An Application of Graph Theory in the Process of Mutual Debt Compensation

Vladimír Gazda, Denis Horváth, Marcel Rešovský

Technical University of Košice, Faculty of Economics; Němcovej 32, 040 01 Košice, Slovakia, vladimir.gazda@tuke.sk

Pavol Jozef Šafárik University in Košice, Faculty of Science; Park Angelinum 9, 040 01 Košice, Slovakia, denis.horvath@upjs.sk

Technical University of Košice, Faculty of Economics; Němcovej 32, 040 01 Košice, Slovakia, marcel.resovsky@tuke.sk

Abstract: In our paper, the class of graph tractable problems of mutual debt compensation (MDC) among firms is introduced. We demonstrate that the debt compensation is related to the optimization problems maximizing the circulation of the consecutive compensation in the digraph of debts. Each optimization problem is firstly formulated by linear programming methods. It is subsequently redeveloped in order to apply the more efficient Klein's cycle-cancelling algorithm. The class of the formulated compensation problems consists of the following models: (i) the model maximizing the returns of the MDC organizer; (ii) the model of profitable subsidies and (iii) the model minimizing the claims exposition of the MDC organizer.

Keywords: mutual debt compensation; digraph of debts; Klein's cycle cancelling

1 Introduction

The increasing complexity of the economic environment and the large number of heterogeneous agents participating in the market necessitate a new approach in modeling creditor–debtor relationships. The seminal papers of Diamond and Dybvig [6] and Allen and Gale [2] [3] offer the graph theory as an efficient tool for modeling in this field. The subsequent boosting research performed mainly during the last decade was primarily motivated by the existing financial crises, the development of the complexity theory and the limitations of the traditional econometric techniques. The research results demonstrate a significant impact of the financial network structure on the financial risks given by the potential extent of the insolvency contagion spreading through the net. Here, Allen and Gale [3], comparing cycle and complete graph structures, conclude that the latter is more

resilient to liquidity shocks. On the other hand, the application of more realistic network structures using different methodologies leads to ambiguous results or even controversy. For instance, Nier et al. [14] introduce the idea that "the effect of the degree of connectivity is non-monotonic, i.e. initially a small increase in connectivity increases the contagion effect; but after a certain threshold value, connectivity improves the ability of a banking system to absorb shocks." On the other hand, according to Vivier-Lirimont [19], "financial fragility clearly depends on several network characteristics. In particular, the higher the network connectivity, the larger the number of banks involved in the contagion process, and the quicker the contagion phenomenon." Despite the above-given controversies, we base our approach on the intuition that the increasing number of debt relations has a positive impact on the danger of financial insolvency. That is also a reason why the reduction of the debt relations becomes the primary goal of our article.

However, determining a way to find possibilities for debt reduction in financial networks, is not a new topic. Activities mobilizing economic agents to settle debts and cut losses from the unpaid debts have been discussed in financial research for a long time. Here, despite having many common features, the approaches differ in some details. Rotemberg [15] deals with the problem of debt repayment between ultimate lenders, ultimate borrowers and intermediaries. He considers rather consecutive repayment of the debts that is restricted by the number of transactions per time unit than debt compensation in the structure. The presence of a cycle in a debt structure represented by a directed graph demonstrates the necessity for an outer source of liquidity. Settling all the debts in the structure depends on the order of the debt repayments made by particular subjects. Rotemberg [15] is concerned with the amount of additional liquidity provided by some subjects from outside the system to settle all the obligations inside the structure. Verhoeff [18] uses graph representation to solve the problem of the settling of multiple debts between agents. The goal is to find the optimal solution with an efficient method that minimizes the number of links (transfers) between creditors and debtors and the total amount of money transferred. An important assumption about the zero total balance of transfers has been made. The efficiency of the transfer process may be expressed by the limitation that no more than $N-1$ transfers are needed in the graph with N vertices (N represents the number of indebted agents). Moreover, if the cycle sub-graph exists in the structure of debts, the minimum amount of owed money can be subtracted from each graph link. Boerner and Hatfield [4] analyze the clearing process of debts between agents in the balanced (the debts vs. claims position is zero for each agent) and the unbalanced financial position. In the balanced case, the cycle removal mechanism that clears all debts is used, while in the case of unbalanced positions of agents, consecutive removal of debts in the chain is applied.

Our model is built on the aforementioned concepts; however, it differs in many important aspects and motivations. The approach we use is devoted to the

compensation of debts among the agents. Our methodology focuses on the application of the standard graph theory algorithm to solve the debt compensation process among firms, although the problem can be rewritten as an application of the linear programming method with the minimization of debts in the given structure as an objective function. The novelty can be seen in the introduction of the following models: (i) the model maximizing the returns of the MDC organizer; (ii) the model of profitable subsidies and (iii) the model minimizing the claims exposition of the MDC organizer.

Our paper is organized as follows. In Section 2, the concept of a digraph of debts is introduced. The formalism of the mutual debt compensation (MDC) process is presented in Section 3. The standard graph approach of Klein's cycle cancelling algorithm, which solves MDC, is applied in Section 4. Modification of this algorithm leading to maximal return of the MDC organizer is described in Section 5. Sections 6 and 7 introduce the concept of a mutual debt compensation organizer in the form of a subsidy centre and the Ministry of Finance. Finally, the conclusions are presented.

2 Digraph of Debts

Let $V = \{1,2,\ldots,n\}$ stand for a set of firms involved in a process of mutual debt compensation. Let the debt that firm i owes to firm j be denoted as $y_{i,j}$. We assume that, at most, one of two mutually exclusive situations $y_{i,j} > 0$ and $y_{j,i} > 0$ may occur. The debtor–creditor relationships are defined by a set

$$E = \{[i,j] \in V^2 \mid y_{i,j} > 0 \quad as \quad i \neq j\}. \tag{1}$$

Here, the function $y : E \to \mathbb{R}^+$ assigns a real positive debt amount to each debtor–creditor relation. Then, using the previous structures, we may introduce formally a digraph of debts as follows:

$$G = (V, E, y). \tag{2}$$

As set $V^-(i)$ denotes the in-neighborhood of vertex i (representing the list of its debtors) and set $V^+(i)$ denotes the out-neighborhood of vertex i (the list of its creditors), the receivables/debts balance of firm i may be expressed as

$$b(i) = \sum_{j \in V^-(i)} y_{j,i} - \sum_{j \in V^+(i)} y_{i,j} \qquad for \qquad i \in V. \tag{3}$$

In this formula, the total receivables of firm i, represented by $\sum_{j\in V^-(i)} y_{j,i}$, are in part balanced by its total debts, represented by $\sum_{j\in V^+(i)} y_{i,j}$. It is important to note that we assume fairness of MDC. This means that the balance between the receivables and the debts expressed by Eq. (3) have to remain constant during the MDC process. Quite obviously, the debt digraph property

$$\sum_{i=1}^{n} b(i) = 0 \tag{4}$$

states that the total sum of receivables/debts is balanced.

3 The MDC Process as Cycle Elimination

This section may be perceived as an intuitive introduction to the methodology based on the cyclic properties of digraphs. The presence of a cycle in the digraph means the existence of a cyclical substructure of mutual debts in the world of interconnected firms. Let $C \subseteq G$ be a cycle in G (G defined in Eq. (2)). Again, C may be written as ordered triple $(V(C), E(C), y)$ comprising the set of vertices $V(C)$ together with a set $E(C)$ of arcs quantified by the debt strength described by the corresponding y weights. By defining the capacity of cycle C as

$$y(C) = \min_{e\in E(C)} y_e, \tag{5}$$

the weights of the edges of C may be reduced by subtracting the cycle capacity $y(C)$ from the weights of all the arcs lying in the cycle[1]. The assignments

$$y_e := y_e - y(C), \qquad \forall e \in E(C) \tag{6}$$

yield debt compensation since they cause a reduction in the total debts in the original digraph G. At the same time, the debtor–creditor relations represented by at least one arc with zero weight are removed from cycle C (as well as from

[1] Further, we are using double notation of the arcs. First, the notation $y_{i,j}$ with subscripts comprising the ordered pair i, j is used if we want to stress the role of the respective vertex; second, the general notation of the arc e is used to indicate an element of E.

digraph G) and thus the cycle is eliminated from G. Eq. (3) rewritten as $b(i) = (\sum_{j \in V^-(i)} y_{j,i}) - y(C) - [(\sum_{j \in V^+(i)} y_{i,j}) - y(C)]$ demonstrates that the receivables/debts balance of each firm before and after the compensation remains conserved. The graph becomes acyclical due to consecutive application of the procedure to all the cycles present in G.

3.1 Example: The Change from a Cyclical to an Acyclical Graph

Consider a simple structure of five firms with the debt structure given in Figure 1A. The debt structure creates a cycle given by a sequence of vertices and arcs 1, [1,2], 2, [2,3], 3, [3,4], 4, [4,5], 5, [5,1] (i.e. a cycle with vertices $V(C) = \{1,2,3,4,5\}$ and arcs $E(C) = \{[1,2],[2,3],[3,4],[4,5],[5,1]\}$).

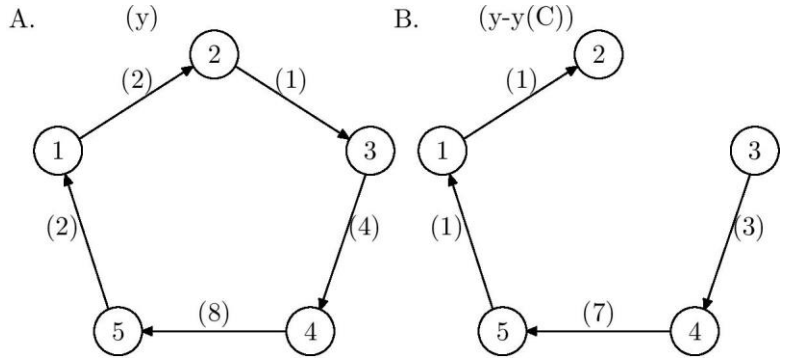


Figure 1
A) The debt structure before compensation; B) The debt structure after compensation

The minimum debt in the cycle is a debt of size 1 between 2 and 3. The debt structure in Figure 1B is obtained by the cycle elimination, i.e. subtracting 1 from all the debts in the cycle. The resulting digraph becomes acyclical.

4 MDC Based on Maximum Circulation in the Digraph of Debts

Consider a strongly connected digraph $G = (V, E, y)$ [2]. In this graph, we define a circulation function [see [7] p. 146] $x : E \to \mathsf{R}_0^+$, which satisfies inequality

$$0 \le x_{i,j} \le y_{i,j}, \tag{7}$$

and constraint

$$\sum_{j \in V^-(i)} x_{j,i} = \sum_{j \in V^+(i)} x_{i,j}, \quad i \in V. \tag{8}$$

Further, we will demonstrate the suitability of the circulation concept for MDC modeling. The variable $x_{i,j}$ is interpreted as part of a debt $y_{i,j}$, which may be compensated between firms i and j. Therefore, the total amount of debts compensated due to the pair-wise effect of circulation may be written as follows:

$$\sum_{[i,j] \in E} x_{i,j}. \tag{9}$$

In the MDC process, the receivables/debts balance of any firm, must not be worsened or improved. The relation in Eq. (8) reports the compensation of the receivables and debts on the level of single firm i. Therefore, it is quite obvious to expect that respecting this condition should not violate the balance of the receivables/debts. Assume that before the compensation process starts, we have $b(i)$ expressed in the form of Eq. (3). Subsequently, adding zero in the form of the right-hand side minus the left-hand side of Eq. (8) leaves $b(i)$ unchanged:

$$b(i) = \sum_{j \in V^-(i)} \left(y_{j,i} - x_{j,i} \right) - \sum_{j \in V^+(i)} \left(y_{i,j} - x_{i,j} \right). \tag{10}$$

The original debts $y_{i,j}$ are now expressed in the compensated form. From this it follows that the minimization of debts becomes equivalent to the identification of maximal circulation

$$\max_x \sum_{[i,j] \in E} x_{i,j} \tag{11}$$

[2] The digraph $G = (V, E, y)$ is strongly connected if for every pair of vertices $i, j \in V$ a direct path exists from i to j as well as from j to i.

subject to the constraints expressed by Eq. (7) and Eq. (8). The main reason why we decided to pass to the maximum circulation problem is that it can be tackled by means of the standard linear programming methods, as well as, standardized special-purpose graph algorithms. For example, the application of Klein's cycle cancelling (see [12]), originally applied to find the minimum cost circulation, might be possible[3]. The algorithm is based on the identification of all the cycles and the determination of the optimal sequence of the elimination in the digraph respecting the trade-off among the cycle capacities (see [7]). Based on Eq. (7), we can conclude that if $y_{i,j} = 0$ then $x_{i,j} = 0$. Thus, the total debt minimization does not cause the creation of new creditor debtor relations, which is considered as an important aspect of the proposed model.

4.1 Example: Cycle Cancelling – from a Cyclical to an Acyclical Graph

Figure 2 describes the mutual debts of five firms. The total amount of mutual debts is equal to the sum of arc weights, i.e. \$25 mil. There are three cycles in the digraph. The first one is given by the debt sequence 1, [1,2], 2, [2,3], 3, [3,4], 4, [4,5], 5, [5,1], 1, the second one by the sequence 1, [1,2], 2, [2,5], 5, [5,1], 1 and the third one by the sequence 1, [1,3], 3, [3,4], 4, [4,5], 5, [5,1], 1. All these cycles contain a common arc [5,1], which determines the minimum capacity in the second and the third cycle. Consequently, the elimination of the second cycle causes the elimination of arc [5,1] and thus the elimination of the first and third cycles, too. On the other hand, the elimination of the third cycle causes the elimination of the first and the second cycle (see Figure 3). In both approaches, the digraph becomes acyclical, meaning that no more debts may be compensated by the cycle cancelling. The total amounts of uncompensated debts are different in the two cases (see Figure 3A, Figure 3B) – \$19 mil. in the first case and \$17 mil. in the second one. It is quite obvious that the most appropriate method is to eliminate cycle 1, [1,2], 2, [2,3], 3, [3,4], 4, [4,5], 5, [5,1], 1 first and then to eliminate 1, [1,3], 3, [3,4], 4, [4,5], 5, [5,1], 1. We can see that elimination applied in a suitable order causes the total amount of the uncompensated debts to remain at the level of \$16 mil. (see Figure 4).

[3] We choose Klein's algorithm for demonstrative purposes; more efficient circulation algorithms were proposed later (Goldberg and Tarjan [10] is a good example).

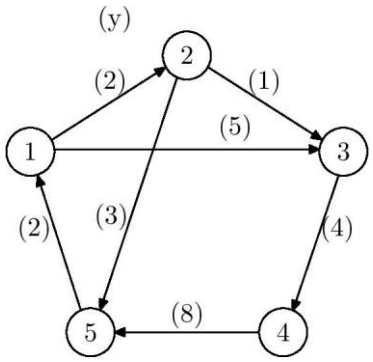


Figure 2
The structure of debts before their compensation

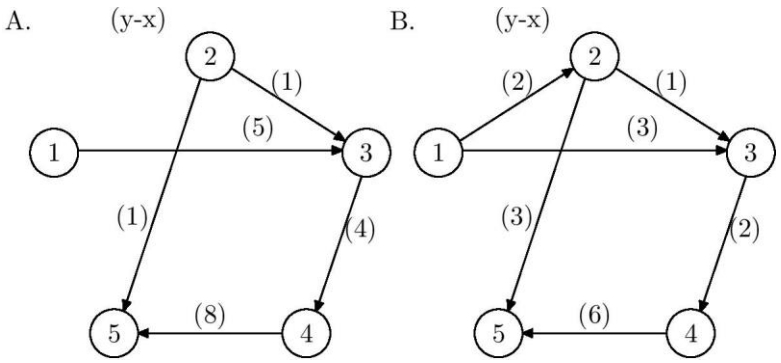


Figure 3
The remaining debt structure if: A) cycle 1, [1,2], 2, [2,5], 5, [5,1], 1 is eliminated; B) cycle 1, [1,3], 3, [3,4], 4, [4,5], 5, [5,1], 1 is eliminated

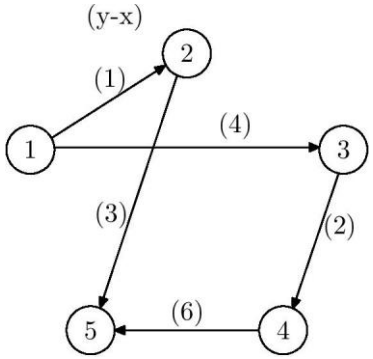


Figure 4
The structure of debts that stems from the application of cycle cancelling

The above-formulated MDC heuristic is rather intuitive and does not provide the optimal ordering of the cycle elimination. That is why it is inapplicable in the case of large debt networks with a high level of density. Here, we decided to maximize the graph circulation by the modified Klein's cycle-cancelling method (see [12]), which is extremely effective in enabling us to solve the extended graph problems. However, we avoid the usage of large graphs in this article because our pictures are intended to be easy to understand. We focus mainly on the economic aspects of the debt problems. We did not want to spend too much time on some of the finer details of the standard algorithm.

5 Return of the Mutual Debt Compensation Organizer

Companies themselves are mostly unable to solve the problem of debts, which needs extensive multi-company projects that can help to solve the problem by initiating and organizing the process of debt compensation. In practice, the organizer of MDC expects to be rewarded. Its return function $c : E \to \langle 0;1 \rangle$ assigns a unit price per monetary unit of each compensated debt. More detailed information may be captured by digraph $G = (V, E, y, c)$. The total income of the organizer can be written as

$$\sum_{[i,j]\in E} c_{i,j} x_{i,j}. \tag{12}$$

The maximal return due to application of the circulation may be defined as follows:

$$\max_{x} \sum_{[i,j]\in E} c_{i,j} x_{i,j}. \tag{13}$$

Again, the optimization is considered in common with the constraints defined by Eq. (7) and Eq. (8).

6 The Subsidy Center and the Problem of Restricted Subsidies

In practice, there are a few examples in which a non-participating subject enters a debt structure to endow a highly indebted firm. It enables it to cover its debts, which consequently allows the elimination of the whole chain of induced debts caused by the insolvency of a highly indebted firm. Now, we formulate a model in

which the subsidy centre does not know a priori which firm should be subsidized. Conversely, the list of firms will be determined by the algorithm. In addition, we assume that the sources available for providing the subsidies are restricted. The vertex $(n+1)$ is added to substitute the effect of the subsidy centre. All of its relations will be represented by $[n+1,i], i \in V$ arcs. Assuming this structure, the linear programming model from Eq. (13) takes the form

$$\max_{x} \left[\sum_{[i,j]\in E} c_{i,j} x_{i,j} - \sum_{i\in V} x_{n+1,i} \right], \tag{14}$$

$$\sum_{j\in V^-(i)} x_{j,i} - \sum_{j\in V^+(i)} x_{i,j} = 0 \quad for \quad i \in V = \{1,2,\ldots,n\}, \tag{15}$$

$$\sum_{i\in V} x_{n+1,i} \le B, \tag{16}$$

$$0 \le x_{i,j} \le y_{i,j} \quad for\ [i,j] \in E, \tag{17}$$

$$0 \le x_{n+1,i} \le y_{n+1,i} \quad for \quad i \in V. \tag{18}$$

According to Eq. (16), the total amount of the subsidies provided is restricted by constant B. The subsidy provided to the i^{th} firm is restricted by its upper bound, expressed by Eq. (18). Now, we finish the specific MDC formulation for the purposes of linear programming. This approach is, however, less efficient for the solving of many vertex graph problems. Greater efficiency may be gained by Klein's cycle-cancelling method. In the next section, we will discuss its adaption to the given problem.

6.1 Klein's Cycle Cancelling Adapted to the Model of the Subsidy Centre

Vertex $(n+1)$ does not include any incident input edges; no circulation through this vertex is possible. As a solution, we propose a modified debt digraph $\overline{G} = (\overline{V}, \overline{E}, \overline{y}, \overline{c})$ with modified topology including subsidy vertex $(n+1)$ as well as auxiliary vertex 0. Formally,

$$\overline{V} = V \cup \{n+1\} \cup \{0\}. \tag{19}$$

All the edges $\{[n+1,i], \forall i \in V\}$ weighted by $y_{n+1,i}$ represent the possible support provided by the subsidy centre localized at the vertex $(n+1)$. The arc $[0,n+1]$ enables us to create the flow condition (see constraint Eq. (8))

balancing the subsidy centre. Vertex 0 is balanced via additional edges $\{[i,0], \forall i \in V\}$. Then, the edges of $\overline{G}$ may be specified as follows:

$$\overline{E} = E \cup \{[i,0] \,|\, i \in V\} \cup \{[0,n+1]\} \cup \{[n+1,i] \,|\, i \in V\}. \tag{20}$$

We now specify function $\overline{y} : \overline{E} \to \mathsf{R}^+$, which consists of the additional rules with respect to the original y

$$\overline{y}_{i,j} = y_{i,j} \quad for \quad [i,j] \in E, \tag{21}$$

$$\overline{y}_{i,0} = y_{n+1,i} \quad for \quad i \in V, \tag{22}$$

$$\overline{y}_{0,n+1} = \min\left(B, \sum_{i \in \overline{V}^-(n+1)} y_{n+1,i} \right), \tag{23}$$

$$\overline{y}_{n+1,i} = y_{n+1,i}. \tag{24}$$

Eq. (21) shows that the distribution of debts remains the same as in the original debt digraph G, while Eq. (23) describes the total amount of disposable subsidies. The remaining constituent of $\overline{G}$ represents the return $\overline{c} : \overline{E} \to \mathsf{R}$. The costs of debt reduction per monetary unit are expressed by $\overline{c}_{n+1,i}$

$$\overline{c}_{i,j} = c_{i,j} \quad for \quad [i,j] \in E, \tag{25}$$

$$\overline{c}_{i,0} = 0 \quad for \quad i \in V, \tag{26}$$

$$\overline{c}_{0,n+1} = 0, \tag{27}$$

$$\overline{c}_{n+1,i} = -1 \quad for \quad i \in V. \tag{28}$$

After the above modifications, we obtain digraph $\overline{G}$, which enables the application of the cycle-cancelling algorithm for the maximum return circulation.

6.2 Example: The Role of the Subsidy Centre

The structure of the mutual debts among the firms is represented in the digraph in Figure 2. We considered \$25 mil of the total mutual debts. The coefficients $c_{i,j}$ are set to be equal to 0.5 for all the arcs' returns. We assume that the subsidy centre ($(n+1) = \text{vertex}6$) is allowed to provide a subsidy by an amount equal

to \$2 mil. to the firms $1,3,5$. The total sum of the subsidies is restricted by the upper bound $B =$ \$3 mil. The diagram of this situation is presented in Figure 5.

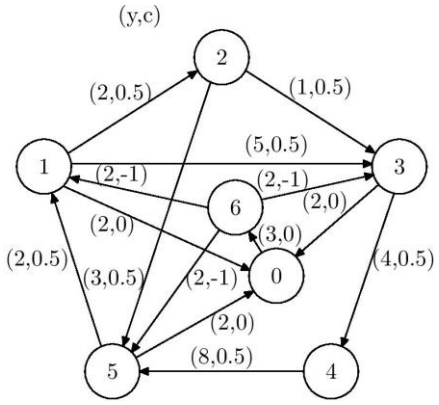


Figure 5
The initial situation before the compensation of debts

The optimal solution of the above-mentioned optimization model is given in Figure 6A. The sum of the compensated debts is \$15 mil. The subsidy center provides a subsidy to firm 1 to the amount of \$2 mil. The residual debts are represented in Figure 6B. It shows that the subsidies available were not completely used and \$1 mil. remained. The MDC organizer achieves a profit of \$5.5 mil.

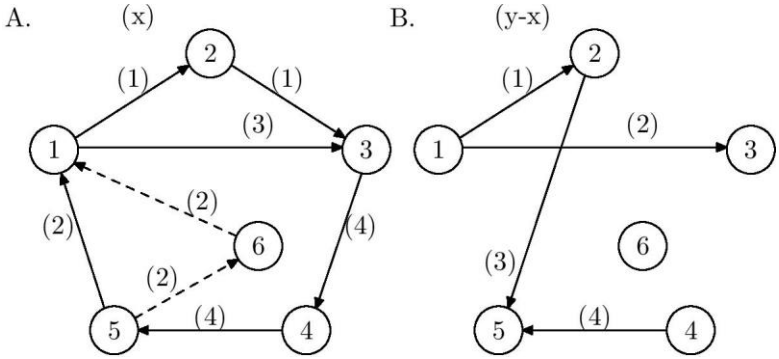


Figure 6
A) The circulation of compensated debts; B) The structure of mutual debts that results from the process of their compensation

7 Minimization of Receivables by the Ministry of Finance

The organizer of the mutual debt compensation is unlikely to be interested in diminishing the total debts in the whole complicated structure of debts; instead, he claims to decrease the receivables/debts of a single particular subject. This motivated us to consider a realistic situation, in which the role of the organizer of debt elimination is the Ministry of Finance. On one hand, the Ministry collects its income from firms (e.g. by way of taxes), but on the other hand it spends its budget by carrying payments to the firms providing services. Here we do not pay attention to the financial relations with another subjects. Let $V = \{1,2,\ldots,n\}$ be a set of firms. By formulation in terms of graphs, the Ministry of Finance, represented by vertex $n+1$, is intended for mutual debt compensation. Then, a set $\overline{V} = V \cup \{n+1\}$ involves all the subjects included in the compensation process. Let $y_{i,n+1}$ represent a debt of firm i towards the Ministry and $y_{n+1,j}$ represent the receivable of firm j towards the Ministry. The comprehensive debt structure is represented by

$$\overline{E} = E \cup \{[n+1, j] \,|\, j \in V,\, y_{n+1,j} > 0\} \cup \\ \cup \{[i, n+1] \,|\, i \in V,\, y_{i,n+1} > 0\}, \tag{29}$$

where the structure of debts E is defined in Section 2. The function $\overline{y}$ assigning a particular debt to each debt relation is defined as follows:

$$\overline{y} : \overline{E} \to \mathsf{R}^{+}, \tag{30}$$

where

$$\overline{y}_{i,j} = y_{i,j} \quad for \quad [i, j] \in E, \tag{31}$$

$$\overline{y}_{i,n+1} = y_{i,n+1} \quad for \quad i \in V, \tag{32}$$

$$\overline{y}_{n+1,j} = y_{n+1,j} \quad for \quad j \in V. \tag{33}$$

The digraph $\overline{G} = (\overline{V}, \overline{E}, \overline{y})$ represents the mutual debt structure among the firms and the Ministry. The receivables/debts balance of the Ministry is expressed as

$$\sum_{i \in \overline{V}^{-}(n+1)} \overline{y}_{i,n+1} - \sum_{j \in \overline{V}^{+}(n+1)} \overline{y}_{n+1,j} = b(n+1). \tag{34}$$

By following the ideas introduced above, we define circulation

$$x:\overline{E}\to \mathsf{R}^{+};\quad 0\le x_{i,j}\le \overline{y}_{i,j};\quad i\in\overline{V}, \tag{35}$$

$$\sum_{j\in\overline{V}^{-}(i)} x_{j,i} = \sum_{j\in\overline{V}^{+}(i)} x_{i,j}, \tag{36}$$

as mutual debt compensation given in a digraph of debts. It should be noted that if the receivables/debts balance of subjects $1,2,\ldots,n$ in digraph $\overline{G}=(\overline{V},\overline{E},\overline{y})$ remains constant in the process of the mutual debt compensation, then the balance of subject $n+1$ remains constant, too. This prevents the Ministry from improving its receivables/debts balance by organizing the mutual debt compensation. The proof is based on the fact that the sum of the receivables/debts balances of all the subjects before debt compensation is (see Section 3)

$$\sum_{i\in\overline{V}}\Big(\sum_{j\in\overline{V}^{-}(i)}\overline{y}_{j,i}\Big)-\sum_{i\in\overline{V}}\Big(\sum_{j\in\overline{V}^{+}(i)}\overline{y}_{i,j}\Big)=0. \tag{37}$$

The balance condition must also be satisfied after the compensation, i.e. for

$$\sum_{i\in V}\Big(\sum_{j\in\overline{V}^{-}(i)}(\overline{y}_{j,i}-x_{j,i})-\sum_{i\in V}\Big(\sum_{j\in\overline{V}^{+}(i)}(\overline{y}_{i,j}-x_{i,j})\Big)+$$

$$+\sum_{i\in\overline{V}^{-}(n+1)}(\overline{y}_{i,n+1}-x_{i,n+1})-$$

$$-\sum_{i\in\overline{V}^{+}(n+1)}(\overline{y}_{n+1,i}-x_{n+1,i})=0. \tag{38}$$

If the fairness condition of the compensated debts is satisfied for each of the n participating firms

$$\sum_{j\in\overline{V}^{-}(i)} x_{j,i} = \sum_{j\in\overline{V}^{+}(i)} x_{i,j}, \tag{39}$$

the following condition must hold, too

$$\sum_{i\in V}\sum_{j\in\overline{V}^{-}(i)} x_{j,i} = \sum_{i\in V}\sum_{j\in\overline{V}^{+}(i)} x_{i,j}. \tag{40}$$

It directly implies the fairness of the compensation of the Ministry of Finance:

$$\sum_{i\in\overline{V}^{-}(n+1)} x_{i,n+1} = \sum_{i\in\overline{V}^{+}(n+1)} x_{n+1,i}. \tag{41}$$

The problem of the minimization of Ministry receivables may be modeled as the maximization of the objective function:

$$\max \sum_{i \in \overline{V}^{-}(n+1)} x_{i,n+1} \tag{42}$$

$$\sum_{j \in \overline{V}^{-}(i)} x_{j,i} = \sum_{j \in \overline{V}^{+}(i)} x_{i,j} \quad for \quad i \in \overline{V}, \tag{43}$$

$$0 \le x_{i,j} \le \overline{y}_{i,j} \quad for \quad [i,j] \in \overline{E}. \tag{44}$$

Such formulation admits the application of the linear programming methods. They are efficient for graphs of a small or moderate size. On the other hand, Klein's cycle-cancelling method, discussed in the next subsection, represents a more suitable tool for larger graphs.

7.1 Receivables of the Ministry of Finance Tackled by Klein's Cycle-Cancelling Method

The implementation of new virtual vertex 0 in order to apply Klein's cycle-cancelling method yields

$$\widetilde{V} = \overline{V} \cup \{0\}. \tag{45}$$

Then, all the arcs entering vertex $(n+1)$ are redirected to the virtual vertex 0, i.e.

$$\{[i,n+1] \,|\, i \in V; y_{i,n+1} > 0\} \rightarrow \{[i,0] \,|\, i \in V\} \tag{46}$$

and a new arc $[0,n+1]$ is added. It yields the set of arcs

$$\widetilde{E} = E \cup \{[i,0] \,|\, i \in V,\, y_{i,n+1} > 0\} \cup \{[0,n+1]\} \cup$$
$$\cup \{[n+1,i] \,|\, i \in V,\, y_{n+1,i} > 0\}. \tag{47}$$

The function $\widetilde{y} : \widetilde{E} \rightarrow \mathsf{R}^{+}$ is defined as

$$\widetilde{y}_{i,j} = y_{i,j} \quad for \quad [i,j] \in E, \tag{48}$$

$$\widetilde{y}_{i,0} = y_{i,n+1} \quad for \quad i \in \overline{V}^{-}(n+1), \tag{49}$$

$$\widetilde{y}_{0,n+1} = \sum_{i \in \overline{V}^{-}(n+1)} y_{i,n+1}, \tag{50}$$

$$\tilde{y}_{n+1,i} = y_{n+1,i} \quad for \quad i \in V. \tag{51}$$

Due to the fact that digraph $\tilde{G} = (\tilde{V}, \tilde{E}, \tilde{y})$ contains vertex 0, the minimization of the receivables becomes equivalent to the finding circulation with the maximum value along arc $[0, n+1]$. This leads us to the application of the aforementioned maximum return circulation method with the return function elements

$$\tilde{c}_{i,j} = 0 \quad for \quad [i, j] \in E, \tag{52}$$

$$\tilde{c}_{i,0} = 0 \quad for \quad i \in V, \tag{53}$$

$$\tilde{c}_{0,n+1} = 1, \tag{54}$$

$$\tilde{c}_{n+1,i} = 0 \quad for \quad i \in V. \tag{55}$$

7.2 Example: The Ministry of Finance as the MDC Organizer

Let the initial situation of mutual debts be represented by the digraph in Figure 7A. The vertex enumerated by 6 represents the Ministry of Finance. The Ministry has receivables towards firm 5 to an amount of \$2 mil. and towards firm 4 to an amount of \$3 mil. The Ministry owes \$5 mil. to firm 1 and \$1 mil. to firm 3. We assume that the Ministry organizes the compensation of mutual debts whereby the minimizing of its receivables is its obvious objective function (see Eq. (42)).

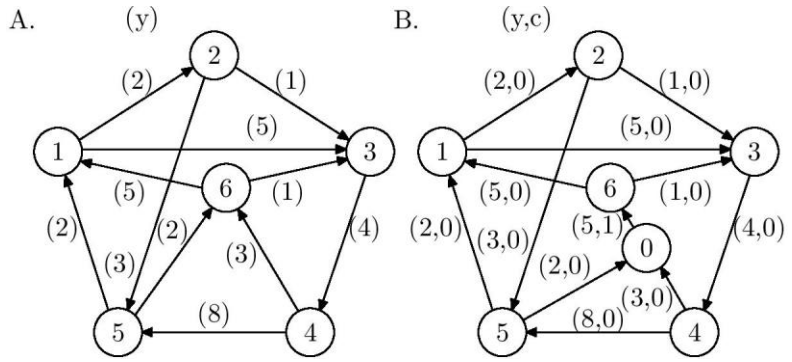


Figure 7

The initial situation before debt compensation

A) $\overline{G} = (\overline{V}, \overline{E}, \overline{y})$ to be solved by the linear programming methods

B) $\tilde{G} = (\tilde{V}, \tilde{E}, \tilde{y})$ to be solved by an appropriate cycle-cancelling method

In order to enable the application of the circulation algorithm, we extend the digraph that includes vertex 0 (see Figure 7B). Accordingly, the arcs [4,6] and [5,6] with weights 2 and 3 in Figure 7A are redirected to the 0 (see Figure 7B). In addition, a new arc [0,6] with debt weight 5 is created. This modified digraph becomes ready for the application of the circulation maximization along [0,6]. After the application of the cycle-cancelling method, the residual debts diminish to 14 (see Figure 8B). All the cycles are eliminated from the final form of the digraph.

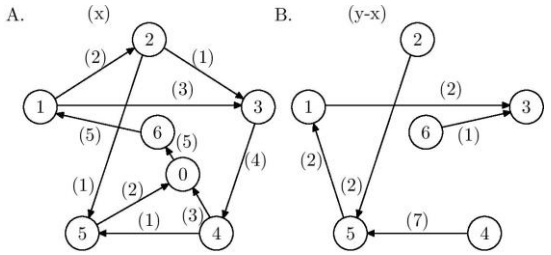


Figure 8

A) The circulation of the compensated debts; B) The resulting structure of debts

Conclusion

In this paper, we discuss the principles of the graph theory that are relevant to performing the MDC procedure. Principally, new formulations are presented, including models with a subsidy center and their conversion into a form tractable by Klein's cycle-cancelling algorithm. Besides the solving of the general task of the cycle cancelling, we demonstrate the method's applicability to the problem of receivables of a single selected firm. Our research shows that both topic and methodology are still far from being exhausted. The approach can be adapted straightforwardly to digraphs of the International relationships among the countries–vertices equipped with the import–export links. The impulses for ongoing research may come from studies on the altering nature of the economic world. Dynamic graph models of debt structures should be introduced that formalize the concept of the flexible links. These can be transformed under the influence of the stochastically varying external conditions represented by the reevaluating/devaluating domestic currency.

Acknowledgement

The paper was elaborated within the project VEGA No 1/1195/12.

References

[1] Ahuja R. K., Magnanti T. L., Orlin J. B.: Network Flows: Theory, Algorithms, and Applications. Prentice Hall, University of Michigan (1993)

[2] Allen F., Gale D.: Optimal Financial Crises. Journal of Finance (1998), 4: 1245-1284

[3] Allen F., Gale D.: Financial Contagion. Journal of Political Economy (2000), 1: 1-33

[4] Boerner L., Hatfield J. W.: The Economics of Debt Clearing Mechanisms. Working Paper No. 2010/27, School of Business, Economics Discussion Paper (2010)

[5] Chartrand G., Oellermann O. R.: Applied and Algorithmic Graph Theory. McGraw-Hill, University of California (1993)

[6] Diamond D. W., Dybvig P. H.: Bank Runs, Deposit Insurance, and Liquidity. Journal of Political Economy (1983), 5: 401-419

[7] Diestel R.: Graph Theory. Springer-Verlag, London (1997)

[8] Fecenko J.: About Optimisation of Receivables and Payables Compensation. Ekonomický časopis (1994), 42: 360-374

[9] Gazda V.: Mutual Debts Compensation as Graph Theory Problem. In: Mathematical Finance (2001)

[10] Goldberg A. V., Tarjan R. E.: Finding Minimum-Cost Circulations by Canceling Negative Cycles. Journal of the ACM (1989), 36(4): 873-886

[11] Harary F.: Graph Theory. Addison-Wesley Publishing Company, London (1969)

[12] Klein M.: A Primal Method for Minimum Cost Flow with Application to the Assignment and Transportation Problem. Management Science (1967), 14(3): 205-220

[13] Jun Q., Wing L. N.: Asymmetric Liquidity Risk Premia in Intraday High Frequency Trading. International J. of Applied Mathematics (2010), 40(1)

[14] Nier E., Yang J., Yorulmazer T., Alentorn A.: Network Models and Financial Stability. Journal of Economic Dynamics and Control (2007), 31(6): 2033-2060

[15] Rotemberg J. J.: Liquidity Needs in Economies with Interconnected Financial Obligations. National Bureau of Economic Research, Working Paper (2008)

[16] Tarjan R. E.: Depth First Search and Linear Graph Algorithms. SIAM J. Computing (1972), 1: 146-160

[17] Upper C.: Simulation Methods to Assess the Danger of Contagion in Interbank Markets. Journal of Financial Stability (2011), 7(1): 111-125

[18] Verhoeff T.: Settling Multiple Debts Efficiently: An Invitation to Computing Science. Informatics in Education (2004), 3(1): 105-126

[19] Vivier-Lirimont S.: Contagion in Interbank Debt Networks. Working Paper (2006)

Developing a Fuzzy Analytic Hierarchy Process for Choosing the Energetically Optimal Solution at the Early Design Phase of a Building

András Szűts, István Krómer

Óbuda University
Bécsi út 96/b, H-1034 Budapest, Hungary
szuts.andras@phd.uni-obuda.hu, kromer.istvan@kvk.uni-obuda.hu

Abstract: Reducing the energy consumption of households and making them more energy efficient are indispensable for fulfilling the European Union's energy strategy since more than one-third of total EU energy use is consumed by households. Our research here focuses on the development of building design methods since, decisions made during the early design phase significantly affect the energy consumption and energy efficiency of the building. In Hungary, current simulation methods are not sufficiently effective in evaluating potential alternatives in the early design process; therefore, in our research, we look for new solutions based on decision support systems. In our view, either a classic AHP or a Hybrid Fuzzy AHP based system could fulfil this task. In this article, we will demonstrate the structure and use of the developed systems, and we will compare them to the simulation methods commonly used nowadays. Further, we will test and evaluate the performance of the decision support system based models developed by us, applying them to four typical buildings in Hungary. Lastly, we will propose opportunities for further development.

Keywords: fuzzy logic controllers; analytic hierarchy process; energy; households

1 Introduction

In Hungary and in the European Union, households consume more than one-third of the total energy usage; thus, it is obvious that in order to achieve EU energy policy objectives, to enhance energy efficiency, and to make more secure the energy supply, it is necessary to reduce household consumption. To reduce residential energy consumption, the European Union has developed multiple strategies, which will be described later. There are several suitable methods for solving the energy efficiency problems related to household energy consumption, as well as combinations of these. In our research, we focus on the development of design methods to increase energy efficiency during this phase.

Currently in Hungarian building design practices, very little attention is paid to energy efficiency, and, unfortunately, building planners generally seek only to meet the requirements of the law. Consequently there are only a few exemplary projects of energy efficiency in the country [1]. This scarcity of attention is due to a lack of regulation, as well as to the very high price sensitivity of Hungarian consumers, as compared to Western Europe. Given this situation, it would be beneficial to improve the design process and to modify the design methods in such a way that the designer is able to effectively and tangibly demonstrate the energy efficiency choices made and the long-term benefits of the different alternatives. With a more effective planning method, it might be possible during the design process to achieve the highest possible energy efficiency without causing an excessive increase in cost to the customer or investor, and without asking the impossible of the designer. It is important to note that in our research we concentrate on the early, conceptual stage of the building design process, because the decisions taken in this phase have the greatest influence on the energy consumption and energy efficiency of the planned or renovated building. The goal of the models developed is to rank the different building concepts developed in the early planning stages.

2 Energy Policy Objectives, the Available Saving Potential in the Sector

Over the past few years and decades, the European Union has announced several targets regarding environmental protection and energy awareness. The first significant action on the matter was Directive 20/20/20, which sets three legally binding targets for 2020: a 20% reduction in greenhouse gas emissions from 1990 levels; raising the share of EU energy consumption from renewable sources to 20%; and a 20% improvement in the EU's energy efficiency [2]. As a further step, the EU developed a long-term energy strategy that seeks by the year 2050 an 80-95% reduction in greenhouse gas emissions compared to the base year 1990 [3]. Accordingly, from the year 2021 the standard expectation for new buildings is that they will be nearly zero energy use buildings [4]. However, these strict requirements for new buildings do not themselves alone ensure the attainment of the climate and energy policy goals. Therefore, it is worth examining what degree of savings potential we have for existing and new buildings in the future, and in what proportion the different fields of application will share these various savings.

In our research, we have examined in detail the structure and the inherent savings potential in household energy consumption. Based on analyses made in the European Union, it can be established that the total household sector up to the year 2050 has a saving potential of 71% [5]. The potential savings can be realized in the following areas:

- Construction of new buildings (14%),
- Renovation of existing building structures (29%),
- Modernization of heating systems (16%),
- Modernization of the domestic hot water supply systems (4%),
- Lighting energy efficiency (4%)
- Energy efficiency of electrical equipment (4%).

Based on the preceding findings, it is clear that a significant proportion of the available savings potential is related to the energy efficient renovation and transformation of existing buildings. This supports the statement made above that strict regulations for new buildings alone will not be sufficient for achieving the desired objectives. It would clearly be useful to encourage investors and owners in complex energy developments to improve energy efficiency in the very conception of these projects, and not only in the changing and/or renovating of certain structures and equipment, as is currently the practice. We would also emphasize the finding that a 4% energy savings could be achieved via the energy requirements of lighting, while another 4% of savings could be achieved by increasing the efficiency of electrical equipment (meanwhile noting that, according to current legislation, the energy demands of lighting and electrical equipment need not be included in the energy balance of residential and housing-type buildings) [6]. Considering this 8% potential, investments made in these areas could yield the earliest returns from energy savings.

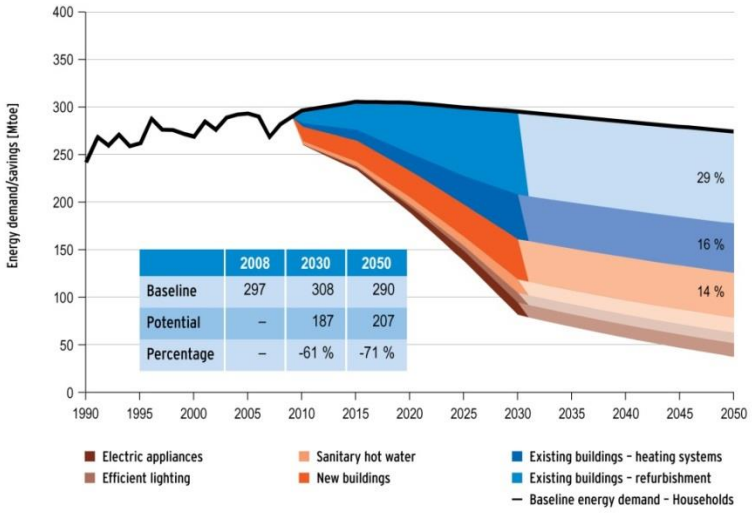


Figure 1
Total final energy saving potentials in the household sector in the EU
[source: Fraunhofer ISI, 5]

3 New Design Methods

Knowing the available savings potential, we should examine the factors that affect a household's energy consumption, factors which – in our opinion – can be divided into two separate groups. One group of factors includes consumer behaviour and habits, an area almost completely ignored in current design practices. The other group is composed of technical and other design variables. In our previous research we have already dealt with the uncertainties arising from consumer habits. We created an artificial neural network based inference system that can help estimate the annual primary energy consumption of a household based on statistics. This system makes manageable the uncertainties arising from consumer habits [1]. This current article therefore focuses on technical issues and other design variables.

