &= U_1 {}^Y {}_{1j} + U_2 {}^Y {}_{2j} + ... + U_s {}^Y {}_{sj} \,.
\end{aligned}
\tag{5}
$$

CCA determines the two vectors of coefficients, $\vec{V}^T = (V_1 \quad V_2 \quad ... \quad V_m)$ and $\vec{U}^T = (U_1 \quad U_2 \quad ... \quad U_s)$, so as to maximize r_{zw}, the coefficient of correlation between the composite input, Z, and the composite output, W. Formally, this is specified in the following

$$
\begin{aligned}
& Max\ \ \mathrm{r}_{zw} = \frac{\vec{V}^T S_{xy}\vec{U}}{\sqrt{\left(\vec{V}^T S_{xx}\vec{V}\right)\left(\vec{U}^T S_{yy}U\right)}} \\
& s.t. \\
& \vec{V}^T S_{xy}\vec{V} = 1, \\
& \vec{U}^T S_{yy}U = 1,
\end{aligned}
\tag{6}
$$

where $S_{xx}, S_{yy},\ S_{xy}$ are the matrices of the sums of squares and sums of products of the variables respectively. Note that the weights $\vec{V}^T$ and $\vec{U}^T$ are determined up to a proportional constant. There is a closed form solution for the weights. However, this ranking is only feasible in the DEA context if all the weights are non-negative.

Friedman and Sinuany-Stern (1997) use the CCA method be defining a scaling ratio score, T_j, as a ratio of linear combinations of the inputs and outputs. They utilize the common weights for the linear combinations that are drawn from the largest eigenvalue of the CCA method, as shown in next equation –

$$
T_j = \frac{W_j}{Z_j} = \frac{\sum_{r=1}^{s} U_r Y_{rj}}{\sum_{i=1}^{m} V_i x_{ij}}, \quad j = 1,2,...,\ n\ ; \tag{7}
$$

note that while the DEA efficiency ratio h_j is bounded above 1, the scaling ratio T_j of the CCA / DEA is unbounded. Thus, it is the rank order of the scaling ratios that is important, rather than their absolute value.

Although the CCA is independent of the DEA results, the same formulation has been applied to the two types of variables; weighted inputs and weighted outputs. Moreover, the scaling ratio score T is also similar to the DEA score, namely it represents the ratio described in the DEA. The use of a single set of weights is possibly the most obvious methodology for ranking and in this case the CCA has been built in the light of the DEA context. Hence, empirically, the results can be tested statistically.

2.4.2. The Discriminant Analysis of Ratios (DR / DEA)

Sinuany-Stern and Friedman (1998a) developed a technique in which discriminant analysis of ratios was applied to DEA (DR / DEA). Instead of considering a linear combination of the inputs and outputs in one equation (as in the traditional discriminant analysis of the two groups), they constructed a ratio function between a linear combination of the inputs and a linear combination of the outputs. In some ways this ratio function is similar to the DEA efficiency ratio, however whilst DEA provides weights for the inputs and outputs, which vary from unit to unit, DR / DEA provides common weights for all units. In principle, DR / DEA determines the weights, so that the ratio score function discriminates optimally between two groups of observations (DMUs) on a one-dimensional scale (in our case, efficient and inefficient units predetermined by DEA). The ratio, T_j, and the arithmetic means of the ratio scores of the efficient and inefficient groups are:

$$
T_j = \frac{\sum_{r=1}^{s} u_r y_{rj}}{\sum_{i=1}^{m} v_i x_{ij}}, \qquad j = 1,...,n;\ \ \bar{T}_1 = \sum_{j=1}^{n_1} \frac{T_j}{n_1} \quad , \quad \bar{T}_2 = \sum_{j=n_1+1}^{n} \frac{T_j}{n_2}\ , \tag{8}
$$

where n_1 and n_2 are the number of efficient and inefficient units in the DEA model respectively. The weighted mean of the entire n units ($n=n_1+n_2$) will be denoted by

$$
\bar{T} = \frac{n_1\bar{T}_1 + n_2\bar{T}_2}{n}\ . \tag{9}
$$

Our problem is to find the common weights v_i and u_r, so that the ratio of the between-group variance of T, ($SS_B(T)$) and the within group variance of T, ($SS_W(T)$) will be maximized, as shown in model as follows

$$\max_{u_r, v_i} \lambda = \max_{u_r, v_i} \frac{SS_B(T)}{SS_W(T)} \quad ,$$

$$SS_B(T) = n_1(\overline{T}_1 - \overline{T})^2 + n_2(\overline{T}_2 - \overline{T})^2 = \frac{n_1 n_2}{n_1 + n_2}(\overline{T}_1 - \overline{T}_2)^2, \qquad 10)$$

$$SS_W(T) = \sum_{j=1}^{n_1}(T_j - \overline{T}_1)^2 + \sum_{j=n_1+1}^{n}(T_j - \overline{T}_2)^2 \ .$$

DR / DEA constructs the efficiency score for each unit *j* as T_j, the ratio between the composite output and composite input. Thus it rank scales the DMUs so that the unit with the highest score receives rank 1 and the unit with the lowest score ranks *n*. If any weight is negative, then non-negativity constraints ought to be added to the optimisation problem. To solve this problem, they used a non-linear search optimisation algorithm; however there is no guarantee that the solution found is globally optimal.

2.4.3. The Global Efficiency Method (GE)

Another criterion to find the best common weights was proposed by Ganley and Cubbin (1992). The criterion is to maximize the sum of efficiency ratios of all the units, namely, if we define the efficiency score E_j^* based on the common weights U_r^*, V^* as

$$E_j^* = \frac{\sum_{r=1}^{s} U_r^* y_{rj}}{\sum_{i=1}^{m} V_i^* x_{ij}}, \qquad (11)$$

then these common weights will be obtained by a single, non-linear program as follows:

$$\begin{aligned} &Max\ \mathrm{Z} = \sum_{\mathrm{j}=1}^{\mathrm{n}} \mathrm{E}_{\mathrm{j}}^{*}, \\ &s.t. \\ &\hat{E}_j^* \le 1; \qquad \mathrm{j} = 1,2,...,\ \mathrm{n}; \\ &U_r \ge 0\,; \qquad \mathrm{r} = 1,2,..,\ \mathrm{s}; \\ &\mathrm{V_i} \ge 0\,; \qquad \mathrm{i} = 1.2,....,\ \mathrm{m}. \end{aligned} \qquad (12)$$

For more details see Friedman et al. (1998), Adler et al. (2002).

3. The Case Study on Restaurants

Economically, restaurants are characterized as a day-to-day high-risk business. Demand is highly randomised, and arrival of customers is highly unpredictable. Moreover, other external factors have to be taken into account, such as seasonal fluctuations (Muller (1999)). Muller suggests the use of several criteria to measure efficiency of restaurants, and states the inherent difficulties in evaluating the efficiency. He suggests using more general, less distinct measures such as average occupation of tables, average price etc.

The literature does not offer results on the use of the DEA methodology in evaluating the efficiency of restaurants. Other methods have been used such as Muller (1999), where an evaluation called "Restaurants Occupation Ratio" is sued. This term is defined as the ratio between the actual number of sales per time unit and the capacity available at that time unit. Kimes (1999) suggests the use of a "Restaurant Revenue Management" (RRM) measure to evaluate restaurant efficiency.

One more point of interest: there are restaurant guides, such as Le Guide Gault&Millau (LGG), that rank restaurants, grading them on a scale of (in case of LGG) 11–20, where 11 is the lowest entry to this guide. The grade is determined based on the quality of the food, the treatment of a customer, service in general, politeness, hygiene in general and in the toilets in particular, price etc. We shall use the ranking of this guide in comparison with the ranking we shall get using the DEA and its extensions.

The data set of this study includes 30 restaurants in Tel-Aviv (Israel). The restaurants that were examined were selected from the restaurant guide "Gault&Millau". An important issue in employing DEA is the selection of inputs and outputs. In this case study the inputs and outputs were chosen by consulting experts in the restaurant business.

3.1. List of Outputs and Inputs

Outputs

Y_1 – the average number of customers in a day.

Y_2 – the price of an average meal.

Inputs

X_1 – the number of seats.

X_2 – the average number of waiters in a shift.

X_3 – the average number of general employees (chef, cooks, cleaning employees, cashier, barman etc.)

X_4 – the area of the restaurant in m^2.

Table 1 below includes the data on the 30 restaurants with 2 outputs and 4 inputs.

Table 1. The numerical data

Unit	INPUTS				OUTPUTS		LGG SCORE
	X_1	X_2	X_3	X_4	Y_1	Y_2	
1	40	2	3	100	90	200	13
2	120	5	17	350	650	125	12
3	60	3	7	100	150	175	12
4	60	1	1	85	20	125	12
5	90	10	15	300	350	75	12
6	55	1	4	100	100	125	13
7	75	3	4	150	110	225	13
8	40	2	4	100	40	200	13
9	70	2	3	65	150	125	12
10	80	5	15	180	120	200	12
11	40	3	6	100	100	75	11
12	45	2	6	100	100	175	11
13	150	12	22	1600	500	225	16
14	40	2	4	120	75	350	15
15	100	7	10	300	200	150	14
16	80	4	8	130	110	250	14
17	110	6	10	450	100	175	12
18	60	2	3	200	50	200	12
19	120	8	25	300	600	200	13
20	150	8	10	350	400	150	13
21	130	7	12	1000	400	175	12
22	70	6	11	100	300	175	12
23	50	2	5	600	70	105	11
24	100	4	12	300	160	125	12
25	64	2	4	100	100	125	12
26	100	3	10	400	200	60	11
27	30	1	5	70	60	75	11
28	75	3	20	80	70	200	12
29	110	4	6	400	80	225	12
30	100	5	12	300	200	200	13

Prior to implementing the DEA method, boundaries were determined for the various weights. These boundaries express the importance of the factors as defined by the expert (the decision maker). The boundaries are given in Table 2 below, in percentage points.

4. The Results

Table 2. The efficiency score of the units

DMU	CCR (TSE)	BCC (TE)	DMU	CCR (TSE)	BCC (TE)
1	100	100	16	69.24	73.68
2	100	100	17	31.76	36.84
3	84.48	89.69	18	73.98	80
4	100	100	19	92.87	100
5	64.79	82.49	20	82.25	100
6	100	100	21	73.57	87.86
7	81.57	83.55	22	100	100
8	67.37	95.34	23	44.32	74.36
9	100	100	24	40.06	48.98
10	46.08	53.48	25	67.85	81.44
11	57.55	91.83	26	51.36	63.57
12	80.14	91.87	27	69.54	100
13	36.68	92.78	28	74.97	100
14	100	100	29	46.29	48.17
15	52.19	56.76	30	50.64	55.21

We ran the DEA on all the 30 units, 7 units came out efficient in the CCR model, and 11 units came out efficient in the BCC model (with no bounds on the weights). The DEA classifies the units into two groups; efficient and inefficient. In order to obtain a full ranking of the restaurants we ran the A&P ranking method, that ranked all the restaurants from 1 (the most efficient restaurant) to 30 (the most inefficient one).

Afterwards, we ran all three ranking methods with common weights CCA, GE and DR / DEA. The common weights for these three ranking methods are given in Table 3. All the other common weights for CCA / DEA, GE and DR / DEA are positive, as we have constrained them. The values of the weights and even the ranking are very different from method to method.

Table 4 includes the scores and the ranking of all the units for all the ranking methods. One can see that unit 14 reached the first place in the average score ranking. It reached the first place for 2 ranking methods, the second place for two others. Unit 17 was placed last by most of the methods. In spite of the different common weights for the inputs and outputs, and therefore the different rank scaling, there are high correlations between the pairs of the ranking methods. Table 5 contains the correlations between all the ranking methods. All the correlations are significantly (with p-value less then 0.01) correlated.

Table 3. The weights of ranking methods

	CCA	GE	DR / DEA
U_1	12.1942	1.8136	0.4774
U_2	2.5305	0.2944	0.1072
V_1	1.2968	0.5217	0.0001
V_2	1.0352	0.1638	0.2872
V_3	1.4618	2.1076	0.1775
V_4	0.0002	0.0001	0.4575

Table 4. Scores and the ranking of the units

UNIT	DEA (CCR)	A&P		CCA		GE		DR / DEA		CE		AVERAGE	
		Score	Rank	Score	Rank	Score	Rank	Score	Rank	Score	Rank	Score	Rank
1	100	1.0061	7	4.5179	3	1	2.5	1.303	5	0.8441	2	0.8108	4
2	100	1.4699	2	5.3183	1	1	2.5	1.5149	2	0.8198	4	0.9151	2
3	84.48	0.8448	9	3.4371	10	0.6737	11	1.0913	8	0.6549	8	0.6230	9
4	100	1.3294	4	1.9277	26	0.5249	19	0.9575	12	0.5263	14	0.5514	14
5	64.79	0.6479	20	2.8232	16	0.6065	16	0.6489	20	0.4440	19	0.4689	18
6	100	1.1578	6	3.4937	8	0.7086	9	1.3809	4	0.6788	6	0.7077	6
7	81.57	0.8157	11	3.234	12	0.7765	8	1.0464	11	0.6378	10	0.6238	8
8	67.37	0.6737	19	2.9198	14	0.5556	17	0.8646	14	0.5831	11	0.5188	15
9	100	1.4274	3	3.9006	6	1	2.5	1.692	1	0.8343	3	0.8807	3
10	46.08	0.4608	26	1.8486	28	0.3122	29	0.5382	25	0.3576	26	0.3211	28
11	57.55	0.5755	21	2.531	21	0.4988	22	0.6744	19	0.4361	20	0.4294	21
12	80.14	0.8014	12	3.4428	9	0.618	15	1.0673	9	0.6475	9	0.6027	11
13	36.68	0.3668	29	3.0418	13	0.6237	14	0.4842	27	0.3818	23	0.4158	22
14	100	1.7135	1	5.2344	2	1	2.5	1.468	3	0.9965	1	0.9704	1
15	52.19	0.5219	22	2.3557	23	0.5319	18	0.5947	23	0.3896	22	0.4044	23
16	69.24	0.6924	17	2.5731	20	0.5135	20	0.8297	16	0.5262	15	0.4840	17
17	31.76	0.3176	30	1.5298	29	0.326	28	0.3702	29	0.2519	30	0.2541	30
18	73.98	0.7398	14	2.7507	17	0.6294	13	0.7756	17	0.5207	16	0.5118	16
19	92.87	0.9287	8	3.9827	4	0.6994	10	1.1039	6	0.6666	7	0.6623	7
20	82.25	0.8225	10	3.3397	11	0.8428	5	0.9371	13	0.5592	13	0.6132	10
21	73.57	0.7357	15	3.6095	7	0.8101	6	0.645	21	0.4798	18	0.5562	13
22	100	1.3078	5	3.9034	5	0.7857	7	1.0946	7	0.7416	5	0.7348	5
23	44.32	0.4432	27	2.3098	24	0.4555	24	0.3279	30	0.2979	29	0.3282	27
24	40.06	0.4006	28	2.0433	25	0.3901	27	0.5842	24	0.3339	28	0.3377	26
25	67.85	0.6785	18	2.8964	15	0.6543	12	1.0657	10	0.5668	12	0.5587	12
26	51.36	0.5136	23	2.4506	22	0.494	23	0.6427	22	0.3616	24	0.3994	24
27	69.54	0.6954	16	2.6143	18	0.4272	26	0.844	15	0.4867	17	0.4624	19
28	74.97	0.7497	13	1.3287	30	0.1829	30	0.4762	28	0.3606	25	0.3027	29
29	46.29	0.4629	25	1.899	27	0.4374	25	0.5053	26	0.3410	27	0.3411	25
30	50.64	0.5064	24	2.6022	19	0.5087	21	0.7163	18	0.4331	21	0.4303	20

The average grade is computed based on the average values of the normalized grades of the different methods.

Table 5. The correlations of ranking methods

	CCA	GE	DR / DEA	CE	AVER.	LGG
A&P	0.7335	0.7737	0.8162	0.8959	0.9221	0.1866
CCA		0.9254	0.8278	0.8803	0.8923	0.2706
GE			0.7987	0.8576	0.9218	0.3349
DR / DEA				0.9724	0.9257	0.1025
CE					0.9751	0.20676
AVER						0.2303

Summary and Conclusions

In this paper, we have applied the DEA methodology (CCR and BCC models) and 5 ranking methods for ranking 30 restaurants in Tel-Aviv (Israel) by efficiency. The rankings arrived at using the different models differ from each other in spite of the high correlation among some of the ranking models we used. The reason is that each ranking model has its own objective function, which is different from the others. In order to select the most appropriate model, the user (decision maker) has to set the criterion (objective function) that is most suitable for his / her purpose (or to his / her opinion), and then decide on the model to be used. In those cases where the user (decision-maker) believes that all ranking models should be used, we have suggested using the average of the rankings, computed from the average ranking grades. Note that there is a high correlation between the ranking methods and the average ranking.

Furthermore, the correlation between the LGG rating and the rankings arrived at in this article has been considered. A low correlation has been found there, and the reasons for that, in our opinion, are two, the first being the difference in scales: The scale in the ranking methods investigated here is 1–30, while the LGG scale is 11–16 (when we grouped the 30 ranks into groups comparable to the other scale, the correlation was even lower). The second reason is the apparent lack of relationship between efficiency and the LGG rating, which expresses a rate of quality and service as per price. In other words: a restaurant could be efficient as a result of charging a high price and providing a poor service. E.g. restaurant number 13, which got the highest mark (16) in LGG, was found inefficient by the CCR model, and placed among the lowest by the different rating models (placed number 22 by the average rating). On the other hand, restaurant number 14, which was graded 15 by LGG, was found efficient by the CRR and BCC models, and was placed very high by average ranking.

Acknowledgment

This paper was partially supported by the Paul Ivanier Center for Robotics and Production Management, Ben-Gurion University of the Negev.