Examining the group of the factors mentioned above, in our research we analyzed the design process, paying particular attention to the decision-making processes during planning. Having examined current Hungarian design practices, we can conclude that, in the early planning stages, decision-makers generally rely on their experience, rather than on facts or data supported by decision support systems. Therefore, it would be beneficial to develop a decision support system that would help both the designer and the developer in the early stages of planning to choose the optimal variants as regards energy efficiency. In order for the decision support model to actually be adopted, it should have relatively low resource demand as well as high speed and wide applicability. As regards energy awareness, in the prevailing design practice, both simple and complex current simulation models can only be used and give accurate results when all technical and design variables are already known. Therefore, the use of these available simulation models during the conceptual design phase is problematic, given their questionable accuracy and effectiveness at this stage and the labour-intensiveness of their use. In the process of our research, in order to identify an applicable decision support model, we examined multiple international, describable decision problems similar to those faced in the building design process, and we have identified a widely used model for multi-criteria decision-making problems, the Analytic Hierarchy Process (AHP).

4 The Classical Analytic Hierarchy Process

The backbone of the classical AHP model is the hierarchical structure of factors, criteria and alternatives [7]. During the decision-making process, it will express the joint conclusion of multiple experts as to the optimal solution. Accordingly, we determined the factors affecting the energy consumption of a household and grouped them into a hierarchical system.

The main criteria are as follows:

- Conceptual efficiency,
- The performance of the building structures,
- The energy efficiency of the equipment,
- Architectural value, design,
- Interior comfort,
- Lifetime and reliability.

The sub-criteria of the conceptual efficiency are: the building's orientation, the window to wall ratio (including the doors also), the shading ratio of the windows, the shape of the roof, and ratio of attic built in. By the performance of the building structures we mean the thermal properties of the different structures. As for the energy efficiency of the equipment, we look at the energy efficiency of the house's total fixtures, as well as the amount of consumption supplied by renewable energy sources. The architectural value, which includes the fit to the environment, and the functional design of the building could be important factors for both the customer and the designer, since it will not matter if the building achieves near zero energy demand if the customer does not like it and chooses a different alternative. Interior comfort refers to the heating, cooling and lighting demand. Lifetime and reliability refer to the longevity and maintainability of all structures and equipment considered in the other main criteria. Another important criterion is investment cost, but at this stage of our work we have not yet integrated this criterion into our research, similarly to other international researchers [8]. The pay-off of the investments strongly depends on energy prices and thus needs to be inspected separately, which is not the subject of this article. At the same time, Hungarian consumers are very price sensitive, and therefore investment costs would weigh extremely high in the AHP hierarchy compared to the other factors, and this would significantly affect and distort the results. Therefore, the question of costs is treated separately from the AHP model during the development. The built AHP structure is shown in Figure 2.

Applying the classical AHP method, the decision makers use the pairwise comparison matrix to evaluate the different alternatives according to the set criteria [7]. In the pairwise comparisons we used the basic Saaty scale, skipping the intermediate values, as using a wider scale than 5 grades could lead to inadequate evaluation for this problem. The elements in the used scale are as follows [7]:

- 1. Equally preferred,
- 3. Moderately preferred,
- 5. Strongly preferred,
- 7. Very strongly preferred,
- 9. Extremely preferred.

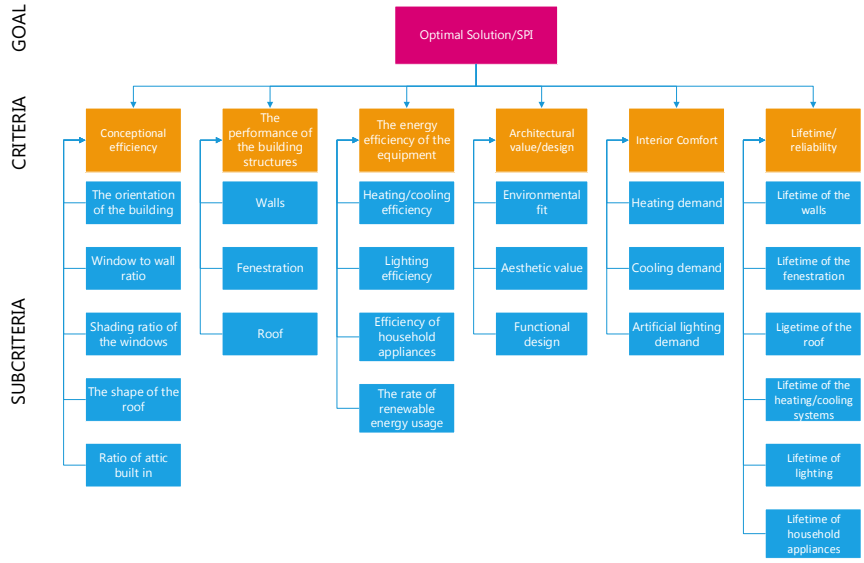


Figure 2
The AHP hierarchy

Table 1
An example for the pairwise comparison matrix

	A1	A2	A3	A4
A1	-	0,333	0,333	0,111
A2	3	-	1	0,111
A3	3	1	-	0,111
A4	9	9	9	-

According to the value in the first column, second row of the sample pairwise comparison matrix seen in Table 1, the specified properties of alternative A2 are moderately preferred compared to the same properties of alternative A1. Considering the elements in the second column, third row, the alternatives A2 and A3 are equal as regards the specified properties.

For the proper application of the model, it is necessary to define the participants in the decision-making process. Examining the process of a construction project, in the concept design stage the following decision-makers need to be considered:

- Investor/Resident,
- Architect,
- Authority.

After, preparing the pairwise comparisons, to determine the weight of the criteria and sub-criteria, as well as to set up the ranking, we use the open source, Java-based "Priority Estimation Tool (AHP)" software, based on the eigenvector method.

5 The Hibrid Fuzzy AHP System

Using the classical AHP method, described in the previous paragraph, it can be concluded that the preparation of the pairwise comparisons is a highly labor-intensive process, and therefore its effectiveness is questionable. The weakness of the method in the described case is the high number of criteria and sub-criteria. If we consider the four different alternatives, a number, which is not unusual during the concept design phase, the three decision-makers must prepare a total of 90 pairwise comparison matrixes, which is not particularly good in terms of effectiveness of decision-making. Therefore, it is necessary to develop the method further for the more general and more efficient applicability.

To eliminate the problems mentioned above, we have developed a Hybrid Fuzzy AHP system on the foundations of the described classical AHP model. In the revised model we have left the above described AHP structure, along with the specified criteria and sub-criteria; however, for the evaluation of the alternatives, we did not use the pairwise comparison matrix as input data, but rather the SPIs (Site Performance Index) calculated from fuzzy inference systems. These SPIs are values between 0 and 1 for standardized input and calculated from the appropriate attributes of the selected alternative. The two-level AHP structure is evaluated according to the classical method with pairwise comparisons, weighted with the aggregated opinion of the experts involved. The formed hierarchical structure is, therefore, weighted in a general sense, and due to this, the alternatives to be ranked with their specified data can be evaluated without expert assistance. Figure 3 shows an example of a fuzzy logic controller input, namely the case of the SPI based on the roof pitch and the amount of attic built in. To create the fuzzy control systems, the Matlab "Fuzzy Logic Toolbox" module was used.

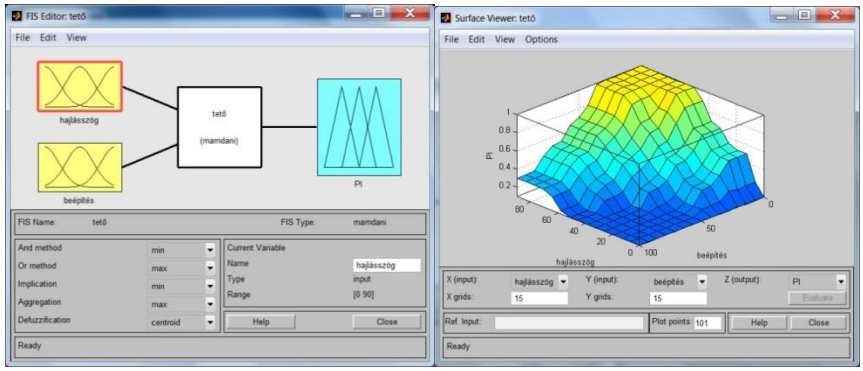


Figure 3

One of the fuzzy logic controllers (FLC)

The full Hybrid Fuzzy AHP model for easy application was created in the Matlab "Simulink" module, and thus hundreds of different alternatives can easily be evaluated and ranked. The "Simulink" model pattern is illustrated in Figure 4. As shown in the figure, in the case of the architectural value criteria, we did not apply the fuzzy controllers at the input side since during the subjective evaluation, system-matching scores had already been generated, and therefore there was no need to convert the data and manage the uncertainties. In our case, it is highly probable that the opinion of the decision-maker with regard to the defined criteria will be measured precisely.

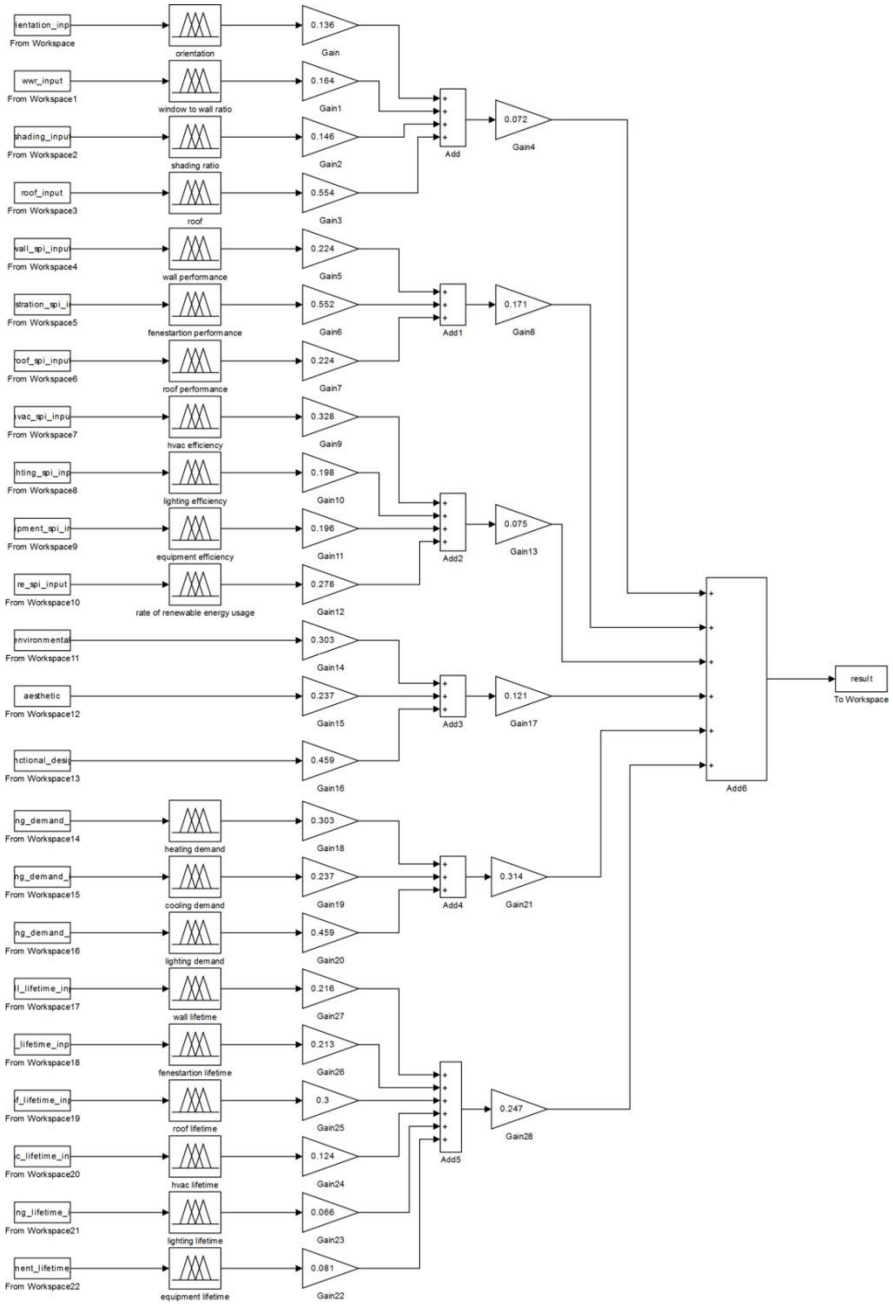


Figure 4
The Simulink modell of the hibrid Fuzzy-AHP system

6 Validation

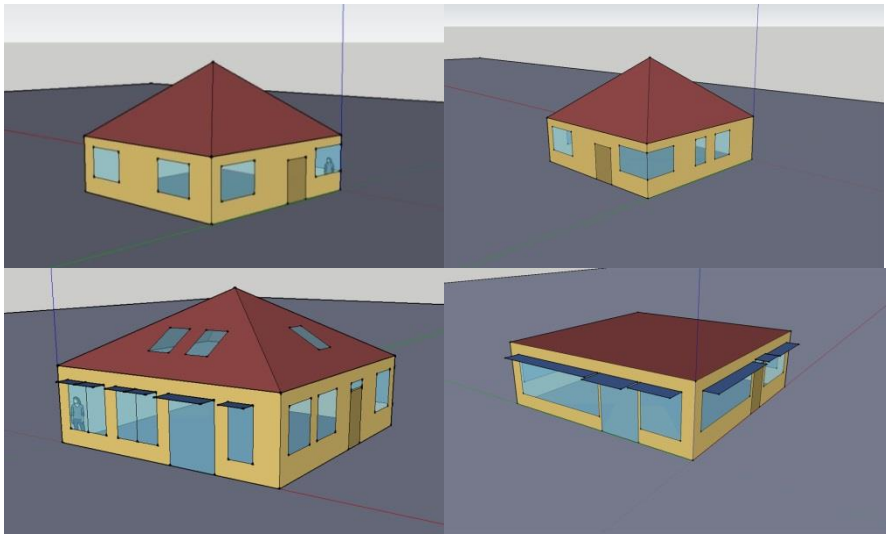


Figure 5
The 3D models of the A1-A4 alternatives (Google SketchUp, OpenStudio)

In order to evaluate the models, we prepared a detailed energy simulation of four typical buildings in Hungary and we evaluated them with the proposed methods and systems. We did the detailed simulation of the alternatives with Open Studio and EnergyPlus software developed by the National Renewable Energy Laboratory, which are widely used in international research [9, 10]. During modelling we prepared realistic, three-dimensional models of the buildings, and we performed the energy simulations in accordance with the given criteria, based on the actual annual meteorological data. The data of the different alternatives systematized according to the AHP structure are included in Table 2. We determined the orientation of the building by the location of the main rooms of the building (living room and dining room). As for the lifetime of the different structures and fixtures, we mean the achievable maximum durability with appropriate and necessary maintenance.

Table 2
The 4 alternatives

	Alternatives				
Criteria	**A1**	**A2**	**A3**	**A4**	
Conceptional efficiency					
The orientation of the building	0	90	180	270	*[°] N - 0°*
Window to wall ratio	0,18	0,18	0,25	0,43	
Shading ratio of the windows	0,00	0,00	0,52	1,00	

The shape of the roof	38,83	38,83	30,00	5,00	*[°]*
Ratio of attic built in	0,00	0,00	0,50	0,00	
The performance of the building structures	**A1**	**A2**	**A3**	**A4**	
Performance of the walls	1,790	0,391	0,242	0,136	*[W/m^{2}K]*
Performance of the fenestration	2,200	1,600	1,600	0,800	*[W/m^{2}K]*
Performance of the roof	1,690	1,690	0,740	0,115	*[W/m^{2}K]*
The energy efficiency of the equipment	**A1**	**A2**	**A3**	**A4**	
Heating/cooling efficiency	0,80	0,80	1,00	1,50	
Lighting efficiency	10,50	10,50	7,00	5,25	*[W/m^2]*
Efficiency of household appliances	3,45	3,45	2,30	1,80	*[W/m^2]*
The rate of renewable energy usage	0,00	0,00	0,10	0,20	
Architectural value/design	**A1**	**A2**	**A3**	**A4**	
Environmental fit	0,20	0,20	0,50	1,00	
Aesthetic value	0,20	0,20	0,30	0,80	
Functional design	0,50	0,50	0,40	0,60	
Interior comfort	**A1**	**A2**	**A3**	**A4**	
Heating demand	22	21	20	21	*[°]*
Cooling demand	-	-	25	24	*[°]*
Artificial lighting demand	13	13	11	8	*[hour]*
Lifetime/reliability	**A1**	**A2**	**A3**	**A4**	
Lifetime of the walls	10	15	25	50	*[year]*
Lifetime of the fenestration	5	10	10	25	*[year]*
Lifetime of the roof	5	15	25	25	*[year]*
Lifetime of the heating/cooling system	5	10	10	15	*[year]*
Lifetime of lighting	2	2	5	10	*[year]*
Lifetime of household appliances	10	10	5	10	*[year]*

After the energy simulation of the above-described alternatives we evaluated them with the classical AHP model and with the proposed Hybrid Fuzzy AHP model. With the classical AHP model, the specified decision-makers prepared the pairwise comparison matrix with regard to all the specified criteria and alternatives, and then with the already mentioned software we set up the ranking of the various alternatives. When the Fuzzy AHP hybrid system was applied, the decision-makers only prepared the pairwise comparison matrix in case of the criteria, while the evaluation of the alternatives were managed with the SPIs defined by fuzzy inference system in the hierarchical system, where the weights had already been determined. Based on the resulting ranking and weighting of the alternatives in the AHP, we could easily evaluate the results.

7 Results

Using the methods described above, during the evaluation of the alternatives A1 to A4, we obtained the results shown in Table 3. In order to get a better overview, the simulated energy consumption of the alternatives is shown on the table according to the different consumption areas. The primary energy conversion factors have been applied according to current Hungarian legislation. The estimated costs in the case of the evaluated alternatives were based on the currently valid Hungarian construction statistics data.

Table 3
The results

Simulation results	**A1**	**A2**	**A3**	**A4**	
Energy consumption					
HVAC systems (natural gas)	1 093	798	648	0	[MJ/m^2]
Water systems (natural gas)	224	224	110	143	[MJ/m^2]
Total (natural gas)	1 317	1 022	757	143	[MJ/m^2]
HVAC systems (electric)	14	8	54	179	[MJ/m^2]
Lighting (electric)	109	54	71	46	[MJ/m^2]
Other (electric)	72	37	48	37	[MJ/m^2]
Total (electric)	194	99	173	262	[MJ/m^2]
Total (net)	*1 511*	*1 121*	*930*	*405*	*[MJ/m^2]*
Total (primary energy)	**1 803**	**1 270**	**1 070**	**638**	**[MJ/m^2]**
Investment costs	**661**	**981**	**1 384**	**3 181**	**[€/m^2]***
Energy scores	0,14	0,20	0,24	0,41	[normalized results]
Classic AHP results	**0,110**	**0,143**	**0,240**	**0,506**	
Deviation from the energy scores	-23,78	-30,17	-1,30	24,10	[%]
Fuzzy AHP results	**0,370**	**0,487**	**0,575**	**0,818**	
Fuzzy AHP normalized results	0,164	0,216	0,256	0,364	[normalized results]
Deviation from the energy scores	13,95	5,69	5,10	-10,83	[%]

*based on 310 HUF/€ exchange rate

Based on the results obtained, it can be seen that all three methods of considering the alternatives result in the same ranking; so by simply considering these rankings, we can conclude that both the classical AHP method and the hybrid Fuzzy AHP method are suitable for evaluating the alternatives that arise in the early design phase. Considering the accuracy of the method, and in particular analysing the results achieved in energy efficiency, it can be stated that our proposed hybrid Fuzzy AHP system differed by only 13.95% from the detailed simulation in the worst case, while the classic AHP differed by more than 30% in one case. This may be explained by the fact that with the classic AHP, the

subjective (lay) evaluation is emphasised more than in the proposed Fuzzy-AHP based expert system. Thus, the system with the fuzzy controllers also appears to have the advantage of eliminating any lack of knowledge on the part of the individual decision-makers.

Conclusions

The characteristics of the used methods are summarized in Table 4. Based on the comparison, it can be seen that our proposed Fuzzy-AHP hybrid system is the most efficient, as regards the labour requirement. The system created is sufficiently precise in the early, concept stage of the design, and therefore it can constitute an alternative to energy simulations made at this stage. There is the slight disadvantage of the need to design the system for local conditions, with the consequent high resource demand, but in our opinion, the fast running speed and the low labour and resource requirements during application compensate for this disadvantage.

Table 4
The comparasion of the applied systems

Criteria	Simulation	Classic AHP	Fuzzy AHP
Ability to evaluate the subjective criteria	Not possible	Adequate	Adequate
Labor demand during application	Very high	High	Low
Model complexity	Very complex	Simple	Complex
Running speed	Slow	Fast	Fast
Accuracy	Very accurate	Moderately accurate	Accurate
Labor demand to develop the system	High	Low	High

It is important to note that neither the development of design systems to reduce the energy consumption of buildings and households nor increasing a building's energy efficiency is enough to achieve EU energy objectives. In our opinion, in addition to appropriate regulation and effective design methods, it is also necessary to increase the interest of consumers (tenants, investors), encouraging their demand for energy efficiency, since, only the combination of these three factors can achieve the desired change in energy efficiency.

Acknowledgement

This work was supported by Doctoral School of Applied Informatics and Applied Mathematics of Óbuda University.

References

[1] András Szűts, István Krómer: Estimating Hungarian Household Energy Consumption Using Artificial Neural Networks, Acta Polytechnica Hungarica, Vol. 11, No. 4, pp. 155-168, 2014

[2] DIRECTIVE 2009/28/EC OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL: on the promotion of the use of energy from renewable sources and amending and subsequently repealing Directives 2001/77/EC and 2003/30/EC, 23.04.2009

[3] EUROPEAN COMMISSION: A Roadmap for Moving to a Competitive Low Carbon Economy in 2050, Brussels, 25.5.2011

[4] DIRECTIVE 2010/31/EU OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL: on the energy performance of buildings, 19.05.2010

[5] Federal Ministry for the Environment, Nature Conservation and Nuclear Safety, Fraunhofer Institute for Systems and Innovation Research ISI: Policy Report, Contribution of Energy Efficiency Measures to Climate Protection within the European Union until 2050, 2012

[6] Statute 7/2006. TMN (V.24): Regarding the Definition of Building Energy Characteristics, 2006 (in Hungarian: Az épületek energetikai jellemzőinek a meghatározásáról)

[7] Thomas L. Saaty: Decision Making with the Analytic Hierarchy Process, Int. J. Services Sciences, Vol. 1, No. 1, pp. 83-98, 2008

[8] Christina J. Hopfe, Godfried L. M. Augenbroe, Jan L. M. Hensen: Multi-Criteria Decision Making under Uncertainty in Building Performance Assessment, Building and Environment, Vol. 69, pp. 81-90, 2013

[9] https://www.openstudio.net/, download time: 20.05.2014

[10] http://apps1.eere.energy.gov/buildings/energyplus/, download time: 20.05.2014

[11] Johnny K.W. Wong, Heng Li: Application of the Analytic Hierarchy Process (AHP) in Multi-Criteria Analysis of the Selection of Intelligent Building Systems, Building and Environment, Vol. 43, pp. 108-125, 2008

[12] Alessio Ishizaka, Nam Hoang Nguyen: Calibrated Fuzzy AHP for Current Bank Account Selection, Expert Systems with Applications, Vol. 40, pp. 3775-3783, 2013

[13] Guozhong Zheng, Youyin Jing, Hongxia Huang, Guohua Shi, Xutao Zhang: Developing a Fuzzy Analytic Hierarchical Process Model for Building Energy Conservation Assessment, Renewable Energy, Vol. 35, pp. 78-87, 2010

[14] G. Kanagaraj, Ashwin Mahalingam: Designing Energy Efficient Commercial Buildings - A Systems Framework, Energy and Buildings, Vol. 43, pp. 2329-2343, 2011

[15] R. Ramanathan: A Note on the Use of the Analytic Hierarchy Process for Environmental Impact Assessment, Journal of Environmental Management, Vol. 63, pp. 27-35, 2001

[16] Yang Yu-lan, Tai Hui-xin, Shi Tao: Weighting Indicators of Building Energy Efficiency Assessment Taking Account of Experts' Priority, J. Cent. South Univ., Vol. 19, pp. 803-808, 2012

[17] Yu-Ting Lai, Wei-Chih Wang, Han-Hsiang Wang: AHP- and Simulation-based Budget Determination Procedure for Public Building Construction Projects, Automation in Construction, Vol. 17, pp. 623-632, 2008

[18] Pablo Aragonés-Beltrán, Fidel Chaparro-González, Juan-Pascual Pastor-Ferrando, Andrea Pla-Rubio: An AHP (Analytic Hierarchy Process)/ANP (Analytic Network Process)-based Multi-Criteria Decision Approach for the Selection of Solar-Thermal Power Plant Investment Projects, Energy, Vol. 66, pp. 222-238, 2014

[19] R. Judkoff, D. Wortman, B. O'Doherty, and J. Burch: A Methodology for Validating Building Energy Analysis Simulations, National Renewable Energy Laboratory, Technical Report NREL/TP-550-42059, 2008

[20] Christina Diakaki, Evangelos Grigoroudis, Dionyssia Kolokotsa: Performance Study of a Multi-Objective Mathematical Programming Modelling Approach for Energy Decision-Making in Buildings, Energy, Vol. 59, pp. 534-542, 2013

[21] Mehmet Kabaka, Erkan Köseb, O˘guzhan Kırılmaza, Serhat Burmao˘gluca: A Fuzzy Multi-Criteria Decision Making Approach to Assess Building Energy Performance, Energy and Buildings, Vol. 72, pp. 382-389, 2014

[22] Building Information Centre Ltd.: Construction Cost Estimation Guide, 2011, (in Hungarian: Építésügyi Tájékoztatási Központ Kft.: Építőipari költségbecslési segédlet, 2011)

Uncertainty and Sensitivity Analyses of PISA Efficiency: Distance Based Analysis Approach

Marina Dobrota, Veljko Jeremić, Milica Bulajić, Zoran Radojičić

University of Belgrade, Faculty of Organizational Sciences,
Jove Ilića 154, 11000 Belgrade, Serbia
E-mail: dobrota.marina@fon.bg.ac.rs, jeremic.veljko@fon.bg.ac.rs,
bulajic.milica@fon.bg.ac.rs, radojicic.zoran@fon.bg.ac.rs

Abstract: The evaluation of education systems and the measurement of their quality, have become increasingly important research subjects for all stakeholders. This paper measures the efficiency of countries' educational performance and ranks countries according to the obtained results. The research is complemented by an examination of how uncertainty and sensitivity analyses affect the changes in rankings. Efficiency of countries' educational performance is presented through a single composite indicator that consists of PISA indicators and is measured using the Distance Based Analysis (DBA) methodology. Uncertainty and sensitivity of a composite indicator are results of different normalization methods. One of the contributions of this paper is the analysis of uncertainty and sensitivity impact on the efficiency measurement.

Keywords: uncertainty analysis; sensitivity analysis; composite indicators; normalization; educational performance; efficiency; PISA; Distance Based Analysis (DBA)

1 Introduction

An economic and social analysis of efficiency in education and training systems significantly contributes to the increase of education system quality. Raising the education level of the population has been recognized as a significant factor of societal economic development [1]. Although the concept of educational efficiency is proposed in several papers [2, 3] and is yet to be improved and discussed, a significant question on how stable these efficiency measurements are remains. Since it is virtually impossible to observe this issue as a one-dimensional problem, multi-criteria efficiency measurement is used and formed as a particular composite indicator. Consequently, its stability ensures an amount of safety of the observed system. Hereupon, it is essential to elaborate more on the importance of securing the safety of a complex system; the issue that has been recognized by various risk analysts in industrial and nonindustrial sectors [4, 5, 6, 7, 8, 9, 10].

The selection of an appropriate methodology is essential to any effort to capture and summarize interactions among the individual indicators included in one composite indicator or ranking system [11, 12].

According to Paruolo, Saisana and Saltelli [13], composite indicators, within social sciences, aggregate individual variables with the aim to capture relevant and possibly latent dimensions of reality. These authors claim that composite indicators have been increasingly adopted by many institutions, both for specific purposes and for providing a measurement basis for shaping broad policy debates. Composite indicators are applied and constructed everywhere [14, 15]. Saltelli et al. [16] characterize the question of composite indicators as follows: "Composite indicators tend to sit between advocacy (when they are used to draw attention to an issue) and analysis (when they are used to capture complex, multidimensional phenomena)." The topic of composite indicators has been described by the OECD [17], which defines, in six steps, why they should be created and what the goals that should thus be fulfilled are.

Many authors emphasize the need for an explicit conceptual framework for a composite indicator and the practical use of multivariate analysis prior to the aggregation of an individual indicator [12]. Authors point to the methods for assessing the robustness of the indicator using uncertainty and sensitivity analyses [12]. Sensitivity analysis is a study of how uncertainty in the models output can be apportioned to different sources of uncertainty in a models input [18]. A related practice is an 'uncertainty analysis' that focuses more on quantifying uncertainty in a model output. Ideally, uncertainty and sensitivity analyses should run in tandem, with an uncertainty analysis proceeding in the current practice [16]. Sensitivity analysis can serve a number of useful purposes in the economy of modeling [16]. Uncertainty quantification studies, which may include sensitivity analysis and uncertainty analysis, are essential in risk assessments. In this context, sensitivity analysis can be used to select the most relevant parameters and to reduce the number of parameters included in the risk assessment [6].

In this regard, the usual conclusion of studies, is that multi-criteria methodology definitions suffer from a ranking instability syndrome [8, 19]. Some authors offer conflict rankings as to what is "best" [20, 21]. According to Keung, Kocaguneli and Menzies [8], given different historical datasets, different sets of best ranking methods exist under various situations.

2 Instruments

In economic terms, the efficiency can be defined as the relationship between inputs and outputs while economic efficiency is increased by gain in units of output per unit of input. In relation to education, different educational outcomes can result from a variety of combinations of inputs (teachers, buildings, class size,

curriculum, etc.) [2, 61]. The problem that confronts economists and educators is how to mix the inputs in the suitable proportions to achieve the most efficient outcome [3]. Apparently, education serves many outcomes, and some of them cannot be measured by using econometric techniques of the orthodox economic theory. Historically, there have been efforts to increase educational efficiency. In order to explain the problems inherent in the measurement of educational efficiency [59], it is necessary to examine some research done on the input-output production functions. Cooze [2] for example, suggests a class size versus student achievement, which is one of the guidelines used in our research.

This paper measures the uncertainty and sensitivity of efficiency measurement of countries' educational performance. The educational performance has been described by the three main lifelong learning indicators as defined by the Program for International Student Assessment (PISA). PISA includes comparative international tests of students' achievements. It has become increasingly popular over the last ten years [22], with a major impact on the educational community, public, and national and international policy makers [23]. The relevance of knowledge and skills measured by PISA is confirmed by recent studies [23] that have been keeping track of the young people in years after they were assessed. Studies in Australia, Canada, and Denmark show a strong interconnection between a student's performance in reading and a chance to complete secondary school and carry on with post-secondary studies at the age 19 [3]. There are countries that share similarities in education, such as universal public systems and diversity in student population. Cross-cultural studies on PISA results [23, 24, 25, 26] find that some are performing at the top level while some are average performing countries [24]. According to Lounkaew [27], the differences in achievements between urban and rural students can be explained by intangible and immeasurable school characteristics. The impact of a student's family as well as school particular influences on student achievements are features that vary along the test achievement distributions [27]. However, so far PISA has been the only international assessment to incorporate the measures of multiple components in order to measure educational performance [24].

Attitudes to PISA significance differ from country to country. For example, while Germany substantially reformed its education system in a response to its average PISA results, England has made almost no change [28, 29, 30]. However, Shanghai's performance in PISA 2009 produced a global 'PISA-shock', which has repositioned this system as a significant new 'reference society' [31]. It shifted the global gaze in education from Finland to the 'East' at the beginning of the so-called 'Asian century' [31].

Measuring countries' educational efficiency is based on the set of input indicators, which envelope the appropriate financial and non-financial indicators, covering and defining the causes of a lifelong learning performance. Financial indicators are used because with a reduction of public expenditure and general rationalization policies developed countries are increasing their expenditures on

research and education [1]. At the same time, they are increasing demands for greater efficiency and effectiveness of their education system [1]. Non-financial indicators are mainly related to a class size and student-teacher ratio [2]. The input indicators used in this study are:

- Annual expenditure per student by educational institutions for all services relative to GDP per capita
- Expenditure on educational institutions as a percentage of GDP
- Ratio of students to teaching staff in the educational institution
- Starting salary/minimum training

Annual expenditure per student by educational institutions for all services relative to GDP per capita provides a measure of the cost of education. The access to education at lower levels of schooling in most OECD countries is universal. Spending per student by educational institutions in terms of GDP per capita can be interpreted as a mean of school's expenditure on the population relative to a country's ability to pay. This measure is difficult to explain on the higher levels of education, as enrolment varies from country to country. For example at the tertiary level, OECD countries can be relatively highly ranked according to this measure if a large part of their wealth is spent on education of a relatively small number of students [32, 1, 2].

Expenditure on educational institutions as a percentage of GDP is a measure of expenses on educational institutions in relation to the wealth of the country. National wealth is estimated based on GDP while the expenditure on education includes spending by governments, enterprises and individual students and their families. The share of education expenditure in GDP depends on the preferences of different public and private actors. In 2009, OECD countries spent on average 6.2% of GDP on educational institutions. In general, OECD countries spend 6.4% of GDP on educational institutions, having in mind both public and private sources of funding [32, 33].

The ratio of students to teaching staff, in educational institutions, compares the number of students to the number of teachers for a given level of education and in similar types of educational institutions. Class size and student-teacher ratios are often discussed aspects of education topics. They have a significant impact on the amount of money spent on education. Smaller departments often allow teachers to focus on the needs of a student as an individual, rather than to deal with disturbances in the department. Ratio of students to teaching staff shows how to allocate resources for education. A small student-teacher ratio is often in conflict with the possibility of higher salaries for teachers, their professional development and training, and greater investment in technology [33, 34]. It can also be in conflict with a larger number of teaching assistants whose salaries are often much lower than the salaries of qualified teachers [33, 34].

Starting salary/minimum training - teachers' salaries are the highest item of expense in school education. Together with alternative employment opportunities, salaries of teachers have a significant impact on how alluring the teaching profession is. They influence the decisions of individuals to engage in an educational profession after graduation (such as graduate career choices related to relative earnings in teaching and non-teaching jobs and their growth over time) [34]. They also impact one's decision to return and remain in the teaching profession after a career break, such as higher wages, fewer people who choose to leave the profession [34].

The countries performance (output) is made of the key competencies in reading, mathematics and science, established by PISA, adopted in 1997 by OECD countries. It represents a commitment of the governments of OECD member countries, to audit the outcomes of education systems in terms of students' achievement within a standard international framework. PISA is a collaborative effort that brings together scientific experts from the participating countries [35]. Their governments jointly steer it on the basis of shared, policy-driven interests [35]. PISA is designed to collect information through triennial assessments and presents data on domain-specific knowledge and skills in reading, mathematics and science of students, schools and countries. PISA combines the assessment of science, mathematics, and reading with information on students' home background, their approach to learning, learning environment, and their familiarity with computers. These are the background factors that student outcomes are associated with. By that, PISA provides an understanding of factors that influence the development of skills and attitudes at home and school, and examines how these factors interact and what the implications for policy development are.

The three domains assessed in PISA are:

- Reading
- Mathematics
- Science

Reading literacy presents an individual's ability to: understand, use, reflect on, and collaborate with written texts, in order to achieve goals. It allows students to develop their knowledge and potential and to participate in society [35].

Mathematical literacy is defined as an individual's ability to identify and to understand the role that mathematics has in the world, to make reasonable judgments. It allows the usage of mathematics in such a way that meets the needs of a student as a constructive, concerned, and reflective citizen [35].

Scientific literacy represents a student's scientific knowledge and its use to identify questions, acquire new knowledge, and explain scientific phenomena. It allows students to make conclusions based on scientific matters and to understand the characteristics of science as a form of human knowledge. It makes students

aware of how science and technology shape our material, intellectual, and cultural environment, and willing to engage in scientific issues [35].

The primary domains of PISA (Science, Reading and Mathematics) are a crucial part of lifelong learning. PISA indicators assess the level of performance of the adolescents. PISA provides information on teaching and learning results in schools and also demonstrates the features of the development of an educational system. The main focus of PISA is not set on the congruence of items within the national curricula of participating countries but to record the core competencies in different real-life tasks that matter in everyday situations. The concept of literacy used in PISA is functional as following: 15-year-old students should apply competencies they have learned at school in the context of authentic tasks that are part of our everyday life.

3 Methodology

Commonly, the ranking of entities is performed in a way that can vigorously affect the general evaluation issue. It can affect sports competitions, exams, medicine selection, university ranking, and many other areas [36, 37, 38, 39, 40]. For creating an objective image of ranking of the observed entities we propose I-distance method [39, 40, 41, 42].

3.1 I-distance

I-distance measures the distance in an n-dimensional space. This method has recently made a significant breakthrough in a considerable number of scientific achievements. Originally it was proposed and defined by B. Ivanovic in 1963 [37]. Affirmation of this method has been made in University ranking [39] and evaluating the socio-economic development of countries [41].

Ivanovic has initially proposed this method to rank countries according to their development level, but based on several indicators. He considered many socio-economic development indicators, but the problem was their usage in order to calculate a single indicator, which can be represented as the rank.

Let $X^T = \left(X_1, X_2, ...X_k\right)$ be set of variables chosen to characterize the entities. I-distance between two entities $e_r = \left(x_{1r}, x_{2r}, ...x_{kr}\right)$ and $e_s = \left(x_{1s}, x_{2s}, ...x_{ks}\right)$ is defined as

$$D(r,s) = \sum_{i=1}^{k} \frac{\left|d_i\left(r,s\right)\right|}{\sigma_i} \prod_{j=1}^{i-1} \left(1 - r_{ji.12...j-1}\right) \tag{1}$$

where $d_i(r,s)$ is the discriminate effect - distance between the values of a variable X_i for e_r and e_s,

$$d_i\left(r,s\right)=x_{ir}-x_{is} \quad i\in\left\{1,...k\right\} \tag{2}$$

σ_i is the standard deviation of X_i, and $r_{ji.12...j-1}$ is a partial coefficient of the correlation [43,44,45,46] between X_i and X_j, $\left(j<i\right)$ [35, 40].

The I-distance construction is an iterative process, which can consist of several steps [47, 50]. The method is performed by calculating the mutual distances between the entities which are then compared to one another to create a rank. First step calculates the amount of discriminate effect of the first variable (the most significant variable that provides the most information on the education evaluation). Second step computes the value of the discriminate effect of the second variable, not covered by the first. This procedure is repeated for all the variables. The ranking of entities in the whole set is based on the distance from the referent entity [41, 42]. In order to perform the ranking, it is necessary to fix one entity as a referent. This is usually the entity with the minimal rate for each indicator or a fictive entity with minimal, maximal or average values for each variable. In this analysis, object with minimum values for each variable was set as a referent [47, 49].

If negative correlation coefficients and negative coefficients of partial correlations occur [48], it is more suitable to use the square I-distance, given as:

$$D^2\left(r,s\right)=\sum_{i=1}^{k}\frac{d_i^2\left(r,s\right)}{\sigma_i^2}\prod_{j=1}^{i-1}\left(1-r_{ji.12...j-1}^2\right) \tag{3}$$

This paper also presents a Distance Based Analysis (DBA) approach as a new measure of efficiency, where the I-distance method is applied to several Input indicators ($I-distance_{input}$) and several Output indicators ($I-distance_{output}$). After I-distance values are calculated, it is necessary to normalize the values using L_∞ metrics (and other metrics). The efficiency of a country is given as follows [51, 52, 53]:

$$Ef=\frac{\text{I-distance}_{output}}{\text{I-distance}_{input}} \tag{4}$$

If the *Ef* value equals 1, the entity produces an equal amount of output, for given amount of input. Any country with an efficiency ratio of at least 1 is considered to be efficient [41, 48].

3.2 Sensitivity Issues

Composite indicators are filled with normative assumptions in variable selection and weighting, where 'normative' means 'related to and dependent upon a system of norms and values' [13]. Results and values of composite indicators significantly depend on a normalization method. This is why composite indicators are often the subject of controversy [54, 13].

In this paper, we propose four methods of normalization: Chebyshev distance (L_∞ norm), Manhattan distance (L_1 norm), Euclidean distance (L_2 norm), and recently proposed 1-2 distance. 1-2 distance implies the normalization of all values among values 1 and 2 thus avoiding the problems with zero values and large, poorly explainable efficiency measurement values.

Performing the sensitivity and uncertainty analyses of the PISA educational performance, in this paper we compared the results gained by these methods and pointed out the changes in rankings, in the process.

4 Sensitivity Analysis Results

In order to gain the DBA efficiency measurement values, we have calculated I-distance ranking measurement for the Input and Output variables. These are used as a basis for creating the DBA efficiency measurement of countries' educational performance.

Input and Output values are firstly corrected with an L_∞ norm so that they could be mutually comparable, subsequently corrected with an L_1 norm, L_2 norm, and 1-2 distance. The results are given in Table 1.

When DBA method was introduced, an initially proposed normalization method was Chebyshev distance (L_∞ norm) and thus is the first to be presented in the results. Table 1 gives the values of educational performance efficiency measurement, as well as the ranking results. As can be seen from Table 1, Finland, Japan, France, Germany, Korea, Slovakia, Netherlands, New Zealand, Estonia, Australia, and Belgium are shown to be efficient in educational performance according to each normalization method. It is interesting to note that Hungary seems inefficient by Chebyshev normalization method. However, if we look at three other normalization methods, despite the fact that the values are on the verge, Hungary shows efficiency in educational performance. In addition, the Czech Republic shows even more confusing results. It is inefficient according to Chebyshev (L_∞) and Euclidean (L_2) distances, and far better positioned according to other Manhattan (L_1) normalization method. Czech Republic is even the most efficient country in educational performance according to Manhattan distance.

Table 1
Efficiency and ranking results gained by different normalization techniques

	Chebyshev (L_∞)		**Manhattan** (L_1)		**Euclidean** (L_2)		**1-2**	
Country	**DBA**	**Rank**	**DBA**	**Rank**	**DBA**	**Rank**	**DBA**	**Rank**
Finland	2.958	1	3.204	2	3.144	1	1.527	1
Japan	2.833	2	3.069	3	2.701	2	1.436	2
France	1.754	3	1.900	4	1.862	3	1.163	7
Germany	1.676	4	1.815	5	1.781	4	1.195	5
Korea	1.662	5	1.801	6	1.767	5	1.239	4
Slovakia	1.562	6	1.692	7	1.660	6	1.122	8
Netherlands	1.532	7	1.660	8	1.628	7	1.172	6
New Zealand	1.223	8	1.325	9	1.300	8	1.095	9
Estonia	1.219	9	1.321	10	1.296	9	1.090	10
Australia	1.183	10	1.282	11	1.257	10	1.079	11
Belgium	1.175	11	1.273	12	1.249	11	1.069	12
Hungary	0.945	12	1.024	13	1.004	12	1.002	14
Switzerland	0.910	13	0.986	14	0.968	13	0.977	15
United Kingdom	0.831	14	0.900	15	0.883	15	0.949	17
Czech Republic	0.814	15	9.357	1	0.918	14	1.360	3
Israel	0.808	16	0.875	16	0.859	16	1.016	13
Ireland	0.749	17	0.811	17	0.796	17	0.919	18
Norway	0.578	18	0.627	18	0.615	18	0.838	20
United States	0.571	19	0.619	19	0.607	19	0.833	21
Poland	0.526	20	0.569	20	0.559	20	0.786	27
Spain	0.524	21	0.568	21	0.557	21	0.849	19
Austria	0.513	22	0.556	22	0.546	22	0.825	23
Sweden	0.510	23	0.552	23	0.542	23	0.810	24
Portugal	0.486	24	0.526	24	0.516	24	0.806	25
Italy	0.450	25	0.487	25	0.478	25	0.797	26
Chile	0.360	26	0.390	26	0.383	26	0.956	16
Mexico	0.300	27	0.000	28	0.000	28	0.828	22
Luxembourg	0.242	28	0.262	27	0.257	27	0.621	28

Answers to this puzzle could be found in the very structure of PISA results gained by the Czech Republic. The Czech Republic scores 500 in Science, 478 in Reading, and 492 in Mathematics. When normalized, it scores 0.029 in Output according to Manhattan distance while only 0.00315 in Input. This makes Czech Republic extremely efficient according to Manhattan normalization method. Might be because L_1 metrics produces solutions that have a few large, and a lot of very insignificant residuals; the residual distribution is quite unequal (spiky).

L_2 metrics produces a few large (but less than L_1), and a lot of small, but still significant residuals. The residual distribution is more balanced than within L_1 metrics.

Table 2
Changes in Rankings

	Rankings								
Country	**1-3**	**4-6**	**7-9**	**10-12**	**13-15**	**16-18**	**19-21**	**22-24**	**25-28**
Finland	4								
Japan	4								
France	1	2	1						
Germany		4							
Korea		4							
Slovakia		1	3						
Netherlands		1	3						
New Zealand			4						
Estonia			1	3					
Australia				4					
Belgium				4					
Hungary				1	3				
Switzerland					4				
United Kingdom					3	1			
Czech Republic	2				2				
Israel					1	3			
Ireland						4			
Norway						3	1		
United States							4		
Poland							3		1
Spain							4		
Austria								4	
Sweden								4	
Portugal								3	1
Italy									4
Chile						1			3
Mexico								1	3
Luxembourg									4

The uncertainty analysis, emphasizing the changes in positions and rankings, is given in Table 2 and Figure 1. Table 2 shows in detail the changes in rankings and dispersion. We can see that Finland and Japan are ranked between the 1st and the 3rd place in all four cases while France is a country with a greater dispersion of the results. Again, Germany and Korea are between the 4th and 6th place in all the

cases. It could also be noted that the dispersion is slightly increasing as we move towards lower ranking results, yet these increases are quite slight.

Figure 1 shows where the greatest differences among ranks are. If we analyze Figure 1, we can see that Manhattan and Euclidean rankings are entirely consistent with one another. As mentioned above, the largest differences are with the Czech Republic. It is ranked 15th according to Chebyshev, 14th according to Euclidean, 3rd according to 1-2, and even the 1st according to Manhattan distance normalization method. As for other countries, Chebyshev, Manhattan, and Euclidean rankings do not differ significantly from one another. On the other hand, a newly proposed 1-2 normalization method has some other differences. The most significant is Poland that is under ranked according to 1-2 norm, and Chile, which is over ranked.

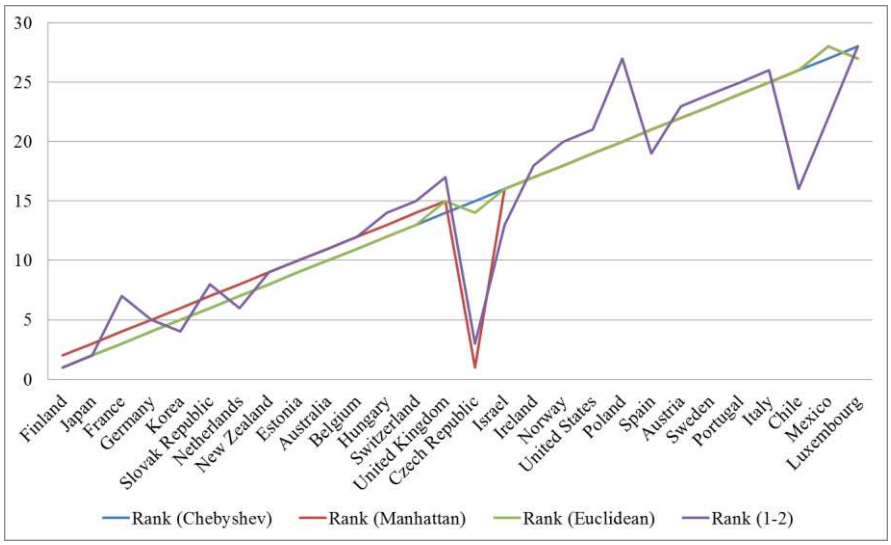


Figure 1
Ranking differences

The differences in DBA values are given in Table 1 and Figure 2.

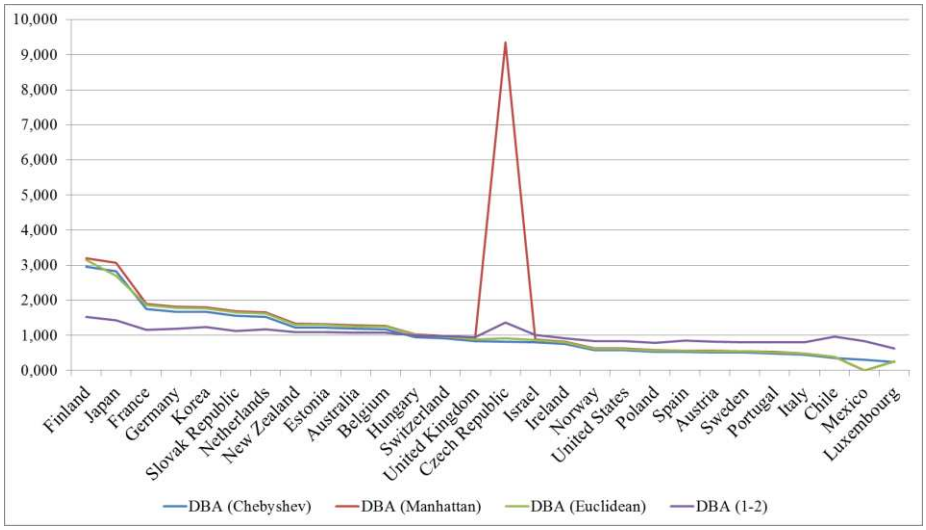


Figure 2
Changes in DBA values

Conclusions

The evaluation and measurement of certain phenomenon are not quite as simple in terms of choosing proper components to be a part of an assessment [55]. It can also be challenging, in terms of creating an appropriate methodology, which would perform the appropriate weighting or normalization [56, 57, 58]. This paper shows how uncertainty and sensitivity analyses affect the shift in rankings of countries educational performance efficiency.

There are many ways to measure educational performance and competitiveness and to evaluate countries accordingly [59, 60]. In this paper, countries educational performance is based on the PISA indicators, because it is the most exhaustive and rigorous international program to assess and explain differences in students' performance [3].

The assessments scope, nature and collective background information are decided by leading experts in participating countries and are steered jointly by governments on the basis of shared and policy-driven interests. PISA focuses on young people's ability to use their knowledge and skills to meet real-life challenges [23, 24, 26]. This orientation reflects the change in the goals and objectives of curricula themselves. The main consideration is what the students can do with what they have learned in school and not merely whether they have mastered a particular curricular content [3].

The paper has illustrated how uncertainty and sensitivity analyses, based on the changes in a normalization method, affect rankings and evaluation. Uncertainty and sensitivity analyses were applied to gain useful insight into the reliability of a

country's ranking. Changes in ranking results mainly occur due to the application of different normalization equations. While Manhattan distance measures the distance between two points in a grid based on a strictly horizontal and/or vertical path (absolute distance values), Euclidean distance uses the root of square distances (the length of the line segment connecting them). Chebyshev distance, on the other hand, is a metric defined on a vector space, where the distance between two vectors is the greatest of their differences along any coordinate dimension.

Furthermore, although the uncertainty and sensitivity analysis were performed with numerous measurement and composite indicator definitions, their impact on efficiency measurement has not yet been examined. The examination of the uncertainty influence, on efficiency, is one of the main contributions of our paper.

Acknowledgment

We are grateful for reviewers' comments and suggestions which helped us reshape and improve our paper.

References

[1] SEEECN (2013). *Efficiency and Effectiveness of the Education System*. South East European Educational Cooperation Network, Retrieved on February 14th, 2013, from http://www.see-educoop.net/education_in/pdf/chal-educ-reform-ser-07-ser-enl-t02.pdf

[2] Cooze, J., (1991). The Elusive Concept of Efficiency in Education. Retrieved on January 26, 2013 from http://www.mun.ca/educ/faculty/mwatch/vol1/cooze.html

[3] Dobrota, M., Jeremic, V., & Martic, M. (2013). Educational Performance Efficiency of OECD Countries: PISA Lifelong Learning Indicators. In *Proceedings of the XL Symposium on Operational Research (SYM-OP-IS 2013),* (pp. 903-908). ISBN: 978-86-7680-286-9

[4] Arndt, S., Acion, L., Caspers, K., & Blood, P. (2013). How Reliable Are County and Regional Health Rankings? *Prevention Science, 14*(5), 497-502. doi:10.1007/s11121-012-0320-3

[5] Monferini, A., Konstandinidou, M., Nivolianitou, Z., Weber, S., Kontogiannis, T., Kafka, P., Kay, A. M., Leva, M. C., & Demichela, M. (2013). A Compound Methodology to Assess the Impact of Human and Organizational Factors Impact on the Risk Level of Hazardous Industrial Plants. *Reliability Engineering & System Safety, 119*, 280-289. doi:10.1016/j.ress.2013.04.012

[6] Wainwright, H. M., Finsterle, S., Zhou, Q. L., & Birkholzer, J. T. (2013). Modeling the Performance of Large-Scale CO_2 Storage Systems: A Comparison of Different Sensitivity Analysis Methods. *International*

Journal of Greenhouse Gas Control, 17, 189-205, doi:10.1016/j.ijggc.2013.05.007

[7] Mahsuli, M., & Haukaas, T. (2013). Sensitivity Measures for Optimal Mitigation of Risk and Reduction of Model Uncertainty. *Reliability Engineering & System Safety, 117*, 9-20. doi:10.1016/j.ress.2013.03.011

[8] Keung, J., Kocaguneli, E., & Menzies, T. (2013). Finding Conclusion Stability for Selecting the Best Effort Predictor in Software Effort Estimation, *Automated Software Engineering, 20*(4), 543-567. doi:10.1007/s10515-012-0108-5

[9] Guttorp, P., & Kim, T. Y. (2013). Uncertainty in Ranking the Hottest Years of U.S. Surface Temperatures, *Journal of Climate, 26*(17), 6323-6328. doi:10.1175/JCLI-D-12-00760.1

[10] Saisana, M., D'Hombres, B., & Saltelli, A. (2011). Rickety Numbers: Volatility of University Rankings and Policy Implications. *Research policy*, *40*(1), 165-177. doi:10.1016/j.respol.2010.09.003

[11] Saisana, M., & Tarantola, S. (2002). *State-of-the-art Report on Current Methodologies and Practices for Composite Indicator Development*, EUR Report 20408 EN, European Commission, JRC-IPSC, Italy

[12] Saisana, M., & D'Hombres, B. (2008). *Higher Education Rankings: Robustness Issues and Critical Assessment.* EUR23487, Joint Research Centre, Publications Office of the European Union, Italy. ISBN: 978 82 79 09704 1. doi:10.2788/92295

[13] Paruolo, P., Saisana, M., & Saltelli, A. (2013). Ratings and Rankings: Voodoo or Science? *Journal of the Royal Statistical Society: Series A (Statistics in society), 176*(3), 609-634. doi:10.1111/j.1467-985X.2012.01059.x

[14] Brüggemann, R., & Patil, G. P. (2011). *Ranking and Prioritization for Multi-indicator Systems*, Dordrecht, Springer. ISBN:978-1-4419-8477-7

[15] Giambona, F., & Vassallo, E. (2014). Composite Indicator of Social Inclusion for European Countries, *Social Indicators Research, 116*(1), 269-293. doi: 10.1007/s11205-013-0274-2

[16] Saltelli, A., Ratto, M., Andres, T., Campolongo, F., Cariboni, J., Gatelli, D., Saisana, M., & Tarantola, S. (2008). *Introduction to Sensitivity Analysis*. In *Global Sensitivity Analysis. The Primer*, by John Wiley & Sons, Ltd, Chichester, UK. doi: 10.1002/9780470725184.ch1

[17] OECD (Nardo, M., et al.). (2008). Handbook on Constructing composite indicators –methodology and user guide (pp. 1-158). Ispra: OECD

[18] Saltelli, A., Tarantola, S., Campolongo, F., & Ratto, M. (2004). *Sensitivity Analysis in Practice: A Guide to Assessing Scientific Models*. John Wiley & Sons, Ltd.

[19] Jovanovic, M., Jeremic, V., Savic, G., Bulajic, M., & Martic, M. (2012). How does the Normalization of Data Affects the ARWU Ranking? *Scientometrics, 93*(2), 319-327. doi: 10.1007/s11192-012-0674-0

[20] Shepperd, M., & Kadoda, G. (2001). Comparing Software Prediction Techniques Using Simulation. *IEEE Transactions on Software Engineering, 27*(11), 1014-1022

[21] Myrtveit, I., Stensrud, E., & Shepperd, M. (2005). Reliability and Validity in Comparative Studies of Software Prediction Models. *IEEE Transactions on Software Engineering, 31*, 380-391

[22] Feniger, Y., Livneh, I., & Yogev, A. (2012). Globalisation and the Politics of International Tests: the Case of Israel, Comparative Education, 48, 3, 323-335, doi:10.1080/03050068.2011.622539

[23] Kankaras, M., & Moors, G. (2014). Analysis of Cross-Cultural Comparability of PISA 2009 Scores. *Journal of Cross-Cultural Psychology, 45*(3), 381-399. doi:10.1177/0022022113511297

[24] Lin, E., & Shi, Q.M. (2014). Exploring Individual and School-Related Factors and Environmental Literacy: Comparing US and Canada Using PISA 2006. *International Journal of Science and Mathematics Education, 12*(1), 73-97. doi:10.1007/s10763-012-9396-2

[25] Eklof, H., Pavesic, B. J., & Gronmo, L. S. (2014). A Cross-National Comparison of Reported Effort and Mathematics Performance in TIMSS Advanced. *Applied Measurement in Education, 27*(1), 31-45. doi:10.1080/08957347.2013.853070

[26] Fernandez, A. A., & Del Valle, A. R. (2013). Educational Inequality in Costa Rica: the Gap between Students in Public and Private Schools. Analysis with the Results of the PISA International Assessment. *Cepal Review, 111*, 37-57

[27] Lounkaew, K. (2013). Explaining Urban-Rural Differences in Educational Achievement in Thailand: Evidence from PISA Literacy Data. *Economics of Education Review, 37,* 213-225. doi:10.1016/j.econedurev.2013.09.003

[28] Martens, K., & Niemann, D. (2013). When Do Numbers Count? The Differential Impact of the PISA Rating and Ranking on Education Policy in Germany and the US. *German Politics, 22*(3), 314-332. doi:10.1080/09644008.2013.794455

[29] Knodel, P., Martens, K., & Niemann, D. (2013). PISA as an Ideational Roadmap for Policy Change: Exploring Germany and England in a Comparative Perspective. *Globalisation, Societies and Education*, doi: 10.1080/14767724.2012.761811, in press

[30] Meisenberg, G., & Woodley, M. A. (2013). Are Cognitive Differences between Countries Diminishing? Evidence from TIMSS and PISA. *Intelligence, 41*(6), 808-816, doi: 10.1016/j.intell.2013.03.009

[31] Sellar, S., & Lingard, B. (2013). Looking East: Shanghai, PISA 2009 and the Reconstitution of Reference Societies in the Global Education Policy Field. Comparative Education, 49(4), 464-485. doi:10.1080/03050068.2013.770943

[32] Education at a Glance (2011). *Financial and Human Resources Invested in Education*. Retrieved on November 7th, 2013, from http://www.oecd.org/education/skills-beyond-school/48630868.pdf

[33] Education at a Glance (2011). What is the Student-Teacher Ratio and How Big are Classes? Retrieved on November 7th, 2013, from http://www.oecd.org/education/skills-beyond-school/48631144.pdf

[34] EAG (2012). *Education at a Glance*. Retrieved on March 15th, 2013, from http://www.uis.unesco.org/Education/Documents/oecd-eag-2012-en.pdf

[35] PISA (2013). *Programme for International Student Assessment*. Retrieved on April 6th, 2013, from http://www.oecd.org/pisa/pisaproducts/44455820.pdf

[36] Al-Lagilli, S., Jeremic, V., Seke, K., Jeremic, D., & Radojicic, Z. (2011). Evaluating the Health of Nations: a Libyan Perspective. *Libyan Journal of Medicine, 6*, 6021. doi:10.3402/ljm.v6i0.6021

[37] Ivanovic, B. (1973). *A method of Establishing a List of Development Indicators*. Paris: United Nations Educational, Scientific and Cultural Organization

[38] Ivanovic, B., & Fanchette, S. (1973). *Grouping and Ranking of 30 Countries of Sub-Saharan Africa, Two Distance-based Methods Compared*. Paris: United Nations Educational, Scientific and Cultural Organization

[39] Jeremic, V, Bulajic, M., Martic, M., & Radojicic, Z. (2011a). A Fresh Approach to Evaluating the Academic Ranking of World Universities. *Scientometrics*, 87(3), pp. 587-596, doi:10.1007/s11192-011-0361-6

[40] Radojicic, Z., & Jeremic, V. (2012). Quantity or Quality: What Matters More in Ranking Higher Education Institutions? *Current Science, 103*(2), 158-162

[41] Jeremic, V., Markovic, A., & Radojicic, Z. (2011d). ICT as Crucial Component of Socio-Economic Development. *Management, 16*(60), 5-9

[42] Jeremic, V., Bulajic, M., Martic, M., Markovic, A., Savic, G., Jeremic, D., & Radojicic, Z. (2012a). An Evaluation of European Countries Health Systems through Distance Based Analysis. *Hippokratia, 16*(2), 175-179

[43] Fisher, R. A. (1924). The Distribution of the Partial Correlation Coefficient. *Metron, 3*(3-4), 329-332

[44] Guilford J. P., & Fruchter B. (1973). *Fundamental Statistics in Psychology and Education*. Tokyo: McGraw-Hill Kogakusha, LTD

[45] Rummel, R. J. (1976). *Understanding Correlation*

[46] Baba, K.; Ritei, S., & Masaaki, S. (2004). Partial Correlation and Conditional Correlation as Measures of Conditional Independence. *Australian and New Zealand Journal of Statistics, 46*(4), 657-664. doi:10.1111/j.1467-842X.2004.00360.x

[47] Seke, K., Petrovic, N., Jeremic, V., Vukmirovic, J., Kilibarda, B., & Martic, M. (2013). Sustainable Development and Public Health: Rating European Countries. *BMC Public Health, 13*(77)

[48] Jeremic, V., Jovanovic-Milenkovic, M., Martic, M., & Radojicic, Z. (2013). Excellence with Leadership: the Crown Indicator of SCImago Institutions Rankings IBER Report. *El Profesional de la Informacion*, *22*(5), 474-480

[49] Dobrota, M., Jeremic, V., & Markovic, A. (2012). A New Perspective on the ICT Development Index. *Information Development*, *28*(4), 271-280. doi:10.1177/0266666912446497

[50] Ivanovic, B. (1977). *Classification Theory*. Belgrade: Institute for Industrial Economic

[51] Coelli, T. J., Prasada Rao, D. S., O'Donnell, C. J., & Battese, G. E. (2005). *An introduction to efficiency and productivity analysis*. New York: Springer

[52] Charnes, A., Cooper W. W., & Rhodes E. L. (1978). Measuring the Efficiency of Decision Making Units. *European Journal of Operational Research, 2*(6), 429-444

[53] Guan, J., Chen, K. (2012). Modeling the Relative Efficiency of National Innovation Systems. *Research Policy, 41*(1), 102-115

[54] Saltelli, A. (2007). Composite Indicators between Analysis and Advocacy. *Social Indicators Research, 81*(1), 65-77

[55] Kiss, G. (2012). Measuring Hungarian and Slovakian Students' IT Skills and Programming Knowledge. *Acta Polytechnica Hungarica, 9*(6), 195-210

[56] Munda, G. (2012). Choosing Aggregation Rules for Composite Indicators. *Social Indicators Research, 109*(3), 337-354. doi:10.1007/s11205-011-9911-9

[57] Shen, Y. J., Hermans, E., Brijs, T., & Wets, G. (2013). Data Envelopment Analysis for Composite Indicators: A Multiple Layer Model. *Social Indicators Research, 114*(2), 739-756. doi:10.1007/s11205-012-0171-0

[58] Tate, E. (2013). Uncertainty Analysis for a Social Vulnerability Index. *Annals of the Association of American Geographers, 103*(3), 526-543. doi:10.1080/00045608.2012.700616

[59] Kabók, J., Kis, T., Csüllög, M., & Lendák, I. (2013). Data Envelopment Analysis of Higher Education Competitiveness Indices in Europe. *Acta Polytechnica Hungarica, 10*(3), 185-201

[60] Szoboszlai, V., Velencei, J., & Baracskai, Z. (2014). Post-Experiential Education: from Knowledge to 'Knowing'. *Acta Polytechnica Hungarica, 11*(10), 235-247

[61] Tóth, Z. E., & Jónás, T. (2014). Enhancing Student Satisfaction Based on Course Evaluations at Budapest University of Technology and Economics. *Acta Polytechnica Hungarica, 11*(6), 95-112

Spatio-Temporal Shape Parameterization of the Human Ventricles

Sándor M. Szilágyi

Department of Informatics, Petru Maior University
Str. Nicolae Iorga Nr. 1, 540088 Tîrgu Mureş, Romania
sandor.szilagyi@science.upm.ro

Abstract: This paper gives a solution for improving the geometric estimation of the human ventricles, by reducing their shape estimation error. The parametric description of the studied organ can be performed at arbitrary resolution during the whole visualization process. After the problem description, the paper presents each main step of the proposed shape estimation algorithm. The presented method determines not only the general shape of the ventricles, but the internal tissue direction, too. The estimation error decreases more than 10 times if the resolution is increased by 3 times. The biological parameters like gender or age also affect the estimation performance. The obtained accuracy of the method was slightly higher in the case of female hearts and for the left ventricle. Finally, concluding remarks emphasize some important features of the given approach and the future research directions.

Keywords: ventricle modeling; heart geometry; shape estimation; interpolation techniques

1 Introduction

Cardiac disease, despite the continuous progress of medical health care and intensive research, still represents the main mortality factor in many low- and middle-income countries. The main reason of slow progress represents the partially understood heart functioning [1].

Sudden cardiac death, mostly caused by ventricular fibrillation (VF), represents an important cause of mortality. In spite of this consideration, the electric activity and mechanical mechanisms responsible for VF are only partially discovered. It would be utmost important to understand how the onset of arrhythmias that cause fibrillation depends on details of the cardiac activity, such as the size of ventricles [2], geometry [3], mechanical [4] and electrical state [5], anisotropic fiber structure [6] and inhomogeneity [7].

A proper presentation of the cardiac activity involves the description of its electric and mechanic properties. The spread of electric excitation wave in the cardiac

muscle determines the nature of contraction, but several important aspects, such as pumping strength and volume cannot be deduced without a detailed mechanical analysis of the heart [8].

One of the most important cardiac pumping factors is the structural architecture of the ventricles [9]. Albeit several general geometric factors such as shape and size of ventricles, wall thickness and structure play a critical role in medical diagnosis, the proper orientation of cardiac muscle fibers is crucial for an adequate contraction [10]. There are several cardiac affections and traumatic events such as cardiac infarction [11] and myocardial fibrosis [12] that may induce the misalignment of ventricular fibers, developing reduced cardiac pump activity that may lead to the rise of further cardiac disorders [13].

Taking the abovementioned considerations into account is imperative for a computational ventricular model that is able to predict accurately the electric and mechanic function of the heart for normal and pathological cases, to involve not only the general description parameters but the distribution of fiber orientations too.

There exist several methods to determine the shape and structure of the ventricles. The anatomy of the whole heart and the overall placement of the cardiac muscle fibers is described by Streeter and Hanna [14], while the orientation changes from epicardium to endocardium is presented by Arts et al [15] and Geerts et al [16]. From anatomical measurements it is concluded that the septal and left ventricle walls fiber orientations are similar.

In spite of wide agreement on fiber directions in ventricular tissue there is no generally accepted concept on how to assign these orientations to realistic computerized ventricular models. Göktepe and Kuhl have used a simplified representation, where the ventricles are considered nested and truncated ellipsoids [17]. Takayama proposed a geometrical-based approach to generate a layered ventricular structure [18].

A geometric approximation of the heart may introduce significant errors. A proper mapping of a non-ellipsoidal form to an idealized ellipsoidal one raises serious architectural deformations especially in pathological cases. Moreover, inherent singularities may appear in the approximation of the septal and apex region [19]. To solve these problems, Kotikanyadanam et al have used various 3D imaging techniques during the development of a patient-specific ventricular geometry model [20], where the fiber orientation in the patient-specific ventricular geometry was determined by a Poisson equation-based interpolation technique [8].

Biological systems represent the most complex studied architectures. Their proper description is possible via models that always represent only a simplification of the investigated object. The reductionist approach applied in the case of complex heart modeling demands a simplified description of the whole ventricle. The cell level geometric description of the ventricles is difficult due to the immense

parameter number, obscure biological information and low computational power, necessitating the usage of biological model parameters.

Our goal in this paper is to present the development of a spatio-temporal ventricular geometry estimation method that involves the anatomical structure of the human ventricles. The rest of the paper is organized as follows: Section 2 gives a detailed description of the geometry estimation method, involving the effect of electric and mechanic properties of the ventricles, studied for normal and pathological cases. We outline the effect of the various parameter alterations on the generated activation potential (AP). Section 3 presents and discusses several aspects of the model functionality and the results of simulations carried out using the model. Finally, in Section 4, the conclusions are formulated.

2 Materials and Methods

2.1 Motivation of Shape Parameterization

The real-time visualization and analysis of the human heart, using up-to-date representation techniques is partially solved. The main problem is the inability of current imaging systems to yield both visual information and internally state parameters. Moreover, several imaging techniques are considered invasive.

The above mentioned problem can be partially solved using computational models that beside the visual representation may give the necessary state parameters. All computational model representation strategy uses an indirect way to obtain the investigated parameters. Physicians have no possibility to measure in real time all requested data, so they have to predict these values using computational models.

Many physiological deficiencies are related to the geometry of heart. These geometric alterations are reflected in the modified values of several medical parameters, thus obtained medical parameters are not always suitable for recent computerized heart model-based visualization techniques.

Computerized data visualization is based on voxels, pixels and triangles. The determined medical information has to be transformable to these basic geometry-related computational data. By using a heart shape parameterization method, we want to realize a transformation of medical data into basic geometry elements via a set of engineer-related parameters. This process is visualized in Fig 1. The adaptive aspect of these inner-level parameters, independently from the used resolution of the geometric representation, allows a detailed description of the heart.

In Fig. 1 the dashed line from the anatomical knowledge box and visualization data element toward the intermediate parameters box represent the possibility of

determination of several intermediate parameters. In addition, the calculated internal parameters may allow the estimation of several medical parameters. This complete loop may be used for modeling or system validation purposes.

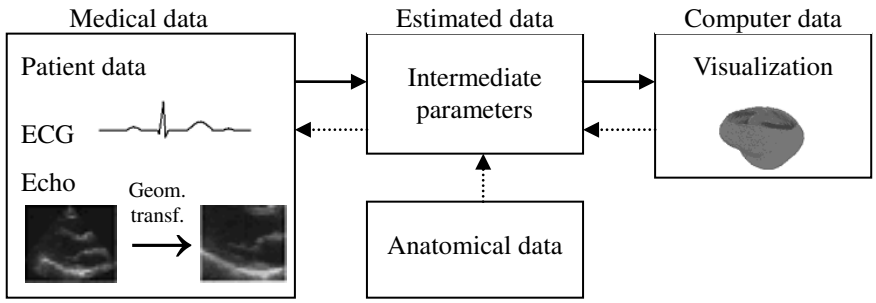


Figure 1

The medical data and graphical visualization data are connected to the estimated engineer-related parameters. The anatomical considerations and visualization data may be used to determine several medical parameters

A real time spatial heart visualization method uses an engineered parametric model, and is able to show all geometrical changes that take place during a whole normal or pathologic cardiac cycle. The lack of information about the effect of various clinical conditions, such as ischemia, on the heart functionality demands the development of intuitive modeling tools that are capable to handle multiple cardiac layers simultaneously. This modeling approach may enhance the detection rate of various phenomena that are related to the irregular contraction of the heart.

An important aspect of this parameter organization is the possibility to determine many internal parameters from raw computerized geometrical data. The set of voxels are organized in triangles, and the connection between them determines the inner and outer surface of the ventricles. Using our anatomical knowledge concerning the structure of human ventricles, allows the determination of many internal parameters.

2.2 Methodology of Cardiac Data Processing

The determination of the geometry of heart ventricles properly represents the solution many physiological problems. Normally, the left ventricle can be estimated by an ellipsoidal shaped object, but neither the wall thickness nor tissue orientation is homogenous. Moreover, the possible alterations from the normal shape due several pathological injuries may dramatically increase the estimation errors. Several papers describe different approximation methods [8], but none of the proposed methods are suitable for all pathological cases.

In our approximation, proper heart geometry estimation is based on the followings:

- Anatomical data of the ventricles, obtained from medical data libraries
- Shape estimation based on direct measurements (for example echocardiography measurements)
- Introduction of several auxiliary medical data such as age, gender, position, ECG signal, blood pressure and respiration

The combination of these data may enhance the precision of approximation of the ventricles boundaries. In 1991 Nielsen et al. [21] have determined the fiber orientation in ventricular tissue. Based on their study it is assumed that all fibers in the epicardium and right endocardium (considered the epicardium of the left ventricle) are inclined forward at 70°, and the endocardial fibers of the left and right ventricles (except the septal region) are inclined backward at 80°. The inclination of the middle regions takes intermediate values, depending on their position.

The key idea in the applied fiber orientation determination algorithm is to generate a smooth coordinate-free interpolation of the spatial directions of the fibers involved in all clusters [22, 23]. The intermediate ventricular tissue layers do not have a precisely determined structure, so all internal compartments should be described by a guiding parameter that determines how the studied compartment behaves related to the endo- and epi-cardial tissue. We assume that a given compartment has a p_{endo}, p_{epi} and p_{middle} probabilities to belong to the endocardium, epicardium or intermediate region, so the spatial direction d of the compartments tissue is:

$$d_{comp.}(x, y, z) = p_{endo}(f_r, x, y, z) + p_{epi}(f_r, x, y, z) + p_{middle}(f_r, x, y, z), \quad (1)$$

where f represents the distance rate of the compartment from the base plane related to the whole base-apex distance, the index r defines the region of the compartment (for example left or right ventricle) and the symbols x, y, z show the three perpendicular spatial directions.

Since the visual interpretation of the heart using ultrasound (US) representation techniques cannot yield as accurate data as CT and MRI images, the ventricular wall detection procedure demands a deeper usage of the a priori information. The sensibility of US imaging to noise and reflections implies an intelligent filtering. In spite of these problems, a large number of images may yield adequate results, but it is necessary to use an automatic border detection (ABD) algorithm that contains the following steps [24]:

- Pre-processing (smoothing, contrast checking)
- Edge or region detection (thresholds, edge detectors)

- Geometric object models (for example radial search)
- Anatomical structure model
- Interpretation (based on high-level knowledge)

The enrolled disadvantages convinced us to use an Active Appearance Model (AAM) that was presented in [24]. Its main steps are:

- Extract the average organ shape
- Extract the principal shape variations
- Create the appearance model
- Generate probable echocardiography image
- Find the desired structure by error minimization technique

2.3 Estimation of Geometrical Data

The estimation of geometric data is a multi-level process. The ventricular tissue has a multi-layered structure, so we have to estimate the barrier of each layer. Most geometric determination methods define the inner and outer surface of the ventricles.

All determined surfaces are constituted by a list of description points (DP) described by spatial coordinates, so the list L can be presented as:

$$L_{Surf.} = \{P_1(x_1, y_1, z_1), \ldots, P_k(x_k, y_k, z_k)\}, \tag{2}$$

where k represents the number of points in the list and the symbols x, y, z show the three perpendicular spatial directions. We have to use separate lists for the inner and outer surfaces, however they may describe the surface of a large object, for example the outer surface of the whole heart. It is important to mention that the lists can be partially combined, so from the list of outer surfaces of the whole heart, we can only partially deduce the outer surface of the left ventricle.

In the classical approach, the resolution of the simulation highly depends on the number of points from the surface list. In order to create a high-resolution simulation from a relatively low number of DPs, we have to interpolate the surface of the studied organ. The interpolation is performed for both the inner and outer surfaces.

To accomplish the interpolation it is necessary to execute the following data processing steps iteratively:

- Triangulation
- Division of the junction lines
- Estimation of the spatial coordinates on the divided lines

We used the Delaunay triangulation method [25], that can generate a set of triangles, in such a way, that P is not defined beforehand. It is supposed, that the points in the surface list are not in a line. The Delaunay triangulation method maximizes the minimum angle of all generated triangles, so it tends to avoid narrow triangles.

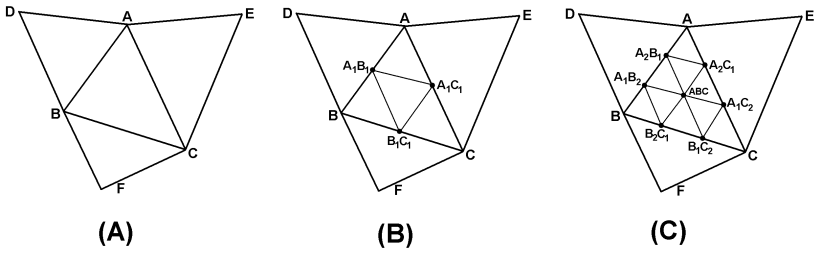


Figure 2
Generation of additional spatial triangles on the surface of the modeled object applying various resolutions. For image (A) $df = 1$, for image (B) $df = 2$ and for image (C) $df = 3$

The next step represents the generation of a high resolution triangulated surface of the spatial object. Now the goal is the construction of a surface generated by q DP-s, where $q >> p$. In the following, we will determine the proper division factor (df). It is important to mention that usually df will not be an integer value, so the processing method must handle these cases.

Figure 2 presents the decomposition of the inner triangle into sub-triangles (A). The sub-figure (B) present the decomposition for df=2, while sub-figure (C) shows the same processing for df=3. It can be observed that for a given positive integer df, on the edges of the inner triangle there will appear $EP = 3 \cdot (df - 1) = 3 \cdot df - 3$ external points (EP). The number of internal points (IP) in the inner triangle for the same integer df will be given by:

$$IP = [(df - 2)^2 + (df - 2)]/2 = [df^2 - 3 \cdot df + 2]/2 . \tag{3}$$

Each EP is a member point of two triangles, so the necessary points to perform at integer df will be approximated as:

$$\begin{aligned} q &= p + t \cdot (EP/2 + IP) \approx p \cdot \left(\frac{3 \cdot df - 3 + df^2 - 3 \cdot df + 2 + 2}{2} \right), \\ &= p \cdot \frac{df^2 + 1}{2} \approx p \cdot \frac{df^2}{2} \end{aligned} \tag{4}$$

where t represent the number of triangles.

During the reverse problem, the necessary *df* in presence of *p* DP-s and *q* aimed DP's will be $df \approx \sqrt{2q/p}$. It is observable that the relation yields proper results for $q >> p$.

For a large, closed-surfaced normal spatial object, where *p* DP-s exist that form *t* triangles, it can be considered that the number of triangles and points are approximately the same, so $p \approx t$. In this case the total number of side-edges (*se*) will be $se = p + t - 1 \approx 2 \cdot p$.

In the likely case of non-integer *df*, the division of some triangles may be performed partially. This partial division may create ambiguous results, where the selected solution possesses the more balanced angles and sizes. The fractional division of a triangle is presented in Figure 3.

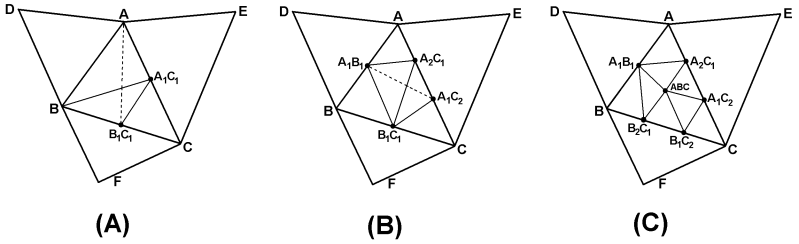


Figure 3
Several cases of fractional division factor. Image (A) and (B) present a simple ambiguous case, while (C) present a unique solution

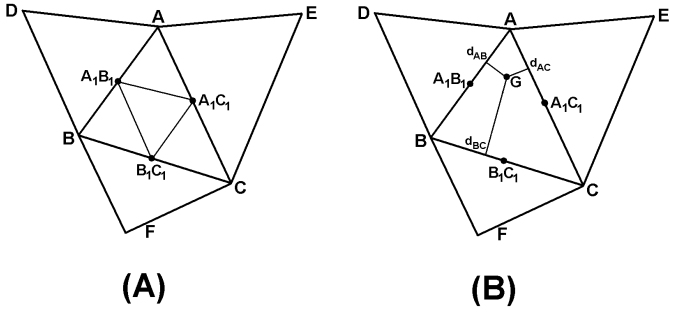


Figure 4
Determination of spatial coordinates of the internal points. Image (A) shows the case of coplanar *A*, *B*, *C* and *F* points. Image (B) visualizes the distances of the point G from the three side-edges

It is important to reduce the number of narrow triangles, so the division is performed more often on the longer side-edges of the triangles. In case of multiple solutions we selected the shorter diagonal line of the trapezoids.

The estimation of spatial coordinates of each created division point is realized by a spherical interpolation method, based on the fact that every non-coplanar four points are situated at the surface of a sphere.

The situation of coplanar points is quite simple. Suppose that points *A*, *B*, *C* and *F* from Figure 4(A) are coplanar. In this case we can affirm that all points of segment *BC* are in the same plane, and the centre of the circumscribed sphere is in the infinite.

If the *A*, *B*, *C* and *D* points are non-coplanar, the centre of the circumscribed sphere obeys the rule: $d_{OA} = d_{OB} = d_{OC} = d_{OD}$, so the spatial coordinates of the unique *O* can be determined by solving the equation:

$$\begin{vmatrix} x^2+y^2+z^2 & x & y & z & 1 \\ x_A^2+y_A^2+z_A^2 & x_A & y_A & z_A & 1 \\ x_B^2+y_B^2+z_B^2 & x_B & y_B & z_B & 1 \\ x_C^2+y_C^2+z_C^2 & x_C & y_C & z_C & 1 \\ x_D^2+y_D^2+z_D^2 & x_D & y_D & z_D & 1 \end{vmatrix} = 0 . \tag{5}$$

Once the spatial centers of spheres $S_{ABCD}, S_{ABCE}, S_{ABCF}$ are determined, it is possible to determine the sphere centre that contains point *G*. As presented in Figure 4(B), the *X* coordinate of sphere centre on whose surface point *G* is satiated will be:

$$S_{G_X} = \left(\frac{1}{d_{AB}} + \frac{1}{d_{AC}} + \frac{1}{d_{BC}} \right)^{-1} \left(\frac{S_{ABCD_X}}{d_{AB}} + \frac{S_{ABCE_X}}{d_{AC}} + \frac{S_{ABCF_X}}{d_{BC}} \right). \tag{6}$$

The *Y* and *Z* coordinates of the sphere center may be determined in an analogous way.

3 Results

The level of geometry estimation error depends on the granularity of representation and the shape of modeled object (see Figures 5 and 6). In our case, the reference constitutes a high-resolution double layered triangle mesh object that involves both ventricles, containing 20000 triangles. For interpolation purposes, a subset of them was selected that must contain more than 1800 triangles. The distances of the compartments were determined from the middle of the AV-node. Figure 5 presents the simulated geometric estimation error plotted against the division factor. The applied granularity has a linear scale, while the estimation error is presented in logarithmic range.

Figure 6 indicates the geometric estimation error for a healthy male (M) and female (F) heart. The solid line represents the left, while the dashed line shows the right ventricle estimation error plotted against the distance from the AV-node (base of ventricles).

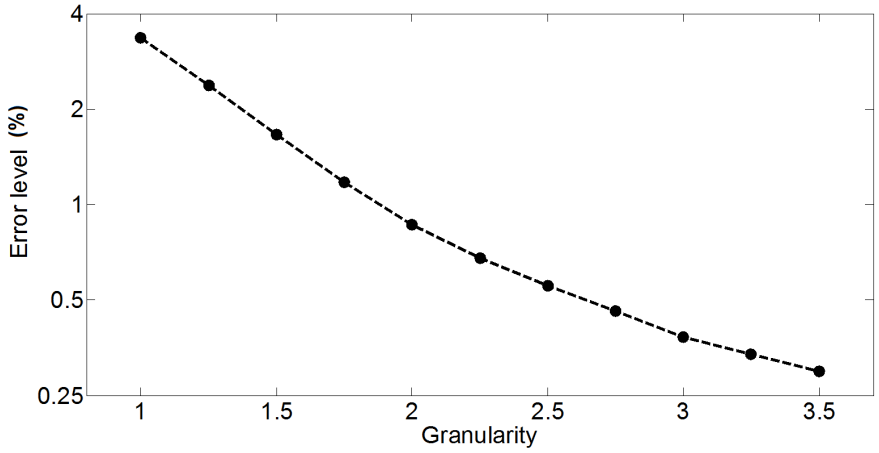


Figure 5

The geometric estimation error as a function of the applied resolution (granularity represent the applied average division factor)

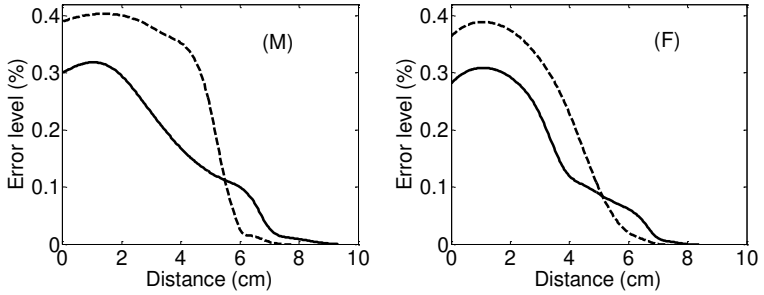


Figure 6

The geometric estimation error for a male (M) and a female (F) heart. The solid line represents the left, while the dashed line shows the right ventricle estimation error as a function of distance from the AV-node (base of ventricles)

Figure 7 exhibits the tissue orientation modeling: the image (A) visualizes the tissue fibers orientation in a sagittal segment, while images (B-D) shows the intersection of the ventricles using a horizontal plane. In all cases the applied planar resolution was 0.5 mm, while the modeled slices (images (B)-(D)) were situated at 4 cm distance.

Table I presents four important medical parameters (maximal and minimal interior volume, ejection fraction and stroke volume) of the modeled left ventricle. The simulation was performed for male and female hearts.

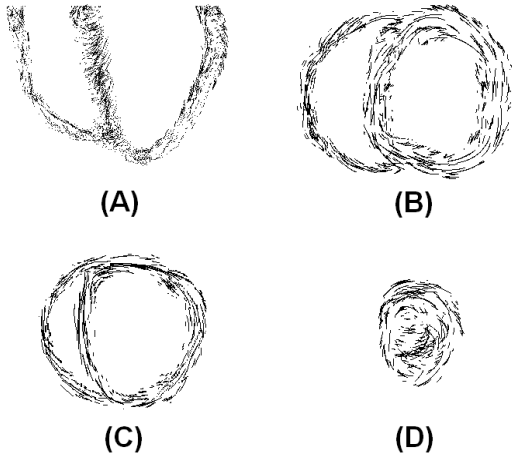


Figure 7
The simulated fiber orientations in the ventricular tissue. Image (A) visualizes the tissue fibers orientation in a sagittal segment, while images (B-D) show consecutive intersection of the ventricles from AV-node to apex using a horizontal plane

Table I
Left ventricle parameters in function of gender

	Male	Female	Average
Max. volume (ml)	105 ± 31	87 ± 25	96 ± 28
Min. volume (ml)	48 ± 15	42 ± 13	45 ± 14
Ejection percentage	57 ± 15	54 ± 12	55 ± 13
Stroke volume (ml)	61 ± 23	49 ± 21	55 ± 22

Discussion and Conclusion

An efficient ventricle geometry estimation method has to solve the continuous and flat transform among the basis voxels. As presented in Section 2.3, the aimed resolution can be acquired using fractional magnification factor. In this case the selected triangles have to be divided in such a way that the resulting new ones are as close to equilateral form as possible. This task can be solved by dividing the longer side-edges.