References

1. Adler, N., Friedman, L., Sinuany-Stern, Z. Review of Ranking Methods in the DEA Context, *European Journal of Operational Research*, Vol. 140, 2002, pp. 249–265.
2. Anderson, P., Peterson, N. C. A Procedure for Ranking Efficient Units in DEA, *Management Science*, Vol. 39, No 10, 1993, pp. 1261–1264.
3. Anderson, R. I., Fish, M., Xia, Y., Michello, F. Measuring Efficiency in the Hotel Industry: a Stochastic Frontier Approach, *Int. J. of Hospitality Management*, Vol. 18, 1999, pp. 45–57.
4. Banker, R. D., Chang, H. A Simulation Study of Hypothesis Tests for Differences in Efficiencies, *Int. J. Production Economist*, Vol. 39, 1995, pp. 37–54.
5. Banker, R. D., Charnes, A., Cooper, W. W. Some Models for Estimating Technical and Scale Inefficiencies in Data Envelop Analysis, *Management Science*, Vol. 30, 1984, pp. 1078–1092.
6. Charnes, A., Cooper, W. W. The Non-Archimedean CCR Ratio for Efficiency Analysis: a Rejoinder to Boyyd and Fare, *Eur. J. of Operational Research*, Vol. 15, 1984, pp. 333–334.
7. Charnes, A., Cooper, W. W., Rhodes, E. Measuring the Efficiency of Decision Making Units, *Eur. J. of Operational Research*, Vol. 2, 1978, pp. 429–444.
8. Charnes, A., Cooper, W. W., Wei, Q. L., Huang, Z. M. Cone Ratio DEA and Multi-Objective Programming, *Int. J. of System Science*, Vol. 20, No 7, 1989, pp. 1099–1118.
9. Cooper, W. W., Tone, K. Measures of Inefficiency in DEA and Stochastic Frontier Estimation, *Eur. J. of Operational Research*, Vol. 99, 1997, pp. 72–88.
10. Doyle, J. R., Green, R. H. Efficiency and Cross-Efficiency in DEA: Derivatives Meaning and Uses, *J. of Operational Research Society*, Vol. 45, No 5, 1994, pp. 567–578.
11. Friedman, L., Sinuany-Stern, Z. Combining Ranking Scales and Selecting Variables in the DEA Context. The Case of Industrial Branches, *Computers and Operational Research*, Vol. 25, No 9, 1998, pp. 781–791.
12. Friedman, L., Sinuany-Stern, Z. Scaling Units Via the Canonical Correlation Analysis in the DEA Context, *Eur. J. of Operational Research*, Vol. 100, No 3, 1997, pp. 629–637.
13. Ganely, J. A., Cubbin, S. A. *Public Sector Efficiency Measurement: Applications of Data Envelopment Analysis*. North-Holland, 1992.

14. Israeli, A., Hadad, Y., Mehrez, A. Employing DEA for Ranking Hotels Advertisement: and a Case of Analyzing Advertising Supplement of Hotels in Israel, *J. of Management Science and regional Development,* Vol. 4, 2002, pp. 47–63.
15. Kimes, S. E. Implementing Restaurant Revenue Management: A Five-Step approach, *The Cornell Hotel and Restaurant Administration Quarterly,* Vol. 40, 1999, pp. 116–121.
16. Morrison, T. *Multivariate statistical methods.* New York: McGraw Hill, 1976.
17. Muller, C. C. A Simple Measure of Restaurant Efficiency, *The Cornell Hotel and Restaurant Administration Quarterly,* Vol. 40, 1999, pp. 31–37.
18. Seiford, L. M. The Evolution of the State-of-Art (1978–1995), *J. of Productivity Analysis,* Vol. 7, 1996, pp. 99–137.
19. Sexton, T. R. The Methodology of DEA. In: *Measuring Efficiency: An Assessment of DEA/ R. H. Silkman (ed.).* San Francisco: Jossey-Bass, 1986, pp. 7–29.
20. Sexton, T. R., Sleeper, S., Taggart, I. E. Improving Pupil Transportation in North Carolina, *Interfaces,* Vol. 24(1), 1994, pp. 87–104.
21. Sharma, K. R., Leung, P. S. Technical Efficiency of Carp Production in Nepal: an Application of Stochastic Frontier Production Function Approach, *Aquaculture Economics and Management,* Vol. 2, 1998, pp. 129–140.
22. Sharma, K. R., Leung, P. S., Chen, H., Peterson, A. Economic efficiency and optimum in fish Polyclone: an application of Data Envelopment Analysis (DEA) to Chinese fish farm, *Aquaculture Economics and Management,* Vol. 180, 1999. pp. 207–221.
23. Sherman, H. D. Improving the Productivity of Service Businesses, *Sloan Management Review,* Vol. 23**,** No 3, 1984, pp. 11–23.
24. Shiuh-Nan, H., Te-Yi, C. Using DEA to Measure Hotel Managerial Efficiency Change in Taiwan, *Tourism Management,* Vol. 24, 2003, pp. 357–369.
25. Silkman, R. H., Hogan, A. J. DEA Critique and Extension. In: *Measuring Efficiency: an Assessment of DEA/R. H. Silkman (ed.).* San Francisco: Jossey-Bass, 1986, pp. 73–104.
26. Sinuany-Stern, Z., Mehrez, A., Barboy, A. Academic Departments Efficiency Via DEA, *Computers and Operations Research,* Vol. 21, No 5, 1994, pp. 543–556.
27. Sinuany-Stern, Z., Friedman, L. DEA and the Discriminant Analysis of Ratios for Ranking Units, *Eur. J. of Operations Research,* Vol. 111, 1998a, pp. 470–478.
28. Sinuany-Stern, Z., Friedman, L. Rank Scaling in the DEA Context, *Studies in Regional&Urban Planning,* Vol. 6, 1998b, pp. 135–140.
29. Sueyoshi, T. Measuring the Industrial Performance of Chinese Cities by Data Envelopment Analysis, *Sociol. Econ. Planning. Sci.,* Vol. 26, No 2, 1992, pp. 85–88.
30. Sueyoshi, T. DEA Non-Parametric Ranking Test and Index Measurement. Slack adjusted DEA and an Application to Japanese Cooperatives, *Omega,* Vol. 27, 1999, pp. 315–326.
31. Sueyoshi, T., Kirihara, L. Efficiency Measurement and Strategic Classification of Japanese Banking Institutes, *Int. J. of Systems Science,* Vol. 29, 1998, pp. 1249–1263.
32. Trout, M. D. Derivation of the Maximum Efficiency Ratio Model from the Maximum Decisional Efficiency Principle, *Ann. of Operations Research,* Vol. 73, 1997, pp. 323–338.

Received on the 12th of November 2007

Computer Modelling and New Technologies, 2007, Vol.11, No.4, 36–39
Transport and Telecommunication Institute, Lomonosov 1, LV-1019, Riga, Latvia

OPTIMIZATION OF INVENTORY MANAGEMENT OF A PERISHABLE PRODUCT IN THE CONDITIONS OF FLUCTUATING DEMAND

A. Babitski

Belarusian State University
Independence Ave. 4, Minsk, 220050, Belarus
E-mail: babitski@hotmail.com

An inventory system with perishable products with deterministic shelf life and demand represented by Poisson process with different intensities for different time intervals is studied. Formulae for calculation of expected storage cost and income are obtained. Algorithms for determining optimal order size for maximization of net income and unit income are proposed.

Keywords: *inventory control optimisation, perishable product, stochastic demand*

1. Introduction

This paper studies an inventory model with zero replenishment lead time and perishable items with deterministic shelf life. The period of shelf life is broken in a set of intervals, so that, time between inventories a consumption moment is exponentially distributed with intensity value changing from interval to interval.

Importance of systems with fluctuating demand is explained by the fact that they provide more accurate calculation of expected storage costs, comparing to the systems with constant parameters of consumption process. In the majority of studied models with fluctuating demand, e.g., [2, 8], change in the parameters of consumption process is bound to the moment of inventory replenishment. Deterministic demand with consumption intensity changing between replenishment moments is investigated in [7]. Literature on inventory management of perishable products is quite extensive, e.g., [3, 4, 5, 6], where various assumption about deterioration of products are considered. A model with exponentially distributed demand with changed intensities is studied in [1], in the current paper these results are extended to the case of perishable products.

This paper provides formulae for calculation of expected storage cost and income in the aforementioned system. Algorithms for solving optimisation tasks for net income maximization during shelf life and for unit income maximization are proposed.

2. Model

The system implies instant replenishment when inventory level reaches zero. All purchased items are sold at price s and can be stored up to moment T, after moment T all unsold items are utilized at price q $(q < s)$. Incurred costs include per order cost c, item price p $(s > p > q)$ and the cost h for storing one item per unit time. Time between inventory consumption moments is exponentially distributed with intensity $\mu(t) = \mu_i, t_i \le t < t_{i+1}$, where t_i are the moments when intensity changes $t_0 = 0 < t_1 < ... < t_M = T$.

Cost for storing l items can be expressed as

$$S(l) = h\sum_{i=0}^{M-1}\sum_{j=0}^{l-1}\psi_{i,j}\varphi_{i,l-j}(t_{i+1} - t_i), \tag{1}$$

where $\psi_{i,j}$ is probability to consume j items by moment t_i, expressed as

$$\psi_{i,j} = \begin{cases} 0, i = 0,\ j > 0 \\ 1, i = 0,\ j = 0 \\ \dfrac{\left(\sum_{m=0}^{i-1} \mu_m (t_{m+1} - t_m)\right)^j}{j!} e^{-\sum_{m=0}^{i-1} \mu_m (t_{m+1} - t_m)}, 0 < i \le M. \end{cases} \tag{2}$$

and $\varphi_{i,j}(x)$ is expected storage cost in the system during time x from moment t_i if inventory level at moment t_i was j, expressed by recursive formulae for μ_i > 0:

$$\varphi_{i,1}(x) = \frac{h\left(1 - e^{-\mu_i x}\right)}{\mu_i},$$

$$\varphi_{i,j}(x) = hjxe^{-\mu_i x} + \int_0^x (hjt + \varphi_{i,j-1}(x-t))d\left(1 - e^{-\mu_i t}\right), j > 1 \tag{3}$$

Applying Laplace transform $\Phi_{i,j}(\tau) = \int_0^\infty \varphi_{i,j}(x) e^{-\tau x} dx$ to (3), obtain

$$\Phi_{i,j}(\tau) = \frac{hj(\mu_i - \tau)}{\tau(\mu_i + \tau)^2} + \frac{\mu_i}{\mu_i + \tau}\Phi_{i,j-1}(\tau),$$

what can be expressed as direct formula

$$\Phi_{i,j}(\tau) = h\sum_{k=1}^{j} \frac{k\mu_i^{j-k}}{\tau(\mu_i + \tau)^{j-k+1}}.$$

So that, direct formula for $\varphi_{i,j}(x)$ can be found as

$$\varphi_{i,j}(x) = \begin{cases} hjx, \mu_i = 0 \\ \dfrac{h}{\mu_i}\displaystyle\sum_{k=1}^{j} k\left(1 - \sum_{n=0}^{j-k} \frac{(\mu_i x)^n}{n!} e^{-\mu_i x}\right), \mu_i > 0. \end{cases} \tag{4}$$

Net income accumulated in the system during time *T* can be calculated with formula

$$I(l) = \sum_{j=0}^{l-1} \psi_{M,j}(sj + q(l-j)) - pl - c + \left(1 - \sum_{j=0}^{l-1} \psi_{M,j}\right) sl - S(l). \tag{5}$$

3. Net Income Optimisation

Optimation task is formulated as

$$I(l) \to \max_l . \tag{6}$$

Search for the optimal value of l relies on convexity of $I(l)$. As can be seen, increments of $I(l)$

$$\Delta I(l) = I(l+1) - I(l) = s - p - (s-q)\sum_{i=0}^{l} \psi_{M,j} -$$

$$-h\sum_{i=0}^{M-1}\sum_{j=0}^{l} \frac{\psi_{i,j}}{\mu_i} \sum_{k=0}^{l-j}\left(1 - \sum_{n=0}^{l-j-k} \frac{(\mu_i(t_{i+1} - t_i))^n}{n!} e^{-\mu_i(t_{i+1} - t_i)}\right) \tag{7}$$

are decreasing monotonically with growth of l, because

$$\Delta I(l+1)-\Delta I(l)=\psi_{M,l+1}(q-s)-h\sum_{i=0}^{M-1}\sum_{j=0}^{l+1}\frac{\psi_{i,j}}{\mu_i}\left(1-\sum_{k=0}^{l-j}\frac{(\mu_i(t_{i+1}-t_i))^k}{k!}e^{-\mu_i(t_{i+1}-t_i)}\right)<0.$$

So that, value l^* that results in maximum value of net income, can be found with the use of formula (7) in binary search. This is the value that satisfies the conditions $\Delta I(l^*-1)>0,$ $\Delta I(l^*)<0.$

4. Unit Income Optimisation

Maximization of unit income $U(l)=I(l)/l$ can be useful in the situation if all items were consumed before time T and inventory can be replenished immediately. Optimation criteria assumes form

$$U(l)=s-p-\frac{1}{l}\left(c+S(l)+(s-q)\sum_{j=0}^{l-1}(l-j)\psi_{M,j}\right)\to\max_{l}. \tag{8}$$

Not all components of increments of unit costs are monotonic:

$$\Delta U(l)=U(l+1)-U(l)=\frac{1}{(l+1)l}\left(c-(s-q)\sum_{j=1}^{l}j\psi_{M,j}-\right.$$

$$\left.-h\sum_{i=0}^{M-1}\sum_{j=0}^{l}\frac{\psi_{i,j}}{\mu_i}\sum_{k=0}^{l-j}(l-k)\left(1-\sum_{n=0}^{l-j-k}\frac{(\mu_i(t_{i+1}-t_i))^n}{n!}e^{-\mu_i(t_{i+1}-t_i)}\right)\right). \tag{9}$$

Therefore binary search for optimal value of l is not applicable in general case. However taking into account the fact that $c-(s-q)\sum_{j=1}^{l}j\psi_{M,j}$ decreases monotonically to negative value $c-(s-q)\sum_{m=0}^{M-1}\mu_m(t_{m+1}-t_m)$ (it is negative because per order cost should be less than expected profit from selling average number of items during time T, otherwise net income would be negative) and the fact that the remaining component is always negative, search for optimal value l can be stopped when condition $c<(s-q)\sum_{j=1}^{l}j\psi_{M,j}$ is satisfied. So that, binary search can be used for this range of l values to find those that satisfy the conditions $\Delta U(l-1)>0,$ $\Delta U(l)<0,$ solution of (8) can be found among them by determining the largest value of $U(l).$

5. Numerical Example

The impact of fluctuations in demand on the optimal number of inventory level can be demonstrated with the following example. Let $h=2.7,$ $T=5,$ $q=6,$ $c=19,$ $s=28,$ $p=15$.

If $t_1=1,$ $\mu_0=16,$ $\mu_1=1,$ (demand type #1 in Table 1), optimal solution of task (6) is 17 and optimal solution of task (8) is 11. But if intervals and the correspondent intensities are reversed ($t_1=4,$ $\mu_0=1,$ $\mu_1=16,$ demand type #2 in Table 1), optimal solution of task (6) is 12 and optimal solution of task (8) is 5.

Table 1. Numerical example

	demand#1		demand#2	
l	$I(l)$	$U(l)$	$I(l)$	$U(l)$
1	-6.17	-6.17	-8.65	-8.65
2	6.49	3.25	-0.78	-0.39
3	18.99	6.33	5.01	1.67
4	31.31	7.83	9.18	2.30
5	43.47	8.69	12.25	**2.45**
6	55.45	9.24	14.61	2.43
7	67.25	9.61	16.51	2.36
8	78.84	9.86	18.10	2.26
9	90.19	10.02	19.45	2.16
10	101.21	10.12	20.54	2.05
11	111.77	**10.16**	21.32	1.94
12	121.66	10.14	**21.71**	1.81
13	130.61	10.05	21.55	1.66
14	138.28	9.88	20.63	1.47
15	144.28	9.62	18.71	1.25
16	148.21	9.26	15.51	0.97
17	**149.73**	8.81	10.76	0.63
18	148.53	8.25	4.22	0.23
19	144.46	7.60	-4.29	-0.23
20	137.49	6.87	-14.83	-0.74

References

1. Babitski, A., Lee, M. H. Inventory Level Optimization in the Conditions of Fluctuating Demand. In: *The 6th International Workshop on MSPT Proceedings*, 2006, pp. 15–16.
2. Cheng, F., Sethi, S. P. Optimality of State-Dependent (s,S) Policies in Inventory Models With Markov-Modulated Demand and Lost Sales, *Production and Operations Management*, Vol. 8(2), 1999, pp. 183–192.
3. Kalpakam, S. and Sapna, K. P. A (s,S) Perishable System with Arbitrary Distributed Lead-Times, *Opsearch*, Vol. 33, 1996, pp. 1–19.
4. Nahmias, S. Perishable Inventory Theory, *Operations Research*, Vol. 30, 1982, pp. 680–708.
5. Pal, M. An Inventory Model for Deteriorating Items When Demand Is Random, *Calcutta Statistical Association Bulletin*, Vol. 39, 1990, pp. 201–207.
6. Ravichandran, N. Probabilistic Analysis of a Continuous Review Perishable System with Markovian Demand. Erlangian Live and Non-Instantaneous Lead Times, *OR Spectrum*, Vol. 10, 1988, pp. 23–27.
7. Teng, J. T., Chern, M.-S., Chang Y.-L. Deterministic Inventory Lot-Size Models with Shortages for Fluctuating Demand and Unit Purchase Cost, *International Transactions in Operational Research*, Vol. 12, No 1, 2005, pp. 83–100.
8. Yin, R., Rajaram, K. Joint Pricing and Inventory Control with a Markovian Demand Model, 2005 – *http://repositories.cdlib.org/anderson/dotm/KR19*

Received on the 22nd of October 2007

Computer Modelling and New Technologies, 2007, Vol.11, No.4, 40–47
Transport and Telecommunication Institute, Lomonosov 1, LV-1019, Riga, Latvia

METHOD OF FINITE STATE MACHINE OPTIMAL IMPLEMENTATION TARGETING LOOK-UP-TABLE ARCHITECTURE

I. Lemberski [1]***, V. Gopeyenko*** [2]

[1]*Baltic International Academy*
Lomonosov 4, Riga LV-1019, Latvia
E-mail: igor.lemberski@bsa.edu.lv

[2]*Information Systems Management Institute*
Ludzas 91, Riga LV-1003, Latvia
E-mail: viktors.gopejenko@isma.lv

The method of state encoding for finite state machine optimal look-up-table implementation is offered. It is oriented on summary support minimization of logic functions. Encoding length is not restricted (may be non-minimal). As a criterion, the number of Configurable Logic Blocks (CLBs) is considered. Our method includes: 1) seed encoding; 2) step-by-step merging state variables to decrease summary support. Heuristic procedures are offered the reduce CPU time of the task to be solved. Experiments for XILINX 3090 architecture have been performed and compared with the results of NOVA, JEDI and LAX, where minimal encoding is accepted. It is shown that, in some cases, better implementation may be obtained for non- minimal encoding generated by our procedure.