The increased curvature values demands a higher resolution for the starting base object, otherwise the resulting object may contain much higher estimation errors (see Figure 5). In this study all measurement tests were done with optimal conditions, so an improper division of the triangles may produce higher estimation errors than the measured ones.

The multi-layer structure of the ventricular muscle was considered as a fuzzy-like system, where the different layers cannot be separated clearly, as the different cardiac muscle cell types may mix among themselves. In this mixed representation all generated compartments contain a unique cell type. The probability of a compartment to become a small cluster of a certain cell type is determined by the place of compartment and the ventricles geometry. Once its cell type is determined, it remains constant during the whole simulation. The movement of compartments was performed by maintaining the elementary connections among adjacent elements, maintaining the law of minimal internal tension.

In the case of healthy adult subjects, the shape of ventricles becomes less complicated, at the apex region, thus the level of estimation error decreases progressively with the distance from the AV-node (see Figure 6). In the case of infant heart modeling or in presence of several pathologies, these estimation errors become higher. It is also observable that in the case of male heart the organ size and estimation uncertainty is significantly higher. The increased incertitude level of male tissue estimation is due its higher mass, large size and less uniform shape. This form may produce higher pumping power, but at the price of higher risk of heart attack or formation of arrhythmias.

In agreement to the fact that the right ventricle is shorter than the left one, the estimation error for the right ventricle becomes zero at 7 cm or higher distance. The estimation error level of the right ventricle is significantly higher than the estimation error of the left ventricle, as the shape of the right ventricle may vary more intensely due to its thinner structure and significantly lower internal blood pressure. The estimation of apex region can be performed much efficiently because its shape is determined by the cardiac muscle and the level of internal blood has a relatively lower impact on it.

The fiber orientation of the ventricles may be determined more precisely than atria fiber direction (see Figure 7). Anatomically the ventricles possess a higher importance than atria and the genetic selection pressure induces a much more standardized ventricular muscle. The fiber direction of the left ventricle has an approximately circular structure because that form may produce an almost optimal mechanical contraction and pumping efficiency.

It was considered important to provide a resolution independent description of the heart, to make connections between medical parameters and graphical representations. The presented method may produce a much more compact representation of the ventricles and can enhance the performance of heart modeling and visualization algorithms.

Acknowledgements

This paper is a result of the project “Transnational Network for Integrated Management of Postdoctoral Research in Communicating Sciences. Institutional building (postdoctoral school) and fellowships program (CommScie)” -

POSDRU/89/1.5/S/63663, financed under the Sectoral Operational Program Human Resources Development 2007-2013.

References

[1] Hill, J. A., Olson, E. N.: Muscle: Fundamental Biology and Mechanisms of Disease, Academic Press, London, Elsevier Inc. 2012

[2] Winfree, A. T.: Electrical Turbulence in Three-Dimensional Heart Muscle, Science 266 (1994) 1003-1006

[3] Panfilov, A. V.: Three-Dimensional Organization of Electrical Turbulence in the Heart, Physical Review E 59 (1999) R6251-R6254

[4] Sainte-Marie, J., Chapelle, D., Cimerman, R., Sorine, M.: Modeling and Estimation of the Cardiac Electromechanical Activity, Computers & Structures 84 (28) (2006) 1743-1759

[5] Coghlan, H. C., Coghlan, A. R., Buckberg, G. D., Cox, J. L.: The Electrical Spiral of the Heart: its Role in the Helical Continuum. The Hypothesis of the Anisotropic Conducting Matrix, European Journal of Cardio-Thoracic Surgery 29 (1) (2006) S178-S187

[6] Caillerie, D., Mourad, A., Raoult, A.: Toward a Fiber-based Constitutive Law for the Myocardium, in: Proceedings of Modeling and Simulation for Computer-Aided Medicine and Surgery (2002) 25-30

[7] Antzelevitch, C., Shimizu, W., Yan, G. X., Sicouri, S., Weissenburger, J., Nesterenko, V. V.: The M Cell: Its Contribution to the ECG and to Normal and Abnormal Electrical Function of the Heart, Journal of Cardiovascular Electrophysiology 10 (9) (1999) 1124-1152

[8] Wong, J., Kuhl, E.: Generating Fibre Orientation Maps in Human Heart Models using Poisson Interpolation, Computer Methods in Biomechanics and Biomedical Engineering, (2012), DOI:10.1080/10255842.2012.739167

[9] Kumar, V., Abbas, A. K., Fausto, N.: Robbins and Cotran Pathologic Basis of Disease. Philadelphia: Elsevier Saunders. 2005

[10] Opie, L. H.: Heart Physiology: from Cell to Circulation. Philadelphia: Lippincott Williams & Wilkins. 2003

[11] Sosnovik, D. E., Wang, R., Dai, G., Wang, T., Aikawa, E., Novikov, M., Rosenzweig, A., Gilbert, R. J., Wedeen, V. J.: Diffusion Spectrum MRI Tractography Reveals the Presence of a Complex Network of Residual Myofibers in Infarcted Myocardium. Circ: Cardiovasc Imaging. 2(3) (2009) 206-212

[12] Schmitt, B., Fedarava, K., Falkenberg, J., Rothaus, K., Bodhey, N. K., Reischauer, C., Kozerke, S., Schnackenburg, B., Westermann, D., Lunkenheimer, P. P. et al.: Three-Dimensional Alignment of the Aggregated Myocytes in the Normal and Hypertrophic Murine Heart. J Appl Physiol. 107 (2009) 921-927

[13] Libby, P., Bonow, R. O., Mann, D. L., Zipes, D. P.: Braunwald's Heart Disease. Philadelphia: Saunders. 2007

[14] Streeter, D. D., Hanna, W. T.: Engineering Mechanics for Successive States in Canine Left Ventricular Myocardium: II Fiber Angle and Sarcomere length. Circ Res. 33(6) (1973) 656-664

[15] Arts, T., Costa, K. D., Covell, J. W., McCulloch, A. D.: Relating Myocardial Laminar Architecture to Shear Strain and Muscle Fiber Orientation. Am J Physiol Heart Circ Physiol. 280(5) (2001) H2222-H2229

[16] Geerts, L., Bovendeerd, P., Nicolay, K., Arts, T.: Characterization of the Normal Cardiac Myofiber Field in Goat Measured with MR-Diffusion Tensor Imaging. Am J Physiol Heart Circ Physiol. 283(1) (2002) H139-H145

[17] Göktepe, S., Kuhl, E.: Electromechanics of the Heart – a Unified Approach to the Strongly Coupled Excitation Contraction Problem. Comput. Mech. 45 (2010) 227-243

[18] Takayama, K., Ashihara, T., Ijiri, T., Igarashi, T., Haraguchi, R., Nakazawa, K.: A Sketch-based Interface for Modeling Myocardial Fiber Orientation that Considers the Layered Structure of the Ventricles. J Physiol Sci. 58(7) (2008) 487-492

[19] Toussaint, N., Stoeck, C. T., Sermesant, M., Kozerke, S., Batchelor, P. G.: Three-Dimensional Prolate Spheroidal Extrapolation for Sparse DTI of the in Vivo Heart. In: Book of Abstracts of the International Society in Magnetic Resonance in Medicine. Sweden: Stockholm. 2010

[20] Kotikanyadanam, M., Göktepe, S., Kuhl, E.: Computational Modeling of Electrocardiograms – a Finite Element Approach towards Cardiac Excitation. Int J Numerical Methods Biomed Eng. 26 (2010) 524-533

[21] Nielsen, P. M., Le Grice, I. J., Smaill, B. H., Hunter PJ. 1991. Mathematical Model of Geometry and Fibrous Structure of the Heart. Am J Physiol Heart Circ Physiol. 260(4) H1365-H1378

[22] Zolgharnein, E., Mirsalimov, V. M.: Nucleation of a Crack under Inner Compression of Cylindrical Bodies. Acta Polytechnica Hungarica 9(2) (2012) 169-183

[23] Taheri, R., Mazaheri, K.: Hydrodynamic Optimization of Marine Propeller Using Gradient and Non-Gradient-based Algorithms. Acta Polytechnica Hungarica 10(3) (2013) 221-237

[24] Szilágyi, S. M., Szilágyi, L., Benyó, Z.: Volumetric Analysis of the Heart using Echocardiography. Lecture Notes in Computer Science 4466:81-90 (2007), ISSN: 0302-9743

[25] Delaunay, B.: Sur la sphère vide, Izvestia Akademii Nauk SSSR, Otdelenie Matematicheskikh i Estestvennykh Nauk, 7 (1934) 793-800

Analysis of the Information Flow within the Information System of Car Parks

Zoran Nesic[1], Leon Ljubic[2], Miroslav Radojicic[1] and Jasmina Vesic Vasovic[1]

[1]University of Kragujevac, Faculty of Technical Sciences
Svetog Save 65, 32000 Cacak, Serbia

[2]JKP "Parking Service Kragujevac "
Vojislava Kalanovica b.b., 34000 Kragujevac, Serbia

E-mail: zoran.nesic@ftn.kg.ac.rs, leon.ljubic@parkingservis.rs, miroslav.radojicic@ftn.kg.ac.rs, jasmina.vesic@ftn.kg.ac.rs

Abstract: In this paper, the major elements of the information flow in car parks information system has been presented. A detailed analysis of information flow in one of the key elements for a parking service company, thus, providing the basis for developing the information system in this area. The analysis of decomposition diagrams and charts, in the context of the information flow, has also been presented here, as well as, the formation of a logical and relational database model. The focus is on the performance of a detailed analysis of the information flow, with the aim of improving the overall functioning of a parking service company. The development of the information system of car parks is the most important step in both the development of the overall information system of this company and its improvement. The results presented in this paper might be a starting point in the development of information systems for parking service companies, in general. The obtained results help in the formation of a clear and consistent information system in this area.

Keywords: information flow; information systems; databases

1 Introduction

The information system of a parking service company is a tool by which service information, as well as, information about the decision-making processes, at all levels, is provided to citizens/users. Accordingly, it should be an integral part of the information system of the city. From that aspect, this information system is actually a tool for decision-making in the areas of service and control; therefore, it is a part of a wider information unit.

Development of the information system of a parking service company is based on the legal requirements [1] - [3] and the general management methodology [4] - [6].

The first and most important advantage of such an information system, is that it provides information to the user, preferably in real time. This generally refers to the information concerning the availability of parking spaces and areas but also about simplified parking payment and other useful services (working/non-working days, zones where parking is allowed, etc).

Another advantage is that the operating processes of a parking service company are integrated into an optimized system, by which management irrationality is eliminated. Processes become tightly linked with no possibility for interruption; therefore, there is no loss of information and no risk of unauthorized access to information. Also, the processes, tasks and activities within individual processes are clearly distinguished in this way; as a consequence, the existing organizational model of processes, tasks and activities remains constant.

The third advantage is optimal management of company's resources, so that it is possible to predict the future state of the company based on the current state. This is very important for determining future goals, strategies and plans and for determining the overall business policy on which the entire management will be based. The term "resources" refers primarily to the area of land that is or will be a part of the parking area. The information system provides information which represents the basis for making good decisions on optimized management of the existing parking areas and on the use of new parking areas. The term "optimized management" refers to the optimal use of parking areas but also the optimal engagement of the personnel and resources involved in the work process. In addition, we should mention the transport infrastructure resources which are very important for the implementation of parking rules. Other resources are people and labour force, capital assets, financial resources, etc.

The fourth advantage is the integration of such an information system into a unique future information system of the city, which will take a decision-making process to the higher level.

The fifth advantage, which unites all other advantages, is provision of simplified decision-making process at all levels of management. It will give the decision maker one or more alternatives which are the result of realistic and precisely defined state of the system, i.e. company.

This paper is based on the fundamental concepts of information systems [7] - [11] and databases [12] - [17] development.

2 Car Park Management Function

This function includes the processes, activities and operations used in managing parking areas in terms of meeting the users' / citizens' needs for parking. Therefore, the emphasis is not on the parking area as a resource, but on a parking area as means for satisfying the needs of users. All parking areas are divided into two main types – open car parks and special car parks.

Open car parks are the unconfined urban public areas where parking is allowed. These parking areas may be in the street, on the sidewalk or around the buildings. Parking spaces are marked with vertical and horizontal signals. Open car parks are divided into zones; there are three zones – 0, 1 and 2. Each zone has special parking rules due to which the users' needs are met in optimal way.

Zone 0 is the zone that covers the city centre; in this zone, parking is limited to 2 continuous hours after which the user must remove the vehicle and may re-park there again after one hour break and stay there for another two hours. For zone 1 and 2 there are no limits regarding parking hours, i.e. the user can park for an indefinite period of time regardless of the time of day. Zones 1 and 2 include the wider city centre and parts between wider city centre and periphery.

By paying for parking via a SMS payment system or by purchasing a paper ticket, the user becomes entitled to a parking space and he/she undertakes the obligation to respect the rules of the particular parking zone.

Control is performed by parking attendants who use a computer device by which they send an enquiry to the central computer and receive information for the specific vehicle regarding payment and parking hours. If the attendant detects that the vehicle owner has exceeded the number of paid parking hours or perhaps has not paid for the parking, he or she must issue a special daily ticket which is a part of the price list and is valid for 24 hours excluding non-working days and holidays.

That means that the user is obliged to pay that special daily ticket in the following period and in return he or she can park in all parking zones in the next 24 hours without any restrictions. Naturally, a parking service company plans to make the citizens use the SMS payment system or scratch tickets more often, with the purpose of reducing the parking congestion in central city zones and directing the vehicles towards city periphery. On the other hand, a special daily ticket is a form of penalty for irresponsible users; it should improve their awareness of the need to respect the parking rules.

Special car parks are all parking areas that are confined and the entrance to the area is paid per hour of parking. There are no limitations for parking at special car parks, but the user must respect the same general traffic rules which apply to open car parks.

Upon entering the special car park, the user is issued a parking confirmation which includes the date and time of entry, license plate number, car park index, price per hour, name of the employee who has issued the parking confirmation and some other status information.

When leaving the special car park, the user submits the parking confirmation and makes the payment; after that, he/she is issued a receipt similar to the parking confirmation, with the additional information about the number of hours, as a basis for payment, and the total amount to be paid. By this activity the user, i.e. vehicle, is relieved of debt. If for some reason the user fails to pay, it will be forbidden to park that vehicle at any other special car park until the due debt is settled.

In addition to special car parks, there are also truck terminals where the rules are similar to those of special car parks.

3 Analysis of Information Flow in Car Parks Management Function

Figure 1 shows a context diagram that displays all the inputs, outputs, controls and mechanisms essential for the best possible functioning of the parking process.

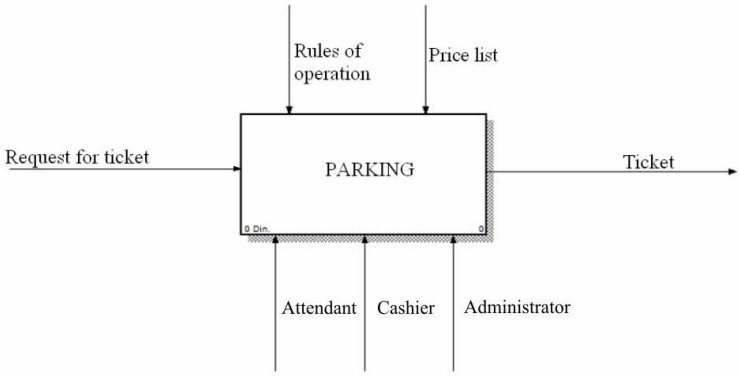


Figure 1
Context diagram for car park management function

A request for the ticket represents the input information. It is not a standardized document, but instead a general entry information by which the user expresses the need to buy a ticket, i.e. to use the parking service. This request usually has the form of an SMS or a ticket bought in the parking service company office or in one of the shops. Also, there is a category of disabled parking tickets that are free of charge, but are issued to the user through a special request.

Output information is a ticket that can have the form of a paper ticket or an SMS ticket. There are many types of tickets which a parking service company can sell:

- Electronic SMS tickets which can be hourly and daily
- Scratch tickets which can also be hourly and daily
- Monthly season tickets which are valid for a longer period of time
- Residential tickets issued to residents of objects located in parking sectors
- Tickets that are issued to users at special car parks

Internal Rules of Procedures, formed by a parking service company for the purpose of unifying work processes and activities, are applied as a mechanism. A price of the ticket is another mechanism. This mechanism is based on the price list of a parking service company and it is approved by the City Assembly.

As shown below, the elements of control are a parking attendant, a cashier and an administrator.

Figure 2 shows that this function consists of two processes: parking payment and parking payment control. The process of parking payment is the process by which parking conditions are assessed and money for parking is collected. For the process of parking payment, the input is a request for the ticket, and the output is a confirmation of ticket purchase. This confirmation may have the following forms:

- Electronic SMS confirmation that the user receives from the electronic payment system on the device from which SMS request for a ticket was sent. In this way the user is informed that the request was successfully completed and that the required SMS ticket is in the ticket database.
- For a scratch ticket for open car parks it is a scratch ticket itself, which is later cancelled by the attendant who puts it back into the system to be reused.
- For monthly season and disabled parking tickets, it is a plastic card which is issued to the user. After this type of ticket had been issued to the user, it is inserted into the database so that the attendants can check it.
- For residential tickets in open car parks, it is a sticker which represents both the ticket and the receipt. It is placed on the windscreen and it serves as a visual confirmation which facilitates the control.
- Confirmation at special car parks has the form of slip receipt which represents both the ticket and the confirmation of purchase.

The elements of the mechanism are various Rules of Procedures and valid price list, which are explained in the section on the context diagram.

Elements of control are also an attendant, a cashier and an administrator who issues monthly season, disability and residential tickets.

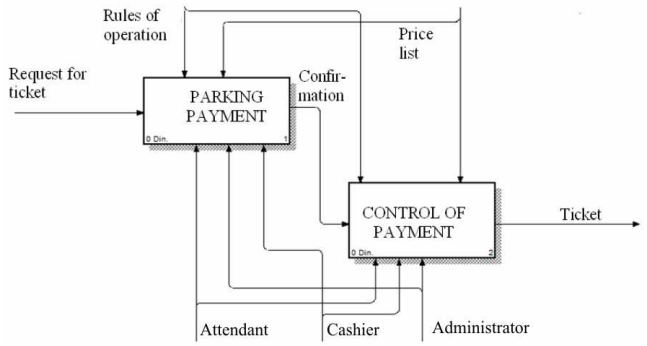


Figure 2
Decomposition diagram for the parking management function

Parking payment control is the process which includes control of payments, i.e. correctness of the payment elements.

The input for this process is the confirmation generated in the process of parking payment, and output is the ticket itself. Same as in parking payment process, control is exercised through a variety of regulations and price lists. Executive mechanisms are an attendant, a cashier and an administrator.

In Figure 3 all the activities used in parking payment process are presented. There are six activities that take place at the same time. They all have their specific characteristics reflected in the various procedural steps of the parking payment process. However, they all have the same input, i.e. request for a ticket that is generated by the user and can have different forms.

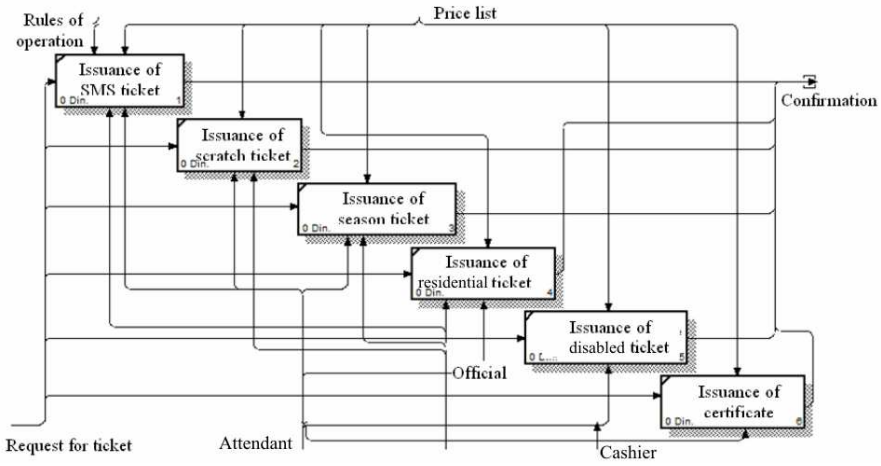


Figure 3
Diagram of the first level in the parking payment process

The output for all the activities is a confirmation issued to the user. Sometimes the ticket is used as a confirmation and sometimes the confirmation is used as the ticket.

The controls are performed according to a variety of regulations that are generally called the Rules of Procedures, as well as according to the valid price list.

Mechanisms are an attendant, a cashier and an administrator. The attendant is a mechanism that performs the control of SMS and scratch tickets payment. The administrator is a mechanism engaged in charging for parking via monthly season tickets and residential tickets and in issuance of the disabled parking tickets. The cashier at special car parks is a mechanism that issues confirmations which are also the tickets.

Figure 4 shows that confirmations generated in parking payment process are the input to the activity of parking payment control. This activity involves the paid tickets control, but not the payment itself. Elements of control are also Rules of Procedures and the valid price list. The mechanisms are attendants, cashiers and administrators.

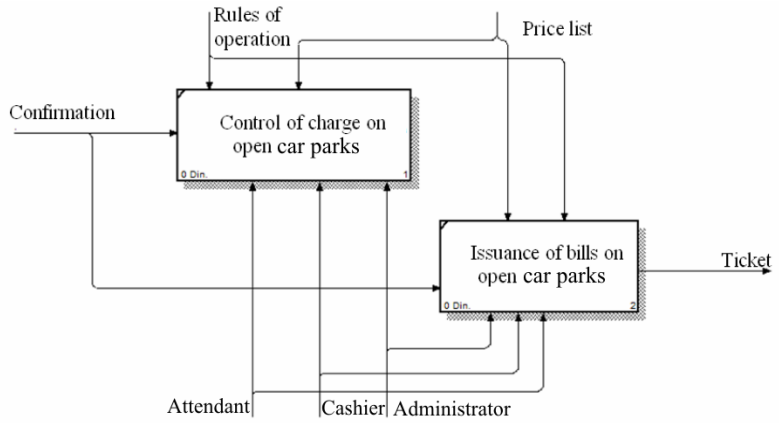


Figure 4
Diagram of the first level for parking payment control process

Issuance of receipts at special car parks is an activity by which procedures for issuing receipts are grouped, since the receipt issuing at special car parks is specific – the issued ticket is at the same time the confirmation.

At special car parks, the output of this activity is the ticket itself. Elements of control are Rules of Procedures and the valid price list. Mechanisms are cashiers and administrators.

4 Results and Discussion

A preliminary analysis of information flow was the most important step in the creation of the information system. That analysis has provided the basis for the creation of a relational database model, as well as for the following major phases of information system development. The creation of the relational database model is the basis for further analysis of information from different aspects [18] and for the use of various business intelligence methodologies [19] - [21]. There are five tables in this model. Some tables are code-books, such as Worker, Zone, Street.

All entities from the group of independent entities are candidates for code-book tables. A dependent entity – the ticket – is a candidate for generated tickets table. Derived entity Zone / Street is an entity via which available data on Street - Zone relation are generated, since these relations are changeable and depend on the activities of a parking service company. The tables that appear in the relational model for this function are presented in Figure 5: Worker, Ticket, Zone, Price list, Street and Zone / Street. Independent entities are Employee, Zone, Price list, Street. A dependent entity is a Ticket. Derived entity is Zone / Street.

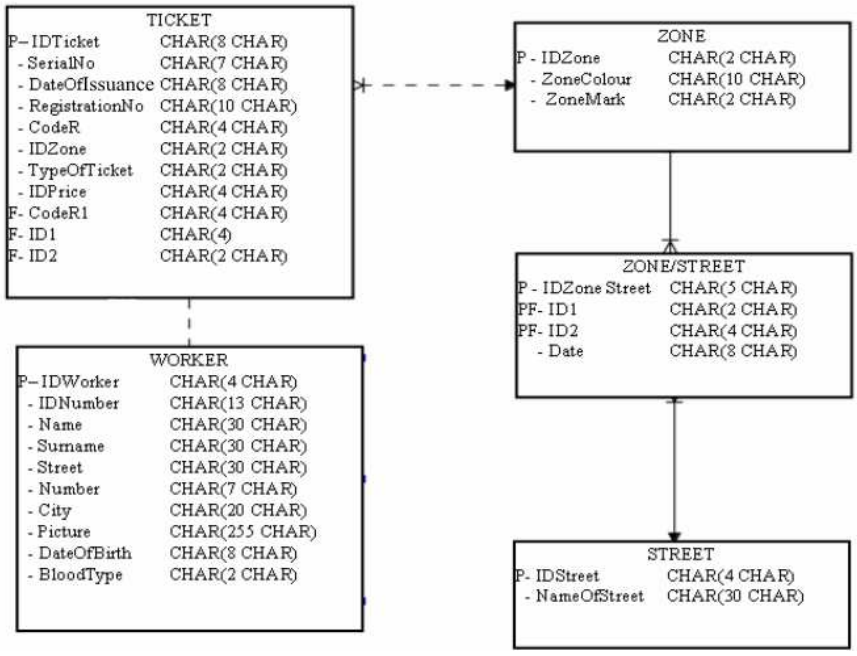


Figure 5

The relational data model for car park management function

Creation of the relational database model has enabled obtaining of numerous database analyses. Figure 6 presents results from the main table for this function.

The most important information on issued tickets for open and special car parks is as follows:

- ID Ticket, identifier of a record in a set of all records. It is necessary for internal processing in the future application.
- Serial number, mark of the ticket issued to the user. This mark serves for identifying the document TICKET and is typically located on the physical document TICKET. In case of SMS, disabled parking or season tickets, it is a unique number that is generated by the system.
- Issuance date, the date of issuance of the document TICKET.
- Registration mark, license plate number of the vehicle for which the ticket is issued.
- Code of the employee obtained from the outer table – employees' code-book.
- Zone code from the outer table – zone code-book
- Type of ticket, code of the type of ticket issued by the employee. 01 - special day ticket, 02 - residential ticket etc.
- Identification of prices from the price list, i.e. pricelist code-book

WORKER : Table

IDWorker	IDNumber	Name	Surname	Street	Number	City	Picture	DateOfB
0001	0510968134124	Petar	Petrović	Ibarskih rudara	10/3	Kragujevac		19680412

IDTicket	SerialNo	DateOfIssuance	RegistrationNo	CodeR	IDZone	TypeOfTicket	IDPrice	CodeR1	ID2
00000001	42343424	20140512	KG954OU	0001	01	01	01	0001	01
00000002	43412540	20140512	KG434TX	0001	01	01	01	0001	01
00000152	45745787	20140514	UE344EW	0001	01	01	01	0001	01
00000153	45745788	20140514	SG434232	0001	01	01	01	0001	01
00000154	45745789	20140514	KG001OP	0001	01	01	01	0001	01
00000168	45745788	20140515	KG002WW	0001	01	01	01	0001	01

0002	1205975342342	Dejan	Stanković	Moravska	12	Kragujevac	19750512
0003	1909955723423	Dragan	Radaković	Jefimijina	4	Kragujevac	19550919
0004	3003982723232	Ivan	Marjanović	Vojislava Kalanovića	11/24	Kragujevac	19820330
0005	2708980730023	Katarina	Radivojević	Sofije Ristić	7/3	Kragujevac	19800827
0006	0605976742095	Biljana	Jovanović	Jugoslovenska	5	Kragujevac	19760506
0007	1504966720012	Dragutin	Radosavljević	Kolubarska	9	Kragujevac	19660415
0008	0111964720055	Jasna	Katić	Zmaj Jovina	22/13-4	Kragujevac	19641101

Record: 1 of 6

Figure 6
Results of the database analysis

Based on the relational database model, Figure 6 shows the corresponding data about employees who have issued the tickets:

- IDWorker - field for the identification of records on a particular employee. It is used only for internal purposes of the future application.

- Personal identification number, a field which contains data on personal identification number (PIN). This is necessary because PIN is gaining importance increasingly as a unique property of a person, i.e. an employee. Most operations in the future application are based on a PIN number (or rather the mark).
- Name – Employee's name
- Last Name – Employee's last name
- Street - Street in which an employee resides
- Number - mark of the residential unit in the street in which an employee resides.
- City - identification of the employee's city of residence
- Photo - a field which contains the photo of an employee or a hyperlink to a location on the disk where the employee's photo is located.
- Date of birth - information on an employee's date of birth in the format YYYY-MM-DD, because it facilitates data processing; therefore, there is no dependence on the local settings of date and time.

Further use of the relational database model allows obtainment of more information about the zones. It plays the role of a code-book in the data model. The most important information is identification of records on a particular zone. It is used solely for internal purposes of the model, as descriptive information of a certain zone. It is used for easier zone identification by the user and as official zone mark.

Based on this, it can be concluded that there is an upward trend regarding income which is the result of the introduction of Information subsystem for car parks management. Based on the analysis of data from Business plan of JKP Parking Servis Kragujevac for 2008-2013 [1], there has been an increase of over 75% since 2008, when the IS was introduced, Table 1 and Figure 7.

Table 1

Income from open and special car parks

Year	2007	2008	2009	2010	2011	2012	2013
Income from open car parks	24.479	42.990	55.321	52.051	56.696	65.533	69.124
Income from special car parks	19.296	21.409	27.249	35.470	47.521	54.067	57.704
Note : Values are in RSD ($x10^3$)							

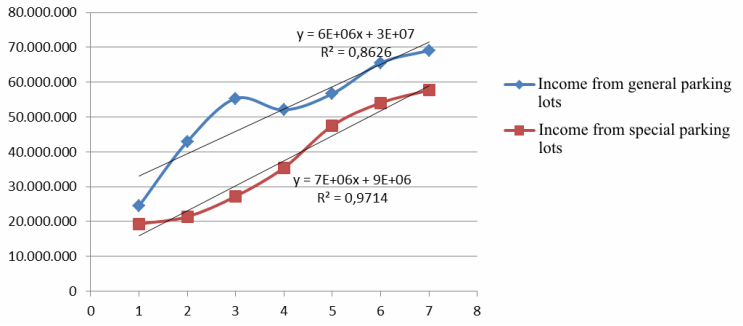


Figure 7
Upward trend of income from open and special car parks

Income from open car parks has increased due to three reasons:

- Users are able to pay for the parking in a less complicated way, through SMS system, which shortens the payment time.
- Payment control, performed by the employees of the Parking Service Company, is more effective; the psychological effect of the necessity of payment is enhanced.
- The number of available parking spaces has been increased, as a result of the introduction of the information system.

The payment of parking at special car parks is also better, because before the introduction of the information system it was practically impossible to control parking attendants at special car parks. As a result of its introduction, payment control, primarily the control of parking attendants by the management, has become more effective. A more detailed structure of sold tickets is shown in Table 2 and Figure 8.

Based on everything herin, we can see that there is a positive trend, i.e. that the introduction of the information subsystem for parking management has led to a better use of the existing parking areas at special car parks . It must also be noted that in this period (2007-2013) two special car parks were closed, which reduced the number of effective parking spaces which is another indicator of the positive effects of the IS introduction.

Table 2
The structure of sold tickets

Year	2007	2008	2009	2010	2011	2012	2013
1-hour parking tickets	720	961	1.274	1.353	1.456	1.181	1.133
Scratch parking tickets	130	193	151	96.	87	47	79
Residential tickets	3.48	3.52	3.45	3.34	3.54	3.57	3.60
Note: Values are expressed in units (number of tickets x10^3)							

The structure of sold tickets indicates both some positive and some negative trends. The sale of classic one-hour SMS tickets has increased, which is in line with previously given facts; however, there is also a negative trend related to multi-hour tickets sale. The reason for that are some additional parking limitations in payment zones (which is not directly connected to the information system), as well as the reduction of purchase of classic paper multi-hour scratch tickets and more frequent use of SMS charging system which is a part of the information system – therefore, it is a positive effect, basically, regardless of the negative trend, because it directs users to the application of more efficient payment system.

As for the residential tickets, the status is slightly positive, but it is not the consequence of IS itself, but of some other factors which, are not the subject of this work.

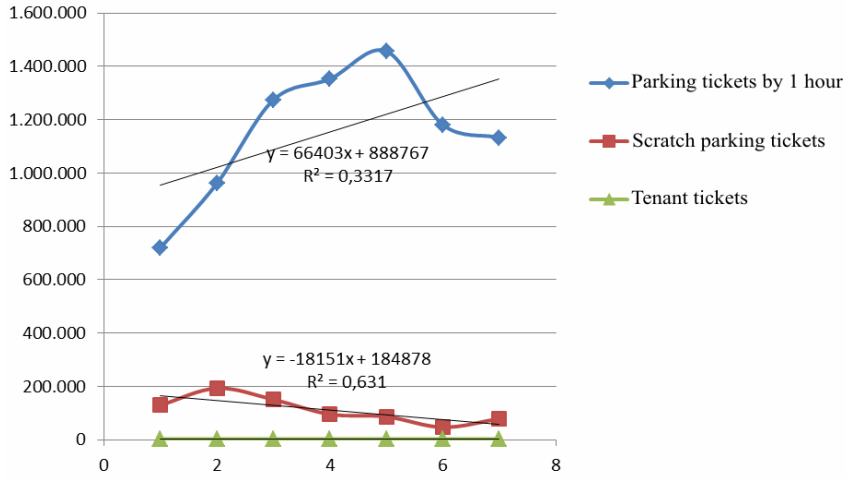


Figure 8
The structure of sold tickets

Conclusion

The analysis presented in this paper is a basic and important step in the development of the information system for car parks management. The advantages of applying the information system (IS) in the parking service company management are obvious, because they help the creation of a rational framework for justifying the existence of a parking service company. Setting up of an information system in this type of company can be considered on three levels, with each level allowing for the improvement of procurement and analysis of information: level of service, level of process management and level of integration in the communal system.

The main purpose of this paper is to provide an overview of the key segments in the development of the information system for a parking service company – analysis and determining of information flow. In that way a precise definition of information flow is obtained and expressed in decomposition diagrams. The presented analysis has enabled the formation of the relational database model, which is also given in the paper. The results presented here are a fundamental and crucial element in the development of the information system of a parking service company. Based on detailed analysis, it is possible to form the physical level of database and software solutions, which will allow access to all the information from the observed business activity. The presented segment of the information system is an integral part of the overall information system of a parking service company.

In this paper, the results have been presented in a methodological fashion, enabling a detailed understanding of these issues and, in general, the creation of the basis for the development of an information system for a parking service company.

Acknowledgement

The research presented in this paper was supported by the Ministry of Education and Science of the Republic of Serbia, Grant III-44010, Title: Intelligent Systems for Software Product Development and Business Support based on Models.

References

[1] The public company „Parking Service“ Kragujevac, Available at: www.parkingservis.rs/ (Accessed: 30.09.2013)

[2] Assembly of the City of Kragujevac, Available: www.kragujevac.rs/Skupstina_grada-40-1 (Accessed: 30.09.2013)

[3] Law on Traffic Safety on Roads, Available at: www.parkingservis.rs/images/pdf/zakon_o_bezbednosti_saobracaja_na_putevima.pdf (Accessed: 30.09.2013)

[4] Parking Lot Service System, Oki Electric Industry Co., Ltd. Available at: www.oki.com/ (Accessed: 30.09.2013)

[5] Caicedo, F. - Blazquez, C. - Miranda P.: Prediction of parking space availability in real time, Expert Systems with Applications, Vol. 39, No. 8, 2012, pp. 7281–7290

[6] Chrest, A. P.: Parking Structures: Planning, Design, Construction, Maintenance, and Repair, Kluwer Academic Publishers, Massachusetts, 2001

[7] Stair, R. - Reynolds, G.: Principles of Information Systems, Cengage Learning, Boston, 2013

[8] Olson, D. L. - Kesharwani, S.: Enterprise Information Systems: Contemporary Trends and Issues, World Scientific, Singapore, 2010

[9] Curtis, G. - Cobham, D.: Business Information Systems: Analysis, Design and Practice, Pearson Education, Harlow, England, 2008

[10] Maria van der Heijden, J. G.: Designing Management Information Systems, Oxford University Press, Oxford, 2009

[11] Johannesson, P. - Söderström, E.: Information Systems Engineering: From Data Analysis to Process Networks, Idea Group Inc (IGI), Hershey, 2008

[12] Narang, R.: Database Management Systems, PHI Learning Pvt. Ltd., New Delhi, 2011

[13] Silberschatz, A. - Korth, H. F. - Sudarshan, S.: Database System Concepts, McGraw-Hill, New York, 2011

[14] Coronel, C. - Morris, S.: Database Systems: Design, Implementation, and Management, Cengage Learning, Stamford, USA, 2014

[15] Ritchie, C.: Database Principles and Design, Cengage Learning IMEA., London, 2008

[16] Garcia-Molina, H. - Ullman, J. D. - Widom, J.: Database systems: the complete book, Pearson Education, UK, 2011

[17] Elmasri, R.: Fundamentals of Database Systems, Pearson Education, UK, 2008

[18] Sandor, Z. P. - Csiszar, C.: Development Stages of Intelligent Parking Information Systems for Trucks, Acta Polytechnica Hungarica , Vol. 10, No. 4, 2013, pp. 161-174

[19] Sabherwal, R. - Becerra-Fernandez, I.: Business intelligence, Wiley, Hoboken, NJ, 2011

[20] Loshin, D.: Business Intelligence: The Savvy Manager's Guide, Morgan Kaufmann, Waltham, USA, 2012

[21] Vitt, E. - Luckevich, M. - Misner, S.: Business Intelligence, Microsoft Press, Redmond, USA, 2010

Main Concepts, State of the Art and Future Research Questions in Sentiment Analysis

Orestes Appel, Francisco Chiclana, Jenny Carter

Centre for Computational Intelligence, Faculty of Technology, De Montfort University, The Gateway, Leicester, LE1 9BH, United Kingdom
orestes.appel@email.dmu.ac.uk , chiclana@dmu.ac.uk, jennyc@dmu.ac.uk

Abstract: This article has multiple objectives. First of all, the fundamental concepts and challenges of the research field known as Sentiment Analysis (SA) are presented. Secondly, a summary of a chronological account of the research performed in SA is provided as well as some bibliometric indicators that shed some light on the most frequently used techniques for addressing the central aspects of SA. The geographical locations of where the research took place are also given. In closing, it is argued that there is no hard evidence that fuzzy sets or hybrid approaches encompassing unsupervised learning, fuzzy sets and a solid psychological background of emotions could not be at least as effective as supervised learning techniques.

Keywords: sentiment analysis; opinion mining; social media; fuzzy sets; supervised learning; unsupervised learning

1 Introduction

Sentiment Analysis (SA) – or Opinion Mining (OM) – is a discipline that has seen a lot of activity since about 2000 [41]. The main reason for this, so it seems, is the proliferation of social media and its tools (e.g. Twitter, Facebook, LinkedIn, etc.), that has made the accessibility to information about *how people feel about things* more readily available to the masses. In addition, companies and other profit and non-profit organisations have accumulated a vast amount of data on how their employees or customers *feel* about the products and services they receive from the aforementioned organisations. Even Human Resources divisions are keen on understanding whether potential employees will be loyal and become a long-term member of the company or whether they would leave after receiving training and benefits.

In a way, a discipline that started as a research topic in Natural Language Processing (NLP) in Computer Science schools around the world, has now made a transition to other departments in academia and the industry, like those more related to business and management schools. The reason is very simple: everyone

wants to maximise their profits, and getting to understand what people think about oneself and one's company products could make a big difference business-wise.

According to some respected researchers [41, 44], there are many challenges lying ahead for SA as will be elaborated in the following sections. The reasons are many, but the fact that NLP has been around for a long time and has only focused on SA recently suggests the intrinsic difficulties with this discipline. Indeed, SA combines the application of NLP, Computational Linguistics and Text Analysis in its own way. A definition of SA attributed to Michelle de Haaff, in her article "Sentiment Analysis, Hard But Worth It" published in her CustomerThink (2010) *blog*, is as follows: ". . . *classifying the polarity of a given text at the document, sentence, or feature or aspect level, whether the expressed opinion in a document, a sentence or an entity feature or aspect is positive, negative, or neutral".* Michelle de Haaff goes ahead to loosely define Advanced SA, as the one that goes 'beyond polarity' sentiment classification, and it looks, for instance, at emotional states such as 'angry', 'sad', and 'happy'.

Understanding the emotions being conveyed by a given source – may it be a tweet, a document, a report, a blog, a segment of a politician's speech, etc. – has proven to be an important activity for humans. However, when volumes of *opinions* are very high, human processing becomes a challenge, hence the need for automated processes to extract sentiments from a variety of sources that keep growing in volume, complexity and diversity.

This article aims to present the fundamental concepts and challenges of the SA research field. To do this, an insight into the basics of SA and a summary of a chronological account of the research performed in SA will be provided. Additionally, some bibliometric indicators are included to shed some light on the techniques, and their geographical origins, most often used in addressing the central aspects of SA. Finally, the following research hypothesis is put forward: "there is no hard evidence to suggest that *fuzzy sets* or *hybrid approaches* encompassing unsupervised learning, fuzzy sets and a solid psychological background of emotions could not be as effective – or even better – as supervised learning methods."

2 Sentiment Analysis Basics

A sentiment (opinion) lexicon is defined as 'a list of positive and negative opinion words or sentiment words for English' [29]. It is assumed that such a lexicon could be built as well for any other language that one desires to use. According to Feldman [22], *the sentiment lexicon* "is the most important resource for most crucial analysis algorithms". Weichselbraun et al. [65] highlight the importance of context when producing sentiment lexicons. Indeed, they claim that "the limited ability of automated systems to resolve ambiguities and process context

information represents a major challenge". Thus, as the reader has probably guessed already, the importance of producing an accurate sentiment lexicon is that any polarity/sentiment evaluation to be performed will be based on the lexicon.

Opinions are easy to understand for human beings, but it is not that easy for a computer to have the same level of understanding. The notion of opinion as given by Liu [40] consists of the following items: (1) *Opinion targets*: entities and their features/aspects: (2) *Sentiments*: positive or negative; (3) *Opinion holders*: people who hold the opinions; (4) *Time*: when opinions are expressed.

Formally, an opinion is represented as a quintuple (e_j , a_{jk} , so_{ijkl} , h_i , t_l), where: e_j is a target entity, a_{jk} is an aspect/feature of the entity e_j; so_{ijkl} is the sentiment value of the opinion from the opinion holder h_i on feature a_{jk} of entity e_j at time t_l; so_{ijkl} is positive, negative or neutral, or more granular ratings; h_i is an opinion holder; and t_l is the time when the opinion is expressed. Liu provides a number of caveats to this definition [40], though: (i) although introduced using a product review, the definition is generic enough, in the sense that is applicable to other domains, e.g. politics, social events, services, topics, etc.; (ii) (e_j, a_{jk}) is also called the *opinion target* – opinion without knowing the target is of limited use; (iii) the five components in (e_j , a_{jk} , so_{ijkl} , h_i , t_l) must correspond to one another; (iv) the five components are essential – without any of them, it can be problematic in general.

Liu continues to describe the SA task as requiring to "structure the unstructured" [40] because Natural Language is regarded as unstructured data, and therefore the problem definition should provide *a structure* to the unstructured problem on the following three areas: (1) *Key tasks*: identify key tasks and their interrelationships; (2) *Common framework*: provide a common framework to unify different research directions; (3) *Understanding*: help us understand the problem better.

2.1. Key Tasks to Perform in SA

In general terms the problem of SA has two different abstraction aspects [40]: (1) *Opinion definition*, which has been addressed above, and (2) *Opinion summarisation*: opinions are subjective, and are needed from a significant number of people. Hence, some kind of summarisation will be required. The main components of the process of extracting sentiment from a given source, as taken from Kumar & Sebastian [36], are:

1. **Subjectivity Classification:** A document is a collection of sentences that may, or may not, express the author(s) opinion, which are called *subjective*. The sentences that are factual in nature are called *objective*. Usually, both types are present in a document. Subjectivity classification is "the task of classifying sentences as opinionated or not opinionated" [36].
2. **Sentiment Classification:** After finishing the task of identifying whether a text is opinionated, the polarity of the opinion must be found. Usually, classifying an opinion as either positive or negative is enough, i.e. *Values* = [*positive,*

negative]. However, sometime a multi-class classification might be used, with a range of values possible as exemplified with *Values* = [*extremely negative, negative, neutral, positive, extremely positive*].

3. **Complimentary Tasks: (a) Opinion Holder Extraction:** Depending on the type of application of SA, it would be necessary to identify the *opinion holder*. In some types of documents, there could be multiple opinion holders expressing their opinions about different subjects, hence the need to identify in those cases who is the opinion holder in every case. **(b) Object/Feature Extraction:** A task that may be necessary to execute – or not depending on the type of document being processed – is the identification of the target entity about which opinions are being issued. For instance, in social media it is not uncommon that a number of issues may be addressed (e.g. in blogs), so it is key to get to know about which object/feature opinions are being expressed.

2.2. Level of Analysis Issues

SA can be performed at many levels and at different complexity standards. According to Liu [41] there are commonly three different levels of analysis: Document level; Sentence level; and Entity Feature or Aspect level. Notice that these three levels of analysis are contained in Michelle de Haaff's definition of SA given before. Using a slightly different approach, Kumar & Sebastian [35] differentiate between the following levels of SA:

- **Document Level:** The whole document being analysed is the basic unit whose sentiment orientation will be determined.
- **Sentence level:** At this level, research focuses on detection of subjective sentences in a document containing a mixture of objective and subjective sentences.
- **Word Level:** At this level, usually the focus is to look for adjectives. However, verbs, adverbs and nouns could just as well convey a sense of subjectivity and carry opinions.
- **Feature based:** The common example to illustrate this level of SA is to consider a *review* containing positives and negatives of a product reflecting the reviewer liking and disliking of some of its features.

2.3. NLP Issues

We must always keep in mind that SA is a NLP problem and, consequently, many of the issues in NLP are also problems that must be addressed when dealing with SA problems. In [8], Bird, Loper and Klein address the need for a NLP toolkit that could be used efficiently in education, research and industry, and they provide the so-called Natural Language Toolkit (NLTK) as a platform for building *Python* programs to work with human language data. According to Liu [41], some of the

sub-problems that still are the object of further research attention by the NLP community are: (1) Coreference resolution; (2) Negation handling; (3) Word sense disambiguation; (4) Meaning extraction; and (5) Optimised parsing. Obviously, the proper resolution or any added improvements to any of these challenges above will have a positive effect in advancing the understanding of the SA problem. Dale [19] considers the following stages of analysis in processing natural language (in order of execution, from 1 to 5). The input to the process is *text* and the output is the *speaker's intended meaning*:

1. **Tokenization:** Converting a string of characters into words, symbols, sentences or other items conveying some sort of meaning, called *tokens.*
2. **Lexical analysis:** Usually deals with generating a *lexicon* and with applying *tagging* to the tokens already generated in the previous step. Most often, the tagging process is called *Parts of Speech (PoS) tagging.*
3. **Syntactic Analysis:** Provides a structure for every single sentence in a given text, including *parsing.*
4. **Semantic Analysis:** Aims to find the *literal meaning* of sentences or text.
5. **Pragmatic Analysis:** Attempts to determine the *meaning in context* of the sentence.

As *Tokenization* is rather mature, any affirmative contribution on the remaining 4 steps above would translate into improvements in the SA process.

2.4. Present Key Challenges in SA

The following list summarises the sub-topics that are considered to be key challenges for the SA discipline according to [12, 22, 29, 38, 39, 41, 51]: (a) Named Entity Recognition, (b) Anaphora Resolution, (c) Parsing, (d) Sarcasm & irony identification, (e) Subjectivity classification, (f) Polarity and graduality of opinions, (g) Use of abbreviations, poor spelling, punctuation or grammar, etc., (h) Sentiment (Opinion) Lexicon acquisition, (i) Negation handling, (j) Aspect-based & Comparative Sentiment Analysis, (k) Effective Classification of multiple opinions (aggregation).

3 SA Research Approaches

It is clear that the challenges present in the SA discipline are many. Liu [39, 41], Feldman [22], Pang & Lee [51] and Manning et al. [44], among others, consider that the future of the research on this area lies on exploring as many options as possible among the many challenges available and explore many sub-domains: customer reviews, politicians' blogs, marketing sites, company's opinion boards, etc. According to Cambria et al. [12] "mining opinions and sentiments from

natural language is challenging because it requires a deep understanding of the explicit and implicit, regular and irregular, and syntactical and semantic language rules. Sentiment analysis researchers struggle with NLP's unresolved problems: co-reference and anaphora resolution, negation handling, named-entity recognition, and word-sense disambiguation. Opinion mining is a very restricted NLP problem because the system only needs to understand the positive or negative sentiments of each sentence and the target entities or topics. Therefore, sentiment analysis is an opportunity for NLP researchers to make tangible progress on all fronts of NLP, and potentially have a huge practical impact."

Liu [39] claims that the main technical challenges for the multi-faceted problem that SA represents can be found among the following topics: (1) Object identification; (2) Feature extraction and synonym grouping; (3) Opinion orientation classification: and (4) Integration. The paragraph that follows will be utilised to define the topics of discussion.

(1) Yesterday, I bought a Nokia phone and my girlfriend bought a moto. (2) We called each other when we got home. (3) The voice on my phone was not clear. (4) The camera was good. (5) My girlfriend said that the sound on her phone was clear. (6) I wanted a phone with good voice quality. (7) So I was satisfied and returned the phone to BestBuy yesterday.

- **Object identification:** Discovering what the object is, about which an opinion has been provided. In the paragraph used as an example the objects are *Motorola*, abbreviated as 'moto' and *Nokia*. The noun 'BestBuy' corresponds to the name of the store; hence, it is *not* part of the comparison processed that the reviewer is providing and is not an object in terms of the products' comparison.
- **Feature extraction and synonym grouping:** The features commented on in the example are 'voice', 'sound' and 'camera'. According to Liu [39] "although there were attempts to solve this problem, it remains to be a major challenge". In addition, a feature can be referred to in different ways (i.e. 'voice' and 'sound' refer to the same feature in our example above.
- **Opinion orientation classification:** The objective of this task is to find out whether there is an opinion on a feature in a given sentence. If there is one, is it positive, negative or neutral? Here again, Liu [39] claims that the "existing approaches are based on supervised and unsupervised methods. One of the key issues is to identify opinion words and phrases, which are instrumental in sentiment analysis. The problem is that there are seemingly an unlimited number of expressions that people use to express opinions, and in a different domains they can be significantly different. Even in the same domain, the same word may indicate different opinions in different contexts".
- **Integration:** As the main objective of SA is to discover all quintuples (e_j, f_{jk}, so_{ijkl}, h_i, t_l), given as an input to an opinionated document, there is a need to integrate the above tasks, which is complex because the five pieces of information in the quintuple are to be matched. The quote just presented

corresponds to the definition of a direct opinion. In addition, Liu [38, 39] mentions that the fundamental problem here is that NLP techniques that still need improvement must be applied to resolve challenges like parsing, word-sense disambiguation, and co-reference resolution. In regard to the example provided, we observe the following issues: (i) understanding, depending on the context, what is meant by 'my phone' and 'her phone' in sentences (3) and (5); (ii) to which phone does the camera belong to?; (iii) in (4), "The camera was good", we do not have a pronoun and neither the sentence mentions a specific phone. According to Liu [38, 39] these are classical examples of **co-reference resolution**, the latter being a problem that despite the fact that the NLP community has studied it for a long time, it is still not accurately resolved.

In the following section, we will briefly describe the main approaches applied in the SA discipline.

3.1. Machine Learning

Machine Learning (ML) has played a fundamental role in NLP, and therefore it is extensively applied in the field of SA. Kumar & Sebastian [51] claim that "most researchers have defined the SA problem as essentially a *text classification problem* and machine learning techniques have proved their dexterity in resolving the sentiment analysis tasks". The main two learning approaches in ML are:

- *Supervised Learning:* "Machine Learning classification relies on the training set used, the available literature reports detail classifiers with high accuracy, but they are often tested on only one kind of sentiment source, mostly movie review, thus limiting the performance indication in more general cases" [36].
- *Unsupervised Learning:* "Use sentiment driven pattern to obtain labels for words and phrases" [36].

3.2. Fuzzy Sets/Logic Contribution to NLP and SA

In [20] Dzogang et al. go beyond the most common motivation for SA – to automate the classification of social media opinions, books reviews, film rankings, etc. – and attempt to "review methods taking account of *intrinsic psychological models components of graduality* as well as extrinsic components issued from computational intelligence approaches. In particular, beyond psychological models of sentiments that define affective states as multidimensional vectors in affective continuous spaces, we identify three components of graduality, namely composition or blending, intensity and inheritance". Basically, these authors perform a deeper analysis of the origins of emotions and sentiments and investigate among the technical tools at hand, which are closer to the nature of the problem being analysed. They highlight the nature of emotions and their graduality and fuzziness, and claim that "...it must be underlined that some

appraisal based approaches make use of graduality through fuzzy inference and fuzzy aggregation for processing affective mechanisms ambiguity and imprecision...". The caveat they make, though, is that these so-called appraisal-based methods are *not* great at *sentiment discrimination*. Nevertheless, these arguments make us think that even if not great for deep psychological and physiological analysis, the fact that fuzzy sets can be used successfully to model the ambiguity and imprecision of *affective states*, will make them an acceptable tool for modelling sentiments.

There have been some successful applications of Fuzzy Sets/Logic theory to both NLP and SA. In the literature, the use of Fuzzy Logic is found in *Anaphora Resolution* – given an expression S_i, its interpretation depends upon another expression S_j in context – in the work of Witte & Bergler [71]. Analysis of affect in text using fuzzy sets by means of the concept of *fuzzy semantic typing* can be found in [57]. Named Entity recognition has been addressed as well using fuzzy techniques in [32]. In summary, there seems to be evidence that a fuzzy approach could be applied in a number of sub-topics of sentiment analysis, and that some researchers are considering this research avenue worthy of further exploration.

3.3. Sentic Computing

As defined in Cambria & Hussein [11], "sentic computing is a multi-disciplinary approach to sentiment analysis that exploits both computers and social sciences to better recognise, interpret, and process opinions and sentiments over the Web. The approach specifically brings together lessons from both affective computing and common sense computing because in the field of opinion mining, not only common sense knowledge, but also emotional knowledge is important to grasp both the cognitive and affective information (termed semantics and sentics) associated with natural language opinions and sentiments". In a way, this approach is fairly new, and at this point it is uncertain whether this research avenue will be appealing enough to gain the attention of researchers in the near future.

4 State of the Art and Research time-line in SA

As SA sits at the confluence of several sub-disciplines – fundamentally NLP/ Computational Linguistics, Text Data Mining and AI – its origins cannot be tracked down to a specific date, but rather to a collection of moments in time that defines progress in the sub-areas mentioned above. Most of the important work in syntax and formal languages is attributed to Chomsky [15, 16] and his revolutionary work that occurred between the late 1950s and the late 1960s. Chomsky laid down the bases for modern languages & grammar theory, syntax theory and also for the concept of transformational grammar. In turn, these advances led to improvement in the automatic processing of syntax and grammars

by using productions and recursive calls. Parsing and Compiling theory, that today is taken for granted by many, was positively influenced by the work of Chomsky and others that followed. By 1872, Charles Darwin had already published his work '*The Expression of the Emotions in Man and Animals*', where he mainly addressed aspects of behaviours that are genetically determined. This is probably the first work related to determining the origin and characteristics of emotions. Many other authors in the Psychology camp have since then augmented the knowledge we have today about emotions as a fundamental human trait. This section aims to show how the SA discipline has evolved chronologically since 1970.

4.1. 1970 through 1979

The 1970s witnessed a lot of progress related to the refining of syntactic techniques and the generation of more advanced parsing and compiling ideas, together these resulted in more efficient algorithms. Making sure the proper parsing tree is generated is a fundamental step before more complex tasks can be started. In this arena, the work by Hopcroft [28] and Aho & Ullman [3, 4] is decisive, despite the fact that it concentrates in programming languages instead of natural languages, this is because it either brought rigour and formality to the parsing techniques or presented the works of others in a digestible way.

4.2. 1980 through 1989

It is possible to argue that no remarkable work applicable to SA was done until the 1980s. The work of Banfield [7] seems to have been instrumental in proposing the use of subjective and objective sentences as indicators, as well as in searching the text by means of simple queries. In 1983, Winograd [70] published work on language as a cognitive process that started a wave of further research into the cognitive aspects of emotions. In 1987, Ortony et al. published a landmark article [49] along with his 1988 book [50] that are a common reference to building an affective lexicon and have become important parts of the puzzle in SA.

4.3. 1990 through 1999

In 1990, Miller et al. gave to the research community *WordNet,* "a useful tool for computational linguistics and natural language processing" [46]. In 1991, Miller and Charles [47] discussed what they called the 'contextual correlates of semantic similarity', advancing the field even more when they researched the basis of semantic similarities in a given context, which would prove to be instrumental in the progress of the SA discipline. In 1994, the concept of Part-of-Speech was pushed forward and new ideas were put on the table in order to improve methods for part-of-speech tagging [10]. This technique would become key to properly identifying the different parts and component of sentences in order to build algorithms that would focus on extracting meaning and orientation out of

sentences. The use of statistical methods in Natural Language Parsing, Linguistics and more generically into NLP as a whole was brought by [1, 13, 45]. 'Text-Based Intelligent System' that focuses on the concept of directionality (e.g. is the agent in favour of, neutral, or opposed to a given event?) was presented in [27]. The authors claim that with their method, sentence meaning is mapped to a metaphorical model that is self-contained as no external references are required to find the directionality of a given sentence or paragraph. In [66, 67] the concept of extraction of subjective words is articulated properly, to the point that the author even proposed a method to determine the beliefs of the characters in the narrative, once the subjective terms have been identified. The possibility of predicting the semantic orientation of adjectives, which carry an important weight on determining the semantic orientation of phrases, was addressed in a landmark paper published in 1997 [25]. Towards the end of the 1990s some researchers started looking at the use of fuzzy reasoning in SA [35].

4.4. 2000 through 2014

From the start of 2000, the SA discipline starts an accelerated development process. Up to 2008, Pang & Lee's book [51] was considered the most complete work in the area. In [39], Liu addressed the complexities and multiple faces that this discipline can show, with the most updated version of the discipline later presented and published in 2012 [41]. We have mentioned before the importance of WordNet. In 2006, Esuli & Sebastiani published a lexical resource specific to SA [21], SentiWordNet, that assigns to each synset – sets of synonyms for groups of English words – of WordNet three sentiment scores: positivity, negativity, or objectivity. In 2013, Feldman attempted to bring SA to the front-page with their work [22], which has created a lot or additional attention in the research community. New techniques have been showing up steadily, with Cambria and collaborators [11, 12] proclaiming that new techniques, like the so-called Sentic Computing, could offer some new lights into the SA problem.

Most of the tools utilised in SA to resolve subjectivity identification and polarity extraction are based on some sort of Machine Learning technique. Most of the literature and bench marking established is based in Supervised Learning [36, 52]. However, some Unsupervised Learning techniques have been very successful as well, as it is unsupervised technique based on the PMI-IR algorithm that is used to estimate the semantic orientation of a phrase by measuring the similarity of pairs of words or phrases [60]. Alternative methods have been proposed, like the *Bootstrapping* Method for Building *Subjectivity Lexicons* for Languages with Scarce Resources [6], the techniques for generating a quality lexicon [59], the recognition of contextual polarity in [69] and the gradability of subjective sentences based on adjective orientation [26]. In all these cases, the focus of the research is at the sentence/phrase level. When one attempts to establish the orientation of the sentiment in a document, one is faced with the need for summarising somehow all the contents. Techniques to effectively summarise

opinions are addressed in [29]. On the same topic, Suanmali et al. [56] proposed a fuzzy logic based method for improving the summarisation of text, while Liu in [39] stressed one more time the important aspect of SA/OM of determining subjectivity by differentiating between *objective* and *subjective* sentences. Somehow less common, some researchers have looked into the possibility of applying semi-supervised learning methods like the one used for opinion summarization and classification for online product reviews [17]. Some hybrid methods started to flourish in the late 2000s. In [42], Liu & Tsou focus on using Supervised Machine Learning methods jointly with a qualified sentiment lexicon, while on the same token, Wiebe & Riloff [68] proposed a method to do simultaneously subjectivity analysis and information extraction based on their claim that by doing one enables the other one, somehow. We tend to believe that the future of research in SA is probably bound to lean towards hybrid methods.

In 2001, Subasic & Huettner [57] released the most important contribution we are aware of to the use of Fuzzy Sets/Logic principles in SA. Since then, others have followed their footsteps, but certainly, Fuzzy Sets Theory has been so far a bit of an outsider in the research field for SA. In 2010, Dzogang et al. published their influential article [20] in which, to be successful in SA and related disciplines, they depict how it is necessary, to understand two key factors: (a) the inclusion of the use of some fundamental emotion structure coming from the world of Psychology and (b) the further looking into the potential fitness of fuzzy sets to model *graduality* in a proper way. In 2011, the interest in fuzzy methods flares up again, with van der Heide et al. [61] addressing the topic of modelling affect through applying fuzzy logic; and by Kar & Mandalof's [33] study on using fuzzy logic for determining the strength of pre-established opinions in web reviews. In 2013 we see a cluster of interest in applying *fuzzy techniques* to SA [31], either to assist in resolving ambiguity in text as in [46], in using fuzzy sets to model sentiment classification at the sentence-level (for the Chinese languages) in [23], to drive the semantic understanding of general linguistic items [34], or a *fuzzy linguistic hedges*-based method for opinion mining in online user product reviews [18]. At the moment of writing this paper, this one seems to be most recent contribution of fuzzy methods to SA/OM.

For some time, the focus of identifying subjectivity by analysing adjectives was the standard. However, verbs and nouns are capable as well of conveying emotions and sentiments. The analysis of verbs to create a *verb lexicon* that would aid with establishing sentiments in opinion mining applications is explored in [43], while a combination of adjective, verbs and adverbs in an effort to improve subjectivity classification is presented in [42]. Using all of these three components of part-of-speech simultaneously may contribute to a more reliable subjectivity classification process. In 2012, Nguyen et al. [72] combine the new features with conventional ones obtained from already established research lines of work, and as such their method can somehow be considered as a *hybrid* approach. A system combining together concept-level sentiment analysis and opinion mining lexicon-

based and learning-based approaches is proposed in [48]. Dealing with metaphoric language is hard, and some researchers have spent some time suggesting how to address the problem [9, 53]. Recognising irony and sarcasm are tough topics as well, and some work has been done in this area [62]. These topics are very important when dealing with opinions, particularly when the text being analysed has politics content. A departure from most of the methods we have presented above, and an alternative possibility that some researchers are currently considering in SA is that of addressing the space of entity-related opinion detection and sentiment ontology trees [64].

Lotfi Zadeh put forward the concept of Precisiated Natural Language (PNL) for describing perceptions [73]. What Zadeh does mean by precisiated? As per [73], "...precisiated in the sense of making it possible to treat propositions drawn from a natural language as objects of computation?". Despite the fact that the idea is very tempting, not too much additional research has been conducted in PNL, as far as SA goes, although we believe that it puts forwards a concept that could possibly blossom in the SA arena.

4.5. Bibliometrics and References Distribution

This section attempts to 'see' the portion of work in SA that has been carried out using supervised machine learning methods versus those based on fuzzy sets theory. The years of publishing as well as the country where the work has been conducted are presented, for which a simple search based on the keywords *machine learning* or *fuzzy sets* in the larger context of SA was used. Although it is true that this search will exclude articles written before the term SA saw the light and became fully accepted – later on in the middle 2000s – it will, however, provide us with good indicators of the numbers of publications on the topic since the middle 2000s as well as the country where the research initiative was conducted and published. Likewise, articles indexed by other sub-topics of SA, like subjectivity classification or identification, polarity extraction, etc. could have been partially excluded, too. Nevertheless, we believe that based on the review of the literature carried out, the potential exclusion of those articles will not create a deviation in the results already obtained. For consistency, we are including only articles written in the English language. The sources used for the search is **Scopus,** a large abstract and citation database of peer-reviewed literature by Elsevier B.V., and the **Web of Knowledge** (WoK), an academic citation indexing and search service by Thomson Reuters. However, as the results obtained using both databases are equivalent, only the results obtained using Scopus are included below:

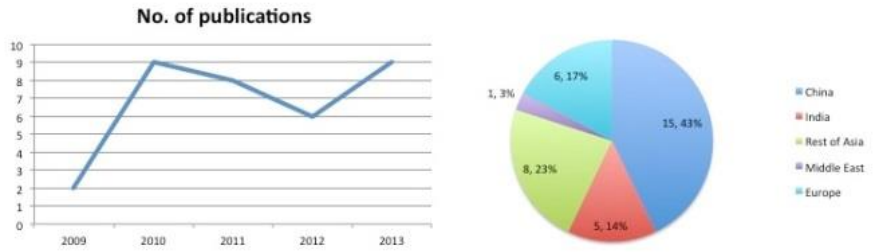


Figure 1
Research using keywords *Fuzzy Sets* and *Sentiment Analysis*

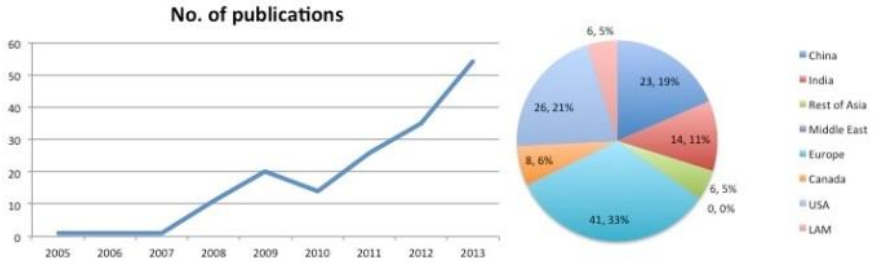


Figure 2
Research using keywords *Machine Learning* and *Sentiment Analysis*

As we take a closer look to the graphics provided above, we immediately note two characteristics:

1. Research in SA using Machine Learning (ML) techniques depicts a curve that shows primarily a clear trend towards sustained growth, whilst the one representing the utilisation of Fuzzy Sets (FS) shows no material growth and a clear lesser number of publications.

2. When we look at the countries where the articles have been published we notice that the utilisation of ML techniques – more specifically Supervised Learning – is high across the board, showing the USA, China, India and Europe as clear leaders. If we observe the utilisation of FS to model the SA problem, the first thing that we realise is that it is mainly an affair mostly pursued in China, India and Europe. In addition, research using ML techniques started earlier as well.

Why is this? One dares to venture that it is perhaps because a significant number of the researchers with a Computer Science/Mathematics educational background involved nowadays with researching SA/OM come from the Text Data Mining/Processing and NLP fields where the use of statistical methods have been a well-established tool for some time now. Hence, it would be natural to export the same knowledge, skills and techniques and apply them to a new domain that, nevertheless, is somehow related. Or is it perhaps that the utilisation of FS

techniques have been proven to be not successful? If so, how has *success* been measured? It is interesting to see that in the cases of China and India, and to a certain extent Europe as well, research efforts are present in both camps (ML and FS). However, for the USA, at least for the period of time chosen and the search keywords and data sources utilised, the focus is clearly on the ML camp, despite the fact that one of the most influential papers supporting the use of FS was written in 2001 in the USA [57]. One may think as well that two of the most reputed researches in SA/OM, Bing Liu and Bo Pang, have made ML their fundamental tool. For instance, Liu's early research was in data mining, Web mining and *machine learning*, fields in which he published abundantly (as appears in his Biography in [41]). At this point, we can only draw some conjectures, but we believe that it will not be completely nonsensical to think that the primary research interest of some authors may have migrated to this newer field of SA/OM. Moreover, the use of statistical techniques in NLP/Computational Linguistics are common and have been aptly utilised since at least 1996 [13, 14].

Fuzzy Sets have been used extensively to model uncertainty and ambiguity, traits that are undoubtedly inherent to Natural Languages and as a consequence part of the challenges inherited by SA/OM. Somehow, Fuzzy Sets may be seen as alien to the community of Linguistics, with the exception perhaps of the utilisation of Fuzzy Grammars [4]. We conclude then that there are a number of potential reasons that could explain why the use of Supervised Machine Learning techniques has been favoured. However, so far, we have not been able to find hard evidence that the utilisation of Fuzzy Sets, perhaps in combination with some other syntactic techniques and even Unsupervised Learning tools, could not yield favourable results. Bing Liu, one of the most world-wide recognised experts in the area of SA/OM and one the researchers that has attempted to push the limits in the field of SA, has mentioned that "we probably relied too much on Machine Learning" [38, 39, 41], when referring to how limited our understanding is about the SA problem, despite the recent progress that has been achieved.

As a result of the discussion presented in this section and other arguments to be presented in the next sections, we do believe there is merit in investigating further the potential use of *Fuzzy Sets* in the *Sentiment Analysis problem*; especially in the research sub-areas of *subjectivity, polarity and graduality* [20, 32, 61, 73, 74].

5 Potential Future Research Path in SA

Traditionally in ML we think of unsupervised, semi-supervised or supervised learning. Supervised learning, as we well know, relies heavily on training, which implies counting with the adequate data sets. To avoid, if possible, having to count on prior data for training purposes would also somehow, disqualify as well until certain extent the use of semi-supervised methods. In the context of SA, an

unsupervised strategy would rather "measure how far a word is inclined towards positive and negative" [63]. Somehow, this makes out of an unsupervised method a *semantic orientation approach* or a *lexicon-based method*.

Ultimately, the problem of SA/OM is basically a NLU/NLP problem with emphasis in finding when a sentence reveals an opinion – as opposed to a fact – and extracting the polarity of the opinion (Negative, Positive, Neutral, etc.). Being successful at determining if a sentence is objective or subjective will predetermine by far how accurate the establishing of polarity on subjectivity will be. Banea [6] claims that "the problem of distinguishing subjective versus objective instances has often proven to be more difficult than subsequent polarity classification, so improvements in subjectivity classification promise to positively impact sentiment classification". This translates into the idea that getting the differentiation between objective and subjective sentences correctly would guide one through the right path.

Kanaga [32], in discussing ideas presented by Lotfi Zadeh in [74], says: "The semantics of natural languages and information analysis is best handled by the epistemic facet of Fuzzy Logic. In the epistemic facet, natural language is viewed as a system for describing perceptions and an important branch of the same is possibility theory and computational theory of perceptions". Hence, is it worthy to take a new look to Fuzzy Sets/Logic as a potential effective tool in SA/OM? Would it be helpful as well to stick to a strong psychological foundation of emotions and feelings to assist us in modelling the problem at hand? The recipe that we would like to pursue will include a solid foundation of emotions theory, unsupervised learning (semantic approach) and *fuzzy sets/logic* as fundamental components of a *hybrid approach* towards SA/OM.

5.1. Where shall We Go from Here?

Based on the information, references and discussions shown above, possible research directions that deviate from the current most followed path – Supervised Machine Learning – are suggested:

1. In SA the most utilised approach is text classification relying heavily on ML techniques; especially Support Vector Machine (SVM) and Naïve Bayes
2. Fuzzy Sets and Fuzzy Logic have been used to a lesser extent and the literature about it is rather reduced when compared to (1) above.
3. If determining subjectivity properly contributes to a more accurate polarity identification, then it is worth it to spend additional time on the topic, before attempting to conclude on polarity.
4. One of the main objections of the use of Fuzzy Logic/Sets in SA is [5] as follows: "...we can show that while the fuzzy models of emotion perform well for a series of cases that fit the described patterns, they remain weak at the time

of acquiring, combining and using new information". However, we believe that some of these shortfalls can be minimised by combining together fuzzy methods and some semantic and linguistic techniques. See, for example, the progress reported on acquiring new information in a given lexicon in Kruse et al. [35] (using neuro-fuzzy modelling) and Hüllermeier [30] (applying learning fuzzy rules).

5. Hatzivassiloglou et al. [25, 26] proposed a methodology to predict the semantic orientation of adjectives. This strategy – so it seems – could be extended to nouns, adverbs, and verbs, as discussed in [58]. As such, predicting the semantic orientation of certain parts of speech can greatly help on suggesting the semantic orientation of sentences and documents.
6. Grammatical dependencies may play a significant role in a proper understanding of a sentence. As quoted from [55]: "In any sentence, words are arranged in a proper sequence to communicate information. The complete meaning of a sentence is not only determined by the meaning of words, but also by the pattern in which words are arranged".
7. Supervised Learning (SL) has proven to be a strong classification technique. However, SL will depend enormously on the availability of training data. In a way, we would like to move towards a system that is less dependent on *pre-existing annotated data*. To rely more on the richness of fuzzy sets as a modelling apparatus – perhaps using *hedges* as well – and in syntactic analysis techniques.

5.2. Future Research Questions to be Addressed

The fundamental research question we are posing is whether a hybrid approach, combining together the *psychological foundations of emotions, linguistics tools, unsupervised learning and fuzzy sets/logic,* is well equipped to model subjectivity and polarity determination in SA/OM. By well equipped we mean for it to be capable of delivering the same or better results than the most commonly used techniques whilst remaining faithful to the original sources of emotions and to modelling tools that are akin to the inherent ambiguity present in natural languages. For simplicity, we will split this question into four sub-questions:

1. Is *Unsupervised Learning* capable of delivering similar accuracy to the one provided by *Supervised Learning* techniques in the determination of subjectivity in Sentiment Analysis?
2. Is *Fuzzy Reasoning* adequate to support subjectivity determination and to model polarity in SA by introducing gradualness (graduality)?
3. Can we represent with more accuracy sentiments expressed in natural languages by using as a bedrock concepts of emotions that originate in psychology [50]? Can we get closer to the heart of the matter by using this foundation and looking into the *cognitive model of emotions* or is doing so futile?

4. Is our model flexible enough to attempt to accommodate afterwards the recognition and understanding of metaphors? Can this be achieved without the use of supervised or semi-supervised machine learning approaches?

Is there going to be synergy among all these elements? Currently, most of research performed has been conducted using Supervised Methods in Machine Learning (mostly SVM, Naïve Bayes and others). Hence, our comparison base will be defined by the results already obtained using the latter methods. In a way, we must try to determine whether good results in the sub-questions will have an aggregated *positive* effect when all of them get combined together. The key performance indicators that will be chosen for the comparison will be decisive in understanding how successful the research journey has been.

Conclusions

In this article we have attempted to cover a few fundamental aspects related to Sentiment Analysis/Opinion Mining. Firstly, we wanted to address the basics of the topic and its main challenges. Secondly, we have provided a chronological account of the research that has been conducted to date as well as some bibliometric aspects showing a distribution of articles published based either on *machine learning* or *fuzzy sets* as the main tools to model the SA/OM sub-problem. Finally, we have ventured to suggest that there is not enough evidence to justify abandoning other potential research paths that may rely on hybrid mechanisms combining a number of foundations, strategies and techniques.

References

[1] S. Abney. *Statistical Methods and Linguistics*. In The Balancing Act: Combining Symbolic and Statistical Approaches to Language. The MIT Press, 1996

[2] A. V. Aho and J. D. Ullman. *The Theory of Parsing, Translation and Compiling, Vol. I: Parsing*. Prentice Hall, 1972

[3] A. V. Aho and J. D. Ullman. *The Theory of Parsing, Translation and Compiling, Vol. II: Compiling*. Prentice Hall, 1973

[4] O. Appel. *Fuzzy Grammars: What They Are and What Their Potential Applications Could Be*. Unpublished; Final Assignment for the course Applied Computational Intelligence at De Montfort University, UK, 2012

[5] A. Balahur. *Methods and Resources for Sentiment Analysis in Multilingual Documents of Different Text Types*. PhD thesis, Department of Software and Computing Systems, University of Alicante (Spain), 2011

[6] C. Banea, R. Mihalcea, and J. Wiebe. *A Bootstrapping Method for Building Subjectivity Lexicons for Languages with Scarce Resources*. In LREC 2008, 2764-2767, 2008

[7] A. Banfield. *Unspeakable Sentences: Narration and Representation in the Language of Fiction*. Routledge and Kegan Paul, 1982

[8] S. Bird, E. Loper, and E. Klein. *Natural Language Processing with Python*. O'Reilly Media Inc, 2009

[9] D. Bollegala and E. Shutova. *Metaphor Interpretation Using Paraphrases Extracted from the Web*. PLoS ONE e74304, 8(9):1614-1617, 2013

[10] E. Brill. *Some Advances in Transformation-based Part of Speech Tagging*. In AAAI '94, 722-727, 1994

[11] E. Cambria and A. Hussain. *Sentic Computing: Techniques, Tools and Applications*. Springer Briefs in Cognitive Computation, 2012

[12] E. Cambria, B. Schuller, Y. Xia, and C. Havasi. *New Avenues in Opinion Mining and Sentiment Analysis*. IEEE Intelligent Systems, 28(2):15-21, 2013

[13] E. Charniak. *Statistical Language Learning*. The MIT Press, 1996

[14] E. Charniak. *Statistical Techniques for Natural Language Parsing*. AI Magazine, 18(4):33-44, 1997

[15] N. Chomsky. *Syntactic Structures*. Mouton de Gruyter (formerly Mouton, The Hague), 2nd revised (2002) edition, 1957 (1st edition)

[16] N. Chomsky. *Aspects of the Theory of Syntax*. The MIT Press, 1969

[17] M. K. Dalal and M. A. Zaveri. *Semisupervised Learning-based Opinion Summarization and Classification for Online Product Reviews*. Appl. Comp. Intell. Soft Comp., 2013(Article ID 910706), 2013

[18] M. K. Dalal and M. A. Zaveri. *Opinion Mining from Online User Reviews Using Fuzzy Linguistic Hedges.* Appl. Comp. Intell. Soft Comp., 2014:1-9, 2014

[19] R. Dale. *Classical Approaches to Natural Language Processing*. In Handbook of Natural Language Processing, Chapter I, pp. 3-7, 2010

[20] F. Dzogang, M.-J. Lesot, M. Rifqi, and B. Bouchon-Meunier. *Expressions of Graduality for Sentiments Analysis - A Survey*. In FUZZ-IEEE2010, 1-7, 2010

[21] A. Esuli and F. Sebastiani. *SentiWordNet: a High-Coverage Lexical Resource for Opinion Mining*. Technical Report Institute of Information Science and Technologies of the Italian National Research Council, 2006

[22] R. Feldman. *Techniques and Applications for Sentiment Analysis*. Communications of the ACM, 56(4):82-89, 2013

[23] G. Fu and X. Wang. *Chinese Sentence-Level Sentiment Classification Based on Fuzzy Sets*. In COLING2010, 312-319, 2010

[24] T. Galli, F. Chiclana, J. Carter and Helge Janicke: *Modelling Execution Tracing Quality by Means of Type-1 Fuzzy Logic.* Acta Polytechnica Hungarica, Volume 10, Issue 8, pp. 49-67, 2013

[25] V. Hatzivassiloglou and K. R. McKeown. *Predicting the Semantic Orientation of Adjectives*. In ACL1997, 174-181, 1997

[26] V. Hatzivassiloglou and J. M. Wiebe. *Effects of Adjective Orientation and Gradability on Sentence Subjectivity.* In ACL2000

[27] M. A. Hearst. *Direction-based Text Interpretation as an Information Access Refinement*. Text-Based Intelligent Systems, 1-13, 1992

[28] J. Hopcroft and J. D. Ullman. *Introduction to Automata Theory, Languages, and Computation*. Addison Wesley, 1979

[29] M. Hu and B. Liu. *Mining and Summarizing Customer Reviews*. In ACM SIGKDD, 22-25, 2004

[30] E. Hüllermeier. *Fuzzy Methods in Machine Learning and Data Mining – Status and Prospects*. Fuzzy Sets and Systems 156(3):387-406, 2005

[31] S. Jusoh and H. M. Alfawareh. *Applying Fuzzy Sets for Opinion Mining*. In (ICCAT2013, 1-5, 2013

[32] V. R. Kanagavalli and K. Raja. *Detecting and Resolving Spatial Ambiguity in Text Using Named Entity Extraction and Self Learning Fuzzy Logic Techniques*. CoRR, abs/1303.0445, 2013

[33] A. Kar and D. P. Mandal. *Finding Opinion Strength Using Fuzzy Logic on Web Reviews*. International Journal of Engineering and Industries, 2, 2011

[34] R. Khoury, F. Karray, Yu Sun, M. Kamel, and O. Basir. *Semantic Understanding of General Linguistic Items by Means of Fuzzy Set Theory*. IEEE Transactions on Fuzzy Systems, 15(5):757-771, 2007

[35] R. Kruse, D. Nauck, and C. Borgelt. *Data Mining with Fuzzy Methods - Status and Perspectives*. In EUFIT99, 1999

[36] A. Kumar and T. M. Sebastian. *Sentiment Analysis: A Perspective on Its Past, Present and Future*. International Journal of Intelligent Systems and Applications, 4(10):1-14, 2012

[37] A. Kumar and T. M. Sebastian. *Machine Learning assisted Sentiment Analysis*. In International Conference on Computer Science and Engineering, 123-130, 2012

[38] B. Liu. *Sentiment Analysis and Subjectivity*. In Handbook of Natural Language Processing, Chapter 26, pp. 627-666, Chapman & Hall CRC, 2010

[39] B. Liu. *Sentiment Analysis: A Multifaceted Problem*. IEEE Intelligent Systems, 25(3):76-80, 2010

[40] B. Liu. *Sentiment Analysis Tutorial, given at AAAI-2011*

[41] B. Liu. *Sentiment Analysis and Opinion Mining*. Morgan and Claypool Publishers: Synthesis Lectures on Human Language Technologies, 2012

[42] B. Lu and B. K. Tsou. *Combining a Large Sentiment Lexicon and Machine Learning for Subjectivity Classification*. In Ninth IEEE International Conference on Machine Learning and Cybernetics, 3311-3316, 2010

[43] I. Maks and P. Vossen. *A Verb Lexicon Model for Deep Sentiment Analysis and Opinion Mining Applications*. In 2nd Workshop on Computational Approaches to Subjectivity and Sentiment Analysis, 10-18, 2011

[44] C. D. Manning, P. Raghavan, and H. Schütze. *Introduction to Information Retrieval*. Cambridge University Press, 2008

[45] C. D. Manning and H. Schütze. *Foundations of Statistical Natural Language Processing*. The MIT Press, 1999

[46] G. A. Miller, R. Beckwith, C. Fellbaum, D. Gross, and K. J. Miller. *Introduction to WordNet: an On-Line Lexical Database*. International Journal of Lexicography, 3(4):235-244, 1990

[47] G. Miller and W. Charles. *Contextual Correlates of Semantic Similarity*. Language and Cognitive Processes, 6(1):1-28, 1991

[48] A. Mudinas, D. Zhang, and M. Levene. *Combining Lexicon and Learning-based Approaches for Concept-Level Sentiment Analysis*. In 2012 International Workshop on Issues of Sentiment Discovery and Opinion Mining, 51-58

[49] A. Ortony, G. L. Clore, and M. A. Foss. *The Psychological Foundations of the Affective Lexicon*. Journal of Personality and Social Psychology, 53:751-766, 1987

[50] A. Ortony, G. L. Clore, and A. Collins. *The Cognitive Structure of Emotions.* Cambridge University Press, 1988

[51] B. Pang and L. Lee. *Opinion Mining and Sentiment Analysis*. Foundations and Trends in Information Retrieval 2 (1-2), 1-135, 2008

[52] B. Pang, L. Lee, and S. Vaithyanathan. *Thumbs up? Sentiment Classification using Machine Learning Techniques.* In Conference on Empirical Methods in Natural Language Processing, 10:79-86, 2002

[53] V. Rentoumi, G. A. Vouros, V. Karkaletsis, and A. Moser. *Investigating Metaphorical Language in Sentiment Analysis: A Sense-to-Sentiment Perspective.* ACM Trans. Speech Lang. Process., 9(3):6:1-6:31, 2012

[54] E. Shutova. *Models of Metaphor in NLP*. In 48th Annual Meeting of the Association for Computational Linguistics, 688-697, 2010

[55] R. Srivastaval, MPS Bhatia, H. K. Srivastava, and C. P. Sahu. *Effects of Adjective Orientation and Gradability on Sentence Subjectivity*. In IEEE International Conference on Computer & Communication Technology, 768-775, 2010

[56] L. Suanmali, N. Salim, and M. S. Binwahlan. *Fuzzy Logic-based Method*

for Improving Text Summarization. International Journal of Computer Science and Information Security, 2(1), 2009

[57] P. Subasic and A. Huettner. *Affect Analysis of Text Using Fuzzy Semantic Typing*. IEEE Transactions on Fuzzy Systems, 9(4):483-496, 2001

[58] V. S. Subrahmanian and D. Reforgiato Recupero. *AVA: Adjective-Verb-Adverb Combinations for Sentiment Analysis*. IEEE Intelligent Systems, 23(4):43-50, 2008

[59] M. Taboada, J. Brooke, M. Tofiloski, K. Voll, and M. Stede. *Lexicon-based Methods for Sentiment Analysis*. Computational Linguistics, 37(2):267-307, 2011

[60] P. D. Turney. *Thumbs Up or Thumbs Down? Semantic Orientation Applied to Unsupervised Classification Reviews*. In 40th Annual Meeting of the Association for Computational Linguistics, 417-424, 2002

[61] A. van der Heide, D. Sánchez, and G. Triviño. *Computational Models of Affect and Fuzzy Logic*. In EUSFLAT 2011, 620-627, 2011

[62] A. A. Vanin, L. A. de Freitas, R. Vieira, and M. N. Bochernitsan. *Some Clues on Irony Detection in Tweets*. In 22nd International Conference on World Wide Web Companion, 635-636, 2013

[63] G. Vinodhini and RM. Chandrasekaran. *Sentiment Analysis and Opinion Mining: A Survey*. International Journal of Advanced Research in Computer Science and Software Engineering, 2(6):282-292, 2012

[64] W. Wei. *Analyzing Text Data for Opinion Mining*. In 16th International Conference on Natural Language Processing and Information Systems, NLDB'11, 330-335, 2011

[65] A. Weichselbraun, S. Gindl, and A. Scharl. *Extracting and Grounding Contextualized Sentiment Lexicons*. IEEE Intelligent Systems, 28(2):39-46, 2013

[66] J. Wiebe. *Identifying Subjective Characters in Narrative*. In International Conference on Computational Linguistics, COLING '90, 1990

[67] J. Wiebe. *Tracking Point of View in Narrative*. Comput. Linguist., 20(2):233-287, 1994

[68] J. Wiebe and E. Riloff. *Finding Mutual Benefit between Subjectivity Analysis and Information Extraction*. IEEE Transactions on Affective Computing, 2(4):175-191, 2011

[69] T. Wilson, J. Wiebe, and P. Hoffmann. *Recognizing Contextual Polarity: An Exploration of Features for Phrase-Level Sentiment Analysis*. Comput. Linguist., 35(3):399-433, 2009

[70] T. Winograd. *Language as a Cognitive Process, Volume I: Syntax*. Addison-Wesley, 1983

[71] W. and S. Bergler. *Fuzzy Coreference Resolution for Summarization*. In 2003 International Symposium on Reference Resolution and Its Applications to Question Answering and Summarization (ARQAS), 43-50, 2003

[72] H. Nguyen, T. Xuan, A. Cuong Le, and L. M. Nguyen. *Linguistic features for subjectivity classification*. In International Conference on Asian Language Processing (IALP), 17-20, 2012

[73] L. A. Zadeh. *Precisiated Natural Language (PNL).* AI Magazine, 25(3):74-91, 2004

[74] L. A. Zadeh. *Is there a Need for Fuzzy Logic?* Information Sciences, 178:2751-2779, 2008

The Influence of Leader-Member Communication on Organizational Commitment in a Central European Hospital

Valentin Kónya[1], Leposava Grubić-Nešić[2], Dejan Matić[2]

[1] Faculty of Economics, University of Novi Sad
Segedinski put 9-11, 24000 Subotica, Serbia

[2] Department of Industrial Engineering and Management, Faculty of Technical Sciences, University of Novi Sad
Trg Dositeja Obradovića 6, 21000 Novi Sad, Serbia

valentink@uns.ac.rs, nesle@uns.ac.rs, dejan.matic@uns.ac.rs

Abstract: It is clear that accomplishing the goals of leadership can be easier and the process is far more successful, if both leaders and their coworkers are fully committed to their organizations, especially in stressful, demanding working environments with high levels of responsibility, such as in the health care sector. The subject of this paper is to investigate the influence of social exchange between leaders and their followers on the organizational commitment of employees. The main research question, to which this paper should give an answer, is whether social exchange between leaders and coworkers have significant influence on the organizational commitment of employees in a large hospital? The research was conducted in a Central European hospital with over 1000 employees. A total number of 359 valid questionnaires were returned. Two questionnaires were used for the research: the quality of leader-member exchange was measured with LMX-7 questionnaire for members and organizational commitment was measured with the Organizational Commitment Questionnaire – OCQ. Identical questionnaires were given to every employee included in the research, so no leader-member distinction was made among them, as the focus was on how employees as members evaluate their exchange with their leaders generally and what impact it has on their commitment. The area of interest of this research was not on specific leader-member relationships. Employees were asked to assess the quality of the exchange relationship with all of their leaders generally. The findings revealed that leader-member communication and organizational commitment have positive connectivity in a non-western environment. It was also revealed that leader-member communication and value commitment have much stronger relationship than the relationship between leader-member communication and commitment to stay.