Keywords: *finite state machine (FSM), look-up-table (LUT), field programmable logic array (FPGA), state encoding*

1. Introduction

Many papers deals with the problem of encoding with the emphasis on implementation. In [4, 6], encoding decreases dependance on present state variables in boolean representation. In [8, 12, 18, 19, 20], encoding methods for optimal two-level (programmable logic array – PLA) implementation targeting product term minimization have been developed. The methods [1, 2] for multi-level CMOS implementation are` oriented on literal minimization in the factored form. In [19], the method primary developed for two-level implementation is also evaluated in terms of literal number in multi-level representation. In [13], encoding method for VLSI implementation to minimize both gate and wiring area is considered.

For look-up-table (LUT) field programmable logic array (FPGA) implementation, different encodings with the goal to minimize number of literals are considered [3, 10, 16]. An optimisation criterion is expressed in terms of number of Configurable Logic Blocks (CLBs). However, typical LUT FPGA consists of the identical cooperating CLBs developed using LUTs, with the restricted number of inputs. Therefore, an important parameter that determines the number of CLBs to implement each function, is a support (set of inputs function depends on). In [9], the approach to support optimization while minimizing literals is considered but it refers to combinational functions. Our method of state encoding is aimed at minimizing summary support of boolean logic representation.

In the literature, there are contradictory conclusions concerning encoding length for which optimal LUT FPGA implementation is achieved. It is stated in [10] that minimal encoding gives better results. However, method in [16] demonstrates that 1-hot leads to better results.

In [12, 18, 20], the dichotomy – based approach is offered. It includes redundant encoding and its further minimization by merging state variables. In [14], redundant encoding compact representation exploiting hypercube symmetry is given.

Following [12, 18, 20], we start with the redundant encoding and try to reduce it paying the main attention to the support minimization. As a result, encoding of non-minimal length may be obtained.

2. Generic State Encoding

The basic idea of our encoding procedure is to construct a final state encoding by an iterative improvement of a generic and highly redundant seed state encoding [18]. The proper selection of such a state encoding is crucial for the success of the procedure.

A dichotomy {A:B} is a disjoint two block partition ($A \cap B = \varnothing$) of the subsets of states S ($A \subset S$, $B \subset S$). Each dichotomy is associated with a state variable in the encoding. It takes the same binary value for all symbols in the left block of the dichotomy, and opposite value for all the symbols in the right block. We arbitrarily assign 1 to left blocks and 0 to right ones. In a seed state encoding, we describe each variable by a dichotomy which contain exactly one state in each block: {a:b} with $a, b \in S$ (the set parentheses are omitted for simplicity).

Assume, we have n states, then we introduce $\binom{n}{2}$ state variables. Although the number of state variables grows with n^2 and is therefore high for practical applications, this encoding is a good starting point for the construction of the final encoding. This comes from the fact that encoding can be constructed by merging dichotomies associated with the state variables.

Two dichotomies d_1 = {A:B} and d_2 = {C:D} are compatible if: 1) $(A \cup C) \cap (B \cup D) = \varnothing$ or 2) $(A \cup D) \cap (B \cup C) = \varnothing$. Compatible dichotomies can be merged to a dichotomy 1) $d_{12} = \{A \cup C : B \cup D\}$ or 2) $d_{12} = \{A \cup D : B \cup C\}$. It is equivalent to merging two variables in the right or right and invert forms respectively.To construct a final encoding we sequentially merge compatible dichotomies which are selected according to the value of a suitable evaluation function.There is an equivalence between merging of dichotomies and state variables. The following strategies become more clear if we consider merging of variables. Clearly, the high number of variables to be merged requires low complexity heuristics.

3. Method Overview

As we don't restrict the final encoding length (may be non-minimal), it is expediently to start with the seed encoding and to reduce its length in such a way to minimize a cost function:

cost = $\sum_{\forall F_t \in F} |\text{sup}(F_t)|$, where $F = Y \cup Z$ – united set of FSM next state (Y) and output (Z) functions, sup (F_t) – support of funcion F_t (Fig. 1). The best (but impractical due to much computation time) way is the global optimization based on obtaining all possible encodings, both minimal and non-minimal, support computation for each one and choosing encoding which gives cost function minimal value.

Heuristic procedures reducing computation time are considered below.

1. We replace the global optimization is by the local one. It means that, in each step, we merge a pair of compatible variables to minimize (locally) the cost function. However, the total number of compatible pairs (we have to check) may be enormous, especially for large FSMs.

For instance, in the initial step, each pair of variables from the seed encoding is compatible. Therefore, $\frac{k(\text{seed})}{2}(k(\text{seed})-1)$, pairs should be considered and evaluated in terms of cost function, where $k(\text{seed}) = \binom{n}{2}$ – length of seed encoding. Furthermore, most pairs can be merged both in right and right and invert forms.

2. To reduce the number of pairs – candidates for merging – to be considered in each step, we give priority to compatible pairs belonging (simultaneously) to the supports of the maximal number of functions.

The higher this value is the more supports are reduced by merging variables. To find pairs, the weighted graph G is built. Its i-th node is associated with the state variable $y_i \in y$. Two nodes i, j are connected if y_i, y_j are compatible and: y_i, $y_j \in sup(F_t)$, $F_t \in F$. The edge (i, j) is weighted with w_{ij} determining the number functions F_t, such, that y_i, $y_j \in sup(F_t)$.

We start with a seed encoding (Fig. 1) and assign it to FSM symbolic states. As a result, boolean functions are generated (Fig. 1c). For each function, we compute a support (procedure based on solving a covering task will be described) (Fig. 1d). The graph G is built (Fig. 1e). It is easy to see that following pairs are candidates for merging: (y_1,y_2), (y_2,y_3), (y_2,y_4), (y_2,y_5), (y_3,y_4), (y_3,y_5), (y_4,y_5).

If $G = \varnothing$, remaining compatible pairs are considered. As candidates for merging, we extract those pairs, where, at least, one variable, belongs to the supports of the maximal number of function. In this step, weighted graph G′, where nodes are associated with the state variables is built. Two nodes i, j are connected if y_i,y_j are compatible. Value w_i determines the number of functions F_t, $F_t \in F$, such that $y_i \in sup(F_t)$.The edge (i, j) is weighted by value w_{ij} ,where $w_{ij} = max(w_i,w_j)$. State variables assigned to nodes connected by the edge (-s) with the maximal weight are considered for merging. In Fig. 1f, graph G′ is shown. (NOTE. In our example, $G \neq \varnothing$. Therefore, graph G′ is not used within the procedure of merging. However, we created G′ to illustrate its description).

3. The best way to find (among the pairs extracted using above heuristics) a pair minimizing the cost function is to merge each pair , compute support and compare values of the cost function obtained. However, this approach requires much CPU time. Therefore, we replace direct support computation by implicit evaluation of the effectiveness of merging (in terms of summary support). Let $F_t = (C_1,C_0)$ be a function, $F_t \in F$, described by two sets C_1, C_0 of ON – and OFF – cubes respectively. For instance, $F_4 = (C_1,C_0)$, where $C_1 = \{c_3,c_4,c_8\}$, $C_0 = \{c_6,c_7,c_{11}\}$ (Fig. 1c). Let $c_{l,v}(c_{f,v})$ be the v-th component of the cube $c_l(c_f)$, $c_l \in C_1$, $c_f \in C_0$. We introduce a distinction function dist:

$$dist\,(y_i,F_t) = |\{c_l,c_f\}: (c_{l,|x|+i} = 0\,)\ \&\ (c_{f,|x|+i} = 1)\ \text{or}\ (c_{l,|x|+i} = 1)\ \&\ (\,c_{f,|x|+i} = 0)|$$

In other words, dist determines how many pairs (c_l,c_f) of F_t are distinguished by the present state variable y_i. For instance (Fig. 1c): dist $(y_6,F_4) = 1$ (cubes c_8, c_{11} are distinguished). Let y_{ij} is a result of merging y_i,y_j. We merge a pair to maximize value $\sum_{\forall F_t : y_{ij} \in sup(F_t)} dist(y_{ij},F_t)$.

The motivation is as follows. The more (c_l,c_f) pairs are distinguished by a variable the less variables will be needed to distinguish remaining (c_l,c_f) pairs. Generating variables with distinguishing effect in a greedy manner leads to summary support minimization. Finally, if merging several pairs of variables leads to the same value of summary distinction, we compute support for each merging and accept one minimizing cost function.

4. Support Computation

Support computation is an important task of the procedure of merging. In each step, information on the support is used to extract variables to be considered as candidates for merging.

The idea of the basic procedure of support computation is similar to one presented in [5]. First, the selection matrix M is arranged for each function $F_t \in F$. The rows of M are associated with input and present state variables, but columns – with pairs (c_l,c_f). We put value 1 on the crosspoint of any row and the column associated with (c_l,c_f), if the variable associated with the row accepts opposite values in c_l and c_f. The number of ones in a column is denoted by cv (covering variables) and the number of ones in a row by cc (covered cubes). In Fig. 2, we present a selection matrix for function F_1 (Fig. 1c).

The task of support computation is formulated as one of searching a subminimal set of rows covering all the columns. The procedure from [11] (both exact and heuristic) may be applied.

However, we are aware that even its heuristic version (of polynomial complexity) is computationally expensive [5, 7], especially for tasks of large dimension. In our case, we have highly redundant encoding. The number of variables (matrix rows) is extremely large, especially at the beginning of encoding. Therefore, we need a fast heuristics to compute the support.Our procedure (named compute_support) includes reordering in accordance with values cc and cv and a pass of linear complexity. Within the pass, a support is generated by adding step-by-step a new variable.

```
.i 2               states
.o 1
.p 11                   1      1 1 1 - - -
.s 4                    2      0 - - 1 1 -
-0 1 1 0                3      - 0 - 0 - 1
11 1 1 0                4      - - 0 - 0 0
01 1 2 -
0- 2 2 1                          b)
11 2 1 0
10 2 3 1
1- 3 3 1  x = {x1,x2,...}- set of FSM inputs;
00 3 2 1  y = {y1,y2,...}- set of present state variables;
01 3 4 1  Y = {Y1,Y2,...} - set of next state functions;
0- 4 4 1  Z = {Z1,Z2,...} - set of FSM outputs;
11 4 3 1  F = Y∪Z;
                           a)
```

```
cubes                     F1F2F3F4F5F6F7 ∈F
  ↓     x1x2y1y2y3y4y5y6   Y1Y2Y3Y4Y5Y6Z1
  c1 =  - 0 1 1 1 - - -    1 1 1 - - - 0   sup(F1)={x1,x2,y1,y2}
  c2 =  1 1 1 1 1 - - -    1 1 1 - - - 0   sup(F2)={x2,y1,y2,y3,y4,y5}
  c3 =  0 1 1 1 1 - - -    0 - - 1 1 - -   sup(F3)={x1,y2,y3}
  c4 =  0 - 0 - - 1 1 -    0 - - 1 1 - 1   sup(F4)={x1}
  c5 =  1 1 0 - - 1 1 -    1 1 1 - - - 0   sup(F5)={x2,y2,y3,y4,y5,y6}
  c6 =  1 0 0 - - 1 1 -    - 0 - 0 - 1 1   sup(F6)={x1}
  c7 =  1 - - 0 - 0 - 1    - 0 - 0 - 1 1   sup(F7)={x1,x2,y1,y2,
  c8 =  0 0 - 0 - 0 - 1    0 - - 1 1 – 1            y3,y4,y5}
  c9 =  0 1 - 0 - 0 - 1    - - 0 - 0 0 1
  c10 = 0 - - - 0 - 0 0    - - 0 - 0 0 1
  c11 = 1 1 - - 0 - 0 0    - 0 - 0 - 1 1

          c)                               d)
```

```
yi   i  1 2 3 4 5 6          yi   i  1 2 3 4 5 6

  1  0 3 2 2 2 0               1  0 0 0 0 0 3   w1 = 3
  2  0 0 3 3 3 1               2  0 0 0 0 0 0   w6 = 1
  3  0 0 0 3 3 1               3  0 0 0 0 0 0   w16 = 3
  4  0 0 0 0 3 1               4  0 0 0 0 0 0
  5  0 0 0 0 0 1               5  0 0 0 0 0 0
  6  0 0 0 0 0 0               6  0 0 0 0 0 0

          e)                            f)
```

Figure 1. Example: a) FSM in a .kiss format [17]; b) seed encoding; c) logic functions (after substituting symbolic states for seed encoding and d) their supports; e) weight matrix of graph G; f) weight matrix of graph G′

1. The rows are ordered by decreasing cc and columns – by increasing cv.
2. First ordered and not covered column is extracted and first ordered row p covering it, is selected. The variable associated with the row p is put into sup (F_t).
3. All the columns having 1 on the crosspoint of the row p are marked as covered.
4. Do from item 2 unless all the columns are covered.

However, this procedure may generate a support with the redundant number of variables.

Example. Consider a matrix M.

	1	2	3	4	5	6	cc
x_1	1	1	1	1			4
y_1	1	1			1		3
y_2			1	1		1	3
y_3					1	1	2
cv	2	2	2	2	2	2	

Applying the above procedure, variables x_1,y_1,y_2 are sequentially included in the support.

Nevertheless, variable x_1 is redundant and should be excluded. Therefore, in the second phase, the procedure to exclude redundant variables (named exclude_redundant) is applied. Within it, we exclude from the support a variable and check, if remaining variables cover all the columns of M. If yes, we continue checking for the next variable until all variables are considered. Otherwise, we return a variable into support and continue checking for the next variable. After applying the above procedures to matrix M (Fig. 2), sup(F_1) is obtained (Fig. 1d).

F_1=(C_1,C_0); C_1={c_1,c_2,c_5}, C_0={c_3,c_4,c_8}

cube pairs : (c_1,c_3) (c_1,c_4) (c_1,c_8) (c_2,c_3) (c_2,c_4) (c_2,c_8) (c_5,c_3) (c_5,c_4) (c_5,c_8)

	1	2	3	4	5	6	7	8	9	cc
x_1	0	0	0	1	1	1	1	1	1	6
x_2	1	0	0	0	0	1	0	0	1	3
y_1	0	1	0	0	1	0	1	0	0	3
y_2	0	0	1	0	0	1	0	0	0	2
y_3	0	0	0	0	0	0	0	0	0	0
y_4	0	0	0	0	0	0	0	0	1	1
y_5	0	0	0	0	0	0	0	0	0	0
y_6	0	0	0	0	0	0	0	0	0	0
cv	1	1	1	1	2	3	2	1	3	

Figure 2. Selection matrix for F_1

5. State Encoding Procedure

Our procedure (Fig. 3) starts with assigning a seed encoding to FSM states. In each step of merging, graph G (if empty, graph G′) is built and a pair minimizing a cost function is merged (procedure merge (G)). After that, procedure compute_support is applied followed by the procedure exclude_redundant.The loop is repeated until there is, at least, one compatible pair.

```
encoding_minimization (seed_encoding)

 {
      while(G ≠∅ or G′ ≠ ∅)
        {
          if (G ≠∅ ) then
          merge(G)
          else
          merge(G′);
        }
          return(encoding);
 }
```

a)

```
merge(G)

{
 extract a subset P of pairs (y_i.y_j) with the maximal w_ij;
```

extract a subset $P' \subseteq P$ of pairs maximizing $\sum_{\forall F_t : y_{ij} \in sup(F_t)} dist(y_{ij}, F_t)$;

from a subset P', merge the pair, for which $\sum_{\forall Ft \in F} sup(Ft)$ is minimal;

```
 compute_support;
 exclude_redundant;
 return (∀Ft∈F:sup(Ft));
}
```

b)

Figure 3. Procedures of a) encoding and b) merging

6. Implementation and Experimental Results

The procedure described has been implemented in C language under SOLARIS 2.4 (SunOS 5.4) on Sun SPARCstation 4,85 Mhz processor, as a module MINISUP and incorporated into SIS. Within SIS, MINISUP performs the state encoding with the emphasis on LUT FPGA implementation. We tested our procedure synthesizing examples onto XILINX 3090 architecture. For decomposition and mapping, following script has been used:

```
sweep
so script.boolean
xl_part_coll -m -g 2
xl_coll_ck
xl_partition -m
full_simplify
xl_imp
xl_partition -t
xl_merge -o doc -v
```

To compare result we perfomed encoding of minimal length (because it is stated in the literature that better result is obtained for minimal encoding) using NOVA (option -e ih) and JEDI. The same script has been used for decomposition and mapping.

From the literature, we reproduced the results obtained by LAX [1], where minimal encoding is generated. The results are presented in the Table 1.

Table 1. Experimental results

Example	#state	#enc_m	MINISUP		JEDI	NOVA	LAX
			#enc	#CLB	#CLB	#CLB	#CLB
bbara	10	4	4	**12**	**12**	**12**	14
bbsse	16	4	6	**22**	26	25	27
cse	16	4	7	46	42	**41**	54
s1	20	5	6	**40**	43	46	48
ex1	20	5	7	**40**	48	43	61
keyb	19	5	6	35	47	**33**	42
styr	30	5	7	98	**90**	94	100
dk16	27	5	6	53	53	49	**32**
s820	25	5	9	**51**	68	53	75
s832	25	5	9	**47**	56	56	58
sand	32	5	8	**109**	112	116	117

First two columns give the name of examples and number of states. Next columns present: minimal encoding length (#enc_m) accepted in JEDI, NOVA, LAX, number of CLBs (#CLB) together with encoding length (#enc) for MINISUP. In Table 1, one can see that for almost all examples MINISUP generates non-minimal encoding. However, the results of mapping for many examples are better than ones obtained by JEDI, NOVA, LAX.