Keywords: leader-member communication; LMX; organizational commitment; hospital; large organization; Central Europe

1 Introduction

The twentieth century is characterized with evolutionary and revolutionary changes that changed the way organizations, humans and whole societies act and function. These changes led to a totally new and different approach to the functioning of organizations, in the form of leadership. The phenomenon of leadership is as old as the civilization itself and scholars around the world first started to investigate leadership or some of its forms, about 100 years ago. This means that it is certain that leadership existed since civilizations and organized groups of people emerged. The development of civilization and the society is unthinkable and could not be possible without leadership and leaders. However, scholars are showing extensive interest only in the last 100 years for researching leadership, its characteristics, dimensions and variables and for the development of modern leadership.

It has been some time since scholars and experts recognized the importance of social exchange between leaders and their coworkers in organizations. Graen and his colleagues [1-8] were among the first in research and theorization of this organizational variable, along with the influences that affects it and its outcomes. They developed, what was first known as the Vertical Dyad Linkage theory [8], which has grown to its current well-known form, the Leader-Member Exchange (LMX) theory [9, 10]. LMX has been linked to a wide variety of organizational variables in the past, such as job satisfaction, employee performance, employee behavior and organizational commitment. Organizational commitment is a very important organizational variable, which has been investigated for half a decade, with intensification in the 70s, 80s and 90s [11, 12-17]. The main components of commitment are "strong belief in and acceptance of the organization's goals, willingness to exert considerable effort on behalf of the organization, and definite desire to maintain organizational membership" [11].

Strong leadership and organizational commitment are crucial for the effective working conditions, in which healthcare workers, save lives, every day. Foreign research have given some indications that the quality of LMX is associated with organizational commitment [18, 19-22], but there was generally not enough research and evidence to support this idea in various work conditions. Research is also needed to investigate in which ways LMX affects commitment and to what extent.

This paper departs from the general idea of the importance of strong leadership and commitment of employees in organizations, especially for those in the healthcare sector. For this purpose, research that reflects the connections between LMX and organizational commitment has been designed. A large public general hospital with over one thousand employees was chosen as the subject for this research. The managers of the hospital agreed that it is crucial to conduct research which will show to what extent, the employees are committed, to what extent leaders affect the commitment of employees and which groups of employees are critical and require immediate measures and changes.

2 Theoretical Background

2.1 Leader-Member Exchange Theory

Leader-member exchange theory (LMX) is a relational, entity perspective approach to leadership and relationships between leaders and members with the area of interest and focus on the behaviors and properties of individuals in interactions with each other [23]. At its center is the social exchange which occurs between leaders and members [10]. According to LMX, leadership is composed of "three primary components: the characteristics of the leader; those of the follower; and the maturity of the leadership relationship" as it "occurs within the context of the leadership relationship [9]. Early LMX scholars [7, 8] argued that managers treat different subordinates differently, generally putting them into two groups, which are called 'IN' and 'OUT' groups. The members of the in-group form stronger and closer relationships with their leaders that are known as high quality exchange relationships, while members of the out-group form low quality exchange relationships with their leaders. High quality relationships are accompanied with high levels of mutual trust, respect and obligation which results in subordinate job execution beyond job descriptions and improved performance, as opposed to low quality relationships which are accompanied with low level of trust, respect and obligation and subordinate execution of only what is requested from them in job descriptions [6, 10]. This is, in part, due to additional support, attention and information which are given to subordinates from their leaders in higher quality relationships [1]. In newer literature, "the central concept of LMX theory is that leadership occurs when leaders and followers are able to develop effective relationships that result in incremental influence and thus gain access to the many benefits these relationships bring" [23]. These effective relationships are called partnerships [10]. Graen and Uhl-Bien [9] argue that these partnerships are built through three stages in the "life cycle" of leadership. In the first "stranger" stage, the relationship and the exchange between the leader and the follower is entirely contractual with no incremental influence among them. Some relationships do not ever go beyond this stage and they stay undeveloped and of a low quality. The second "acquaintance" stage is characterized by increased, equitable, but not absolute exchange between the leader and the follower in a form of a test for advancing to the final "mature" stage of the relationship, which brings very high quality relationships-partnerships followed with loyalty, support, long time span reciprocation exchanges, emotions and high incremental influence between the leader and the follower [9, 10].

The origins of LMX theory lies in the VDL – Vertical Dyad Linkage theory. With VDL, differentiated dyadic relations between leaders and followers were discovered [1, 8]. In sociology, the term 'Dyad' is used for denoting a group of two people, the smallest possible social group. 'Dyadic' is a word that is used for

denoting interaction and communication of the mentioned groups. VDLs focus is on the reciprocal influence processes inside differentiated vertical dyads of a superior and a subordinate. It means that managers are developing differentiated relationships with direct reports, as opposed to earlier leadership approaches by which managers are using an average leadership style and develop same relationships with all of their followers [1, 8]. The focus was generally on the leaders' individual behaviors. In the next stage of theory development, the focus was moved from individual behaviors to the relationships and its outcomes, which changed the focus from VDL to LMX – Leader-member exchange theory [3]. Important findings from this stage are: leaders' and followers' characteristics and behavior influence the development of the exchange relations taking place in the role creation process; high quality exchange relations have considerable positive effects for leaders, followers, departments, groups, teams and the whole organization; results of high quality exchange relations development and maintenance are also very effective leadership processes [10]. The next phase of development was on a much higher level and the traditional distinction between leaders and followers have been abandoned in favor of 'partnerships' between co-workers. Managers should make an offer to every subordinate to develop a mutual partnership among them, so every employee has equal opportunities for a high quality relationship with their manager, making the whole process of leadership more equitable. It is important that managers are encouraged to make these offers [3, 5]. The decision on whether subordinates accept or this offer not is on them, but an equal partnership offer is crucial, because employees are well aware that if their manager treats them differently, this affects their perception of fairness [24]. Employees who accept the offer and build high quality relationships with their managers as a result have much higher performance than those subordinates who don't accept the offer [3, 5]. Also, higher quality relationships are related to lower turnover rates and higher perceptions of leader support [2]. However, Dunegan, Uhl-Bien and Duchon [25] argue that LMX was in the past indeed connected to various organizational variables, but the proof for the links between LMX and turnover, and LMX and subordinate performance were inconsistent and required further investigation. The fourth, final phase in LMX development is turned towards "systems of interdependent dyadic relationship" [10]. A system level perspective was adopted to answer the "question of how differentiated dyadic relationships combine to form larger systems of network assemblies" [10].

2.2 Organizational Commitment

Organizational commitment is a work related attitude [11, 17]. Because attitudes influence our behavior toward objects, situations, persons or groups, the most simple way to define organizational commitment is to say it is an attitude that reflects the strength of the relation between an organization and its employees [26], or the extent to which an employee is loyal to his/her organization [27].

Kanter [16] was one of the first do define commitment as the willingness of a social actor to give his/her energy and loyalty to a social system. In terms of organizational commitment, the term actor refers to employees and the term system refers to an organization. Porter, et al. [11] defines organizational commitment as "a strong belief in and acceptance of the organization's goals, a willingness to exert considerable effort on behalf of the organization, and a definite desire to maintain organizational membership". Similarly, Bateman and Strasser [28] discuss that commitment is defined as "multidimensional in nature, involving an employee's loyalty to the organization, willingness to exert effort on behalf of the organization, degree of goal and value congruency with the organization, and desire to maintain membership". Rusbult and Farrel [29] discuss commitment as "the likelihood that an individual will stick with a job and feel psychologically attached to it, whether it is satisfying or not". According to Agnew, et al. [30] commitment can be seen as intent to stay and endure in a relationship, including long-term navigation toward involvement and feelings of psychological attachment.

At the beginning of commitment theory and research development, it was considered as a one-dimensional variable, but later approaches recognized the need for the distinction between several types of commitment. The most recognized approach to commitment type distinction is that of Meyer and his colleagues [31, 32-34] which makes a distinction between three types of commitment: Affective, Continuance and Normative commitment in their Three-Component Model.

Meyer and Allen [35] argue that there can be a wide variety of factors which affect commitment development, but the strongest and most common factors are usually situational. Many researchers [15, 17, 35, 36, 37] investigated the influence of personal characteristics, mainly age, education, tenure and similar, on commitment and found connections between these two variables. According to Coe, Zehnder and Kinlaw [38] there are four critical conditions in the mind of people for building commitment in an organization: clear visions about core values and performance goals; influence over the job; competence to perform the job; and appreciation for the demonstrated performance. According to Meyer, et al. [35] age and tenure have mostly weak correlations with commitment. Further, external locus of control negatively correlates, while task self-efficacy positively correlates with affective commitment. Moral and ethics have an important influence on commitment, and this is highly expressed in public sector employees, because they have strong ethics [39].

Job characteristics and work related experiences also have a strong influence on organizational commitment [17, 35, 37]. Meyer, et al. [35] found strong correlations between work-related experiences and commitment, especially affective commitment. The investigated variables were role ambiguity, role conflict and perceived organizational support [35]. Perceived organizational support is very important for building affective commitment, by producing a felt

obligation that helps in achieving organizational goals [40], as well as organizational dependability and trust [17]. Commitment can also be influenced by organizational changes, specifically affective commitment by changes in comfort related and competence related work experiences; continuance commitment by changes in the budget, job security and alternatives; and normative commitment by changes in the perception of the investments that the organization makes in its employees [41]. Johns and Saks [26] discuss that during recessions, a typical scenario is that employees have to stay in an organization they hate, which is related to low affective and high continuance commitment. Job security is an especially important antecedent of continuance commitment for employees in the public sector [39, 42]. Continuance commitment is also often associated with antecedents like investments and alternatives. According to Meyer and Allen [36], there can be no continuance commitment if employees don't recognize the alternatives.

Different levels of commitment can have various outcomes. Meyer, Allen and Topolnytsky [41] argue that "conditions that lead to changes in the nature of commitment can have important implications for employee morale, motivation, performance and, ultimately, organizational success". Angle and Perry [13] discuss that "a committed member's definite desire to maintain organizational membership would have a clear relationship to the motivation to participate". In their research, they also found strong evidence for the claim that there is an inverse relationship between organizational commitment and employee turnover, which was supported and proved by numerous other researches [11, 35, 36, 37, 43]. Low absenteeism is also an important commitment outcome [36, 37, 43], but only for affective commitment, because continuance and normative commitment lead to higher levels of absenteeism [35]. For Steers [37], the most important outcomes of commitment are desire to remain, intent to remain, attendance, employee retention and high job performances. In their meta-analysis Meyer, et al. [35] found correlations between commitment and numerous other work related variables-consequences of commitment: negative correlation between commitment and job turnover; negative correlation between affective commitment and absenteeism; positive correlations between affective and normative commitment and job performance; negative correlation between continuance commitment and job performance; positive correlations between affective and normative commitment and organizational citizenship behavior; negative correlations between affective commitment and stress and work–family conflict; and positive correlations between continuance commitment and stress and work–family conflict.

3 Method

3.1 Sample and Procedure

The research was conducted in February and March 2013 in a large hospital with over 1000 employees, located in Central Europe. According to the confidentiality agreement that the authors of this paper signed with the management of the hospital, the exact name and location of the hospital cannot be revealed. A total number of 530 sets of questionnaires were distributed evenly in all departments of the hospital to all employees, without focusing on a specific type of employee except for leaders in the highest positions in hospital [management], since the research had a member focus, i.e. the interest was on the members' perceptions about the quality of the leader-member exchange relationship and its impact on their commitment. A total number of 359 valid questionnaires were returned. The return rate was 67.74%. This was a very decent return rate, considering the type of activity and employees, its importance and their high level of occupancy at work. The questionnaires were completely anonymous and on every department, one person was in charge for their collection in a specially intended box. The questionnaires were all put together after the completion of the research and no distinctions were made among departments, because it was not the goal of the research.

3.2 Description of the Sample

The majority of the respondents were females (77.20%); the males were in the minority (20.9%), while only 7 employees (1.9%) did not specify their sex. There were two significant age groups among respondents, first with the average age about 30 years and second with about 55 years as shown in Figure 1 (1.67% did not answer this question). The arithmetic mean for age was 41.26 years (SD = 11.045). Regarding tenure with the organization, the most important group had 10 years of tenure (Figure 2), with an average of 16.33 years (SD = 10.739) (3.34% left this item blank). Tenure was given only in full years, months were not taken into consideration, so 0 years in analyses refers to employees with less than a year of tenure. Since the research was conducted in the healthcare sector, with specific types of activity compared to other types of organizations, there were eight types of education offered in the questionnaires, with the results: primary school (3.1%), secondary school (69.9%), higher education (9.7%), faculty-bachelor's degree (5.8%), master's degree (0.8%), doctors of medicine (2.8%), specialist doctors of medicine (7.2%) and other (0.3%). Only one employee (0.3%) did not provide the data for his education level. Considering that most respondents were with secondary school (nurses and administrative workers), it can be concluded that they were far more willing to complete the questionnaire than the others, since

there is much higher percentage of medical doctors and specialists than the percentages in this research. In order for further analyses to be more concise and meaningful, respondents with higher education, faculty-bachelor's degree and master's degree were merged into a single group (with 16.3% participation in the whole sample) for further use in the analyses. Considering its irrelevance, the type "other" was excluded from further analyses.

3.3 Instruments

Two questionnaires were used for the research: The quality of leader-member exchange was measured with the concise LMX-7 questionnaire for members [4, 10] on a standard 5 item Likert Scale. Identical questionnaires were given to every employee included in the research, so no leader-member distinction was made among them, as the focus was on how employees as members evaluate their exchange with their leaders generally and what impact it has on their commitment, so they were all viewed as members. The area of interest of this research was not on specific leader-member relationships. Employees were asked to assess the quality of the exchange relationship with all of their leaders generally. The LMX-7 is a one-dimensional scale and includes seven items with the response anchors differing with each item. This questionnaire was validated in a great number of researches [4, 10, 44-46]. It is the most accepted questionnaire for measuring LMX. Cronbach's alpha for this questionnaire in the present study was very high (α=0.93). The validity of the questionnaire was confirmed using principal components analysis, the statistics are relevant according to Guttman-Kaiser criterion. Considering the quantity of variance that the first component includes (70.653% of the total variance, Λ=4.946), the questionnaire is one-dimensional and homogeneous. All component saturations were above 0.76. The representativeness of the items according to the KMO criterion was significant (0.923).

Organizational commitment was measured with the 15-item Organizational Commitment Questionnaire - OCQ [11], also measured on a standard 5-point Likert Scale from "completely disagree" to "completely agree". This questionnaire is the most widely used instrument for measuring organizational commitment [14, 35], with investigated and proven psychometric characteristics and used in measuring commitment in a wide range of job categories [12]. It includes items concerning the employee's perceptions about their loyalty to the organization, their willingness to highly engage in activities to achieve organizational aims and their acceptance of organizational values [11]. Cronbach's alpha for this questionnaire in the present study was satisfactory (α = 0.881). The representativeness of the items according to the KMO criterion was significant (0.901). The validity of the questionnaire was confirmed using factor analysis, with principal components method. According to Guttman-Kaiser criterion results, two subscales were created, similar to the subscales of Angle and

Perry [13]. The first subscale (Cronbach's alpha =0.913, Λ=6.096, includes 40.639% of the total variance) refers to the respondents value commitment, which reflects their affective commitment and includes items 1, 2, 4, 5, 6, 8, 10, 13, and 14. The second subscale (Cronbach's alpha =0.718, Λ=1.883, includes 12.550% of the total variance) refers to the respondents commitment to stay, which reflects their continuance commitment and includes items 3, 9, 11, 12, and 15. The two subscales are negatively correlated (r=-0.378), which reflects a weak relationship. Normative commitment was not measured in this research. In further analyses of the main scale, the items with reversed directions (items 3, 9, 11, 12 and 15) were re-coded.

3.4 Data Processing Methods

The data in this research was analyzed completely with the SPSS statistics software. Analyzes included descriptive statistics, instruments check (Cronbach's alpha, Guttman-Kaiser, factor analyses, representativeness, validity), analysis of the distribution of scores, descriptive statistics for scores (Mean, SD, Skewness, Kurtosis, Kolmogorov-Smirnov), correlations, ANOVA, Post-hoc test - least significant difference (LSD), Analysis of covariance (ANCOVA) and t-test.

3.5 Research Questions and Hypotheses

Four main research questions were created for this research:

- Does high quality exchange with leaders exist among employees of the hospital?
- Are employees of the hospital committed to the organization and its goals?
- Do personal characteristics of employees affect their LMX and commitment level?
- Does the quality of exchange with leaders influence the commitment of employees?

The hypotheses created on the basis of the research questions:

- H1: Employees of the hospital have high quality leader-member exchange relations with their leaders.
- H2: Employees of the hospital are committed to the organization.
- H3: Personal characteristics of employees influence their LMX and commitment level.

- H3.1: Differences in gender don't affect LMX and commitment level.
- H3.2: Older and employees with longer tenure and higher education have higher levels of LMX and commitment.

- H4: The quality of the employee's LMX is correlated with their commitment levels.
  - H4.1: Employees with high quality LMX are more committed to the organization's values.

4 Results

4.1 Descriptive Indicators for Scores

LMX-7 – Mean = 21.349, SD = 6.2412, Skewness = -0.102, Kolmogorov-Smirnov = 0.070; according to the results, the discriminability was not significantly disrupted.

OCQ – Mean = 48.91, SD = 10.804, Skewness = -0.217, Kolmogorov-Smirnov = 0.052; the discriminability was not disrupted significantly.

OCQ (value commitment subscale) – Mean = 30.24, SD = 8.169, Skewness = -0.268, Kolmogorov-Smirnov = 0.063; the discriminability was not disrupted significantly.

OCQ (commitment to stay subscale) – Mean = 17.40, SD = 4.398, Skewness = 0.212, Kolmogorov-Smirnov = 0.073; the discriminability was not disrupted significantly.

4.2 t-test for Differences

Between Genders

The tests showed that there is statistically significant difference in organizational commitment on the whole ($t=-2.131$, $p\leq0.05$) and in value commitment ($t=-2.699$, $p\leq0.01$) depending on gender. Females had higher scores on the general OCQ and the value commitment subscale. The tests did not find significant differences for LMX (Table 1).

Table 1
t-test for differences between genders

	Levene's test		t test for independent samples						
	F	p	t	df	p	group	N	M	s
LMX-7	2.358	0.126	-0.608	342	0.543	males	74	21.149	6.9216
						females	270	21.644	6.0031
OCQ	0.317	0.574	-2.131	318	0.034	males	70	46.50	11.362
						females	250	49.59	10.533
Value commitment subscale	0.037	0.848	-2.699	327	0.007	males	72	28.14	8.411
						females	257	31.07	8.070
Commitment to stay subscale	2.168	0.142	0.400	332	0.690	males	72	17.53	4.753
						females	262	17.30	4.203

4.3 Age and Tenure

The relationships of age and tenure with LMX and commitment were determined with Spearman's coefficient of correlation. Negative relationships were detected between LMX and age, as well as between LMX and tenure. Older respondents and respondents with longer tenure assess the quality of exchange as lower than younger respondents and respondents with shorter tenure.

There is also negative relationship between commitment level and tenure. Employees with longer tenure are less committed to the organization. Subscale analysis revealed negative correlation between value commitment and tenure, positive correlation between commitment to stay and age, and weak positive correlation between commitment to stay and tenure (Table 2).

Table 2
Spearman's coefficients of correlation

		Age	Tenure
LMX-7	Spearman ρ	-.134*	-.190**
	p (2-tailed)	0.012	0.000
	N	346	340
OCQ	Spearman ρ	-0.044	-0.119*
	p (2-tailed)	0.428	0.035
	N	322	317
Value commitment subscale	Spearman ρ	0.006	-0.093
	p (2-tailed)	0.914	0.095
	N	332	326
Commitment to stay subscale	Spearman ρ	0.105	0.085
	p (2-tailed)	0.054	0.122
	N	335	330

4.4 Education Level

The ANOVA test noticed statistically significant differences among subgroups of respondents with different education levels in LMX evaluation ($F (4; 344) = 3.058$, $p \leq 0.05$). The LSD test showed that respondents with primary education and doctors of medicine more positively evaluate their exchange with leaders compared to other employees. However, LSD is very liberal and typically has a high Type I error rate.

In evaluating their commitment level, the ANOVA test noticed statistically marginal differences among subgroups of respondents with different education levels ($F (4; 320) = 2.200$, $p=0.07$). The LSD test showed that doctors of medicine more positively evaluate their organizational commitment compared to respondents with high school, bachelor's and master's degrees and specialist doctors of medicine.

Value commitment subscale ANOVA analysis revealed statistically significant differences among subgroups of respondents with different education levels ($F (4; 329) = 2.541$, $p \leq 0.05$). The LSD test showed that respondents with primary education and doctors of medicine are more committed to the values of the organization than other respondents.

Commitment to stay subscale ANOVA analysis did not reveal any statistically significant differences among subgroups of respondents with different education levels ($F (4; 334) = 1.927$, $p=0.11$).

4.5 Pearson Correlations between LMX and Organizational Commitment

There are significant connectivity between LMX and organizational commitment evaluation ($r=0.539$, $p \leq 0.01$), LMX and value commitment subscale ($r=0.553$, $p \leq 0.01$), and negative connectivity between LMX and commitment to stay subscale ($r=-0.311$, $p \leq 0.01$).

Table 3

Pearson's correlations between concepts

		Organizational commitment	Value commitment subscale	Commitment to stay subscale
LMX	Pearson Correlation	0.539**	0.553**	-0.311**
	p (2-tailed)	0.000	0.000	0.000
	N	321	330	335

4.6 ANCOVA

In order to get more accurate and useful results, respondents with primary and high school education were merged into one group in this test. Levene's test for equality of error variances showed that the variance is homogeneous, which is a requirement for this analysis. LMX, age, tenure, gender and education were used as predictors, while organizational commitment was used as the criterion (dependent) variable.

Covariance analysis showed that predictors LMX and gender had significant effect on organizational commitment (LMX: ($F(1; 301) = 114.582$, $p \leq 0.01$, Partial $\eta 2 = 0.276$), gender: ($F(1; 301) = 5.587$, $p \leq 0.05$, Partial $\eta 2 = 0.018$)). The corrected value of squared multiple correlation shows that 29% of the total variance of the criterion variable are explained with this set of predictors.

Respondents that positively evaluate the quality of LMX also positively evaluate their organizational commitment ($b = 0.909$, $p \leq 0.01$, Partial $\eta 2 = 0.276$). Females are more committed to the organization in general ($b = -6.738$, $p = 0.08$, Partial $\eta 2 = 0.010$). Doctors of medicine and specialist doctors of medicine are also more committed to the organization compared to other respondents, although the effects of education do not have high significance.

Table 4

Tests of Between-Subjects Effects

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial η^2
Corrected Model	11153.540[a]	9	1239.282	15.301	0.000	0.314
Intercept	7273.178	1	7273.178	89.800	0.000	0.230
LMX	9280.356	1	9280.356	114.582	0.000	0.276
Age	66.305	1	66.305	0.819	0.366	0.003
Gender	452.537	1	452.537	5.587	0.019	0.018
el	491.278	3	163.759	2.022	0.111	0.020
Gender * el	119.362	3	39.787	0.491	0.689	0.005
Error	24378.833	301	80.993			
Total	773352.000	311				
Corrected Total	35532.373	310				
R Squared = 0.314 (Adjusted R Squared = 0.293)						

5 Discussions

Some scholars already emphasized the important connections between LMX and organizational commitment [18, 19-22]. However, most researches do not reveal anything deeper than the simple notation that LMX correlates with organizational commitment. For example, Gerstner and Day [22] conclude that LMX is consistently correlated with commitment, but they don't reveal anything significant about the nature of this relationship. Some of the researchers, who have previously conducted similar studies [19, 20] also noted that it is difficult to find research that incorporates LMX as an antecedent of commitment and that only a few studies deal with this issue. Joo [20] tried to go deeper in explaining the relationship between these two variables, however he measured only affective commitment and provided only the basic analysis and just discussed that LMX is indeed related and has impact on commitment. Kang and his colleagues [19] made an interesting observation in favor of the necessity to conduct research on the impact of LMX on organizational commitment in various working environments that most of the researches incorporating LMX and commitment have been conducted in western developed economies. It also seems that in the majority of these researches LMX as an antecedent of organizational commitment, analysis of their relationship and their importance are lost in the chaos between many other variables measured in these sometimes over complicated studies. Another problem is that these studies often use different instruments for measuring both LMX and commitment, so the findings cannot be homogenous and completely comparable. In some researches, LMX is only used as a mediator variable [21].

This study intended to contribute in solving at least two main issues noticed, the necessity to conduct researches outside western settings, and the unnecessary complication of studies incorporating LMX and commitment. It also had the intention to go further and deeper in explaining the relationships of these variables.

5.1 Discussion of the Results

H1, which proposes that employees have high quality LMX with their leaders, was partially supported with descriptive statistics. H2, proposing that employees are committed to the organization, also received partial support. The analyses showed that there is much empty space for improving the employees LMX and commitment. Personal characteristics have significant influence on both of the variables tested. Therefore, H3, suggesting that personal characteristics of employees influences their LMX and commitment level, received full support. Females are more committed to the organization and to its values. It is known that females are typically drawn to the healthcare sector because of their nurturing attitudes. LMX was not affected by gender, so H3.1, proposing that differences in gender do not affect LMX and commitment level was supported only for LMX.

Age, tenure and education analyses revealed very unexpected results. Regarding age differences and tenure, analysis showed that older respondents and those with longer tenure have lower LMX. Respondents with longer tenure are also less committed to the organization generally and to its values, while their commitment to stay increases with age and tenure. The presumption is that older and employees with longer tenure are more objective in evaluating their LMX and commitment. As time passes, they are more informed about the reality. The shock of facing the reality can have significant impact on people. The results can be perceived through the self-determination theory, which points on the autonomous behavior regulation style through the integration of life experience and personality disposition in a unique system of the self [47, 48]. They also realize that as they get older that finding another job will get harder, so the job security granted to them as to most public service organization employees leads to higher commitment to stay. Higher LMX and value commitment had respondents with primary education and doctors of medicine than others, while doctors of medicine are generally most committed to the organization. The ANCOVA test added specialist doctors of medicine to the list of generally most committed, but with small significance. High levels of LMX and commitment among employees with only primary education was indeed a surprise. As age, tenure and education do not influence LMX and commitment in the expected way, H3.2 is not supported.

The most important analysis in this study is the correlation analysis between LMX and commitment. The tests confirmed the results of many earlier studies [18, 19-22] that LMX and commitment are significantly positively correlated. Further, analysis revealed that respondents with high quality LMX are more committed to the organization's values [positive correlation] and have lower level of commitment to stay [negative correlation]. Employees with higher LMX are more likely to engage themselves in achieving the values of the organization, however they leave easier from the organization, which reflects affective commitment towards the organization. The ANCOVA test also confirmed the connection between LMX and commitment and revealed that 29% of the total variance of organizational commitment is explained with predictors LMX and age. Thus, H4, proposing that LMX correlates commitment is supported with high certainty. H4.1, proposing that employees with high LMX have higher value commitment, is also supported with great confidence.

Conclusions

Different economic and cultural conditions can reveal different facts about the nature of important organizational variables, in this case, concerning the exchange between leaders and members, and organizational commitment. As mentioned before, most of the studies conducted on both LMX and organizational commitment were in Western environments. Thus, non-western environments had little influence on the theory and practice development. Therefore, sometimes the theory does not correspond to the characteristics of these cultural and business environments.

The main findings in this study revealed that LMX and organizational commitment have positive connectivity in a non-western environment. It was also revealed that LMX and value commitment have much stronger relationship than the relationship between LMX and commitment to stay, which is a negative relationship. Critical employee categories were also detected. Measures for improving their LMX and commitment are needed primarily for males, older employees, employees with longer tenure and employees with secondary education, bachelor's and master's degrees. The conclusion is that leaders and leadership have strong influence on the commitment of their employees, so these findings could have significant implications in the process of leadership modernization in stressful and demanding working environments. Knowing how important commitment can be among healthcare workers, greater attention needs to be given to leadership in hospitals.

Further research is needed and data should be collected from many different environments, in the region, to learn more about the influence of LMX on organizational commitment. According to Perry [39], moral and ethics have an important influence on commitment, and this is highly expressed among public sector employees, because they have strong ethics. It would be interesting to incorporate moral and ethics, as moderator variables, in further studies examining the influence of LMX on organizational commitment.

Acknowledgement

The authors wish to thank the reviewers for their constructive comments that were very helpful in improving the quality of this paper.

The work on this paper was supported by the Provincial Secretariat for Science and Technological Development of the Autonomous Province of Vojvodina, through the project "Effects of Organizational Communication on Organizational Behavior of Employees"

References

[1] G. Graen, J. Cashman: A Role-Making Model of Leadership in Formal Organizations: A Developmental Approach. In J. G. Hunt, L. L. Larson (eds.): Leadership Frontiers, Kent State University Press, Kent, 1975, pp. 143-166

[2] G. Graen, W. Hoel, R. C. Liden: Role of Leadership in the Employee Withdrawal Process, Journal of Applied Psychology, 67, 6, 1982, pp. 868-872

[3] G. Graen, M. A. Novak, P. Sommerkamp: The Effects of Leader—Member Exchange and Job Design on Productivity and Satisfaction: Testing a Dual Attachment Model, Organizational Behavior and Human Performance, 30, 1, 1982, pp. 109-131

[4] T. A. Scandura, G. Graen: Moderating Effects of Initial Leader-Member Exchange Status on the Effects of a Leadership, Journal of Applied Psychology, 69, 3, 1984, pp. 428-436

[5] G. Graen, T. A. Scandura, M. R. Graen: A Field Experimental Test of the Moderating Effects of Growth Need Strength on Productivity, Journal of Applied Psychology, 71, 3, 1986, pp. 484-91

[6] M. D. Zalesny, G. Graen: Exchange Theory in Leadership Research. In: Reber G, editor. Encyclopedia of Leadership, Linz University Press, Linz, 1986

[7] G. Graen, T. A. Scandura: Toward a Psychology of Dyadic Organizing. In B. M. Staw, L. L. Cummings (eds.): Research in Organizational Behavior, JAI Press, Greenwich, 1987, pp. 175-208

[8] F. Dansereau Jr., G. Graen, W. J. Haga: A Vertical Dyad Linkage Approach to Leadership within Formal Organizations: A Longitudinal Investigation of the Role Making Process, Organizational Behavior and Human Performance, 13, 1, 1975, pp. 46-78

[9] G. Graen, M. Uhl-Bien: The Transformation of Work Group Professionals into Self-Managing and Partially Self-Designing Contributors: Toward a Theory of Leadership-Making, Journal of Management Systems, 3, 3, 1991, pp. 25-39

[10] G. Graen, M. Uhl-Bien: Relationship-based Approach to Leadership: Development of Leader-Member Exchange (LMX) Theory of Leadership over 25 Years: Applying a Multi-Level Multi-Domain Perspective, The Leadership Quarterly, 6, 2, 1995, pp. 219-247

[11] L. W. Porter, R. M. Steers, R. T. Mowday, P. V. Boulian: Organizational Commitment, Job Satisfaction, and Turnover among Psychiatric Technicians, Journal of Applied Psychology, 59, 5, 1974, pp. 603-609

[12] R. T. Mowday, R. M. Steers, L. W. Porter: The Measurement of Organizational Commitment, Journal of Vocational Behavior, 14, 2, 1979, pp. 224-247

[13] H. L. Angle, J. L. Perry: An Empirical Assessment of Organizational Commitment and Organizational Effectiveness, Administrative Science Quarterly, 26, 1, 1981, pp. 1-14

[14] J. P. Meyer, N. J. Allen: Testing the 'Side-Bet Theory' of Organizational Commitment: Some Methodological Considerations, Journal of Applied Psychology, 69, 3, 1984, pp. 372-378

[15] J. P. Meyer, N. J. Allen, C. A. Smith: Commitment to Organizations and Occupations: Extension and Test of a Three-Component Conceptualization, Journal of Applied Psychology, 78, 4, 1993, pp. 538-551

[16] R. M. Kanter: Commitment and Social Organization: A Study of Commitment Mechanisms in Utopian Communities, American Sociological Review, 33, 4, 1968, pp. 499-517

[17] B. Buchanan: Building Organizational Commitment: The Socialization of Managers in Work Organizations, Administrative Science Quarterly, 19, 4, 1974, pp. 533-546

[18] A. Yousaf, K. Sanders, N. Torka, J. Ardts: Having Two Bosses: Considering the Relationships between LMX, Satisfaction with HR Practices, and Organizational Commitment, International Journal of Human Resource Management, 22, 15, 2011, pp. 3109-3126

[19] D. S. Kang, J. Stewart, H. Kim: The Effects of Perceived External Prestige, Ethical Organizational Climate, and Leader-Member Exchange (LMX) Quality on Employees' Commitments and their Subsequent Attitudes, Personnel Review, 40, 6, 2011, pp. 761-784

[20] B. K. Joo: Organizational Commitment for Knowledge Workers: The Roles of Perceived Organizational Learning Culture, Leader-Member Exchange Quality, and Turnover Intention, Human Resource Development Quarterly, 21, 1, 2010, pp. 69-85

[21] J. Lee: Effects of Leadership and Leader-Member Exchange on Commitment, Leadership and Organization Development Journal, 26, 7/8, 2005, pp. 655-672

[22] C. R. Gerstner, D. V. Day: Meta-Analytic Review of Leader–Member Exchange Theory: Correlates and Construct Issues, Journal of Applied Psychology, 82, 6, 1997, pp. 827-844

[23] M. Uhl-Bien: Relational Leadership Theory: Exploring the Social Processes of Leadership and Organizing, Leadership Quarterly, 17, 6, 2006, pp. 654-676

[24] P. M. Sias, F. M. Jablin: Differential Superior-Subordinate Relations, Perceptions of Fairness, and Coworker Communication, Human Communication Research, 22, 1, 1995, pp. 5-38

[25] K. J. Dunegan, M. Uhl-Bien, D. Duchon: LMX and Subordinate Performance: The Moderating Effects of Task Characteristics, Journal of Business and Psychology, 17, 2, 2002, pp. 275-285

[26] G. Johns, A. M. Saks: Organizational Behaviour: Understanding and Managing Life at Work, Pearson Prentice Hall, 2005

[27] J. R. Schermerhorn, J. G. Hunt, R. N. Osborn, M. Uhl-Bien: Organizational Behavior, John Wiley & Sons, 2010

[28] T. S. Bateman, A. Strasser: A Longitudinal Analysis of the Antecedents of Organizational Commitment, The Academy of Management Journal, 27, 1, 1984, pp. 95-112

[29] C. E. Rusbult, D. Farrell: A Longitudinal Test of the Investment Model: The Impact on Job Satisfaction, Job Commitment, and Turnover of Variations in Rewards, Costs, Alternatives, and Investments, Journal of Applied Psychology, 68, 3, 1983, pp. 429-438

[30] C. R. Agnew, P. A. M. Van Lange, C. E. Rusbult, C. A. Langston: Cognitive Interdependence: Commitment and the Mental Representation of Close Relationships, Journal of Personality and Social Psychology, 74, 4, 1998, pp. 939-954

[31] N. J. Allen, J. P. Meyer: The Measurement and Antecedents of Affective, Continuance and Normative Commitment to the Organization, Journal of Occupational Psychology, 63, 1990, pp. 1-18

[32] N. J. Allen, J. P. Meyer: Affective, Continuance, and Normative Commitment to the Organization: An Examination of Construct Validity, Journal of Vocational Behavior, 49, 3, 1996, pp. 252-276.

[33] J. P. Meyer, D. J. Stanley, T. A. Jackson, K. J. McInnis, E. R. Maltin, L. Sheppard: Affective, Normative, and Continuance Commitment Levels across Cultures: A Meta-Analysis, Journal of Vocational Behavior, 80, 2, 2012, pp. 225-245

[34] J. P. Meyer, N. J. Allen: A three-Component Conceptualization of Organizational Commitment, Human Resource Management Review, 1, 1991, pp. 61-89

[35] J. P. Meyer, N. J. Allen: Commitment in the Workplace: Theory, Research, and Application, SAGE, Thousand Oaks, 1997

[36] J. P. Meyer, D. J. Stanley, L. Herscovitch, L. Topolnytsky: Affective, Continuance, and Normative Commitment to the Organization: A Meta-Analysis of Antecedents, Correlates, and Consequences, Journal of Vocational Behavior, 61, 1, 2002, pp. 20-52

[37] R. M. Steers: Antecedents and Outcomes of Organizational Commitment, Administrative Science Quarterly, 22, 1, 1977, pp. 46-56

[38] C. Coe, A. Zehnder, D. C. Kinlaw: Coaching for Commitment: Achieving Superior Performance from Individuals and Teams, John Wiley & Sons, 2008

[39] J. L. Perry: Antecedents of Public Service Motivation, Journal of Public Administration Research and Theory, 7, 2, 1997, pp. 181-197

[40] L. Rhoades, R. Eisenberger: Perceived Organizational Support: A Review of the Literature, Journal of Applied Psychology, 87, 4, 2002, pp. 698-714

[41] J. P. Meyer, N. J. Allen, L. Topolnytsky: Commitment in a Changing World of Work, Canadian Psychology-Psychologie Canadienne, 39, 1/2, 1998, pp. 83-93

[42] J. L. Perry, J. L. Brudney, D. Coursey, L. Littlepage: What Drives Morally Committed Citizens? A Study of the Antecedents of Public Service Motivation, Public Administration Review, 68, 3, 2008, pp. 445-458

[43] C. Ostroff: The Relationship between Satisfaction, Attitudes, and Performance: An Organizational Level Analysis, Journal of Applied Psychology, 77, 6, 1992, pp. 963-974

[44] J. M. Maslyn, M. Uhl-Bien: Leader-Member Exchange and Its Dimension: Effects of Self-Effort and Other's Effort on Relationship Quality, Journal of Applied Psychology, 86, 4, 2001, pp. 697-708

[45] G. J. Greguras, J. M. Ford: An Examination of the Multidimensionality of Supervisor and Subordinate Perceptions of Leader-Member Exchange, Journal of Occupational and Organizational Psychology, 79, 2006, pp. 433-465

[46] J. A. Zhong, W. Lam, Z. G. Chen: Relationship between Leader-Member Exchange and Organizational Citizenship Behaviors: Examining the Moderating Role of Empowerment, Asia Pacific Journal of Management, 28, 3, 2011, pp. 609-626

[47] E. L. Deci, R. M. Ryan: Self-Determination Theory: A Macrotheory of Human Motivation, Development, and Health, Canadian Psychology-Psychologie canadienne, 49, 3, 2008, pp. 182-185

[48] C. F. Lam, S. T. Gurland: Self-determined Work Motivation Predicts Job Outcomes, but What Predicts Self-determined Work Motivation?, Journal of Research in Personality, 42, 4, 2008, pp. 1109-1115

Effect of Leakage in Electrohydraulic Servo Systems Based on Complex Nonlinear Mathematical Model and Experimental Results

Attila Kovari

College of Dunaújváros, Táncsics M. 1/A, 2400 Dunaújváros, Hungary
kovari@mail.duf.hu

Abstract: This paper looks into the background of the internal leakage effect on the dynamic behavior of an electrohydraulic servo positioning system. The electrohydraulic servo systems are widely used in precise control at high forces, but the overall quality of this servo control could be impaired by faults in the system. There could be several defects which have effect on the system response, but this work points to the effect of internal leakage of the hydraulic cylinder because this error cannot be easily detected. To show the effect of internal leakage on electrohydraulic servo system a complex nonlinear mathematical model considering the leakage of the hydraulic cylinder is used. The work herin demonstrates the relationship between internal leakage and the dynamic behaviour of the servo positioning system. Laboratory test results are used to verify the conclusions drawn from the mathematical model of the complete servo positioning system.

Keywords: servo system; mechatronics; position control; hydraulic servo; leakage

1 Introduction

In control technology, several uses of servo systems can be found, among which the most common applications are motion/rotation speed or position control [1]. Servo systems are mainly used when the advantages of these systems can be exploited to set the technological parameters of the controlled process: higher precision, reliability, better repeatability, shorter settling time, coordinated movement (even more in case of more degrees of freedom) [2], [3]. The advantages previously mentioned can be achieved by precise and appropriate control of technological equipment. In the case of precise and fast control, such as manufacturing equipment, mouldering equipment, mobile machines, material handling equipment and so on, servo systems are widely used [4].

The electrohydraulic servo systems can be applied in various ways in the referred fields, including testing equipment, active shock absorbers, mining machinery, material testing equipment, airplane simulators, paper manufacturing machines,

hydraulic systems on ships, robotics as well as steel and aluminum rolling mills, etc. Hydraulic actuators are widely used in the industrial environment due to their large reaction force on heavy loads, large power/volume ratio, good dynamic behavior, adequate rigidity, good heat removal, simple protection against overload, resistance against exterior impacts and also such actuators can be started under load with maximum acting power [5]. Hydraulic systems are also widely applied on airplanes where the application of hydraulic actuators is an ideal choice because of their large power/weight ratio and precise controllability [5].

The theoretical background and the elements of servo hydraulic systems are introduced in different studies [4]-[12], [23]-[25]. Generally, every regulated hydrostatic drive can be regarded as a hydraulic servo that is a closed loop system and in which the controlled feature of the drive is controlled by the application of some hydraulic element. In case of hydraulic systems, the servo mechanism usually controls large power units with low energy input signal. In the servo system, the servo hydraulic actuator corrects the output signal of the system by negative feedback based on the difference between the reference and real output signal, as well as it amplifies the acting power [7].

In modern hydraulic systems, more types of electromagnetically operated controlling units are used to implement electrohydraulic control. In precise servo technology valves are used and the valve spool operated by proportional magnets or electromagnet adjusted flapper ensures more precise operation [10], [12].

The block diagram of a typical electrohydraulic position control servo system can be seen in Fig. 1.

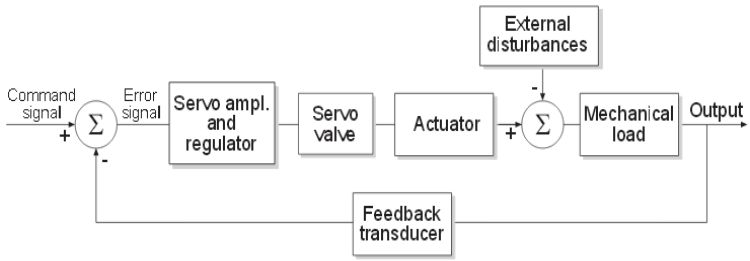


Figure 1
Typical hydraulic servo system

The control quality of hydraulic servo systems depends on the main elements used such as: servo regulator/controller, servo valve, hydraulic cylinder/actuator, feedback position transducer and power supply. The dynamic response of the system mainly builds upon the frequency characteristic of servo valve and load, but this overall quality of servo control could be impaired by faults in the system. There could be several defects which affect the system response. This study points to the effect of internal leakage of the hydraulic cylinders, since this defect and consequently the replacement or refurbishment of the actuator, can cause a longer

downtime of the equipment. This downtime may cause a greater loss of production and revenue in the case of continuously operated manufacturing equipment, for example a hot rolling mill. The fault of internal leakage is caused by wrong wear of piston seal or abrasion. This seal prevents the leakage and closes the gap between piston and the cylinder wall. External leakage can be detected visually, so this leakage can be easily perceived [21]. The internal leakage cannot be detected easily until the actuator seal is almost completely damaged, due to this reason the detection of internal leakage is more important. Several complex models based on the fault detection method are applied for electrohydraulic servo systems, but the leakage is not examined in details by the applied models [13]-[20].

This paper looks into the background effect of the leakage by a complex nonlinear mathematical model and demonstrates the relationship of internal leakage on dynamic behavior of the servo positioning system. Experimental results are presented to verify the conclusions drawn from the model.

2 Electrohydraulic Servo Positioning System

The examined equipment is a hydraulic positioning system which consists of a hydraulic power supply with relief valve and accumulator, flow control servo valve, a linear actuator unit (hydraulic cylinder), a position sensor and an electronic controller unit. The output signal of the system is the electric signal of the position transducer which is proportional with the actuator position. The servo amplifier and controlling unit determine the control signal of the servo valve based on the error signal that shows the difference between the reference and the actual output signal. The servo valve modifies the oil flow of the hydraulic cylinder until it is in appropriate position; I. e. the error signal is zero. Proportional or nozzle-flapper based servo valves are used in general hydraulic servo positioning systems depending on their applications. A simple servo valve controlled positioning system can be seen in Figure 2.