Conclusions

We offered a summary support as a criterion for optimal state encoding while LUT FPGA implementation. Our approach bases on the seed encoding and its minimization, thus to decrease summary support in boolean logic representation.The heuristic procedure is offered to reduce CPU time for a task to be solved. The procedure has been implemented (module MINISUP) and incorporated into SIS. The experiment demonstrates that for many examples, MINISUP gives better results (in the most cases, for non-minimal encoding), than JEDI, NOVA and LAX. However, procedure remains computationally expensive for examples having many states. It is because the number of variables rapidly grows (with n^2) while increasing number n of states.

References

1. Burgun, L, Dictus, N., Prado Lopes, E., Sazwary, C. A Unified Approach for FSM Synthesis on FPGA Architecture. In: *Proc. 20th EUROMICRO Conf., Liverpool, September, 1994*. Liverpool, 1994, pp. 660–668.
2. Devadas, S., et al. MUSTANG: State Assignment of Finite State Machines Targeting Multilevel Logic Implementation, *IEEE Tran. CAD,* Vol. 7, No 12, 1988, pp. 1290–1300.
3. Feske, K., Mulka, S., Franke, G., Koegst, M. *FSM-Synthese Durch Linear Partitionierung für Lookup-Table-basierte FPGAs: Report: SFB 358-B1-1/95*. January 1995.
4. Hartmanis, J., Stearns, R. E. *Algebraic Structure. Theory of Sequential Machines*. New Jersey: Prentice Hall, 1966. 211 p.
5. Fujita, M., Matsunaga, Y. Multi-level Logic Minimization Based on Minimal Support and its Application to the Minimization of Look-up Table Type FPGAs, *Proc. ICCAD*, November, 1991, pp. 560–563.
6. Karp, R. Some Technigue for State Assignment of Synchronous Sequential Machines, *IEEE Trans. Elect. Comput.,* EC-13, 1964, pp. 507–518.

7. Lin, B. Efficient Symbolic Support Manipulation, *Proc. ICCD*, October, 1993, pp. 513–516.
8. Lin, B., Newton, A. Synthesis of Multi-Level Logic from Symbolic High-Level Description Languages, *Proc. Int. Conf. on VLSI*, August, 1989, pp. 187–196.
9. Lu, A., Dagless, E., Saul, J. Tradeoff Literals Against Support for Logic Synthesis of LUT-based FPGAs, *IEE Proc.- Comput. Digit. Tech.,* Vol. 143, No 2, 1996, pp. 111–119.
10. Murgai, R., Brayton, R. K., Sangiovanni-Vincentelli, A. Sequential Synthesis for Table Look Up Programmable Gate Arrays, *Proc. 30th DAC*, 1998, pp. 224–229.
11. Rudell, R. L., Sangiovanni-Vincentelli, A. Multiple-Valued Minimization for PLA Optimitation. *IEEE Trans. CAD*, Vol. 6, No 5, 1987, pp. 727–750.
12. Saldanha, A., Villa, T., Brayton, R. K., Sangiovanni-Vincentelli, A. L. Satisfaction of Input and Output Encoding Constraints, *IEEE Trans. CAD,* Vol. 13, No 5, 1994, pp. 589–602.
13. Saucier, G., Duff, C., Poirot, F. State Assignment of Controller for Optimal Area Implementation, *Proc. EDAC*, March, 1990, pp. 547–551.
14. Saucier, G., Duff, C., Poirot, F. State Assignment Using a New Embedding Method Based on an Intersecting Cube Theory. In: *Proc. 26th DAC, Las Vegas, June, 1989*. Las Vegas, 1989, pp. 321–326.
15. Saucier, G., Crastes de Paulet, M., Sicard, P. ASYL: A Rule-Based System for Controller Synthesis. *IEEE Trans. CAD,* Vol. 6, No 6, 1987.
16. Schlag, M., Chan, P. K., Kong, J. Empirical Evaluation of Multilevel Logic Minimization Tools for an FPGA Technology. In: *Int. Workshop on Field Programmable Logic and Applications*. Oxford, 1991.
17. Sentovich, E. M., et al. SIS: A System for Sequential Circuit Synthesis Electronic Research Laboratory. *Memorandum No. UCB/ERL M 92/41,* 1992.
18. Shi, C. J., Brzozowski, J. A. Efficient Constrained Encoding for VLSI Sequential Logic Synthesis, *EURODAC'92*, 1992, pp. 266–271.
19. Villa, T., Sangiovanni-Vincentelli, A. NOVA_State Assignment of Finite State Machines for Optimal Two-Level Logic Implementation, *IEEE Trans. CAD,* Vol. 9, No 9, 1990, pp. 905–924.
20. Yang, S., Ciesielski, M. J. Optimum and Suboptimum Algorithms for Input Encoding and Its Relationship to Logic Minimization, *IEEE Trans. CAD,* Vol. 10, No 1, 1991, pp. 4–12.

Received on the 1st of November 2007

Computer Modelling and New Technologies, 2007, Vol.11, No.4, 48–51
Transport and Telecommunication Institute, Lomonosov 1, LV-1019, Riga, Latvia

XILINX4000 ARCHITECTURE-DRIVEN BOOLEAN NETWORK RE-DECOMPOSITION TARGETING PERFORMANCE

I. Lemberski [1]***, V. Gopeyenko*** [2]

[1]*Baltic International Academy*
Lomonosov 4, Riga, LV-1019, Latvia
E-mail: igor.lemberski@bsa.edu.lv

[2]*Information Systems Management Institute*
Ludzas 91, Riga, LV-1003, Latvia
E-mail: viktors.gopejenko@isma.lv

Architecture-driven (instead of look-up-table-LUT-driven) method of Boolean functions logic synthesis for the performance is proposed. It takes XILINX4000 architectural features (heterogeneous LUTs of 3 and 4 inputs) into account and includes two-step decomposition. In the first step, two-level logic representation is transformed into a graph of at most 4 fan-in nodes (after this step, each node can be mapped onto 4 input LUT). In the second step, selected 4 fan-in nodes within a critical path are re-decomposed into 3 fan-in nodes to ensure mapping onto 3 input LUTs. The re-decomposition task is formulated as substituting of node two fan-ins for exactly one fan-in, either existing node or one especially created, is considered as a fan-in to be substituted for. The extended PLA format describing a multi-level Boolean network is proposed. Based on this description, substituting is formulated in terms of a covering task.

Keywords: *XILINX, Boolean function, decomposition*

1. Introduction

LUT FPGA-based logic synthesis methods developed don't consider architectural features of LUT FPGAs [1] (as exception, the paper [2] can be mentioned where decomposition methodology is oriented on XILINX3000 and 4000 structure). Instead of it, two-level logic is transformed into the set of sub-functions, where each one depends on the restricted number of inputs. Usually, this number doesn't exceed the number of LUT inputs. Therefore, each sub-function can be implemented onto one LUT. To optimise, methods of minimization of the number of sub-functions are considered (LUT-driven synthesis).

However, FPGAs such approach may be not optimal for FPGAs having several heterogeneous LUTs with specific connections between them.

In this paper, architecture – driven (instead of LUT-driven) approach to logic synthesis for the performance targeting XILINX4000 structure is offered. It takes heterogeneous LUT architecture (two type LUTs of 3 and 4 inputs) into account and includes two-step decomposition to ensure minimization of the number of logic blocks (rather than the number of LUTs) within the critical path.

2. XILINX4000 Architecture and Mapping

XILINX4000 logic structure bases on cooperating Configurable Logic Blocks (CLBs) (Fig. 1) (minimum 64 CLBs per module). A CLB consists of 3 look-up-tables (LUTs) (function generators), namely, LUT of type F, G and H (further, F-LUT, G-LUT, H-LUT).

F-LUT and G-LUT can implement any function of up to 4 variables. H-LUT implements up to 3 input function, where two inputs are F-LUT and G-LUT outputs (internal inputs), but the third input is external. CLB has two external outputs. Any two of three LUT outputs can be connected to CLB outputs either directly or through flip-flops (FF).

Let decomposed logic functions PO = f (PI) be represented by a directed a-cyclic graph D: D = (PI, PO, G), where PI, PO – set of primary input and output nodes, G – set of nodes obtained as a result of decomposition. We call the node $g_{i-1} \in RI \cup G$ as a fan-in of the node $g_i \in G \cup PO$, if there is an arc from g_{i-1} to g_i (Fig. 2). Let sup (g_i), $g_i \in G \cup PO$, be the set of g_i fan-ins. Suppose, that $|\sup(g_i)| \le 4$ for each $g_i \in G \cup PO$. Therefore, each node can be mapped onto F- or G- LUT (Fig. 1). We call such mapping as trivial. Let τ be CLB propagation delay. Then, the critical path propagation delay d: $d = n \times \tau$, where n – number of nodes within the critical path. To increase performance, H-LUTs should be used for mapping. Suppose: $|\sup(g_i)| \le 3$ (Fig. 2) and g_{i-1}, g_i are in the critical path. The nodes g_{i-1} and g_i can be mapped onto the same CLB (F- or G- LUT and H-LUT respectively). As the CLB delay remains the same, the propagation delay within the critical path is reduced: $d = (n – 1)*\tau$.

Therefore, our goal is to develop decomposition technique, to map as many as possible nodes within the critical path onto H-LUTs. The approach may include generating a graph where each node has at most 3 fan-ins. However, in this case, logic capacities of 4-input (F-, G-) LUTs will be not used in the best way. As a result, the total number of nodes may be increased.

To ensure optimal implementation we propose two step decomposition. In the first step, a graph D: $|\sup(g_i)| \le 4$, $g_i \in G \cup PO$ is obtained and optimised for area and speed using SIS tool [4]. In the second step, selected 4 fan-in nodes are re-decomposed into 3 fan-in nodes to map them onto H-LUTs.

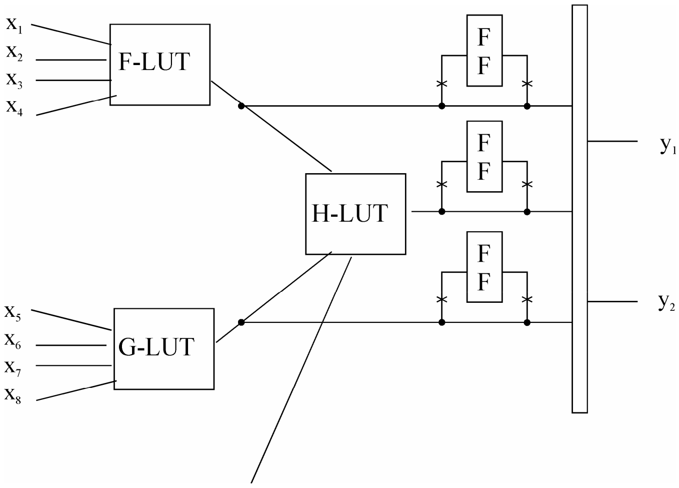


Figure 1. Configurable logic block structure

3. Re-decomposition via Substituting

Consider a 4 fan-in node (Fig. 2). To transform it into a 3 fan-in one, we substitute its 2 fan-ins for an exactly one fan-in while preserving functionality and a-cyclic.

While substituting, two types of nodes may be considered: substituting for an existing node g_j, $g_j \in G$ and a new node: $g_j \notin G$ especially created and introduced into the graph: $G = G \cup \{g_j\}$. Clear, that for a new node: $|\sup(g_j)| \le 4$. However, in the last case, an additional node is introduced and the total number of nodes is increased. Therefore, we give preference to substituting for an existing node. If such a node can't be found, substituting for a new node is considered.

4. Substituting as a Covering Task

Usually, two-level logic is represented by the PLA table (Fig. 3a) of dimension p∗(name), where n = |PI| – number of primary inputs, m = |PO| – number of functions, p – number of input vectors for which functions are described (for simplification, two-valued vectors are considered). The main advantage of the PLA format is joined description of all the functions. It ensures compactness, clearness, convenience for analysis.

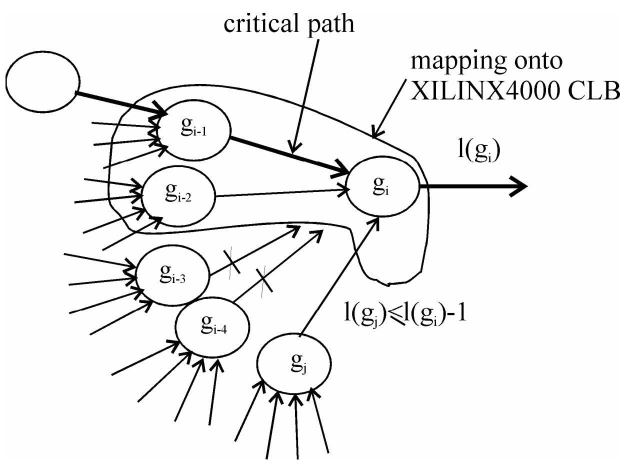


Figure 2. *Substituting for a node*

After decomposition by the SIS, the set G of node sub-functions is generated (Fig. 3b). Each sub-function is described separately, by ON-set vectors (complement vectors belong to OFF-set) together with the list of inputs and output names (keyword .names). Taking the advantage of the PLA format into account, it is expediently to describe sub-functions by the extended PLA table using as a starting point, a table of two-level representation.

We represent the decomposed logic by the extended PLA table of dimension p*(n + |G| + m), where |G| – number of additional columns describing sub-functions from the set G.

The procedure of the PLA table extension is as follows. In each step, a sub-function g_i is described in the PLA table, $g_i \in G/G_i$, $sup(g_i) \subset (PI \cup G_i)$, G_i – set of sub-functions already described in the PLA table during the previous steps. To describe g_i , for each PLA table row, an embedded vector of sup (g_i) variables is checked if it is covered by a vector from the sub-function g_i ON-set. If yes, the value 1 is put on the cross-point of this row and the column g_i. Otherwise, the value 0 is put. For example (Fig. 3), to describe g_1 , we check PLA table embedded vectors of sup (g_1) = {x 1, x 3} variables and put value 1 in the 2nd (embedded vector 01 is covered by ON – set vector: 1) and in the 4–8th rows (Fig. 3c).

Let $l(g_i)$ is a node g_i level (maximal number of nodes between PI and g_i). In case of substituting for an existing node (Fig. 2), we search for a node to be substituted for, among the set $S_i = PI \cup G/g_i$. In addition, this node has to satisfy the conditions as follows: after substituting, the graph remains a-cyclic, node level doesn't exceed $l(g_i) – 1$.

Let $S'_i \subset S_i$ be the set satisfying above conditions. The selection matrix is arranged where columns are associated with the sub-function g_i all possible (ON-, OFF-) seed dichotomies [5], but rows – with the variables from the set S'_i. To substitute, we try to cover columns uncovered by g_{i-1}, g_{i-2} , using exactly one variable from the set S'_I/g_{i-3} , g_{i-4}.. To substitute for a new node, (ON-, OFF-) seed dichotomies not covered by g_{i-1}, g_{i-2} are merged to one dichotomy describing the sub-function g_i and the set S'_i is generated. To compute $sup(g_i) \subset S'_i$, covering procedure [5] is applied and solution is accepted, if $|sup(g_i)| \leq 4$.

x_1	x_2	x_3	F_1
0	0	0	1
0	0	1	0
0	1	0	1
0	1	1	0
1	0	0	1
1	0	1	0
1	1	0	0
1	1	1	1

a)

```
. names x1 x3 g1
1 - 1
- 1 1
. names x2 x3 g2
1 1 1
. names g1 g2 F1
0 0 1
1 1 1
```

b)

x_1	x_2	x_3	g_1	g_2	F_3
0	0	0	0	0	1
0	0	1	1	0	0
0	1	0	0	0	1
1	1	1	1	1	1
1	0	0	1	0	0
1	0	1	1	0	0
1	1	0	1	0	0
1	1	1	1	1	1

c)

Figure 3. PLA table (a), nodes description (b), extended PLA table (c)

5. Experimental Results

XILINX4000 architecture driven re-decomposition procedure for performance is implemented and tested on some FSM examples (see Table 1).

Table 1. Experimental results

XILINX4000				
Example	Re-decomposition		Decomposition	
	#CLB c.p.	#CLB total	#CLB c.p.	#CLB total
ex2	2	10	4	12
ex3	2	5	4	5
ex4	2	13	3	14
ex5	2	4	3	5
ex6	3	26	5	32
ex7	2	4	3	4
beecount	2	7	4	8
dk14	2	15	4	17
coprodes	2	8	3	10
Total:	23	128	43	146
Percentage:	53,5%	87,7%	100%	100%

For state encoding, the procedure [5] is used. In Table 1, one can see that the re-decomposition for speed reduces significantly the CLB number within the critical path (c.p).

Conclusions

The architecture-driven (instead of LUT-driven) approach to logic synthesis for the performance targeting XILINX4000 structure is offered. It takes heterogeneous LUT architecture (two type LUTs of 3 and 4 inputs) into account and includes 1) decomposition for speed and area using SIS tool; as a result, the Boolean network is obtained; 2) within the critical path, selected 4 fan-in nodes are re-decomposed into 3 fan-in nodes to map them onto H-LUTs. The experimental results showed that the re-decomposition for performance reduces significantly both CLB total number and CLB number within the critical path.

References

1. Cong, J, Ding, Y. Combinational Logic Synthesis for LUT Based Field Programmable Gate arrays, *ACM Trans. on Design Automation of Electronic Systems,* Vol. 1, No 2, 1996, pp. 145–204.
2. Abouzeid, P., Babba, B. Crastes de Paulet, M., Saucier, G. Input-Driven Partitioning Methods and Application to Synthesis on Table-Look-up-Based FPGA's, *IEEE Trans. CAD,* Vol. 12, No 7, 1993, pp. 913–925.
3. Murgai, R., Brayton, R., Sangiovanni-Vincentelli, A. *Logic Synthesis for Field-Programmable Gate Arrays*. Kluwer Academic Publishers, 1995.
4. Sentovich, E., et al. SIS: a System for Sequential Circuit Synthesis. In: *Electronic Research Laboratory Memorandum No. UCB/ERL M92/41*, 1992.
5. Grass, W., Lemberski, I. Support-Based State Encoding targeting FSM Optimal LUT FPGA Implementation. In: *Int. Workshop on Logic and Architecture Synthesis*. Grenoble**,** France, 1997, pp. 97–104.