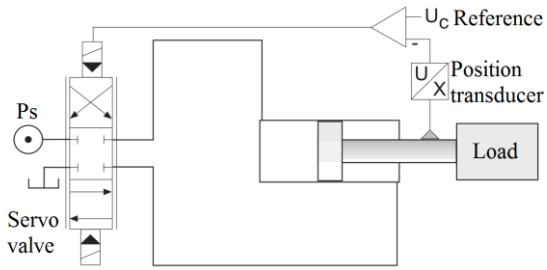


Figure 2
Servo valve controlled hydraulic cylinder

A complete servo hydraulic positioning system with accumulator, relief valve, asymmetric hydraulic cylinder, mass-spring-dumper load and sensors is shown in Figure 3.

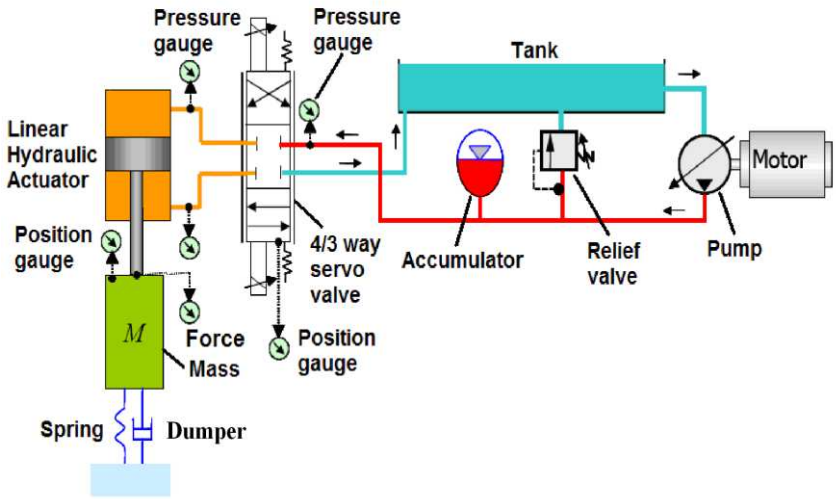


Figure 3
Electrohydraulic servo positioning system with mechanical load

3 Mathematical Model of the Electrohydraulic Positioning System

The electrohydraulic servo system was examined by a close-to-reality mathematical model, taking into account the non-linear and dynamic behaviour of hydraulic elements and the leakage of hydraulic actuator [16]-[20].

A servo-valve has high-order non-linear response and a number of valve parameters are required to calculate an accurate mathematical model [9]. The servo valve construction, models [10], specification standards [12], dynamic response model are described in detail in the literature [9], [17], the construction of torque motor and valve sleeve can be seen in Figure 4 and 5. The torque motor consists of an armature moved sleeve pivot. When current flows in the armature coils, the armature ends rotate because of the torque generated by the magnet fields of the current. This movement changes the oil flow of opposing nozzles, because the differential pressure of spool changes, therefore the spool is moved inside the valve sleeve shown in Figure 5.

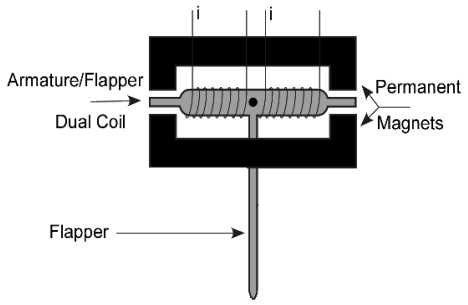


Figure 4
Servo-valve torque motor

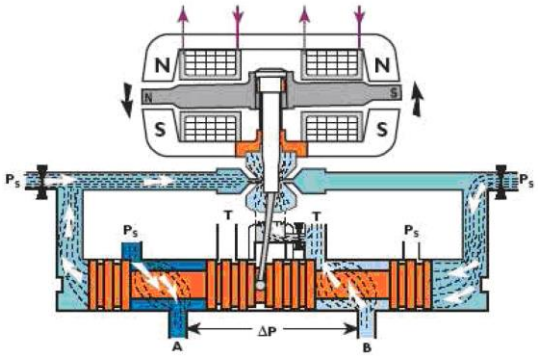


Figure 5
Spool movement inside the valve sleeve

When the spool moves from centre position the nozzles open the oil flow between the pressure port and tank port (P and T), and the two control ports (A and B), so the actuator is operated by the pressured oil (Figs. 6 and 7). In most cases a second order linear model of servo-valve can approximate the physical system properly [9]. The coefficients of second order system can be calculated by the frequency response based on Bode plots included in the servo-valve reference data [9], [22]. The torque motor of the servo-valve can be modeled as a series L-R circuit. The torque and consequently the moving force F_s of the valve spool are proportional to the torque motor current, the approximation of dynamic model of the valve spool movement x_s can be specified by a second order transfer function (ω_s is the natural frequency and ζ_s is damping ratio of the spool, k_t is proportionality coefficient) [9, 18]:

$$\frac{d^2 x_s}{dt^2} + 2 \cdot \zeta_s \cdot \omega_s \cdot \frac{dx_s}{dt} + \omega_s^2 \cdot x_s = \omega_s^2 \cdot k_t \cdot i \tag{1}$$

The oil flow rate at the control ports also depends on the pressure drop across the valve. The *As* flow cross-section is proportional (factor is *w*) to the valve spool displacement (Fig. 6).

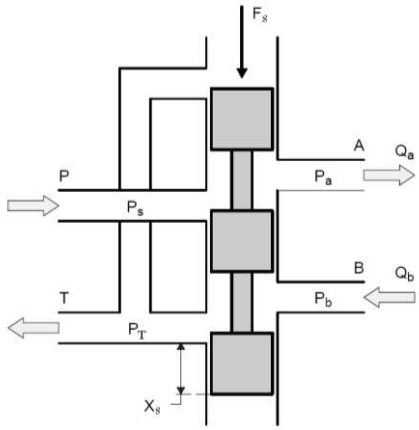


Figure 6
4/3 Flow control valve (Three Land, Four-way)
P_S,P_T: supply and tank pressure; P_a, P_b: actuator chamber pressures; Q_a, Q_b: actuator oil flows

The oil flow is also proportional to the square root of the pressure drop and c_s is the volumetric flow coefficient (ρ volumetric density of the oil) [7], [17], [18]:

$$Q = c_s \cdot w \cdot x_s \cdot \sqrt{\frac{2}{\rho} \cdot \Delta P} \tag{2}$$

A $c_s \cdot w$ product can be approximated by the valve nominal pressure drop ΔP_N and flow Q_N [10]:

$$C = c_s \cdot w = \frac{Q_N}{\sqrt{0{,}5 \cdot \Delta P_N}} \tag{3}$$

The compressibility of the oil creates a "spring" effect in the cylinder chambers which interacts with the piston mass to give a low frequency resonance. The effect can be modeled using the flow continuity equation from fluid mechanics [6] (ρ_a and ρ_b the density of the oil in chamber "a" and "b", Q_a and Q_b the load and return flow of the valve):

$$\frac{d}{dt}(\rho \cdot V) = m_a - m_b = \rho_a \cdot Q_a - \rho_b \cdot Q_b \tag{4}$$

In case of constant oil density the formula is:

$$Q_a - Q_b = \frac{dV}{dt} + \frac{V}{\rho} \cdot \frac{d}{dt} \tag{5}$$

Using the β fluid bulk modulus (mineral oils have 1.4 10^9 N/m, the value is pressure dependent) the formula can be written as [17], [18]:

$$\frac{d\rho}{\rho} = \frac{dP}{\beta} \tag{6}$$

$$Q_a - Q_b = \frac{dV}{dt} + \frac{V}{\beta} \cdot \frac{dP}{dt} \tag{7}$$

where V is the internal fluid volume. The pressure in cylinder chambers can be calculated using the A_a, A_b active areas of the piston annulus and V_a, V_b internal oil volume in chamber "a" and "b", x_d , v_d position:

$$\frac{dP_a}{dt} = \frac{\beta}{V_a} \cdot \left(Q_a - \frac{dV_a}{dt} \right) \tag{8}$$

$$\frac{dP_b}{dt} = \frac{\beta}{V_b} \cdot \left(-Q_b - \frac{dV_b}{dt} \right) \tag{9}$$

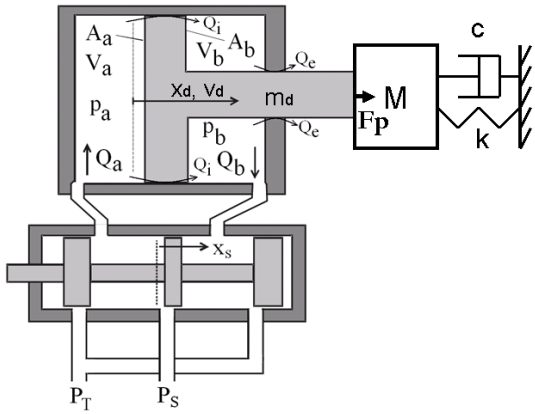


Figure 7
Hydraulic cylinder with 4/3 servo valve and mass-spring-dumper load

Using Merrit's laminar flow leakage model [10], the internal leakage oil flow at piston Q_i and external leakage oil flow at piston rod Q_e (see Fig. 7) can be modeled by the leakage resistance. The formula of leakage model using R_i and R_e the internal and external cylinder's leakage resistance or C_i and C_e the internal and external cylinder's leakage coefficient [17], [18]:

$$Q_i = \frac{P_a - P_b}{R_i} = C_i \left(P_a - P_b \right) \tag{10}$$

$$Q_e = \frac{P_b}{R_e} = C_e \cdot P_b \tag{11}$$

The chamber's pressure values can be written as:

$$\frac{dP_a}{dt} = \frac{\beta}{V_a} \cdot \left(Q_a - \frac{dV_a}{dt} - Q_i \right) \tag{12}$$

$$\frac{dP_b}{dt} = \frac{\beta}{V_b} \cdot \left(-Q_b - \frac{dV_b}{dt} + Q_i - Q_e \right) \tag{13}$$

The net acting force F_P can be calculated by the pressure and annuluses values of the two sides of the piston:

$$F_P = P_a \cdot A_a - P_b \cdot A_b \tag{14}$$

Based on (12-14) the acting force is:

$$\begin{aligned} F_P = &\int \left\{ \frac{\beta}{V_a} \cdot \left(Q_a - A_a \cdot v_d - Q_i \right) \right\} dt \cdot A_a - \\ &- \int \left\{ \frac{\beta}{V_b} \cdot \left(-Q_b + A_b \cdot v_d + Q_i - Q_e \right) \right\} dt \cdot A_b \end{aligned} \tag{15}$$

$$\begin{aligned} F_P(t) = &\int_0^t \left(\frac{\beta \cdot A_a}{V_{a0} + A_a \cdot x_d(\tau)} \cdot Q_a(\tau) + \frac{\beta \cdot A_b}{V_{b0} - A_b \cdot x_d(\tau)} \cdot Q_b(\tau) \right) d\tau \\ &- \int_0^t \left(\left[\frac{\beta \cdot A_a^2}{V_{a0} + A_a \cdot x_d(\tau)} + \frac{\beta \cdot A_b^2}{V_{a0} + A_b \cdot x_d(\tau)} \right] \cdot v_d(\tau) \right) d\tau \\ &- \int_0^t \left(\left[\frac{\beta \cdot A_a}{V_{a0} + A_a \cdot x_d(\tau)} + \frac{\beta \cdot A_b}{V_{b0} + A_b \cdot x_d(\tau)} \right] \cdot \frac{P_a(\tau) - P_b(\tau)}{R_i} \right) d\tau \\ &- \int_0^t \left(\frac{\beta \cdot A_b}{V_{b0} + A_b \cdot x_d(\tau)} \cdot \frac{P_b(\tau)}{R_e} \right) d\tau \end{aligned} \tag{16}$$

The position of the piston can be calculated by the acting force and the load. In case of a mass M (mass of piston m_d) – spring k – dumper c load and friction F_s, the movement of the system is:

$$(M + m_d) \cdot \frac{d^2 x_d(t)}{dt^2} + c \cdot \frac{dx_d(t)}{dt} + k \cdot x_d(t) = F_P(t) - \mathrm{sgn}(v_d) \cdot F_s \tag{17}$$

The complete nonlinear state space model of the electrohydraulic positioning system is:

$$\begin{bmatrix} \dot{x}_d \\ \dot{v}_d \\ \dot{P}_a \\ \dot{P}_b \\ \dot{x}_s \\ \dot{v}_s \\ \dot{i} \end{bmatrix} = \begin{bmatrix} v_d \\ -\frac{k}{M+m_d}\cdot x_d - \frac{c}{M+m_d}\cdot v_d + \frac{1}{M+m_d}\cdot\left(P_a\cdot A_a - P_b\cdot A_b\right) \\ \frac{\beta}{V_{a0}+A_a\cdot x_1}\cdot\left(C\cdot\sqrt{\frac{2}{\rho}\cdot\left(P_S-P_a\right)}\cdot x_s - A_a\cdot v_d - \frac{P_a-P_b}{R_i}\right) \\ \frac{\beta}{V_{b0}+A_b\cdot x_1}\cdot\left(-C\cdot\sqrt{\frac{2}{\rho}\cdot\left(P_b-P_T\right)}\cdot x_s + A_b\cdot v_d + \frac{P_a-P_b}{R_i} - \frac{P_b}{R_e}\right) \\ v_s \\ -\omega_s^2\cdot x_s - 2\cdot\zeta_s\cdot\omega_s\cdot v_s + \omega_s^2\cdot k_t\cdot i \\ -\frac{R_C}{L_C}\cdot i + \frac{K_a}{L_C}\cdot u \end{bmatrix} \tag{18}$$

This state space model can be used to examine the effect of leakage that appears in hydraulic cylinder.

4 Effect of Internal Leakage on Dynamic Performance

In the case of a worn piston seal, the internal leakage causes an oil flow decrease in chamber "a" and an oil flow increase in chamber "b". The external leakage causes the oil to flow out of chamber "b" and decreases the oil flow returning to the tank. The dynamic behavior of hydraulic positioning system depends on the pressure difference of hydraulic cylinder chambers and control force generated by these pressure values. Based on the nonlinear mathematical model of the electrohydraulic servo positioning system the following relation can be observed between the internal leakage resistance and chamber's pressure values:

$$\frac{dP_a}{dt} = \frac{\beta}{V_a}\cdot\left(Q_a - \frac{dV_a}{dt} - \frac{P_a-P_b}{R_i}\right) \tag{19}$$

$$\frac{dP_b}{dt} = \frac{\beta}{V_b}\cdot\left(-Q_b - \frac{dV_b}{dt} + \frac{P_a-P_b}{R_i} - \frac{P_b}{R_e}\right) \tag{20}$$

The differential equations of chamber pressure values clearly show that, while the pressure difference *Pa-Pb* of chambers is low, the effect of internal leakage is negligible as well. When the pressure difference and consequently the net acting force starts to increasing, the internal leakage oil flow also increases [22]. This greater leakage, decreases the pressure alteration speed in chamber "a" according to Eq. (19):

$$\frac{dP_a}{dt} = \frac{\beta}{V_a} \cdot \left(Q_a - \frac{dV_a}{dt} - \frac{P_a - P_b}{R_i} \right) \tag{21}$$

This greater internal leakage causes contrary effect in chamber "b" based on Eq. (20):

$$\frac{dP_b}{dt} = \frac{\beta}{V_b} \cdot \left(-Q_b - \frac{dV_b}{dt} + \frac{P_a - P_b}{R_i} - \frac{P_b}{R_e} \right) \tag{22}$$

The internal leakage modifies the acting force dynamic through its effect on chamber's pressure values. According to Eq. (14) the differential formula of the acting force can be written as:

$$\frac{dF}{dt} = \frac{dP_a}{dt} \cdot A_a - \frac{dP_b}{dt} \cdot A_b \tag{23}$$

Based on this formula, it is recognized that the pressure alteration speed in chamber "a" decreases while, on the opposite side, in chamber "b", this speed increases, resulting in the overall decrease of the acting force alteration speed. These effects impair the dynamic controllability of the acting force. Differential equation of the acting force:

$$\begin{aligned} \frac{dF(t)}{dt} = & \left(\frac{\beta \cdot A_a}{V_{a0} + A_a \cdot x(t)} \cdot Q_a(t) + \frac{\beta \cdot A_b}{V_{b0} - A_b \cdot x(t)} \cdot Q_b(t) \right) \\ & - \left(\left[\frac{\beta \cdot A_a^2}{V_{a0} + A_a \cdot x(t)} + \frac{\beta \cdot A_b^2}{V_{a0} + A_b \cdot x(t)} \right] \cdot v(t) \right) \\ & - \left(\left[\frac{\beta \cdot A_a}{V_{a0} + A_a \cdot x(t)} + \frac{\beta \cdot A_b}{V_{b0} + A_b \cdot x(t)} \right] \cdot \frac{P_a - P_b}{R_i} \right) \\ & + \left(\frac{\beta \cdot A_b}{V_{b0} + A_b \cdot x(t)} \cdot \frac{P_b}{R_e} \right) \end{aligned} \tag{24}$$

The complex mathematical formula of acting force shows the effects of each factor. The third part of this form describes the negative effects of internal leakage. Based on this negative effect of internal leakage, the control speed of acting force decreases, this impacts the acceleration of mechanical system adversely. This decelerating factor causes a damping effect in the dynamic response of this positioning system.

5 Experimental Results

The Laboratory electrohydraulic test system at Óbuda University Donát Bánki Faculty of Mechanical and Safety Engineering, was used to examine the leakage effect on the behaviour of this hydraulic positioning system. The laboratory test hydraulic system and computer-aided control and data acquisition unit are shown in Figures 8-10.

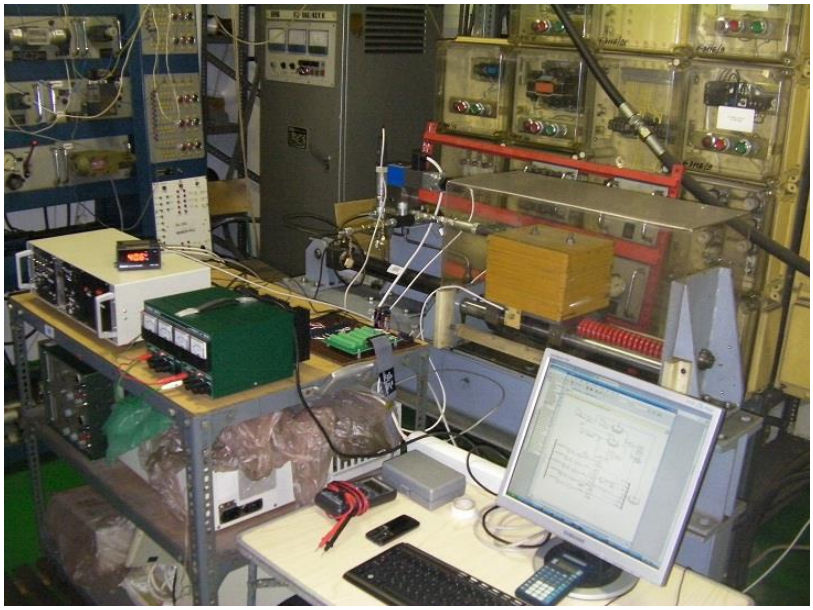


Figure 8
Electrohydraulic positioning system and computer control

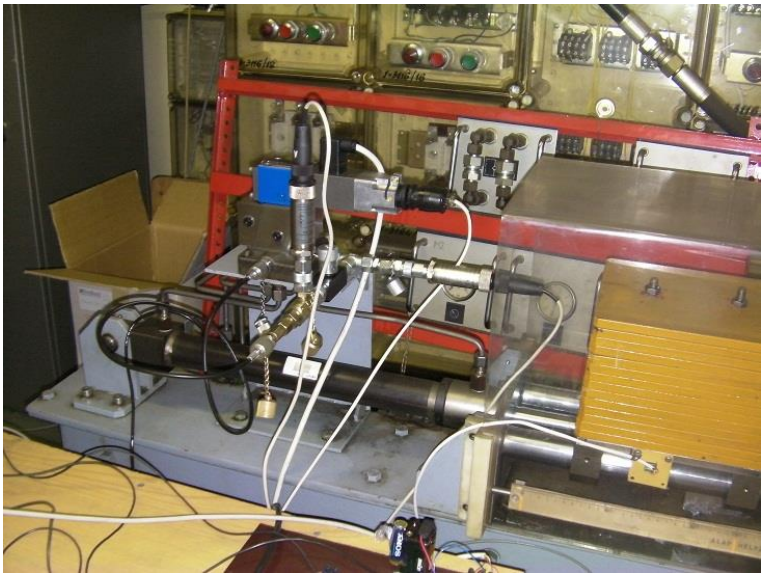


Figure 9
Hydraulic cylinder and proportional valve with pressure transducers

Figure 10
Mass-spring mechanical load of the positioning system

The hydraulic circuit diagram of the electrohydraulic test positioning system is shown in Figure 11.

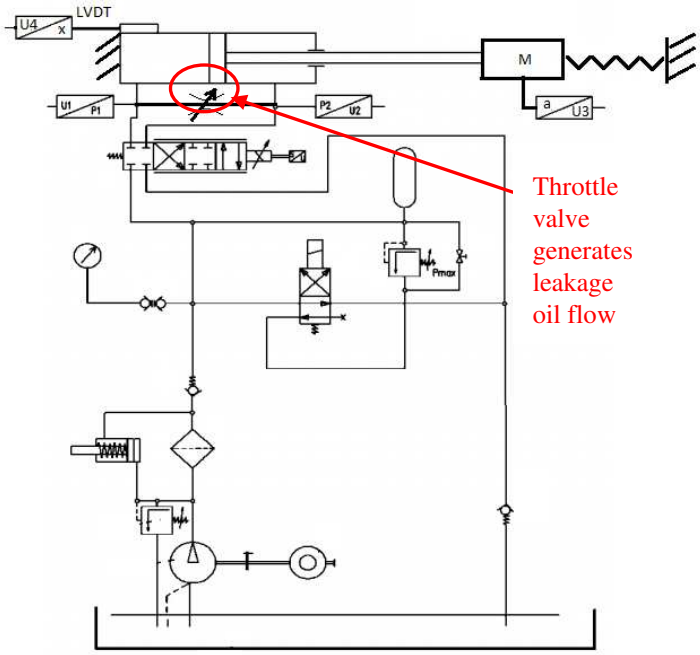


Figure 11
Hydraulic circuit diagram of the electrohydraulic positioning system

Parameters of the positioning system are included in Table I.

Table 1
Contains the result of comparing in pairs with the final result

Component	Type	Main parameter
Asymmetric, double acting hydraulic cylinder	Hagenbuch AG KWS 40/28 - 400	L=400 mm D=40 mm^2 d=28 mm^2 Pos.: I: 4…20 mA
Bosch servo valve with linear magnet	4WRP H6C4B 10L 1X/G24Z4/M 0 811 404 112	Pmax=250 bar Qn: 10 l/min (ΔP=70 bar)
Bosch valve drive	RV45 ±10V	U: -10…10V
Pump	Hydromatik GmbH A2F 10 R4P1	Vg=9,4 ml/rot
Accumulator	Hydac FAB NR 3 226 C 27198 1L	V=1 l
Precision throttle valve	Parker 9N600S4F	Qn=30 l/min Amax=0,22 cm^2

Pressure gauge 1	Hydac Druckfermer Typ. 905632	P=0…100 bar Iki=4…20 mA
Pressure gauge 2	Hydac Druckfermer Typ. 905934	P=0…400 bar Iki=4…20 mA
Accelerometer	Pololu MMA7341L	a=±3/11 g E=440mV/g
Mass	-	M=60 kg
Spring	RECOM SZ 8030 50x254	D_h=50 mm L_0=254 mm k=89 N/mm
Hydraulic oil	HIDROKOMOL P-46	ρ=0,872 g/cm^3 (T=15 °C) ν=46 mm^2/s (T=40 °C)

The internal leakage was generated by a precision adjustable throttle connected parallel with the hydraulic cylinder. An accelerometer provided usable information about the dynamic movement of the positioning system and pressure sensors are used to observe the pressure in actuator chambers.

The Bosch Rexroth industrial servo position controller was changed to a PC aided control with MATLAB and Real-Time Windows Target runtime environment for position control algorithm and National Instruments NI PCI6251 unit for data acquisition. The next figure (Fig. 12) shows the curves of observed displacement parameters in case of start-stop:

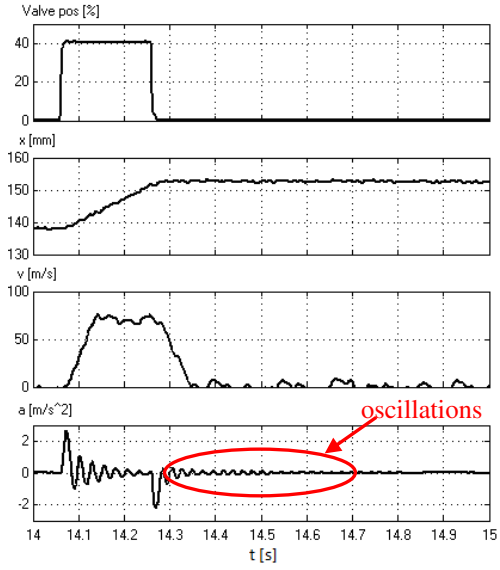


Figure 12
Observed displacement parameters of electrohydraulic positioning system

The mechanical oscillations of the electrohydraulic servo system can be detected more accurately by the accelerometer signal. The oscillations were monitored at quick stop of the system at different throttle valve values, therefore at different leakage resistance [20]. The settling times (ε_a=0,1 m/s^2) were calculated and can be seen in Figure 13.

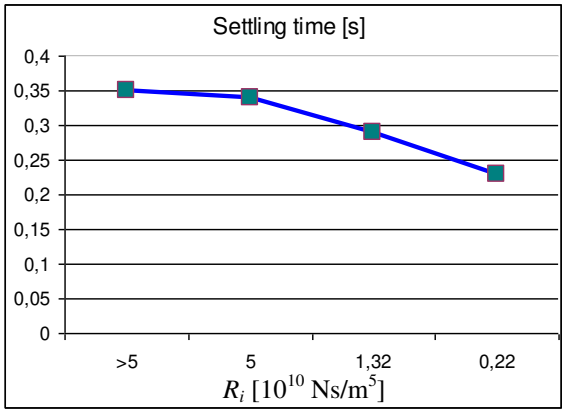


Figure 13
Settling time at different leakage resistance

The decelerating feature of the electrohydraulic servo positioning system at decreasing leakage resistance/increasing leakage oil flow can be seen in Figure 12. When the leakage oil flow increases the settling time of the system is reduced. This effect means that leakage causes a higher dumping factor of the electrohydraulic system.

6 Discussion

The results of an electrohydraulic servo system presented in this paper, can be observed from the point of effect of internal leakage. A complex nonlinear mathematical model was used to see into the background of the effect of internal leakage and how the dependency of the system dynamic response. Laboratory test verification presented in this paper was carried out to verify the conclusions drawn from the mathematical model of the electrohydraulic servo positioning system and the relation between internal leakage and dynamic behaviour of the system.

Based on the description and interpretation of the mathematical model, it can be concluded that the presented model can be adaptable in the design of leakage fault detection method for electrohydraulic servo systems. Early leakage detection could help in the planning of the preventive maintenance of servo hydraulic

actuators, which is a very important aspect, in the case of a continuously working production unit.

Conclusions

In this paper, a detailed nonlinear complex mathematical model was presented for an electrohydraulic servo positioning system. This complex model includes the effects of leakage, based on Merrit's laminar flow leakage model. Using this mathematical representation of the electrohydraulic servo system, an obvious effect can be demonstrated between internal leakage oil flow and the chamber's pressure values and acting force of the hydraulic cylinder. It can be concluded that increased internal leakage causes a negative effect on the dynamic performance of the positioning system with an increased damping feature.

Experimental results of internal leakage were examined in the laboratory for hydraulic tests at Óbuda University Donát Bánki within the Faculty of Mechanical and Safety Engineering. Damping features of the system were observed by the calculated settling time of the system's oscillations. It can be seen that increased leakage leads to lower settling times in free oscillations of the system, because the system's damping effect is increased.

References

[1] Tan Kok Kiong, Andi Sudjana Putra: Drives and Control for Industrial Automation, Advances in Industrial Control, Springer-Verlag London, 2011

[2] George W. Younkin: Industrial Servo Control Systems: Fundamentals and Applications, Second Edition, Marcel Dekker Inc. New-York, 2003

[3] Hubert Maxwell James: Nathaniel B. Nicholas, Ralph Saul Phillips, Theory of Servomechanism, McGRAW-HILL Inc., 1947

[4] Philco Technological Institute: Servomechanism Fundamentals and Experiments, Prentice-Hall, 1964

[5] Karl-Erik R.: Hydraulic Servo Systems, Linköpings universitet, TMHP51, 2008

[6] Jelali M., Kroll A.: Hydraulic Servo-Systems, Modelling, Identification and Control, Springer-Verlag London, 2003

[7] Anderson W.:Controlling Electrohydraulic Systems, Marcel Dekker Inc. New-York, 1988

[8] Backe W.: The Present and Future of Fluid Power, Proceedings of the Institution of Mechanical Engineers, Part I: Journal of Systems and Control Engineering, Vol. 207, 1993, pp. 193-212

[9] Richard P.: DSP Control of Electro-Hydraulic Servo Actuators, Texas Instruments, 2005

[10] Merritt, H. E.: Electro-Hydraulic Servo Valve Construction, Models and Use, Hydraulic Control Systems, John Wiley & Sons, 1967

[11] Edvard D., Uros Z.: An Intelligent Electro-Hydraulic Servo Drive Positioning, Journal of Mechanical Engineering, Vol. 57, 2011, pp. 394-404

[12] William J. T.: Specification Standards for Electrohydraulic Flow Control Servovalves, Technical Bulletin, Vol. 117, 1962

[13] Isermann R: Supervision, Fault-Detection and Fault-Diagnosis Methods - an Introduction, Control Engineering Practice, Vol. 5, 1997, pp. 639-652

[14] Le T T, Watton J, Pham D T: Fault Classification of Fluid Power System Using a Dynamic Feature Extraction Technique and Neural Networks, Journal of System and Control Engineering, Vol. 211, 1998, pp. 307-317

[15] Shi Z, Gu F, Lennox B, Ball A D: The Development of an Adaptive Threshold for Model-based Fault Detection of a Nonlinear Electro-Hydraulic System, Control Engineering Practice, Vol. 13, 2005 pp. 1357-1367

[16] Halnay A, Safta C A, Ursu I, Ursu F: Stability of Equilibria in a Four-dimensional Nonlinear Model of a Hydraulic Servomechanism, Journal of Engineering Mathematics, Volume 49, Number 4, 2004 pp. 391-405

[17] A. Kovari: Influence of Cylinder Leakage on Dynamic Behaviour of Electrohydraulic Servo System, Proceedings of 7th IEEE International Symposium on Intelligent Systems and Informatics, Subotica, Serbia, 25-26 September, 2009, pp. 375-379

[18] A Kovari, D Fodor: ARX Model-based Fault Detection of Rolling Mill's Automatic Gauge Control System, Proceedings of 15th IEEE International Power Electronics and Motion Control Conference, Novi Sad, Serbia, 4-6 September, 2012, pp. DS1d.6-1-DS1d.6-6

[19] Attila Kővári: Observer-based Leakage Detection of Hydraulic Cylinder Applied in Rolling Mills Electro-Hydraulic Gap Adjustment System, Materials Science, Testing and Informatics VI, Materials Science Forum, Vol. 729, 2013, pp. 424-429

[20] A Kovari, D Fodor: Identification-based Leakage Detection of Hydraulic Capsules, Proceeding of the 14th IEEE International Symposium on Computational Intelligence and Informatics, Budapest, Hungary, 19-21 November, 2013, pp. 419-422

[21] Kevan Slater: Detecting and Managing Hydraulic System Leakage, Machinery Lubrication, 2001/7

[22] A Kovari: Dynamic Model of Rolling Mill's Electro-Hydraulic Gap Adjustment System, Materials Science, Testing and Informatics V, Materials Science Forum, Vol. 659, 2010, pp. 411-416

[23] Besancon-Voda, A.: Iterative Auto-Calibration of Digital Controllers - a Systematic Design Approach and Case Study, Control Engineering Practice, Vol. 6, Num. 3, 1998, pp. 345-358

[24] Radu-Emil Precup, Stefan Preitl: PI and PID Controllers Tuning for Integral-Type Servo Systems to Ensure Robust Stability and Controller Robustness, Electrical Engineering, Vol. 88, Issue 2, 2006, pp. 149-156

[25] Samuel John, Jimoh O. Pedro: Neural Network-based Adaptive Feedback Linearization Control of Antilock Braking System, International Journal of Artifical intelligence, Vol. 10, Num. S13, 2013

Introducing the Information Security Management System in Cloud Computing Environment

Laslo Tot

Singidunum University, Danijelova 32, 11000 Belgrade, Serbia
E-mail: laslo.tot.10.dls@singimail.rs

Gojko Grubor

Sinergija University, Raje Baničića bb, 76300 Bijeljina, B&H
E-mail: ggrubor@sinergija.edu.ba

Takacs Marta

John von Neumann Faculty of Informatics, Institute of Applied Mathematics, Óbuda University, Bécsi út 96/b, 1034 Budapest, Hungary
E-mail: takacs.marta@nik.uni-obuda.hu

Abstract: Numerous organizations coordinate and certify their information security systems according to the Information Security Management System (ISMS) standard. Available Cloud Computing Services (CCSs) include new types of vulnerability (management, virtualization, sprawl, etc.) and differ in management requirements from other computational systems. Establishing a consistent security management framework (SMF) and information security management system (ISMS) in CC environment is a complicated, demanding and time-consuming process. Every experience from applying ISMS standard solutions is certainly useful, but not enough to entirely cover all security requirements of the customers and Cloud Service Provider (CSP). Attempts of establishing an integrated and consistent SMF and ISMS in CC environment have not been researched in-depth in recent available literature. In this paper, authors suggested a framework for an establishing quality management system (QMS) of CCSs, including CC SMF and CC ISMS, proactive digital forensic (DF), proactive and predictive security controls and corporate DF investigation process up to the level specified in Service Level Agreement (SLA).

Keywords: CC ISMS; CC SMF; Proactive Digital Forensic; Digital forensics; CCS QMS

1 Introduction

A manager of the clients' information security system should consider the nature of business, degree of information control and security risks within a Cloud Computing (CC) environment, in order to establish consistent and quality management system (QMS), and an information security management system (ISMS). Establishing ISMS in a CC system should gather all of the solutions and experiences in related technologies that are included in CC development. Establishing a Security Management Framework (SMF) and the methods of security risks management are basic requirements for designing all of the ISMS information systems, including a CC one. Establishing ISMS in a CC environment mostly depends on the model of the CC system and the type of the CC service. The greatest number of the ISMS specifications refer to the public CC system, which is, in fact, the subject of this paper. In the model of the public CC system and Software-as-a-Service (SaaS), multi-tenancy is the key to success. An effective and fully integrated simple template-based model transformation and synchronization approach has been proposed for the development of SaaS multi-tenant applications [14]. In the Platform-as-a-Service (PaaS) type of CC service, a user works with their data and with applications on hired platform from Cloud Service Providers (CSP). In the Infrastructure-as-a-Service (IaaS)-type of CC service, the infrastructure is entirely owned by CSP, and the user installs their applications on hired platforms. In most realistic scenarios, the end user is responsible for using hardware and software in the CC environment. In such cases, focus is on security specifications of the main data in the CC environment and on influence of the virtualization and management on establishing SMF and ISMS in CC environment including improvement of the CCSs QMS.

Related Works Review

According to the available references many authors focus their research on the technology solutions for data security of the clients in CC environment, such as: Distributed encryption system [13]; Multi-layer security (network, host and application levels) [1, 2, and 14]; Technical aspects of potential digital forensics in distributed Cloud environments by Dominik Birk, 2011; Possibility of digital signature application in the Cloud suggested by Doug Bannister, Omnivex Insight, 2011; Use of honeypot technology, its advantages and disadvantages and value to the Cloud security proposed by Nithin Chandra S. R and Madhuri T. M. in their article, Cloud Security using Honeypot Systems, International Journal of Scientific & Engineering Research Volume 3, Issue 3, 2012; Applications of end point security in virtualized environment written by Darren Niller, 2002; Detailed analysis of all aspects of CC security and lack of trust in the Cloud, regarding confidentiality, integrity and availability of information, privacy and auditability, and DF investigation, by authors Bharat Bhargava, Anya Kim, and Youn Sun Cho in their article, Research in Cloud Security and Privacy, 2012; Possibility of traditional DF investigation in CC environment suggested by Gartner group, in the

article, Assessing the Security Risks of Cloud Computing, in June 2008, etc. These are some of many technical solutions suggested for improvement of data and privacy security in the CC system.

Generally CC security refers to a broad set of policies, technologies and highly distributed security technical and procedural controls. After all, CC security is an evolving sub-domain of computer security, network security, cyber security and eventually information security. Authors of this paper contribute by assessing and collecting all relevant aspects of the CC system security, and by suggested CC SFM and CC ISMS including not only technical security controls (proactive and predictive), but also procedural controls and contractual obligations to overcome lack of trust in data security on both sides – CSP and customers. Vulnerabilities in virtualized CC environment are quite well described by IBM and many authors [4, 5 and 13], etc. The authors of this paper suggested a comprehensive framework of CC SMF and CC ISMS, including SMF and ISMS recommended by standards ISO/IEC 27001/2:2013, to improve CC management system and QMS of the CCSs and security and privacy of data, and to increase mutual trust in CC environment. Furthermore, the authors suggested the overall strategy of CC QMS including security best practice principles, real time managing risk, continuous monitoring, and proactive and predictive[1] security controls to increase security and enable DF investigation in CC systems up to certain level specified in the SLA.

2 Virtual Infrastructure Management

A basic driver in the development of a CC service is virtualization, which most usually refers to virtualization of computers' and networks' hardware and it is the basis of public CC system configuration. In this way, different applications could be applied on different virtual platforms, but on the same physical machine [12]. The complexity of the virtual infrastructure management includes the following:

- Mandatory control of availability and usage, and approach to the physical and virtual resources;
- Implementation of the solutions for special situations, and
- Security of the customers' virtual machines (VM), tasks, monitoring and reporting on the CC centre usage.

In practice, there is a lack of suitable tools and theory for discovering errors and performances analysis of the ISMS virtual infrastructure. Relative independence of the numerous created VMs from physical machines, results in so-called liquid

[1] Security control that includes artificial intelligence mechanism to detect, prevent and predict attacks.

computing, better known as the VM sprawl [19]. In order to avoid the security risks from VMs sprawl, a management tool limits the creation of new VMs by assigning authorizations. Further, it provides a strong monitoring system which reports about allocated, but unused VMs. Because of the requirements for balancing load of the physical machine in the CC centre, locations of the created VMs and evaluation of the virtualized CC centre efficiency are specific problems.

Consistent ISMS in a CC environment requires a strong monitoring system of the virtual infrastructure that includes monitoring of physical hosting servers, virtual machine monitors (VMM) or hypervisor, and VMs and applications that operate on them. Virtual machine Introspection (VMI), as the process by which the state of a VM can be observed from either the VMM or from other VMs, is a new security risk factor to customers' data. That is why VMM has full access to the resources of all VMs and if it is compromised the customers' data can be misused, too. This phenomenon remains an open research problem.

Monitoring of the physical servers is very important, because of hosting a great number of VMs, and the responsibility of the CSP. The CSP uses not only some of the known software tools, such as Open Manage and IT Assistant (Dell), Open View (HP) or Nagios (open code), but also their own hardware tools with better performances. There is no fundamental difference between the meaning of the terms of VMM and hypervisor. A VMM monitors and directs the VM, and the hypervisor refers to the function of a machine's kernel supervision, while in CC it means that the kernel supervises more than one VM (hypervisors). Therefore, in a CC system, the hypervisor/VMM could be considered the kernel of a virtual infrastructure. Some of the more famous hypervisors are VMware ESX, which includes web interface for ESX – MUI (Management User Interface) for monitoring and management of the current VMM usage, Hyper V in the server Windows 2008 R2, Windows 2012 OSs, etc. Usually, for the monitoring of applications on a VM and for applications on the physical machine, the same tools are used, but the difference is in the management of system usage from the application side.

3 Cloud Computing System Management

In theory, as well in practice of computer science, there are many definitions of service management [2], such as:

- The capacity of the organizations to deliver services to clients
- The set of special organizations' capacities for giving additional values to customers in the form of services
- More than a set of capacities, because it includes a professional team of experts with specific knowledge and skills.

The service delivery model in a CC environment refers to a set of more customers, assets, resources and capacities with flexible payment only, for what was used.

Management of the CC system services is an inherent vulnerability in a CC environment, not only because of the virtualization and accompanying vulnerabilities, but also because of certain other reasons. The main reason for the new approach to the CC system services management lies in the enormous increase of digital information in the world, and an ever-more difficult and more complex management of the so-called digital universe and big data. The organization IDS.com (Intelligent Document Solutions) predicted that the digital universe would grow from 1.2 million PB (petabytes, 1 PB = 1.000 GB) or 1.2 ZB (zeta bytes) in 2010, up to 35 ZB in 2020 [9]. As the CC is already becoming an integral part of the digital universe, more than 34% of the total digital data in the world will be stored or secured with CC services. Until 2020, CC services will become even more important, due to the following reasons [9]:

- They will be cheaper, and the economy of organizations will initiate innovative development
- They will become much more important for individual users because of mobile access possibility
- They will include advanced services (data compression, de-duplication, etc.) for easier management of the big data. They will become part of the solution for more efficient management of the digital universe. In a CC environment, management of the performances and errors should be entirely automated, sophisticated and analytical. The CC operator should adjust, integrate and coordinate management of the all CC functions. Management of the infrastructure change and configuration in the CC system requires fast reaction and the Configuration Management Database (CMDB) must be used for better performance. For the effective management of the CC server's change and configuration, the CMDB must be embedded in this process and integrated with other controlling operations.

The tools for management in physical, virtual and CC environment are comparatively shown in Table 1 [2].

Table 1
Main management activities in physical, virtual and CC environment

Functions of the management	Physical servers	Virtual servers	CC environment
Management of the performances and errors	Manual or procedural approach to the monitoring and management of the resources, based on the events	Greater automation Access based on models and focused on services and applications	Entirely automated, sophisticated and analytical Operator adjusts, integrates and coordinates management functions

Functions of the management	Physical servers	Virtual servers	CC environment
Management of the infrastructure: change and configuration management	Manual processes on demand and change management according to the plan	For better performance a CMDB could be used	Fast change management For better performance, a CMDB must be used
Management of the server: change and configuration management	Manual processes on demand Change management according to the plan	For better performance, a CMDB could be used	Change management requires speed For better performance, a CMDB must be embedded in the process and integrated with other controlling operations

3.1 Infrastructure Management in Cloud

Holistic method of the integrated management of the physical and virtual resources (Iaas) is the main problem of the CSP. The main goal of the CSP management is to secure dynamic and fast provisions of the resources for CC applications. For this purpose, there is a tool already in development, the so-called Manager of Virtual Infrastructure (VIM) [4], which is usually called CC OS (Operating System) [23] or CC engine[2]. This type of software resembles a traditional OS, but instead of managing just one computer, the VIM aggregates resources of numerous computers with a uniform view at users and applications. Two following categories of the tools for management of IaaS were suggested [2]:

- Tool sets for the CC system that provide a distant and secure interface for creating, controlling and monitoring the virtualized resources, but it is not as efficient for VIM as specialized tools.
- The VIM tools with automatic balance of load, consolidation of servers, but without an interface for distant access to the CC system.

Both categories of tools are basic tools for the CC ISMS, or management of VM's vital cycle, but they are not applicable for implementation into all CC systems. The main difference between these two categories of tools is that the CC tools have an interface for remote access and management of the great number of consumers' accounts.

[2] *Amazon Cloud Computing*, www.amazon.com

3.2 Relationship Management among Cloud Service Provider and Consumers

The concept of classic relationships of the CSP and consumers is shown in Figure 1 [2].