Received on the 1st of November 2007

Computer Modelling and New Technologies, 2007, Vol.11, No.4, 52–58
Transport and Telecommunication Institute, Lomonosov 1, LV-1019, Riga, Latvia

MEDIA AND ACTIVE DOTS OPTIMAL DEVELOPMENT STRATEGIES

R. I. Muhamediyev, YU. N. Shunin, V. I. Gopeyenko

Information Systems Management Institute
Dept. of Natural Sciences and Computer Technologies
Lomonosov 1, LV-1019, Riga, Latvia
ravil@inbox.lv, shunin@isma.lv, via@latnet.lv

Media dots are metaphor mean elements that uses for reception and reproduction of the media data. Potentially they can be used in systems of protection and supervision, "smart house" systems, media-content control systems of an educational institution etc. The examples of existing complexes of video-supervision and projects of some systems are considered. The assumption on general character of processes of reception and processing of media data of various purpose systems is done. The opportunity of creation of the unified means of the user interface and systems of parallel working programs of agents processing, which are carrying out a recognition of situations (intellectual detecting) and formation of reactions, are discussed. The tasks and problems of the further researches are formulated.

Keywords: *media dots, active dots, intellectual agent, intellectual detecting, smart house, user graphic interface, natural behaviour strategies, security signal system*

We need men who can dream of things that never were.
John F. Kennedy, 35th President of the USA

1. Introduction

In connection with avalanche growth of volumes of the transmitted information and capacity of carriers of the data there is an urgent task of formation of the general concept of processing of large-scale flows media data. The concept of media dots, which are understood as every possible source of media data, is formulated. At the same time, in discussion of problems of objects management a metaphor of active dots uniting various management executive devices can be useful.

The association of various sources of media data within the framework of the uniform concept allows to generalize a task of construction of contextually dependent systems of the graphic user interface (GUI) both intellectual means of the control and management. These means or, otherwise, intellectual agents, being in media dots space, can solve tasks appropriate to a special-purpose designation of systems and to have by many similar features. Hence, the various systems based on use of means of a video and audio-control can be managed by functionally close program systems. One of kinds of similar technical systems (but not unique), are IP cameras. It is quite possible to consider a community of the program agents, placed in a network and general interfaces and called as required through media dots system. Thus, the systems of recognition (for example, image recognition), systems of video-supervision, systems of the media-content collection in educational institutions, etc. can operate. Alongside with it, the intellectual agents can be characterized by some own behaviour ensuring an achievement of the purposes facing a system. The behaviour of similar systems can have much common with a behaviour of artificial alive essences-animates.

2. Avalanche Growth of Volumes of the Transmitted Information and Capacities of Carriers

The growth of capacities of information carriers is an objective process, which is observed during all computer history. In essence, this process is one more display of the famous Moore's law [1]. Periodically, alongside with an avalanche growth of volumes of the transmitted and stored information, the expansion of number and change of popularity of carriers of the data are observed. The increase of its quality and volume leads to a necessity to store and to process Tbytes and Pbytes [2]. From the one side this is a treasure, but, from the other side, it is a "breed", in which to reach useful "layer" no so easy.

One of particular ways of the task solution of the data large volumes is the segmentation or personification of Internet-space [3]. However, the business is not limited in information searching in the Internet. There are information media-sources, which transfer the data in a real time scale, for example, technological computers, measuring instruments and sensors, IP-cameras, video-servers etc.

3. Media-Active Dots

Let's consider various media data sources.

1. Systems of video-supervision, which can be advanced up to monitoring systems of city or area (Fig. 1). The example of the project of the monitoring system of area of city is described, for example, in [4].
2. The systems of “smart” buildings too can be a rather volumetric media-data source) are considered [5, 6]. The monitoring system and management of a “smart” building can contain some media-data sources intended for supervision over a condition of rooms and technical subsystems, and also management blocks (Fig. 2).
3. The project of a mobile system of supervision (Fig. 3).
4. The systems of security supervision, in which increasing popularity win IP-cameras [7, 8]. The modern variants of systems of video-supervision and security signal system include algorithms of recognition of the people, cars and some situations (sharp acceleration of movement, fall of the man etc.) [9].
5. The monitoring system of warehouses and rooms, with adjustable zones of the control. One of popular tasks in these systems is the systems of the recognition of faces [10].
6. Systems of supervision and control used in educational institutions, for example, at schools [11], where the cameras are used for the control of study rooms and sports grounds.
7. Project of a control system educational content of higher-school (Fig. 4). The general concept of a system is illustrated on Fig. 5. Generally, an educational media-content, formed by the teacher can be transferred directly in a network in a real time scale and saved on server. The isometric projection of building used for control learning content. User selects floor of building by mouse strike (Fig. 4, fragment A; Fig. 6) then he selects media dot on the floor scheme (Fig. 4, fragment B).

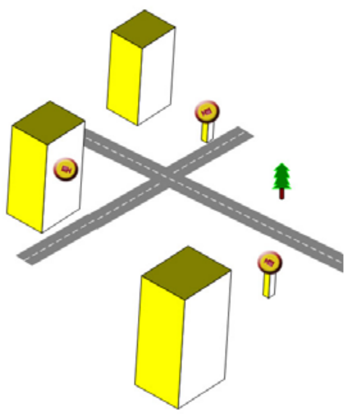


Figure 1. Town control system

Figure 2. Media-active dots in a SMART HOUSE. Green signs are media dots, red ones is active dots

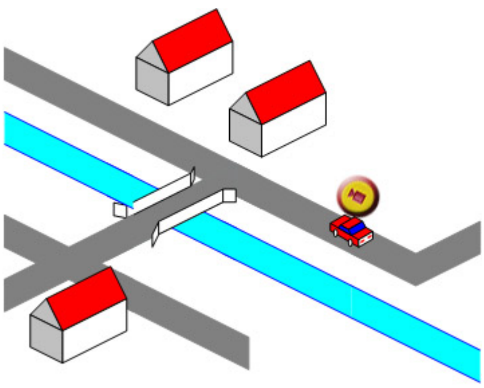


Figure 3. Mobile M-dots

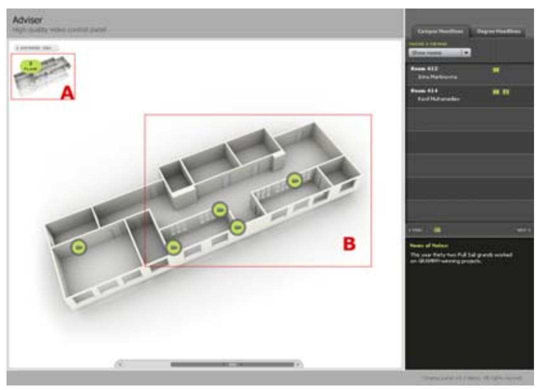


Figure 4. Graphical user interface for control media dots in higher school

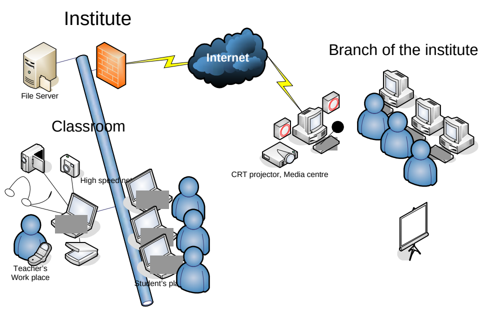


Figure 5. Concept of higher school e-Learning system

Figure 6. Floors schema

The various media-information sources can be generalized, using media-dots concept.

Let's name *media or multi-media dot* (*M-dot*) an element, which is capable to accept and/or reproduce media data. Under media data we understand any kinds of the data, which can be perceived or are reproduced by modern computer and communication means and their combinations. First of all, we shall mean audiovisual data – a video and sound. On the other hand, the *active dot (A-dot*) is an element capable to operate by object or system. This pair we shall name MA- dot or simply MA, at least, two dots of view on MA are possible. A practical variety of media-information sources can be presented by particular examples.

Examples of M- dots

- WEB-pages
- WEB-cameras
- IP-cameras
- Computer (e.g., information on its screen)
- Measuring instruments and sensors

Examples of A-dots

- Control facility of a "smart" house; regulators of heat providing devices, illumination devices
- Computers
- Alarm devices
- Restriction access devices
- Control facility mobile devices

There is a natural question: whether the association of so diverse systems and media-sources is justified within the framework of general paradigm? On our sight, in all listed examples the needs of practical use are rather similar. In all cases the reception and processing of media data, and, as a rule, in a real time scale is meant. Hence, in the first, the contextually dependent means of GUI, ensuring the simplicity of media-information sources are necessary. The example of the project similar GUI for media-content management of higher school is described below. In the second, the intellectual means of the control and managements allowing to lower volume of the process-able data up to a "reasonable" level are necessary by allocating its only that corresponds to a special-purpose systems designation or user needs. Such means or intellectual agents can doing autonomous, prepare timetable of viewing for users and detecting systems.

4. IP-Cameras and Intellectual Detecting

The IP-camera supplied with a microphone to become a source of a video and audio-information. At presence of means of wireless communication the IP-camera is capable to work on distance up to several tens meters from an access point. Overview of cameras considered in [13].

For cameras management the software as commercial [14, 15, 16], and conditional free-of-charge [17] are used. Many manufacturers deliver the advanced software together with IP by cameras. As a rule, the cameras are capable to react to the following events [18]:

- ❖ Change of a level of a sound, perceived by a microphone;
- ❖ Switching a source of a signal (for example, from the infra-red camera on a usual video-camera);
- ❖ Movement;
- ❖ Change of intensity of light;
- ❖ Approach of given time intervals.

The camera makes record of video signal or sequence of the staff. In other words, the modern cameras are capable to execute detecting some signals. The intellectual system should be capable on something greater. *Intellectual detecting* we shall define as an ability of system to react not only to signals, but also on images. Such a task can be solved by the *agent of intellectual detecting* expanding list of recognized situations in a field of sight of the camera. Generally, the agent can have a memory and an ability of recognition both static and dynamic images. In this case media recognition of dynamic images becomes basic. The first step in this direction is the Advanced Video Motion Detection and Unattended Object Detection technology, used in cameras of new generation [19]. As the development of the movement detector (video motion detector) the detector of criminal situations can be offered. The prototype is the system described in [8], where are applied both motionless, and mobile cameras. It is possible to speak also about *detectors technogenic situations*, which could serve for revealing dangerous situations, for example, connected with automobile movement, risk situations at the airports, at factories etc., and also unusual condition distinguished from habitual, that can be an attribute of danger etc. The *agent of intellectual search* in media-dots space can serve a means of recognition of similar images in media-dots space, searching the definite objects etc.

5. Strategies of Media-Data Processing. *Model of Random Work Scanning of Security Space*

The systems of security supervision in which IP-cameras win the increasing popularity. But there are some serious technical problems in data volumes processing. The ways of decisions of these problems, which we can point out as more effective, are as follows:

- Development of optimal strategies of servers' (or processor's) charging, designated for software agents. The choosing of processors, when the charging and distance for a source are taken into account, can diminish the probability of troubles in data processing;
- Self-teaching of agents. Using of various strategies of agents' self-teaching, including an imitation of natural ways of teaching [20, 21], will allow to adopt them for changes in media-ponts space;
- Using of strategies of transmitted media-information volumes (see 3.1.). The first approach can be concluded in dislocation of agents of intellectual detection directly in collection points of media-data. Thus, the translation of a part of algorithm of data processing closer to a point of data receiving can diminish the data amount for the next stage of data processing. Nowadays, there are examples of video-cameras with mounted software, which detect the movement and, recognize the available subject [15];
- The second approach is concluded in using of pseudo-chaotic strategies of media-points activation which is similar to a natural way of vision field of human eye. E.g., in decision of security video-observation we could not fulfil the full-scale processing of the whole data flux, coming from all observation cameras. It can be enough to analyse the particular images of video-data from various cameras, which are chosen in random way. This approach is demonstrated in Fig. 7, 8 and 9. In particular, Fig. 7 shows the "linear set" (security wall)" model. Each camera scans the observation field randomly changing the orientation angle $\varphi(t)$. The same way is used for "radial set" model (Fig. 8). Fig. 9 demonstrates the algorithm of movement a video-camera. The orientation angle $\varphi(t)$ is change according the logistic map law (e.g. the well-known logistic map of Verhulst) [22, 23]: $x_{t+1} = rx_t(1-x_t)$.

For 3 < r <4 the generation of chaos is observed. Thus, the algorithm of random walk scanning of active media dot is: $\Delta\varphi_{t+1} = 90^\circ[(rx_t(1-x_t)-0,5]$.

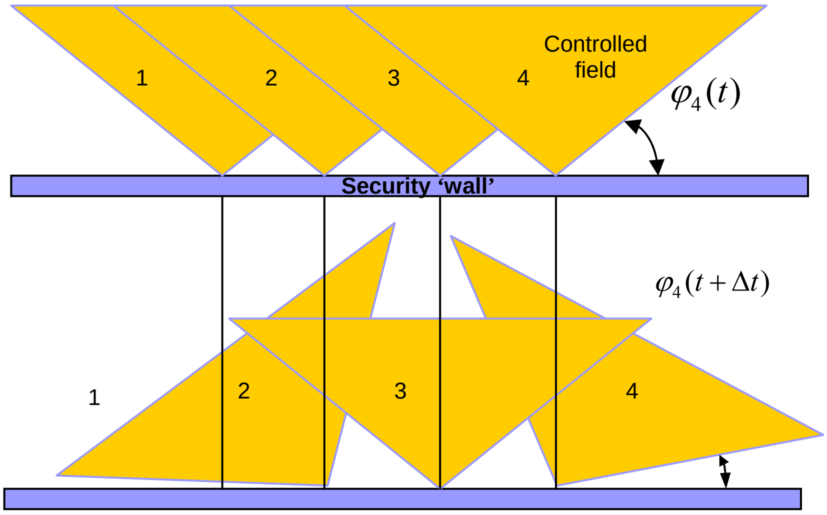


Figure 7. Media and active dots models: dots' linear set

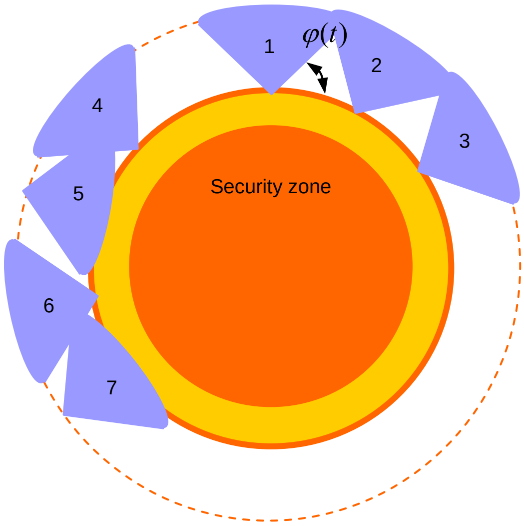


Figure 8. Media and active dots models: dots' radial set

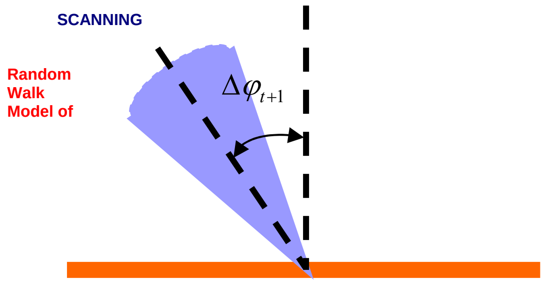


Figure 9. Random Walk Model of Scanning

Random Walk Model of Space Scanning is a chaotically determinate process of rotations of a number of active media dots. After detection of an object a system of observation concentrate it's attention on the object and controls it. It is an intellectual action, based on observed image recognition.

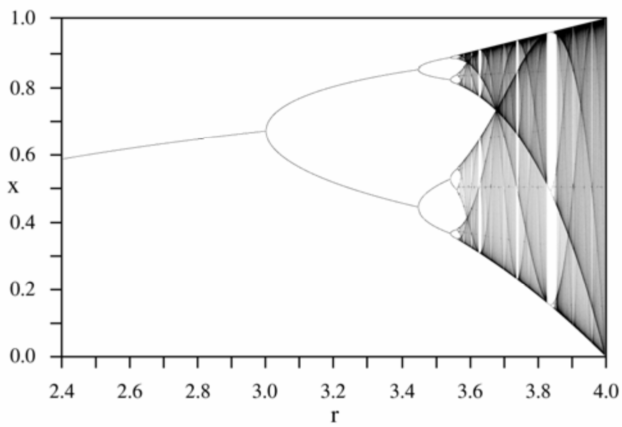


Figure 10. Verhulst's logistic map: the equilibrium point of iterations "x" via the growth parameter "r"

The **Random Walk Model of Space Scanning** is a chaotically determinate process of rotations of a number of active media dots. The task is to optimise the number media-points and their spatial distribution. After detection of an object a system of observation concentrate it's attention on the object and controls it. It is an intellectual action, based on observed image recognition. Various types of image recognition systems can be used. Signal systems include algorithms of recognition of the people, cars and some situations (sharp acceleration of movement, fall of the man, etc.), (see, e.g. Fig. 11, 12).

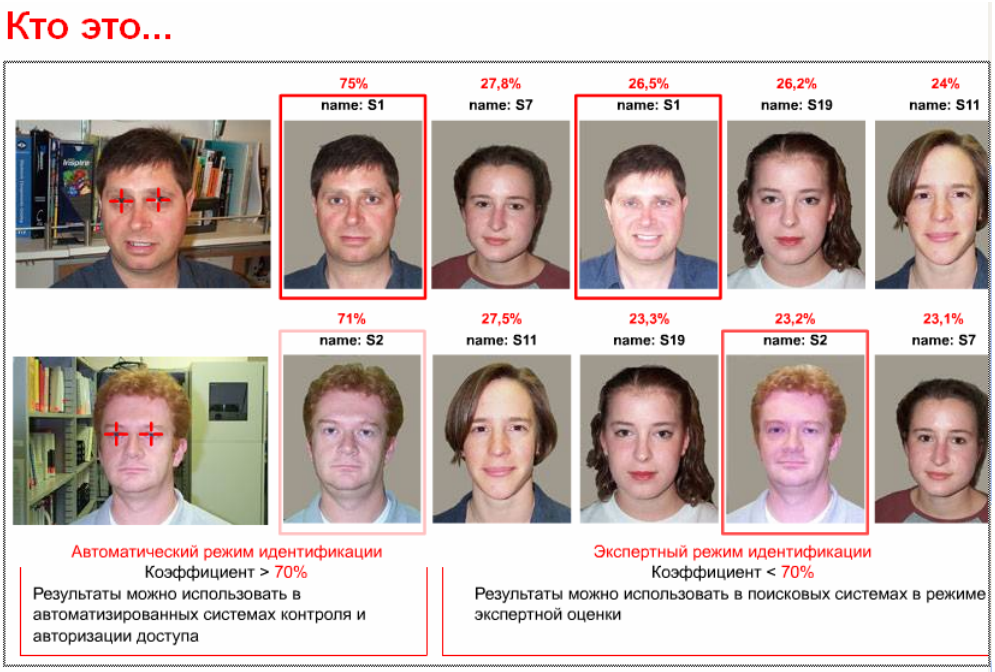


Figure 11. Face control system

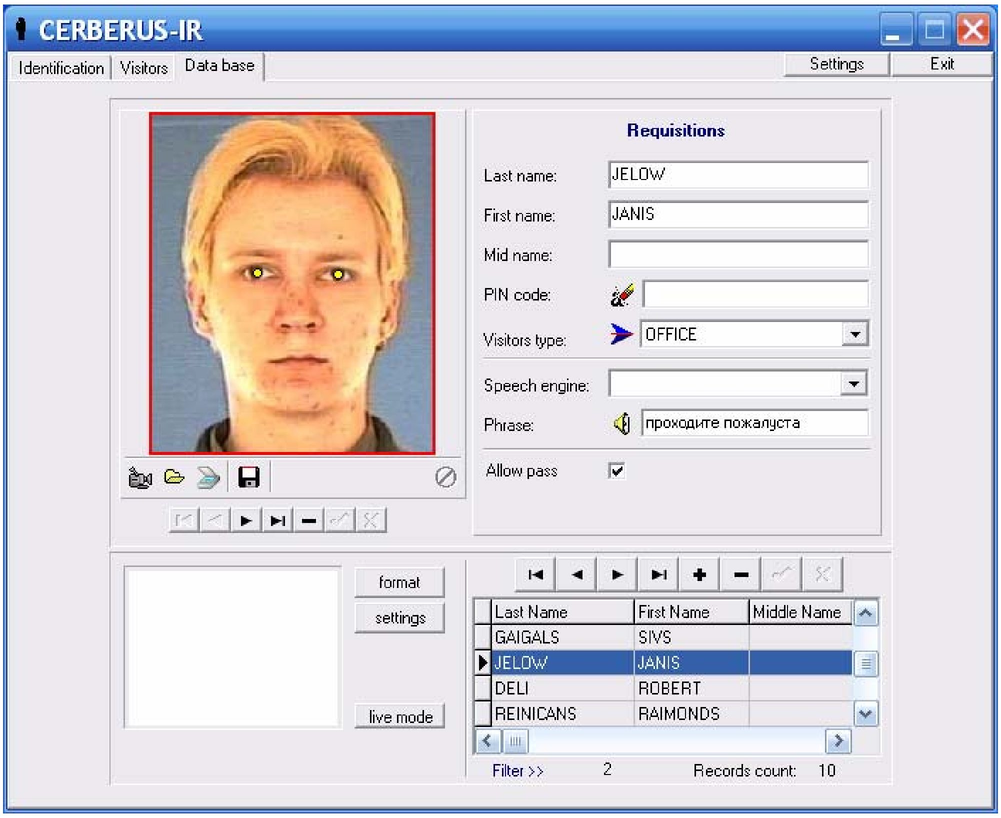


Figure 12. Cerberus image recognition system

Conclusions

An increase of the transmitted data volumes, the growth of speed of transfer and the decrease of a storage cost initiates the development of the general concept of media data processing. Considering various sources of media-information within the framework of a media dots metaphor, it is possible to discuss the unified means of the user interface, system of working in parallel programs of processing agents who are carrying out recognition of situations, and also consider own behaviour of the agents in media-dots space.

References

1. Moore's Law – *http://en.wikipedia.org/wiki/Moore's_law*
2. Hector, G.-M., Jeffrey, U. *Data Bases Systems*. Moscow: Williams, 2004. 1088 p. (In Russian)
3. Jones, M. T. *Programming of Artificial Intelligence in Applications.* Moscow: DMK Press, 2004. 312 p. (In Russian)
4. *http://articles.security-bridge.com/articles/10/11889/ Security TV in the town*
5. The Smart House System. A Technical Overview – *http://www.circuitcellar.com/library/print/hcs-pdf/31-Stauffer.pdf*
6. Bystrov, A. V. Intellectual Building Conception and Innovative Technologies of Insurance Business, *Artificial Intelligence News,* Vol. 6, 2003.
7. *http://www.sec4all.ru/content/view/141/27/ Network IP-cameras*
8. *http://www.bos.dn.ua/view_article.php?id_article=210. IP-systems of video-supervision*
9. Orwell 2k-City. *http://elvees.ru/index.php?id=239#c271. Systems of video-supervision with the computer sight*
10. *http://www.frsservice.lv/ Image recognition.*
11. *http://www.proxim.com/learn/library/appnotes/education.pdf. Wireless Networking Solution For K-12 education*
12. Vorotnikov, S. A. *Information Systems of Robo-technic Devices*. Moscow: MSTU, 2005. 384 p. (In Russian)
13. *www.ravil.astersoft,net MA-dots. IP-cameras.ppt*
14. *http://www.ivimg.com/products/ ViewCommander NVR*
15. LuxRiot is a Digital Video Recording and Remote Surveillance Software Package for the Windows Based PC – *http://www.mondots.lv/lv/luxriot_s.php*
16. *http://www.smartvideo.ru/download/ Digital supervision programmes*
17. *http://www.zoneminder.com/ ZoneMinder the top Linux video camera security and surveillance solution*
18. *http://www.industrialnetworking.com/products/spec_sheets/Mobotix/M1M_Manual_Intro.pdf Mobotix-Concept. Network-CCTV&WebCam.*
19. *http://www.webcam-production.com/pdf/en/sony_snc-cs50p.pdf The Sony SNC-CS50*1 Network Camera*
20. *www.iont.ru/projects/rfbr/90197/Preprint.pdf. Searching animates models*
21. Nepomnyaschih, V. A. Searching of General Principles of Adaptive Behaviour of Alive Organisms and Animates, *Artificial Intelligence News,* Vol. 2, 2002.
22. Gutzwiller, M. *Chaos in Classical and Quantum Mechanics*. New York: Springer-Verlag, 1990.
23. Muhamediyev, R. I., Shunin, Yu. N., Gopeyenko, V. I. Media and Active Dots Strategies. In: *The 1st International Conference Innovative Information Technologies, November 8–10, 2007, Vilnius. Innovative Infotechnologies for Science, Business and Education*: *Programme and Theses, No1.* Riga, 2007, pp. 28–29.

Received on the 21st of October 2007

Computer Modelling and New Technologies, 2007, Vol.11, No.4, 59–63
Transport and Telecommunication Institute, Lomonosov, LV-1019, Riga, Latvia

THE NATURE OF NEGATIVE DIFFERENTIAL RESISTANCE IN *SnO_2-Si* STRUCTURE

D. Fink [1]***, A. Kiv*** [2]

[1]*Hahn-Meitner Institute*
Glienicker Str. 100, 14109 Berlin, Germany

[2]*Ben-Gurion University of the Negev*
Israel, P.O.B. 653, Beer-Sheva, 84105, Israel

The current-voltage correlation of tin dioxide on Silicon structures is characterized by the emergence of a hysteresis. Upon ion irradiation the hysteresis becomes steeper and steeper, indicating a decrease of the sample´s resistance, which eventually even switches to negative values. An equivalent electrical circuit is proposed that consists of both the intrinsic parts of the SnO_2-Si structure and external parts of the electrical loop. The different types of local states are considered. The model gives a qualitative explanation of the nature of the negative differential resistance in the studied structures.

Keywords: *negative differential resistance, SnO_2-Si structure*

1. Introduction

Electronically anisotropic materials have different electronic properties (such as conductivities or dielectricity coefficients) in different directions. Many naturally grown materials such as columnar or fibrous substances and a number of special crystals belong to this class of materials. Also artificially created aligned nanosystems such as porous alumina or insulating foils (e.g. polymers, SiO_2, SiON or mica) containing etched ion tracks can be made electronically anisotropic by filling them with any (semi) conducting material.

Bi-layer structures consisting of electronically anisotropic materials on semiconductors yield a new family of electronic devices, the so-called TEAMS (**T**unable **E**lectronically **A**nisotropic **M**aterial on **S**emiconductor) structures [1–4]. Specifically, (semi)conductor-filled nanoporous layers, combined with a semi-conducting substrate and appropriate contacts, have been denoted as TEMPOS structures (**T**uneable **E**lectronic **M**aterial in **P**ores on **O**xide on **S**emiconductor [5–8]).

In [9] the non-trivial electronic characteristics of SnO_2-Si TEAMS structures are described. In this work the effect of swift heavy ion irradiation on the current-voltage characteristics (CVC) of SnO_2 / Si structures was studied. It was found that a local negative differential resistance (NDR) emerged already after relatively small irradiation doses. The NDR vanished after a short relaxation time.

The existence of a threshold ion dose for the appearance of NDRs led to the conclusion that complex structural defects, formed by interacting ion tracks, may be responsible for the generation of new traps and the appearance of NDR. The instability of these defect configurations causes a small relaxation time of the NDR.

In the figures of the CVC presented in [9], two kinds of hysteresis curves show up: one in the region (– 1…– 2) V for the non-irradiated samples, and a second more pronounced one showing up between about + 7 and + 12 V after the ion irradiation (with gate at ground potential in both cases); for a principle sketch see Fig. 1. It was supposed that both types of hysteresis curves result from transient trapping of electrons at a trap levels. Such trap levels might stem from intrinsic or radiation induced defects. These two types of defects are assigned to the two hysteresis curves found in CVC.

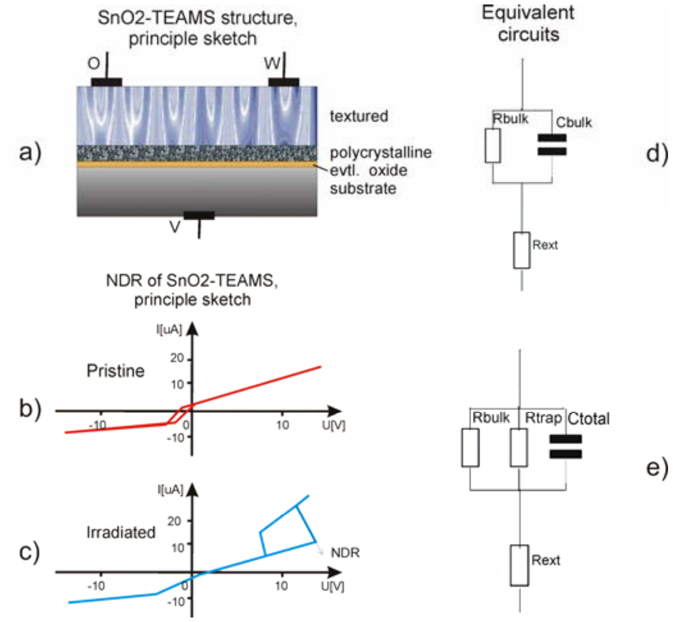


Figure 1. Principle sketch of the NDR in an irradiated SnO_2 / Si TEAMS structure.
a) The structure itself, b, c) schematic current / voltage relation of the b) pristine and c) irradiated structure; d, e) equivalent circuits describing d) the pristine and e) the ion irradiated structure

As a rule interfaces are characterized by complex spectra of intrinsic traps for electrons and holes. As one can see in Fig. 1 of [10], there are three interfaces in the studied sample: (i) between the textured SnO_2 layer and the transition SnO_2 layer, (ii) between SnO_2 and SiO_2 and (iii) between SiO_2 and Si. At the same time the high energy swift heavy ions produce structural defects which act as deep traps. Below we consider mechanisms of appearance of NDR that correspond to unirradiated and irradiated samples.

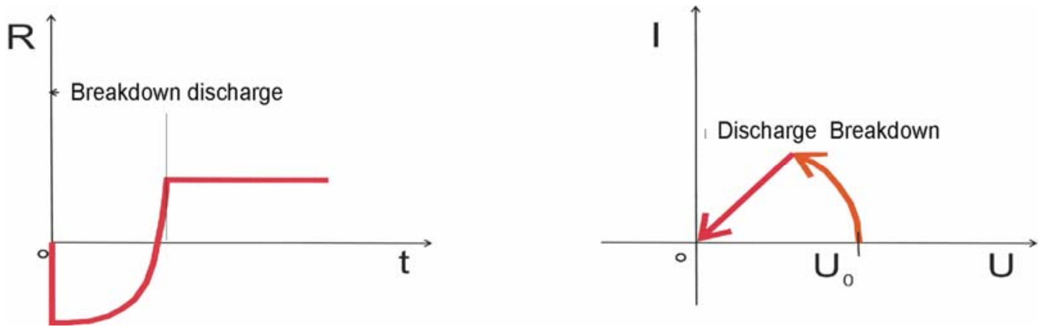


Figure 2. Principle sketch of the NDR in an irradiated SnO_2 / Si TEAMS structure; a) time dependence of the structure resistance; b) principle current / voltage behaviour, for the case that new charge is not supplied from the outside

2. The Model for NDR Explanation

Unirradiated samples. Consider first the small hysteresis which shows up for unirradiated samples. Then one can proceed in the following way:

a) One can describe the interface traps in the corresponding equivalent circuit as resistances R_{trap} in parallel with capacitances C_{trap}. The resistances R_{trap} have an infinitely high value as long as the applied voltage U is lower than the bonding energy, E_{trap} for the charge carriers: $U < E_{trap}$. When U exceeds E_{trap}, then the traps release the electrons (or holes) and R_{trap} becomes zero. This resistance R_{trap} is in series with another resistance R_{ext} which represents the resistance of the external electronic measuring circuit. When an applied external voltage gradually increases from zero upwards, no current will flow as long as $U < E_{trap}$ because $R_{trap} = \infty$. When $U > E_{trap}$, then $R_{trap} = 0$. Hence $R_{total} = R_{ext} + R_{trap} = R_{ext}$, i.e. a current $I_{ext} = U/R_{ext}$ will start flowing which is proportional to the applied voltage U.

b) Now the capacitances C_{trap} play their role as a consequence of the charge storing property of the interface. As long as $R_{trap} = \infty$, the applied voltage $U < E_{trap}$ leads to the accumulation of a charge $Q = C_{trap} U$, with C_{trap} being the interface capacity. Consequently, when "the door is opened" at $U > E_{trap}$ for the electric current to flow (as now $R_{trap} = 0$), the interface capacitor will discharge, i.e. a discharge current $I_{disch} = – dQ/dt = – C_{trap}\ dU/dt$ adds to the external Ohmic current $I_{ext} = U/R_{ext}$, i.e. the total current will be $I_{tot} = I_{ext} + I_{disch}$.

c) As we deal with small E_{trap} and with small applied voltages U, the stored charge is also small and so is the discharge current I_{disch}. Consequently, with increasing U, the total current I_{tot} will rapidly decrease from $I_{ex}t + I_{disch}$ to I_{ext} only. It will come to some kind of Ohmic saturation as seen in the measurements.

d) Decreasing the applied voltage, we return to the regime $U < E_{trap}$ where the traps become active again and force the Ohmic current to vanish. There is, however, still a remanent charge $Q = C_{trap} U \sim C_{trap} E_{trap}$ stored in the interface which leads to the delayed discharge upon further decrease of the applied voltage, i.e. to a hysteresis in the current voltage characteristic (CVC).

e) The model would predict the hysteresis to be symmetric around the origin (i.e. from $– E_{trap}$ to $+ E_{trap}$), but the measured hysteresis is shifted. This may be attributed to the differences in the band structures of SnO_2 and Si, whereby some voltage is generated and this leads to the observed shift. Besides this, it is worth to account for two additional facts; (i) in real systems, as a rule, there are traps with a wide spectrum of ionization energies and (ii) the current flow consists simultaneously of both electrons and holes. These facts also influence the form of the hysteresis curves and their displacement relative to the origin.

From the measured characteristics, we estimated E_{trap} to be around ½ eV, and the generated voltage to be roughly in the order of 1½ V; the fact that this hysteresis curve is shifted to negative currents indicates that we deal here with trapping of negative charge carriers, i.e. of electrons.

Irradiated samples. Now let us consider the pronounced hysteresis curves which we found in CVC of samples after ion irradiation (Fig. 3 in [9]). It is remarkable that this hysteresis sets in only upon application of at least (+ 10…+ 13) V. This points at the very deep level of these traps, and it clearly rules out that these traps are of the same nature as the ones treated above. Nevertheless, for a qualitative explanation of our observations, we can partly overtake the above explanation.

a) Again, the radiation-induced traps in the proposed equivalent circuit are taken as resistances R_{trap} in parallel with capacitances C_{trap}. Again, the resistances R_{trap} have an infinitely high value as long as the applied voltage U is lower than the bonding energy for the charge carriers: $U < E_{trap}$ (with E_{trap} being in this case as large as ~ 13 eV). When U exceeds E_{trap}, then the traps are again assumed to release all the charge carriers, hence R_{trap} will be zero.

b) In contrast to the above case, in the present equivalent circuit we consider R_{trap} as being parallel to some resistance R_{bulk} for that current fraction flowing through the sample which is not trapped at the defects. To this total sample resistance $R_{sample} = (R_{trap}^{-1} + R_{bulk}^{-1})^{-1}$, the resistance R_{ext} of the external electric circuit has to be added: $R_{total} = R_{ext} + R_{sample}$. This means that, in whatever state the radiation-induced traps are (i.e. active or inactive), there is always the current component $I_{bulk} = U/(R_{ext} + R_{bulk})$ flowing, even for $U < E_{trap}$.

c) When the applied external voltage exceeds E_{trap}, a discharge current $I_{disch} = – dQ/dt = – C_{trap}\ dU/dt = U/R_0$ will add to the current I_{bulk} with R_0 describing macroscopically the resistance of traps. Hence, $Q/C_{trap} = R_0\ dQ/dt$. The solution of this differential equation shows that the decreasing trap charge is described by: $Q = C_{trap}U_0 \exp(– t/(R_0C_{trap}))$. Consequently, this current component is given by $I_{disch} = dQ/dt = C_{trap}U_0/(R_0C_{trap}) \exp(– t/(R_0C_{trap})) = U_0/R_0 \exp(– t/(R_0C_{trap}))$. This current will release the overall charge Q stored in all the deep level radiation defects.

In contrast to the above case, we have now a much stronger current due to the larger charge $Q = C_{trap}U$ stored at the higher applied voltage $U \sim E_{trap}$, with C_{trap} being now the capacity which can be ascribed to the multitude of charges trapped all over the irradiated sample. The rapid discharging leads to a rapid drop of the voltage ΔU along the traps: $\Delta U = \Delta Q/C_{trap} = U_o/(R_0C_{trap})^2 \exp(– t/(R_0C_{trap}))$. As long as R_{ext} is large enough so that this voltage drop is not compensated immediately by the external voltage supply, the effective resistance of the equivalent circuit can be described by $R_{eff} = (U_0 – \Delta U)/I$. Hence:

$$R_{eff} = R_0 \exp(t/(R_0C_{trap})) – 1/(R_0C_{trap}^2) . \qquad (1)$$

Eq. (1) signifies that the effective trap resistance during the build-up of the discharge current is initially negative, and then rises exponentially until it obtains positive values.

The discharge current will not rise indefinitely and will gradually come to saturation. From that moment, the following can be considered: by differentiation of $Q = C_{trap}U$ we get: $dU/dt = -1/C_{trap}\ dQ/dt = -I/C_{trap}$; hence: $I = C_{trap}dU/dt$. Due to $I = U/R_{eff}$ we have: $dU/dt = -(C_{trap}/R_{eff})\ U$, with the known solution $U(t) = U_0 \exp(-(C_{trap}/R_{eff})\ t)$. Due to $I = U/R_{eff}$, the discharge current $I(t)$ decreases now proportional to the voltage: $I(t) = (U_0/R_{eff}) \exp(-(C_{trap}/R_{eff})\ t)$. In other words, the effective trap resistance in that second phase, the discharge phase, is positive and constant.

To summarize, the overall current will first rise steeply from $I_{tot} = I_{ext}$ to some maximum $I_{tot} = I_{ext} + I_{disch}$, and thereafter decrease until all the stored charge Q is released. This is accompanied by an initial negative differential resistance which grows rapidly until it reaches a stationary positive value (see also Fig. 2). The model also predicts that the higher the breakdown voltage, the shorter the time constant of discharge, i.e. the sharper will be the discharge pulse.

d) The consequences of that local voltage drop will depend on the magnitude of R_{ext}. Above we had assumed R_{ext} to be sufficiently large so that the external voltage source could not replace the missing charge in the sample rapidly enough. Hence the local voltage drop at the radiation-induced defects leads to an overall voltage drop of the whole sample. In this case, the overall strong current increase is accompanied by a transient fall in voltage, which shows up in the CVC by a negative slope. Only upon further increase of the applied voltage this fall will be compensated again, and the resistance will return to positive values.

e) If, however, R_{ext} is sufficiently small, then the applied external voltage source will replace the missing charge in the sample in due time so that a voltage fall is avoided. In this case the resistance will become quite small but remain positive.

The fact that an increasing applied voltage leads to an increase in current, for whatever sign of resistance, signifies that the released charge carriers are positive. Hence we have trapped holes in the defects. This is in fact what one can see in Fig. 4 of [9]. On increasing voltage applied to the irradiated samples, pronounced negative differential resistances show up. These resistances are quite small but positive on the other side of the hysteresis curves, when the applied voltage is decreased. Fig. 4 in [9] further reveals that the magnitude of the NDR increases with increasing the fluency, starting at around 1×10^{10} cm^{-2} and reaching an optimum at 3.2×10^{11} cm^{-2}.

f) When the applied external voltage is decreased again below E_{trap}, then a number of charge carriers will be trapped again in the radiation-induced traps, thus lowering the overall current. As in the above case, the remanent charge $Q = C_{trap}\ U = C_{trap}E_{trap}$ stored in the traps will lead to a delayed discharge upon further decrease of the applied voltage, i.e. to a hysteresis in the characteristic.

Post-irradiation sample behaviour. The fact that after a few weeks of irradiation the characteristics of Fig. 4 in [9] have completely vanished indicates a dramatic change of the overall sample structure, possibly due to some radiation induced phase change. It is, however, still too early for further speculations till detailed studies are done. The fact that even another ion irradiation does not alter the latter picture any longer indicates that the new structure is apparently more stable than the original one. As mentioned above, during the relaxation of irradiated specimens, they come to a state with a new disordered potential distribution. Such specimens may reveal an abnormally high resistance to the repeated irradiation.

Conclusions

We developed a model to explain the electrical properties of SnO_2-Si TEAMS structures. The effect of swift heavy ion irradiation on the CVC of these structures is considered. The studied structures are characterized by specific hysteresis regions in their CVC. The conditions of the appearance of NDR in CVC are determined. Promising electric characteristics of irradiated structures vanish after relatively short time. Once these degrading processes are understood and it is known how to prevent them, the SnO_2-Si TEAMS structures may have potential application in oscillators, amplifiers, and flip-flops, due to their pronounced NDRs. Furthermore, one could then combine the NDRs with the sensing properties of SnO_2 to build compact structures which switch or give signals upon the appearance of a lot of toxic substances in the ambient. Eventually this study points at a whole class of new electronic elements with NDRs.

References

1. Fink, D. In: *ISL Information*. Berlin: HMI, May 2, 2005.
2. Fink, D., Sinha, D., Opitz–Coutureau, J., Petrov, A. V., Demyanov, S. E., Borisenko, V. E., Fahrner, W. R., Hopee, K., Fedotov, A. K., Chadderton, L. T., Berdinsky, A. S. In: *Physics, Chemistry and Application of Nano-Structures/Edited by S. V. Gaponenko, V. S. Gurin*. World Scientific Publications, 2005. 474 p.
3. Berdinsky, A. S., Fink, D., Yoo, J-B., Chadderton, L. T., Chun, H.-G., Han, J.-H. *Radiation Effects and Defects in Solids,* Vol. 159, 2004. 233 p.
4. Sinha, D., Petrov, A. V., Fink, D., Fahrner, W. R., Hoppe, K., Chandra, A. *Radiation Effects and Defects in Solids,* Vol. 159, 2004. 515 p.
5. Fink, D., Petrov, A., Hoppe, K., Fahrner, W. R., Papaleo, R. M., Berdinsky, A. S., Chandra, A. A., Biswas, A., Chadderton, L. T. *Nucl. Instr. Meth*., Vol. B 218, 2004. 355 p.
6. Fink, D., Petrov, A., Hoppe, K., Fahrner, W. R., Chubaci, J. F. and Tabachnics, M. H. In: *Proceedings MRS Fall Meeting/Edited by Lu-Min Wang, R. Fromknecht, L. Snead, D. F. Downey, and H. Takahasi, Vol. 792, R 7.9.* Boston, 2003.
7. Fink, D., Petrov, A., Fahrner, W. R., Hoppe, K., Papaleo, R. M., Berdinsky, A., Chandra, A., Zrinen, A., Chadderton, L. T. *Int. J. of Nanoscience,* Vol. 4, 2005, p. 965.
8. Fink, D., Chadderton, L. T., Hoppe, K., Fahrner, W. R., Chandra, A., Kiv, A. In: *Proc. of the Int. Conf. on Applications of Accelerators in Research and Industry*. Fort Worth, 2006. 193 p.
9. Fink, D., Kiv, A., Golovanov, V., Chen, J., Chandra, A., Ivanovskaya, M., Khirunenko, L., Fuks, D. Tin Dioxide on Semiconductor as Material for New Electronics, *J. Mater. Sci.: Electronic Materials.* (Submitted)
10. Fink, D., Kiv, A., Fuks, D., Chandra, A., Golovanov, V., Ivanovskaya, M., Khirunenko, L. *Appl. Phys. Lett.*, Vol. 91, 2007, p. 55.

Received on the 1st of November 2007

Computer Modelling & New Technologies, 2007, Volume 11, No. 4
Transport and Telecommunication Institute, Lomonosova 1, Riga, LV-1019, Latvia

Authors' index

Yuri N. Shunin (born in Riga, March 6, 1951)

- Vice-Rector on academic issues (Information Systems Management Institute), professor, Dr.Sc. Habil., Member of International Academy of Refrigeration
- *Director of Professional Study Programme* Information Systems (Information Systems Management Institute)
- *Director of Master Study Programme* Computer systems (Information Systems Management Institute)
- **University study**: Moscow physical and technical institute (1968–1974)
- Ph.D. (physics & mathematics) on solid state physics (1982, Physics Institute of Latvian Academy of Sciences), Dr. Sc. Habil (physics & mathematics) on solid state physics (1992, Ioffe Physical Institute of Russian Academy of Sciences)
- **Publications:** 320 publications, 1 patent
- **Scientific activities**: solid state physics, physics of disordered condensed media, amorphous semiconductors and glassy metals, semiconductor technologies, heavy ion induced excitations in solids, mathematical and computer modelling, system analysis

Igor V. Kabashkin (born in Riga, August 6, 1954)

- Vice-rector for Research and Development Affairs of Transport and Telecommunication Institute, Professor, Director of Telematics and Logistics Institute
- PhD in Aviation (1981, Moscow Institute of Civil Aviation Engineering), Dr.Sc.Habil. in Aviation (1992, Riga Aviation University), Member of the International Telecommunication Academy, Member of IEEE, Corresponding Member of Latvian Academy of Sciences (1998)
- **Publications:** 350 scientific papers and 67 patents
- **Research activities:** information technology applications, operations research, electronics and telecommunication, analysis and modelling of complex systems, transport telematics and logistics

Victor Gopeyenko (born in Siktivkar, Russia, October 16, 1957**)**

- Information System Management Institute
- Vice-Rector on innovative issues. As. Professor of the Department of Natural Sciences and Computer Technologies
- **Publications:** 50
- **Research activities:** mathematical and computer modelling, digital signal processing

Ravil Muhamediyev (born on June 23, 1959)

- Information System Management Institute
- The Head of the Department of Natural Sciences and Computer Technologies, As. Professor, Dr. Sc.
- **Publications:** 85 scientific papers
- **Research activities:** programming, artificial intellect and intellectual systems, computer training system and developing games

Igor Lemberski (born on August 7,1952)

- Dr.Sc.Ing., As. Pprofesor, Baltic International Academy,
  e-mail: igor.lemberski@bsa.edu.lv
- Director of University Program "Microelectronic Circuits Design"
- Research interests: VLSI design, asynchronous circuits design
- **Publications:** 50 papers and patents

Alexander Babitski (born in Minsk, March 7, 1973)

- Senior Researcher, Dept. of Applied Mathematics and Computer Science, Belarusian State University
- University study: Belarusian State University (1995)
- Candidate of Science in Physics and Mathematics (1998)
- **Publications:** 16 publications
- **Research activities:** queuing theory, inventory management, mathematical statistics, decision support systems

Arnold E. Kiv

- Professor, Dr. Sc. Habil. (phys&math), Chief of the Department of Theoretical Physics and Laboratory of Computer Modelling
- South Ukrainian Pedagogical University, Professor of the Department of Materials Engineering, Ben-Gurion University of the Negev P.O.B. 653, Beer-Sheva, 84105, Israel
- **Scientific activities**: sub-threshold radiation effects in solids, computer modelling of processes in physics, psychology and social sciences

CUMULATIVE INDEX

COMPUTER MODELLING and NEW TECHNOLOGIES, volume 11, No. 4, 2007

(Abstracts)

A resource delivery system which supplies several customers by homogenous non-renewable resources is considered. For each customer its minimal and maximal amounts of resources to be delivered, are pre-determined, as well as its preference (importance) rate. A real-time model which at any decision-making moment reallocates optimally the total available amount of resources among the customers, is presented. The values to be optimised are the resource amounts actually delivered to each customer, while the objective to be maximized is the summarized product of the delivered resource amounts and the customers' rates. A numerical example is presented.

Keywords: resource reallocation, homogenous consumable resources, resource delivery model, customer's preference rate, precise resource reallocation algorithm

Some levels of alternativity in long-term network projects by examining both fully divisible and more generalized controlled alternative activity networks will be considered. We will demonstrate that those alternative (multi-variant) projects may be controlled in the course of their actual realization. Moreover, the level of indeterminacy of such projects may influence the process of the project's capital investment, i.e., the procedure of contracting has to comprise consecutive step-wise stages.

Keywords: project management (PM); alternative network project; deterministic and stochastic decision-making; joint variant; outcome tree; long-term construction project; consecutive step-wise contracting

The Data Envelopment Analysis (DEA) is a methodology for evaluating the relative efficiency of units based on multiple inputs and multiple outputs. The efficiency score is defined as the ratio between weighted outputs to the weighted inputs. The weights for the inputs and outputs are chosen in order to maximize this efficiency score. These optimal weights differ from unit to unit, and can have a very large range. Therefore, one cannot rank the units according to these weights. Another reason for not ranking the units based on the DEA is that, according to these optimal weights, one can only distinguish between an efficient unit and an inefficient one. Especially when the number of units is small, relative to the number of inputs and outputs, most of the units will be efficient. In this paper, we refer to several ranking methods in the DEA context: Super Efficiency, Cross Efficiency (CE) Canonical Correlation Analysis (CCA), Discriminant Analysis of Ratio (DR/DEA) and the Global Efficiency (GE). We illustrate all the ranking methods based on data from 30 restaurants. Each restaurant has 2 outputs and 4 inputs.

Keywords: The Data Envelopment Analysis (DEA), Common Weights (CW), Super Efficiency, Multi-criteria Decision Analysis (MCDA), Cross Efficiency (CE) Canonical Correlation Analysis (CCA), Discriminant Analysis of Ratio (DR/DEA), Global Efficiency (GE), Ranking

A. Babitski. Optimization of Inventory Management of a Perishable Product in the Conditions of Fluctuating Demand, *Computer Modelling and New Technologies*, vol. 11, No 4, 2007, pp. 36–39.

An inventory system with perishable products with deterministic shelf life and demand represented by Poisson process with different intensities for different time intervals is studied. Formulae for calculation of expected storage cost and income are obtained. Algorithms for determining optimal order size for maximization of net income and unit income are proposed.

Keywords: inventory control optimisation, perishable product, stochastic demand

I. Lemberski, V. Gopeyenko. Method of Finite State Machine Optimal Implementation Targeting Look-Up-Table Architecture, *Computer Modelling and New Technologies*, vol. 11, No 4, 2007, pp. 40–47.

The method of state encoding for finite state machine optimal look-up-table implementation is offered. It is oriented on summary support minimization of logic functions. Encoding length is not restricted (may be non-minimal). As a criterion, the number of Configurable Logic Blocks (CLBs) is considered. Our method includes: 1) seed encoding; 2) step-by-step merging state variables to decrease summary support. Heuristic procedures are offered the reduce CPU time of the task to be solved. Experiments for XILINX 3090 architecture have been performed and compared with the results of NOVA, JEDI and LAX, where minimal encoding is accepted. It is shown that, in some cases, better implementation may be obtained for non- minimal encoding generated by our procedure.

Keywords: finite state machine (FSM), look-up-table (LUT), field programmable logic array (FPGA), state encoding

I. Lemberski, V. Gopeyenko. XILINX4000 Architecture-Driven Boolean Network Re-Decomposition Targeting Performance, *Computer Modelling and New Technologies*, vol. 11, No 4, 2007, pp. 48–51.

Architecture-driven (instead of look-up-table-LUT-driven) method of Boolean functions logic synthesis for the performance is proposed. It takes XILINX4000 architectural features (heterogeneous LUTs of 3 and 4 inputs) into account and includes two-step decomposition. In the first step, two-level logic representation is transformed into a graph of at most 4 fan-in nodes (after this step, each node can be mapped onto 4 input LUT). In the second step, selected 4 fan-in nodes within a critical path are re-decomposed into 3 fan-in nodes to ensure mapping onto 3 input LUTs. The re-decomposition task is formulated as substituting of node two fan-ins for exactly one fan-in, either existing node or one especially created, is considered as a fan-in to be substituted for. The extended PLA format describing a multi-level Boolean network is proposed. Based on this description, substituting is formulated in terms of a covering task.

Keywords: XILINX, Boolean function, decomposition

R. I. Muhamediyev, Yu. N. Shunin, V. I. Gopeyenko. Media and Active Dots Optimal Development Strategies, *Computer Modelling and New Technologies*, vol. 11, No 4, 2007, pp. 52–58.

Media dots are metaphor mean elements that uses for reception and reproduction of the media data. Potentially they can be used in systems of protection and supervision, «smart house» systems, media-content control systems of an educational institution etc. The examples of existing complexes of video-supervision and projects of some systems are considered. The assumption on general character of processes of reception and processing of media data of various purpose systems is done. The opportunity of creation of the unified means of the user interface and systems of parallel working programs of agents processing, which are carrying out a recognition of situations (intellectual detecting) and formation of reactions, are discussed. The tasks and problems of the further researches are formulated.

Keywords: media dots, active dots, intellectual agent, intellectual detecting, smart house, user graphic interface, natural behaviour strategies, security signal system

D. Fink, A. Kiv. The Nature of Negative Differential Resistance in *SnO_2-Si* Structure, *Computer Modelling and New Technologies,* vol. 11, No 4, 2007, pp. 59–63.

The current-voltage correlation of tin dioxide on Silicon structures is characterized by the emergence of a hysteresis. Upon ion irradiation the hysteresis becomes steeper and steeper, indicating a decrease of the sample´s resistance, which eventually even switches to negative values. An equivalent electrical circuit is proposed that consists of both the intrinsic parts of the SnO_2-Si structure and external parts of the electrical loop. The different types of local states are considered. The model gives a qualitative explanation of the nature of the negative differential resistance in the studied structures.

Keywords: negative differential resistance, SnO_2-Si structure

COMPUTER MODELLING and NEW TECHNOLOGIES, 11.sējums, Nr. 4, 2007 (Anotācijas)

D. Golenko-Ginzburgs, A. Ben-Jiers, J. Hadads, N. Badarna. Resursu sadales modelis klientiem ar dažādām prioritātēm, *Computer Modelling and New Technologies,* 11.sēj., Nr.4, 2007, 7.–14. lpp.

Resursu piegādes sistēma, kas apgādā dažus pircējus ar viendabīgiem neatjaunojamiem resursiem, tiek izskatīta dotajā rakstā. Katram klientam tiek piegādāts to minimālais un maksimālais resursu daudzums, kas tiek arī iepriekš paredzēts, kā arī to privilēģiju (svarīguma) pakāpe. Rakstā tiek piedāvāts reālā laika modelis, kas pie katra lēmuma pieņemšanas optimāli pārdala starp klientiem kopējo pieejamo resursu daudzumu. Rakstā tiek izskatīti arī skaitliskie piemēri.

Atslēgvārdi: resursu pārdale, viendabīgie izmantojamie resursi, resursu piegādes modelis, klienta privilēģiju pakāpe, precīzs resursu sadales algoritms

D. Golenko-Ginzburgs, Z. Laslo, A. Ben-Jears. Ilgtermiņa celtniecības projektu vadīšana ar atšķirīgiem rezultātiem, *Computer Modelling and New Technologies,* 11.sēj., Nr.4, 2007, 15.–24. lpp.

Rakstā tiek izskatīti daži alternativitāšu līmeņi ilgtermiņa tīkla projektos, izpētot kā pilnībā, tā arī vairāk vispārināti kontrolētos alternatīvās darbības tīklus. Autori parāda, ka multi-variantu projekti var būt kontrolēti to faktiskās realizācijas gaitā. Bez tam šādu projektu nenoteiktības līmenis var ietekmēt šī projekta kapitālieguldījumus, t.i., līguma noslēgšanas procedūrai ir jābūt pakāpeniskai.

Atslēgvārdi: projekta vadīšana, alternatīvais tīkla projekts, deterministiskā un stohastiskā lēmumu pieņemšana, saistītais variants

J. Hadads, L. Fridmans, M. Z. Hanani. Restorānu efektivitātes mērīšana, lietojot datu apiešanas analīzes metodoloģiju, *Computer Modelling and New Technologies,* 11.sēj., Nr.4, 2007, 25.–35. lpp.

Datu apiešanas analīze (The Data Envelopment Analysis – DEA) ir metodoloģija, lai novērtētu vienību relatīvo efektivitāti, kas pamatojas uz daudzkārtīgām ieejām un daudzkārtīgām izejām. Efektivitātes rēķins tiek noteikts kā attiecība starp izsvērtām izejām un izsvērtām ieejām. Svari ieejām un izejām tiek izvēlēti, lai maksimizētu šo efektivitātes rēķinu. Šie optimālie svari atšķiras no vienības uz vienību, un var ietvert ļoti lielu amplitūdu. Tādēļ nevar sarindot vienības saskaņā ar šiem svariem. Cits iemesls, kādēļ nevar sarindot vienības, pamatojoties uz DEA, ir sekojošs – saskaņā ar šiem optimālajiem svariem var tikai atšķirt starp efektīvo vienību un neefektīvo vienību. Īpaši, ja vienību skaits ir neliels, relatīvs ieeju un izeju skaitam, lielākais vienību skaits būs efektīvs. Šajā rakstā autori izmanto vairākas sarindošanas metodes DEA kontekstā: Superefektivitāte (Super Efficiency), Savstarpējā efektivitāte (Cross Efficiency – CE), Kanoniskā korelācijas analīze (Canonical Correlation Analysis – CCA), Proporcijas diskriminantā analīze (Discriminant Analysis of Ratio – DR/DEA) un Globālā efektivitāte (the Global Efficiency – GE). Autori rakstā parāda visas sarindošanas metodes, pamatojoties uz datiem no 30 restorāniem. Katram restorānam ir 2 izejas un 4 ieejas.

Atslēgvārdi: Datu apiešanas analīze(The Data Envelopment Analysis – DEA), Savstarpējā efektivitāte (Cross Efficiency – CE), Kanoniskā korelācijas analīze (Canonical Correlation Analysis – CCA), Proporcijas diskriminantā analīze (Discriminant Analysis of Ratio – DR/DEA) un Globālā efektivitāte (the Global Efficiency – GE), sarindošana

A. Babitskis. Ātri bojājošos produktu inventarizācijas vadības optimizācija svārstīga pieprasījuma apstākļos, *Computer Modelling and New Technologies,* 11.sēj., Nr.4, 2007, 36.–39. lpp.

Rakstā tiek izskatīta ātri bojājošos produktu ar noteiktu izlietošanas laiku un pieprasījumu inventarizācijas sistēma, kas tiek pārstāvēta ar Puasona procesu ar dažādām intensitātēm un kas paredzēts dažādiem laika intervāliem. Pētījuma gaitā tiek iegūtas formulas, lai aprēķinātu sagaidāmās uzglabāšanas izmaksas un ienākumu. Algoritms, lai noteiktu optimālā pasūtījuma lieluma tīrās peļņas maksimizēšanu un vienības ienākumu, tiek izstrādāts.

Atslēgvārdi: inventarizācijas kontroles optimizācija, ātri bojājošies produkti, stohastiskais pieprasījums

I. Lemberskis, V. Gopejenko. Mehānisma galīgā stāvokļa optimālās ieviešanas metode, runājot par "*look-up-table*" arhitektūru, *Computer Modelling and New Technologies,* 11.sēj., Nr.4, 2007, 40.–47. lpp.

Rakstā tiek piedāvāta stāvokļa šifrēšanas metode mehānisma galīgā stāvokļa optimālai ieviešanai. Tā ir orientēta uz loģisko funkciju kopsavilkuma atbalsta minimizēšanu. Šifrēšanas garums nav ierobežots (drīkst būt ne-minimāls). Kā kritērijs, *Configurable Logic Blocks (CLBs)* skaitlis tiek izskatīts, Mūsu metode ietver sekojošo: 1) iedīgļa šifrēšanu; 2) pakāpenisku mainīgā lieluma stāvokļa saplūšanu, lai samazinātu kopsavilkuma atbalstu. Rakstā tiek piedāvātas heiristiskās procedūras. Tiek parādīts, ka dažos gadījumos labāka ieviešana var tikt iegūta ne-minimālai šifrēšanai, ko ģenerē mūsu procedūra.

Atslēgvārdi: galējā stāvokļa mehānisms, "look-up-table", stāvokļa šifrēšana

I. Lemberskis, V. Gopejenko. Būla tīkla pār-dekompozīcija darba plānošanā pēc XILINX4000 arhitektūras, *Computer Modelling and New Technologies,* 11.sēj., Nr.4, 2007, 48.–51. lpp.

Rakstā tiek izskatīta Būla funkciju loģiskās sintēzes veikšanai arhitektoniski veidotā metode (*look-up-table-LUT* veidotās vietā). Tā ņem vērā XILINX4000 arhitektoniskās īpašības (heterogēnas *LUT* no 3. un 4. ieejas) un ietver divsoļu dekompozīciju. Pirmajā solī divlīmeņu loģiskā reprezentācija tiek transformēta grafā, kas sastāv vismaz no 4 *fan-in* mezgliem (pēc šī soļa, katrs mezgls var būt attēlots uz 4 ieeju *LUT*). Otrajā solī izvēlētie 4 *fan-in* mezgli, kas kritiskā ceļā tiek pār-dekomponēti 3 *fan-in* mezglos, lai nodrošinātu attēlojumu uz 3 ieeju *LUT.* Pār-dekompozīcijas uzdevums tiek formulēts kā 2. *fan-in* mezglu aizvietošana ar tieši vienu *fan-in*, vai eksistējošais mezgls, vai arī viens speciāli radītais mezgls tiek uzskatīts kā *fan-in*, kam jābūt aizvietotam. Paplašinātais *PLA* formāts, kas apraksta multi-līmeņa Būla tīklu, tiek piedāvāts. Pamatojoties uz šo aprakstu, aizvietošana tiek formulēta dotā uzdevuma izteiksmē.

Atslēgvārdi: XILINX, Būla funkcija, dekompozīcija

R. I. Muhamedijevs, J. N. Šuņins, V. I. Gopejenko. Mediju un aktīvo punktu optimālās attīstības stratēģijas, *Computer Modelling and New Technologies,* 11.sēj., Nr.4, 2007, 52.–58. lpp.

Mediju punkti ir metaforas vidējie elementi, ko lieto mediju datu uztverei un reprodukcijai. Potenciāli tie var būt lietoti aizsardzības un pārskata sistēmās, „smalko māju" sistēmās, mediju-satura kontroles sistēmās u.c. Rakstā tiek izskatīti video pārskata eksistējošo kompleksu piemēri un dažu sistēmu projekti. Rakstā tiek veikta uztveres procesu vispārējā rakstura un dažādu mērķu sistēmu mediju datu apstrāde. Bez tam arī tiek formulētas tālāko pētījumu uzdevumi un problēmas.

Atslēgvārdi: mediju punkti, aktīvie punkti, intelektuālais aģents, intelektuālais uztvērējs, smalkā māja, dabīgās uzvedības stratēģijas, drošības signalizācijas sistēmas

D. Finks, A. Kivs, Negatīvās diferenciālās rezistences daba *SnO_2-Si* struktūrās, *Computer Modelling and New Technologies*, 11.sēj., Nr.4, 2007, 59.–63. lpp.

Alvas dioksīda silikona struktūrās tekošās strāvas korelācija tiek raksturota ar histerēzes parādīšanos. Līdz ar jonu iradiāciju histerēze kļūst arvien stāvāka un stāvāka, norādot uz rezistences piemēra samazināšanos, kas gala rezultātā pārslēdzas uz negatīvām vērtībām. Rakstā tiek izskatīts ekvivalents elektriskais cikls, kas sastāv kā no SnO_2-Si struktūras iekšējām dalām, tā arī no elektriskās cilpas ārējām daļām. Tiek konstatēti lokālo stāvokļu dažādie tipi. Pētāmajās struktūrās modelis dod negatīvās diferenciālās rezistences dabas kvalitatīvu izskaidrojumu.

Atslēgvārdi: negatīvā diferenciālā rezistence, SnO_2-Si struktūra

COMPUTER MODELLING & NEW TECHNOLOGIES

ISSN 1407-5806 & ISSN 1407-5814(on-line)

THE JOURNAL IS DESIGNED FOR PUBLISHING PAPERS CONCERNING THE FOLLOWING FIELDS OF RESEARCH:

- mathematical and computer modelling
- mathematical methods in natural and engineering sciences
- physical and technical sciences
- computer sciences and technologies
- semiconductor electronics and semiconductor technologies
- aviation and aerospace technologies
- electronics and telecommunication
- navigation and radar systems
- telematics and information technologies
- transport and logistics
- economics and management
- social sciences

Articles can be presented in journal in English. All articles are reviewed.

EDITORIAL CORRESPONDENCE

Transporta un sakaru institūts (Transport and Telecommunication Institute)
Lomonosova 1, LV-1019, Riga, Latvia. Phone: (+371)7100593. Fax: (+371)7100535
E-mail: journal@tsi.lv, http:// www.tsi.lv

COMPUTER MODELLING AND NEW TECHNOLOGIES, 2007, Vol. 11, No.4

Scientific and research journal of Transport and Telecommunication Institute (Riga, Latvia)
The journal is being published since 1996.

PREPARATION OF CAMERA-READY TYPESCRIPT: COMPUTER MODELLING AND NEW TECHNOLOGIES

1. In order to format your manuscript correctly, see the Page Layout Guideline for A4 (21 cm x 29,7 cm) paper size. Page Layout should be as follows: Top – 3 cm, Bottom – 3 cm, Left – 3 cm, Right – 3 cm.
2. Maximum length for the article is 10 pages.
3. **Application of other Styles with the exception of Normal is impossible!**
4. ***Articles*** should be Times New Roman typeface, single-spaced.
5. The article should include:
   - title;
   - author's name(s) and information (organisation, city, country, present address, phones, and e-mail addresses);
   - abstract (100–150 words);
   - keywords (max. about six);
   - introduction – clearly explaining the nature of the problem, previous work, purpose and contribution of the research;
   - description of proper research;
   - conclusion section (this is mandatory) which should clearly indicate advantages, limitations and possible applications;
   - references.

   Attention! First name, last name, the title of the article, abstract and keywords must be submitted in the English and Latvian languages (in Latvian it is only for Latvian authors) as well as in the language of the original (when an article is written in different language).
6. The text should be in clear, concise English (or other declared language). Please be consistent in punctuation, abbreviations, spelling (***British English***), headings and the style of referencing.
7. ***The title of the article*** – 14 point, UPPERCASE, style Bold and centered.
8. ***Author's names*** – centered, type size 12 point, Upper and lower case, style Bold Italic.
9. ***Author's information*** – 10 point, Upper and lower case, style Italic, centered.
10. ***Abstract and keywords*** – 8 point size, style Normal, alignment Justify.
11. ***The first level Headings*** – 11 point, Upper and lower case, style Bold, alignment Left. Use one line space before the first level Heading and one line space after the first level Heading.
12. ***The second level Headings*** – 10 point, Upper and lower case, style Bold, alignment Left. One line space should be used before the second level Heading and 1/2 line space after the second level Heading.
13. ***The third level Headings*** – 10 point, Upper and lower case, style Italic, alignment Left. One line space should be used before the second level Heading and 1/2 line space after the third level Heading.
14. ***Text*** of the article – 10 point, single-spaced, alignment Justify.
15. The set of ***formulas*** on application of fonts, signs and a way of design should be uniform throughout the text. The set of formulas is carried out with use of editors of formulas MS Equation 3.0 or MathType. The formula with a number – the formula itself should be located on the left edge of the text, but a number – on the right. Font sizes for equations are: 11pt – full, 7pt – subscripts/superscripts, 5pt – sub- subscripts/superscripts, 16pt – symbols, 11pt – subsymbols.
16. All ***Figures*** – must be centered. Figure number and caption always appear below the Figure, type size 8 point.

Figure 1. This is figure caption

Diagrams, Figures and Photographs – must be of high quality, в in format *.TIFF, *.JPG, *.BMP with resolution not less than 300 dpi. Also formats *.CDR, *.PSD are possible. Combination of Figures in format, for instance, *.TIFF with elements of the in-built Figure Editor in MS Word is prohibited.

17. ***Table Number and Title*** – always appear above the Table. Alignment Left. Type size 8 point. Use one line space before the Table Title, one line space after the Table Title and 1/2 line space after the Table.

Table 1. This is an example of a Table

Heading	**Heading**	**Heading**
Text	Text	Text
Text	Text	Text

18. ***References*** in the text should be indicated by a number in square brackets, e.g. [1].
References should be numbered in the order cited in the manuscript. The correct format for references is the following:

Article: author, title, journal (in italics), volume and issue number, year, inclusive pages
Example: 1. Amrahamsson M., Wandel S. A Model of Tearing in Third – Party Logistics with a Service Parts Distribution Case Study, *Transport Logistics*, Vol. 1, No 3, USA: VSP, 1998, pp. 181-194.

Book: author, title (in Italics), location of publishers, publishers, year
Example: 2. Kayston M. and Fried W. R. *Avionic Navigation Systems.* New York: John Wiley and Sons Inc, 1969.

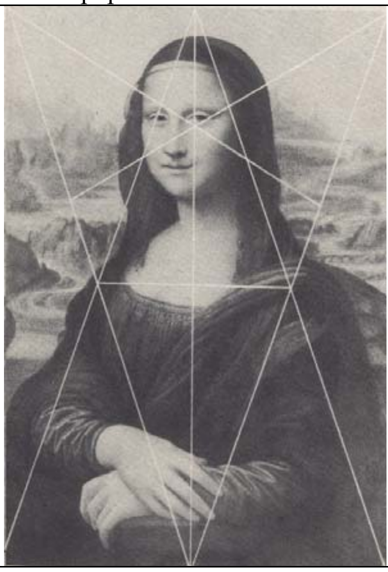
Conference Proceedings: author; title of an article; proceedings (in italics); title of a conference, date and place of a conference; publishing house, year, pages.
Example: 3. Canales Romero J. A First Step to Consolidate the European Association of Aerospace Students in Latvia (Presented by the Munich Local Group). In: *Research and Technology – Step into the Future: Programme and Abstracts. Research and Academic Conference, Riga, Latvia, April 7–11, 2003, Transport and Telecommunication Institute.* Riga: TTI, 2003, p. 20.

19. ***Authors Index***
Editors form the author's index of a whole Volume. Thus, all contributors are expected to present personal colour photos with the short information on the education, scientific titles and activities.
20. ***Acknowledgements***
Acknowledgements (if present) mention some specialists, grants and foundations connected with the presented paper. The first page of the contribution should start on page 1 (right-hand, upper, without computer page numbering). Please paginate the contributions, in the order in which they are to be published. Use simple pencil only.
21. **Articles poorly produced or incorrectly formatted may not be included in the proceedings**.

NATURE AND COMPUTER WORLD
Scientific popular journal

ПРИРОДА И КОМПЬЮТЕРНЫЙ МИР
Научно-популярный журнал

DABA UN DATORPASAULE
Zinātniski populārais žurnāls



PUBLISHERS:
TRANSPORT AND TELECOMMUNICATION INSTITUTE
***TRANSPORTA UN SAKARU INSTITŪTS** (TSI)*
&
INFORMATION SYSTEMS MANAGEMENT INSTITUTE
***INFORMĀCIJAS SISTĒMU MENEDŽMENTA AUGSTSKOLA** (ISMA)*

NATURE AND COMPUTER WORLD,
PUBLISHED SINCE 2001
ISSN 1407-8422
ISSN 1407-8422 **- on-line**

Editorial office:
Nature and Computer World,
Lomonosova 1, Build. 4, LV-1019, Riga, Latvia
shunin@isma.lv, kiv@tsi.lv

N&CW is the new scientific popular journal addressed to senior school students. It pays attention to some actual fields of undergraduate studies:

- Mathematics
- Physics
- Computers
- Programming and Computer Modelling
- Nature Phenomena

We are open for collaboration in these directions in the framework of N&CW and ready to consider corresponding materials. All articles are reviewed.