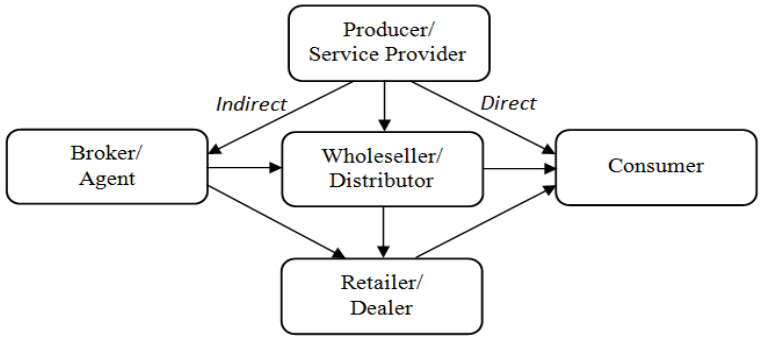


Figure 1
Diagram of classic relationship between the CSP and service consumer [2]

In the CC environment, the CSP performs many activities for service delivery, such as: defines the strategy, designs, invests, implements, transmits the data and operates within the CC infrastructure. The CC services increase the value of the customer's product that pays for what they use and they do not accept responsibility for expenses of service delivery and its risks. To ensure satisfaction of the customers with the quality of the CC services, a CSP must consider numerous factors, such as: nature of business and profiles of the target users, value of the CC services for users, form of use and payment for the services, security of the user's information, possibility of digital forensic (DF) investigation in the case of some incidents, etc. The CSP can increase the value of services on different levels, using different equipment, specialised technologies and additional services, etc.).

Quality of the CC service includes two basic elements: *benefit,* or what the customer gains, and *assurance* of the service delivery – available on time, with sufficient capacity, and continuous and secured. Service value for the customers is a combination of these two elements. Estimation of the CC service quality mostly depends on the two component perceptions: *expected value* – service level that the user expects, and *empiric quality* – experienced service level estimation of the user. There are various models for payment of the CC service, but payment for the used items are mostly applied (e.g. expenses for electricity and water bills, etc.), and subscription (e.g. on an annual level). Security of the user's information is the key factor of distrust in the CC systems. In practice, it is defined as the share of responsibility between the customers and the CSP, through the Service Level Agreement (SLA). In the case of incidents or crime in the CC environment, DF

investigation is naturally expected, but it is almost impossible to perform in practice. Implementation of the proactive forensics [7, 21, and 22] and the definition of some level of forensic investigation in the SLA and CC SMF documents are likely to yield solutions for the potential DF investigation in the CC system. They should be applied to the CC Service's Development Life Cycle (see Figure 2).

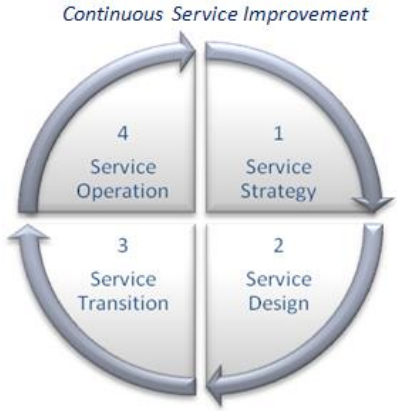


Figure 2
Development Life Cycle of the CC Services

3.3 The Importance of Introducing Proactive Forensics into CC SMF and CC ISMS

Proactive forensics could be defined as a preventive step for locating and collecting key forensically relevant data. In the CC system, with multi-tenancy VMs, the rate of overwriting data is much larger than in an ordinary computer. What the customers and CSP do, and who is responsible to collect forensic evidence proactively is the key question? Besides, it is doubtful whether classic (or post-mortem) DF investigation can be applied in CC system at all [11]? Generally, post-mortem forensic process has four phases: data access and identification, data acquisition, data analysis and evidence reporting. Access to the data sources and collection is the first step. Taking forensic images to the forensically sound media is the next step. In the analysis phase, the forensic image is installed by outside booting processes (usually from forensic tools) onto forensic computers for analysis. If a forensic image is booted onto a physical machine with different hardware configuration, the OS will discover those differences and try to install the missing drivers into image and contaminate it. Obviously, it is unacceptable according to the principles of the DF science. In the CC virtualized environment this approach, including even live forensics, is not sufficient to recreate the original environment. Actually, VM simulates some basic

hardware components and does not support a wide range of hardware devices. Therefore, a forensic image cannot be installed directly into a VM that requires some added files with information on the booted system. Some forensic tools could resolve this problem, including EnCase Physical Disk Emulator (commercial), ProDiscover family (commercial), Live View (free), and others [11].

A clause in the SLA, on some level of forensic investigation, could be a practical solution for the first time. It can provide an obligation for the CSP to proactively collect forensically relevant data. But the question is: who is to be trusted with discovering and collecting data, and identifying the nature of the incident? The security team, without forensic techniques, tools, knowledge and experience may compromise the evidence. Hiring of a third-party consultant could be unacceptable, too. A well-trained forensic investigator in the organization's security team can provide the following services:

- Electronic information discovery
- Forensic and technical consulting
- Storage media forensic investigation
- Proactive network digital forensic
- Internet forensics
- E-mail and other messaging systems forensics
- Incident nature identification
- Data collection and registration
- Malware risk reduction.

Generally, proactive forensics, included into CC SMF and CC ISMS, is a promising concept. It provides protection to an organization's responsibility, competency advantages and forensic evidence collection for future forensic investigations. In the CC ISMS, a strong monitoring system should be a mandatory proactive measure, including monitoring of the CC centres, CCSs availability, access and usage of VMs, data security, security emergency solution, flexibility for new VMs and new task additions. It is a good protection from VMs sprawl vulnerability, too. Monitoring of the VMs in the CC system includes physical servers, VM monitors or hypervisors and applications running on the VMs. Proactive forensic processes of the CC monitoring system include the generation, collection, analysis and reporting of the forensically relevant events instead of huge volumes of security events. It means that security problems could be stopped before critical escalations [11].

In general, computer incidents' proactive forensics in an organization includes proactive identification and collection of digital traces, investigation and malware

risk mitigation. Eventually, in the CC system, security monitoring has a much greater importance since, it simultaneously provides a real-time security monitoring alarms, reliable CC and security system functioning, and fast reaction against incidents. Defining and setting forensic evidence selection criteria is the main challenge in the networks and Internet forensics, including Cloud forensics, too. In the CC environment, Cloud forensics [13] means an application of the DF investigation as a subset of the CC system, computer forensic, and network and Internet forensic (Figure 3).

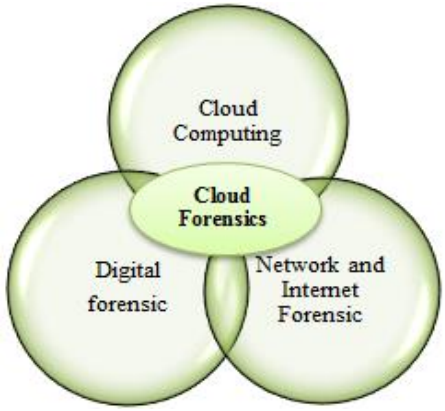


Figure 3
Place and role of cloud forensic

The summary of the critical success factors (CSF) of Cloud forensics is shown in Table 2 [11].

Table 2
Summary of the *Cloud forensics* critical success factors

Cloud forensics CSF	**Descriptions**
Management quality	Provide all of the stakeholder with cooperation and support to the DF investigation [9]
SLA	Oblige CSP on certain level of DF service in the CC systems [15]
Proactive forensics	Provide the traces of malware and direct attacks collection to the CC systems log files of network and security devices [22, 1 and 11]
Anti-forensic tools	Disable anti-forensic activities in the CC systems [11 and 9]
Managers' responsibility	Use ISMS standard (ISO 27001) to assess individual mangers' responsibilities [13]
Data source quality verification	Develop method and criteria for forensic incident data source quality evaluation [8]
Mobile devices forensic	Provides specific requirements for mobile devices forensic and balance them to the requirements for privacy protection [11]

Monitoring the key CC service performances indicators requires metrics establishing in the design phase (see Figure 2) [6]. Examples of the metrics are as follows:

- Service performances according to business and strategic plans
- Financial gain to the business
- Monitoring of the key ICT processes supporting CC services
- Reporting on the level of services Customers' satisfaction, etc.

The final criteria that require many activities to achieve wanted quality of the CC service management must be implemented before service delivery. They include investigation of the CC services acceptance against overload, users' activities, error tolerance, data recovery, network delay, and payment process.

3.4 Management of the SLA Agreement

Agreement between the customers and CSP on the SLA is the main document. It is based on the users' information security requirements and needs for DF investigation of computer misuse or criminal activities in the CC environment. In current practice, potential DF investigation in the CC environment requires a detailed SLA contract between the customers and CSPs that compels the CSPs on forensic investigation services up to the specified level. In this way, the customer knows what to obtain, or not, in the case of a DF investigation. Requests of the customers for this CC service should become a standard part of the SLAs to secure their vulnerable information into the CC system. This implies that the CSP should implement security controls that could enable proactive digital evidence collection and DF investigation later on [22]. Proactive forensics means taking precautionary steps that do not include the need for locating key evidence ("smoking gun") in the DF investigation process. Some CSPs have already offered DF investigation up to the certain level, as a CC service [18].

3.5 Management of the Anti-Forensic Tools and Activities

Legal processes for the DF investigation are increasingly becoming mandatory services in the ISMS systems [10 and 17]. Generally, computer criminals use numerous activities, techniques and tools to disrupt DF investigation process. Apart from classic hacking techniques for erasing of digital traces generated by computer (such as PC Cleaner, Evidence Eliminator, etc.), setting of time bombs, applications of the users for shut-down, etc., some anti-forensic techniques are also popular, including those mentioned in [23 and 16]:

- Modification of the time seal
- Header change and file extension, extracting hash password values without leaving a trace

- Data mule types of attacks on the reserved sectors by the system on hard disk
- Generating random file names (detection based on discovery of file signature is avoided), etc.

3.6 Human Resources Management in Cloud Environment

Quality management of the human resources and cooperation of all stakeholders are the most important factors in the CC system management and for successful corporate DF investigation. Management of the CC system that includes the CSP and customers must explicitly support teams for corporate DF investigation [17]. The organisations of the CSP and customers must establish security policies and procedures for managing computer incidents and corporative DF investigations. In the case of braking into the CC security system, DF forensic expert has to lead DF investigation process and harmonise functions of the team combined by programmers, administrators, engineers of network equipment, etc. Digital forensic experts must provide professionals with various skills, job procedures and protection of digital evidence integrity in the chain of custody. All participants who have access to the forensic data and evidence must maintain records for every activity in these procedures [15].

Carelessness of the managers and neglect of security events are common reasons for numerous court cases. It is particularly easy to prove consequences of carelessness and neglect of the plan and conduction of the technologies management strategy. Documented international ISMS standards, ISO/IEC 27001:2013 and ISO/IEC 27002:2013 provide necessary instructions regarding organization, methods, tools, security controls and procedures for the potential CC ISMS. These standards should be supported by newly developed (or in development process) set of standards for DF investigation, such as ISI/IEC 27033, ISI/IEC 27034, ISI/IEC 27035, ISI/IEC 27037, ISO/IEC 27041, ISO/IEC 27042 and ISO/IEC 27043. They should be used for the ISMS implementation into the processes of CC information security systems, including DF investigative procedures, due to increasing forensic needs for estimating individual responsibility of the security technologies managers.

3.7 Functional Model of Introducing ISMS into CC system

For the reliable identification of the computer incidents, an uninterrupted verification of data source signs and indicators of incident is required [3]. Some standards of good practice in proactive detection of incidents in computer networks recommend the following activities [5]:

- Data users should develop and document the method and criteria for evaluating quality of incident data information source;

- Users and / or the CSP should verify information on an incident before importing them into the database or in a program for incident management, but without the process flow being delayed;
- Information on an incident should be in correlation with external services, refilled and filtrated in order to prevent event duplication, but without the process flow being delayed;
- If the source of information implements mechanisms of return links, the users should use it in the case of request for quality improvement of incident information;
- Data sources for threats and attacks (CERT/CIRT), should develop its sensor networks for malware detection and other incident indicators, implement honeypot and sandbox technologies at customer's side and, if it is possible, implement passive scanners for DNS monitoring [19].

In CC environment expected level of the company security and business continuity management may be achieved by real-time business risk assessment and by compliance to a wide range of security expectations, not only to recommendations of the information security international standards [21]. The basic principles of good practice for introducing SMF and ISMS into CC system are suggested by authors and summarized in Table 3.

Table 3
Basic principles of the SMF and ISMS best practices in CC system

Category	Basic principles of ISMS good practice in cloud system
Risk management	Key factor of the CC security, including identification, location, assessment and security risk mitigation, by choice and implementation of the best practice security controls in the CC environment (from the business point of view).
Information security policy	Includes security risk management methodology, scope and limits that change with the type of CC service and may be overlapping among SaaS, PaaS and IaaS policies; it should be updated and supported by standards, procedures and guides.
Control and change management	Includes procedural controls for configuration management (CM) and change control (CC); CM and CC automation is required in the CC environment.
Audit	Proves standards and security policy conciliation and effectiveness of security controls. Detects new vulnerabilities and requires both automated tools and manual procedures.
Vulnerability scanning	Includes all CC managing platforms, servers and network devices and identifies all new software, hardware and configuration vulnerability. Uses tools like Nessus, Back Track (Kali), etc.
Duty separation	Limits user privileges according to the "need-to-know" principle for duty performance.
Information assets	All information assets, including people, hardware, software

Category	Basic principles of ISMS good practice in cloud system
management	and network devices that make CC system infrastructures, must be protected.
Symmetric and asymmetric cryptography	Key management and digital certificates secure CSP infrastructure protection, but it is not applicable in the CC system.
Cryptography on demand	Proposed solution where customers are obliged to require cryptographic key from CSP before communication with the CCS.
Security of data storage	Identifies the need for encrypted protection of data storage and recognizes that some users need the special data storage's.
End point security	Secured by protection of the CC system resources, end points and access restriction to protocols and device types.
Network security	Secured by network traffic protection on the network devices and data package levels
Proactive forensic	Provides network incident prevention and strong real time monitoring, but lack of expensive and matured technologies, tools and knowledge are its limitations.
Access control and authorization	Secured by policy, strong authentication and effective access control implementation, depends on defined identity, role and privilege of the CCS users
Choice of the CSP	Should be based on CSP relations to the security and complementary security solutions with current organisational security practice.
Transparency	Choice of the CSP should be based on its readiness for transparency in key security items, including DF processes.
Security controls	CSP should implement proactive and predictive security controls to be acceptable and to fulfil user and legal requirements.
Security standards and practice	For the CCSs users the best practice is to define a comprehensive security program or CC SMF, including security polices for all CC service life cycle activities.
Network isolation	To provide isolation among different user services and network information flows, a CSP should separate the controls and flow of the network information.
Use of CMDB	It is most important for CC maturity and effective resources and automated CC system processes management.
Configuration integrity	Provided by software installation for the CC system configuration check and identification of system file platform, non-authorized modification is prevented.
Identity management	CSP must implement scalable and robust identity and access management systems, which establish strong identity authentication, authorization and user platforms management.

Introducing ISMS into CC system is more complex process than doing that in LAN or distributed computer network under full control of the users. However, best practice and experience from information security management system

(ISMS) according to the standards ISO/IEC 27001&2: 2013 could be used. Using system engineering knowledge and process approach, CC users can apply some process model such as PDCA (Plan, Do, Check, and Act), Six Sigma, CMMI, etc. The authors suggested PDCA model due to its simplicity and broader use in the many ISO standards. Implementation of the CC ISMS is the main purpose of the PDCA process model application. Functional model for ISMS implementing into CC environment, suggested by authors, is shown in Fig. 4.

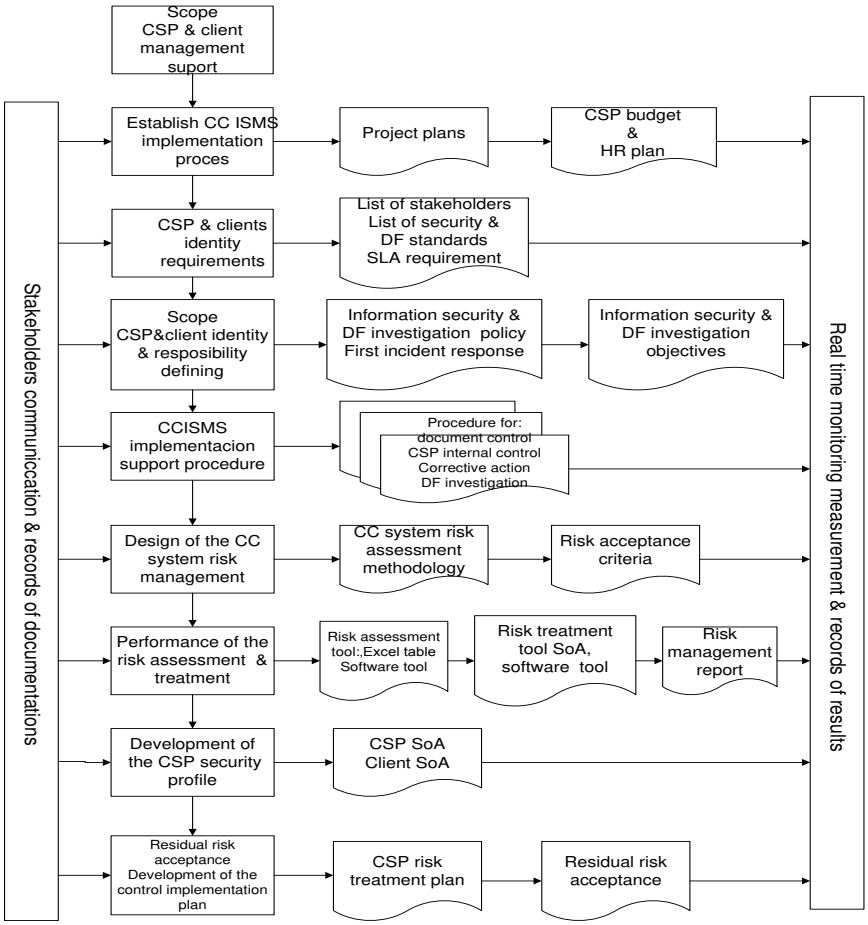


Figure 4
Functional model of the CC ISMS implementation (Part 1)

In Fig. 1 the CC system consists of the CC centre and part of the client ICT system that use service of the cloud service providers (CSPs), such as SasS, PaaS or IaaS. The detailed and completely new form of the SLA document is probably the most important concept in this model. The SLA document must include all

agreed upon activities by the clients and CSPs, throughout every step of the CC ISMS implementation process. In practice, the CSPs, as a rule, design and form SLA document, leaving more responsibilities for information security on the client side and without any word about possible incident, loss of the clients' information and potential digital forensic (DF) investigation. Many authors and organizations (such as Gartner group, NSA, Norton Symantec, etc.) suggested that the DF investigation in CC system has become almost impossible mission. The CSPs usually claims that security posture in CC centre is at highest level of the efficiency and effectiveness in the CC centre. However, attacks and penetration into CC system must be supposed, although there is no much information about that. That is why clients and CSPs must cooperate closely to prevent cloud services misuse, including stealing information of the clients. Functional model of the CC ISMS implementation process (see Fig. 4) is adapted by the authors from best practice standardized ISMS (ISO/IEC 27001:2013). Many phases, steps and activities must be accomplished by consensus of the clients and CSPs, such as: Scope determination and management support; Identity and responsibilities requirements; Establishment of the CC ISMS implementation process and project budget; Development of many documents, e.g. security policy and procedures, including level of the proactive DF infrastructure, etc. Most likely, closest collaboration should be in intention and responsibility defining, including information security and digital forensic policies, objectives and both side responsibilities and obligations. The main problem in this process could be the level of the potential DF investigation in CC system, in case of clients data lost in virtualized CC environment. As a first step a CSP can agree upon some level of DF investigation in CC system in the SLA document (Fig. 5). Implementing real time monitoring system and proactive network DF infrastructure can help to convince clients in highest level of the CC information security system. In developing security profile both sides must align their Statement of Applicability (SoA) documents, at least regarding residual risk acceptance.

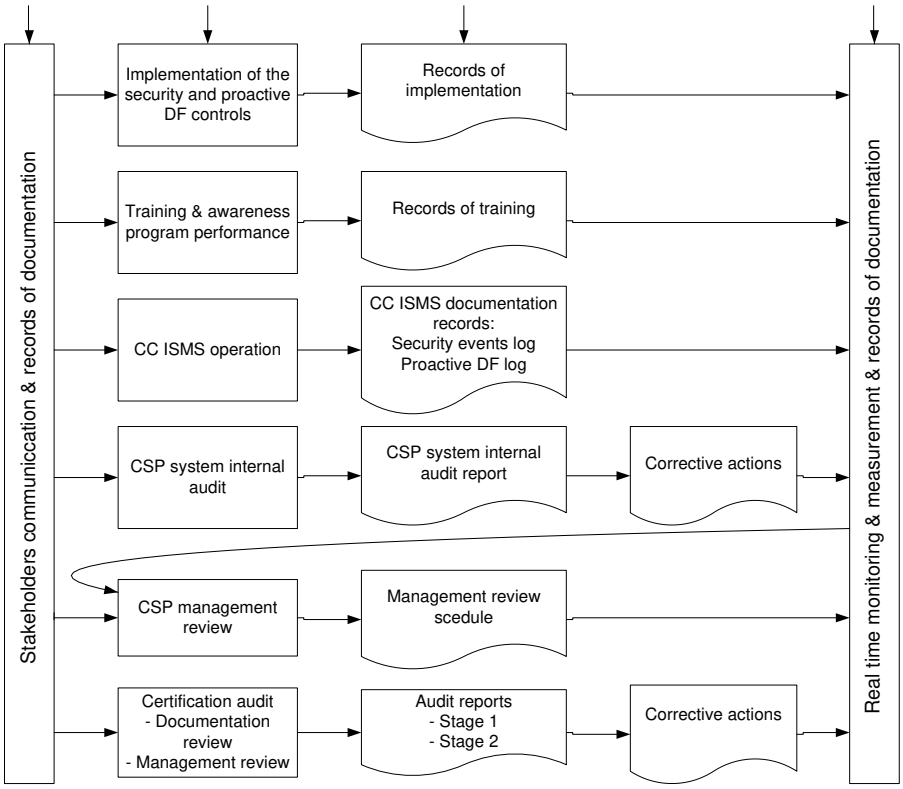


Figure 5
Functional model of the CC ISMS implementation (Part 2)

All other steps in Fig. 5 such as training and awareness, ISMS operation, and internal, management and certification audit, could be performed separately and independently at both side locations. Finally, certified ISMS on both sides could be a starting point to increase trust among clients and a CSP.

Conclusion

The CC system (CCS) is a modern computational type, including informative, human, legal, business, philosophical, and many other aspects. New types of CC vulnerabilities require abandoning the traditional container type protection concept. In the CC environment, especially for specific requirements, management of the CCSs life cycle, including the CC SMF and CC ISMS, must be established and implemented. In order to establish CC ISMS and CC SMF, principles and recommendations of the ISMS standards (ISO/IEC 27K) should be adapted and included in CC management system, and the traditional security management framework (SMF) should be enhanced by proactive and predictive security controls, proactive forensic infrastructure and mandatory DF investigation process

that is specified up to the certain level in the SLA. Quality of the CCSs could be provided by implemented both CC SMF and CC ISMS. Those are especially important due to further ISMS and security field investigations in the CC environment, open up new branches of smart grid computing [8] that uses more efficient network computational power, but requires enhanced information security, and proactive and even predictive protection from potential cyber-attacks. Therefore, security system in CC environment should be embedded, and DF investigation of computer crime in the CC environment should be enabled. Cloud forensic introduces a new approach to the traditional DF investigation methods and applications of the forensic tools and techniques. In this paper, a new framework of CC management system and potential use of proactive and predictive security controls, standardized ISMS and SMF, proactive forensic, and cloud forensic service embedded in the CC SMF and CC ISMS is suggested, in order to improve the QMS of the CCSs and to increase mutual trust in data security between CSPs and customers. Except of the CC security and management systems state-of-the-arts overview, comprehensive security principles and components suggested for establishing and implementing CC SMF, CC ISMS, and CCSs quality management system, are the main contribution of this paper. They are supposed to embed not only well-known security controls, but also some new ones, such as strong authentication, uninterrupted real-time monitoring, proactive network forensic, proactive and predictive security controls, embedded DF investigation infrastructure, etc.

The detailed process model and data flow information of the CC ISMS and CC SMF implementation process suggested by authors, should be subjected to the further research. A new system of legal and quality CC SMF and CC ISMS security policy and procedures should be formally described and defined, and standardized for more trusted, effective and efficient CC systems.

References

[1] Bradford, P. G., Hu, N: A Layered Approach to Insider Threat Detection and Proactive Forensics, 2005

[2] Buyya, R., Broberg, J., Goscinski, A., et al, Cloud Computing: Principles and Paradigms, John Wiley & Sons, Inc., 2011

[3] Cichonski, P., Millar, T., Grance, T. and Scarfone, K., Computer Security Incident Handling Guide, NIST Special Publication 800-61, Rev. 2 Recommendations of the National Institute of Standards and Technology, 2012

[4] ENISA (European Network and Information Security Agency) Report, Cloud Computing Information Assurance Framework, http://www.enisa.europa.eu, 2009

[5] ENISA (European Network and Information Security Agency) Report, Proactive Detection of Network Security Incidents, http://www.enisa.europa.eu, 12.07. 2012

[6] ENISA (European Network and Information Security Agency) Insecure magazine, str. 14-19, no. 35. www.insecuremag.com, September 2012

[7] Erickson, T., IDS, Digital Universe Study for EMC Corp., SearchStorage.com, The Security Help Net News, 2012

[8] Gorzelak, K., et al., Proactive Detection of Network Security Incidents, ENISA, 2011

[9] Gottlieb, J., Key Challenges in Proactive Threat Management, CEO of Sensage, The Security Help Net News, 2012

[10] Grance, T., Chevalier, S., Kent. K. and Dang, H., Guide to Computer and Network Data Analysis: Applying Forensic Techniques to Incident Response, NIST Special Publication 800-86, 2005

[11] Grubor, G. & Njeguš, A., An Application of Proactive Digital Forensic in Cloud Computing Environment, International conference TELFOR. Belgrade, 2012

[12] Hoopes, J., Virtualization for Security, Syngress, ISBN: 1597493058, 2012

[13] Krutz, R. L. and Vines, R. D., Cloud Security: A Comprehensive Guide to Secure Cloud Computing, Wiley Publishing, http://23510310jarinfo.files.wordpress.com/2011/08/ebooksclub-org_cloud_ security_a_comprehensive_guide_to_secure_cloud_computing.pdf, 2010

[14] Kun Ma, Bo Yang and Ajith Abraham, A Template-based Model Transformation Approach for Deriving Multi Tenant, Acta Polytechnica Hungarica, Journal of Applied Sciences, Vol. 9, No. 2, 2012

[15] Leibolt, G., The Complex World of Corporate Cyber Forensics Investigations, Springer's Forensic Laboratory Science Series, 2010

[16] Milosavljević, M. and Grubor, G., Digitalna forenzika računarskog sistema, Univerzitet Singidunum, 2010

[17] Milosavljević, M., Grubor, G., Istraga kompjuterskog kriminala, Univerzitet Singidunum, 2011

[18] MMSC – Morrison Maierle Systems Corp., Norman ASA, 2011

[19] Nacionalni CERT, Cloud computing, NCERT-PUBDOC-2010-03-293, www.cert.hr, 2010

[20] O Gara, M., Virtual srawl, Virtualization Magazine, irtualization.ulitzer.com, 2010

[21] Pál Michelberger Jr., Csaba Lábodi, After Information Security – Before a Paradigm Change (A Complex Enterprise Security Model), Acta Polytechnica Hungarica, Journal of Applied Sciences, Volume 9, Issue Nr. 4, 2012

[22] Taylor, P., Proactive Forensics in the Workplace, Litigation and Forensics, Data Recovery Services, Inc. www.legalforensics.com, 2010

[23] Zimmerman, S., Glavach, D., Cyber Forensics in the Cloud, The Newsletter for Information Assurance Technology Professionals, Vol. 4, No. 1, pp. 4-7, 2011

Sliding Mode Control of a Wind Turbine with Exponential Reaching Law

Abderrahmen Mechter, Karim Kemih, Malek Ghanes*

L2EI Laboratory, Jijel University, BP 98, Oued Aissa, Jijel, Algeria
* ECS-Lab ENSEA, Cergy-Pontoise, 95014, France
a.mechter@ensea.fr, karim.kemih@ensea.fr, ghanes@ensea.fr

Abstract: This paper presents a novel control strategy of a wind turbine based on Doubly Fed Induction Generator (DFIG) to maximize the electrical power generated by the DFIG, for low wind speed. To achieve this objective, the nonlinear sliding mode control with exponential reaching law (ERL), the Bees algorithm and the fuzzy maximum power point tracking MPPT are applied to the system. The ERL and the Bees algorithm are used to obtain a good reference tracking and to suppress the chatter phenomenon. The effectiveness of this control strategy is proven through the simulation results.

Keywords: Wind Energy; Doubly fed induction generator; Sliding Mode Control; Exponential Reaching Law; Bees algorithm; fuzzy maximum power point tracking

1 Introduction

Wind energy conversion is the fastest-growing energy source among the new power generation sources in the world. Advances in wind turbine technology made necessary the design of more powerful control systems. This is in order to improve wind turbines behavior, namely to make them more profitable and more reliable.

Many control strategies have been studied and applied to optimize the power for low wind speed: in reference [2] a fuzzy logic and a second order sliding mode controller were presented, the control scheme based on feedback linearization and gain scheduled linear quadratic regulator (LQR) is applied for the control of a wind turbine in [3]. In reference [4] a nonlinear state feedback PI controller with an estimator was used to optimize the power; a comparative study on the performance between PI and SMC controllers is presented in [5]. Reference [6] has developed an adaptive feedback linearization controller; a fuzzy-PI controller is used to extract an optimal power from the wind turbine system based on the Permanent Magnet Synchronous Generator in [7]. In reference [8] an RST controller is used to optimize the produced power, an LQG controller based on the

linearized wind turbine model was used in [9]. In reference [10], a variable structure control based on a sliding mode control has been presented, and a robust nonlinear controller has been presented in [11].

The objective of this paper is the control of active and reactive powers in the variable speed wind turbine. For this, we must continually adjust the turbine rotational speed according to the wind speed. The nonlinear sliding mode control with an exponential reaching law is used to achieve our objective. In [12], [13] and [14], the authors have used the conventional sliding-mode control to control the power produced by a wind turbine, but in this paper, we will use the same method but by employing an exponential reaching law (ERL) proposed by Fallaha et al. [15] to calculate the control law, the fuzzy maximum power point tracking *MPPT* to determine the rotational speed reference and the Bees algorithm to determine the optimal parameters of the controller. The advantage of the proposed approach is that allows to obtain a good reference tracking and the suppression of chatter phenomenon.

This paper is organized as follows. Firstly, In Section 2, the wind turbine configuration is presented. In Section 3, the sliding mode control is used to establish the control law applied to the system. Finally, the simulation results are given to show the efficacy of the proposed approach.

2 Wind Turbine Systems

A wind turbine generally consistes of a wind rotor, drive shaft, gear system and a generator, as presented in Fig. 1 [16]. The wind turbine consists of several blades (mostly three blades). The wind turbine rotates at a speed which depends on the wind speed. This speed will be adapted to that of the electric generator through a gearbox.

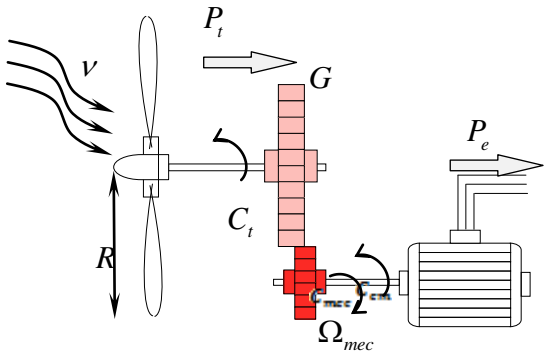


Figure 1
One-mass wind turbine model

The energy recovered by the wind turbine rotor is given by:

$$P_t = C_p\left(\frac{1}{2}\rho\pi R^2 v^3\right) \tag{1}$$

Where, ρ is the air density [Kg/m^3], R is the blade length [m] , v represents the wind speed [m/s]. C_p is the power coefficient; it depends on the tip speed ratio (λ) and the pitch angle of the blades (β) . It is given by [17]:

$$C_p(\lambda,\beta) = 0.5109\left(116\left(\frac{1}{\lambda+0.08\beta}-\frac{0.035}{\beta^3+1}\right)-0.4\beta-5\right)e^{-21\left(\frac{1}{\lambda+0.08\beta}-\frac{0.035}{\beta^3+1}\right)} + 0.0068\lambda \tag{2}$$

The tip-speed ratio (λ) is defined as:

$$\lambda = \frac{R\Omega_t}{v} \tag{3}$$

Ω_t , is the mechanical angular speed of the turbine [rad/sec].

The mechanical torque and the mechanical angular speed are given by:

$$\begin{cases} C_{mec} = \dfrac{1}{G}C_t = \dfrac{1}{G}\dfrac{P_t}{\Omega_t} = \dfrac{1}{G}\dfrac{0.5C_p\rho\pi R^2 v^3}{\Omega_t} \\ \Omega_{mec} = G\Omega_t \end{cases} \tag{4}$$

C_t is the mechanical torque [$N.m$] and G is the multiplication ratio.

The mechanical angular speed is determined by applying the fundamental equation of dynamic:

$$J\dot{\Omega}_{mec} = C_{mec} - C_{em} - f\Omega_{mec} \tag{5}$$

Where:

J is the total inertia of the rotating parts [$Kg.m^2$], f is the coefficient of viscous damping, and C_{em} is the electromagnetic torque of the generator [$N.m$].

The electrical and magnetic equations of the DFIG are given by:

$$\begin{cases} [V_s]_{ABC} = R_s[I_s]_{ABC} + \dfrac{d}{dt}[\varphi_s]_{ABC} \\ [V_r]_{ABC} = R_r[I_r]_{ABC} + \dfrac{d}{dt}[\varphi_r]_{ABC} \end{cases}. \tag{6}$$

$$\begin{cases} [\varphi_s]_{ABC} = [L_{ss}][I_s]_{ABC} + [M_{sr}][I_r]_{ABC} \\ [\varphi_r]_{ABC} = [L_{rr}][I_r]_{ABC} + [M_{sr}][I_s]_{ABC} \end{cases} \quad (7)$$

Where,

$$[L_{ss}] = \begin{bmatrix} l_s & m_s & m_s \\ m_s & l_s & m_s \\ m_s & m_s & l_s \end{bmatrix}, \quad [L_{rr}] = \begin{bmatrix} l_r & m_r & m_r \\ m_r & l_r & m_r \\ m_r & m_r & l_r \end{bmatrix},$$

$$[M_{sr}] = \begin{bmatrix} m_1 & m_3 & m_2 \\ m_2 & m_1 & m_3 \\ m_3 & m_2 & m_1 \end{bmatrix} \text{ and } \begin{cases} m_1 = m_{sr}cos(\theta) \\ m_2 = m_{sr}cos\left(\theta - \frac{2\pi}{3}\right) \\ m_3 = m_{sr}cos\left(\theta + \frac{2\pi}{3}\right) \end{cases}.$$

Where:

V_s is the stator voltage.

V_r is the rotor voltage.

φ_s is the stator flux.

φ_r is the rotor flux.

R_s is the stator resistance.

R_r is the rotor resistance.

l_s is the inductance of stator phase.

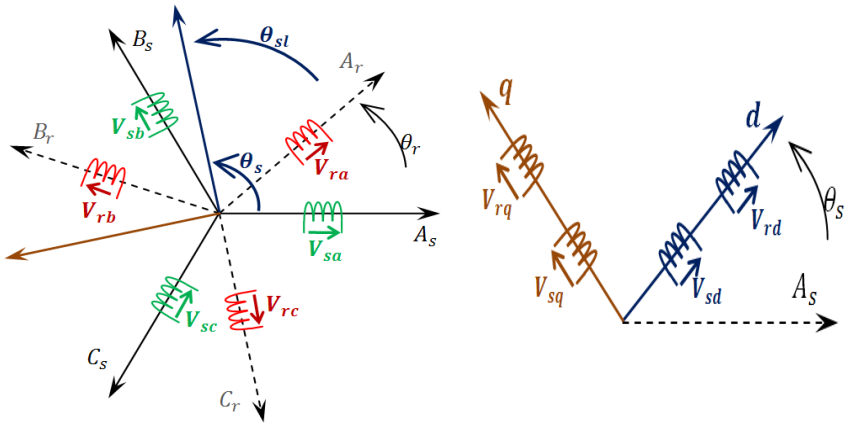
l_r is the inductance of rotor phase.

m_s is the mutual inductance between two stator phases.

m_r is the mutual inductance between two rotor phases.

m_{sr} is the maximum of the mutual inductance between the stator and the rotor.

In the three-phase plan, the representation of asynchronous machine is difficult because it is strongly coupled. So it is necessary to represent the machine behavior in a two-phase plane given by the transformation of Park (Fig. 2).



a) DFIG representation on a three-phase plane b) DFIG representation on a two-phase plane

Figure 2

The Park Transformation

The Park reference frame is given by the following set of equations [18-19]:

$$\begin{cases} V_{sd} = R_s I_{sd} + \dfrac{d}{dt}\varphi_{sd} - \omega_s \varphi_{sq} \\ V_{sq} = R_s I_{sq} + \dfrac{d}{dt}\varphi_{sq} + \omega_s \varphi_{sd} \\ V_{rd} = R_r I_{rd} + \dfrac{d}{dt}\varphi_{rd} - \omega_{sl} \varphi_{rq} \\ V_{rq} = R_r I_{rq} + \dfrac{d}{dt}\varphi_{rq} + \omega_{sl} \varphi_{rd} \end{cases} \tag{8}$$

Where, $\omega_{sl} = \omega_s - P\Omega_{mec}$.

$$\begin{cases} \varphi_{sd} = L_s I_{sd} + L_m I_{rd} \\ \varphi_{sq} = L_s I_{sq} + L_m I_{rq} \\ \varphi_{rd} = L_r I_{rd} + L_m I_{sd} \\ \varphi_{rq} = L_r I_{rq} + L_m I_{sq} \end{cases} \tag{9}$$

Where,

V_{sd}, V_{sq}, direct and quadrature voltages on the stator axis.

V_{rd}, V_{rq}, direct and quadrature voltages on the rotor axis.

I_{sd}, I_{sq}, direct and quadrature stator currents.

I_{rd}, I_{rq}, direct and quadrature rotor currents.

ω_s , the stator currents pulsation.

P , the number of machine poles pairs.

φ_{sd} , φ_{sq} , direct and quadrature stator flux.

φ_{rd} , φ_{rq} , direct and quadrature rotor flux.

L_s , the stator inductance.

L_r , the rotor inductance.

L_m , the mutual inductance between the stator and the rotor.

The stator active and reactive powers of the DFIG are given by [20],

$$\begin{cases} P_s = \dfrac{3}{2}\left(V_{sd} I_{sd} + V_{sq} I_{sq}\right) \\ Q_s = \dfrac{3}{2}\left(V_{sq} I_{sd} - V_{sd} I_{sq}\right) \end{cases} \tag{10}$$

3 Control Strategy

As it is shown in Fig. 1, there are two control parts (mechanical and electrical parts). For each one, we will design a sliding mode controller. The optimal rotational speed of the turbine corresponds to λ_{opt} and $\beta = 0$. It is used as a reference for the controller to determine the command (electromagnetic torque) that will be applied to the generator. The difficulty in controlling the DFIG is due to the coupling between the stator and rotor.

To simplify the task, the model is approximated to that of a DC machine. By orienting *d* axis in the direction of the flux φ_s . We obtain:

$$\begin{cases} \varphi_{sd} = \varphi_s \\ \varphi_{sq} = 0 \end{cases} \tag{11}$$

Taking into account of (9), we obtain:

$$\begin{cases} \varphi_{sd} = L_s I_{sd} + L_m I_{rd} \\ \varphi_{sq} = L_s I_{sq} + L_m I_{rq} = 0 \end{cases} \Rightarrow \begin{cases} I_{sd} = -\dfrac{L_m}{L_s} I_{rd} + \dfrac{V_{sq}}{\omega_s L_s} \\ I_{sq} = -\dfrac{L_m}{L_s} I_{rq} \end{cases} \tag{12}$$

The rotor flux can be written as:

$$\begin{cases} \varphi_{rd} = L_r \sigma I_{rd} + \dfrac{L_m V_{sq}}{L_s \omega_s} \\ \varphi_{rq} = L_r \sigma I_{rq} \end{cases} \tag{13}$$

If we neglect the stator resistance (for medium and high-power machines), and we make the stator flux constant (ensured by a stable grid connected to the stator). We obtain the direct and quadrature voltages on the stator axis:

$$\begin{cases} V_{sd} = 0 \\ V_{sq} = \omega_s \varphi_{sd} \end{cases} \tag{14}$$

Taking into account (10) and (12), we obtain the stator active and reactive powers:

$$\begin{cases} P_s = -\dfrac{3L_m}{2L_s} V_{sq} I_{rq} \\ Q_s = -\dfrac{3L_m}{2L_s} V_{sq} I_{rd} + \dfrac{3{V_{sq}}^2}{2L_s \omega_s} \end{cases} \tag{15}$$

Taking into account equations (8) and (13), we obtain the direct and quadrature voltages on the rotor axis:

$$\begin{cases} V_{rd} = R_r I_{rd} + \left(L_r - \dfrac{L_m^2}{L_s} \right) \dot{I}_{rd} - \omega_{sl} \left(L_r - \dfrac{L_m^2}{L_s} \right) I_{rq} \\ V_{rq} = R_r I_{rq} + \left(L_r - \dfrac{L_m^2}{L_s} \right) \dot{I}_{rq} + \omega_{sl} \left(L_r - \dfrac{L_m^2}{L_s} \right) I_{rd} \\ \qquad\qquad + \omega_{sl} \dfrac{L_m V_{sq}}{L_s \omega_s} \end{cases} \tag{16}$$

The optimal stator active and reactive powers given by (17) are used as reference for controllers to calculate commands (V_{rd} and V_{rq}) that will be applied to the DFIG rotor.

$$\begin{cases} P_s^{opt} = \dfrac{\rho \pi C_p R^5}{2\lambda^3} \Omega_t^3 \\ Q_s^{opt} = 0 \ (To ensure\ unity\ power\ factor) \end{cases} \tag{17}$$

The synthesis of these commands (V_{rd} . and V_{rq} .) is represented thereafter, using the nonlinear sliding mode control with exponential reaching law.

4 Sliding Mode Control

The sliding mode control belongs to the family of a variable structure controllers. The advantage of this method is its simplicity and robustness in spite of uncertainties in the system and external disturbances.

It consists of designing a control law that helps, to attract the state vector toward the sliding surface and to slide on the surface until reaching the equilibrium point (Fig. 3) [21-23].

We consider a nonlinear system defined as:

$$x^{(n)}(t) = f(x,t) + b(x,t)u(x,t) \tag{18}$$

where, x is the state vector and $f(x,t), b(x,t)$ are nonlinear functions and u is the control input.

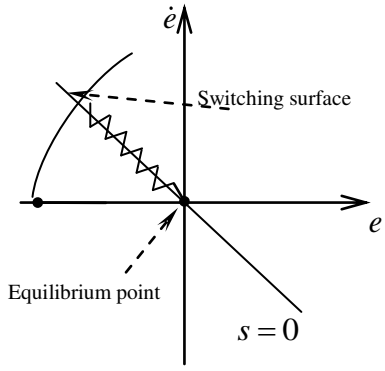


Figure 3
Sliding mode in a phase plane ($s = \lambda e + \dot{e}$)

To design a sliding mode control law, we must firstly, choose the switching surface. We take the general form proposed by Slotine [24]:

$$S = \left(\frac{d}{dt} + \lambda \right)^{(n-1)} e \tag{19}$$

Where, $e = x - x_d$ is the tracking error; x_d is the desired state, n is the system order and λ is a positive coefficient.

After choosing the sliding surface, we must choose the control law where the reaching condition defined by Lyapunov equation satisfied,

$$S\dot{S} < 0 \qquad \forall t$$

The control law has the following form,

$$u = u_{eq} + u_{disc} \tag{20}$$

In order that the condition (19) is verified at any time, $\dot{S}$ is chosen as follows [15] (a complete study of sliding-mode with ERL can be found in [15]):

$$\dot{S} = -\frac{k}{\delta_0 + (1-\delta_0)e^{-\alpha|S|^p}} sat\left(\frac{S}{\phi}\right) \tag{21}$$

Where, k is the discontinuous gain, $t_{rd} = \delta_0 \frac{|S(0)|}{k} + \frac{(1-\delta_0)}{k\alpha^{1/p}}$ is the desired reaching time, $\alpha \geq 0$, $0 < \delta_0 < 1$ and $p > 0$

$sat\left(\frac{S}{\phi}\right)$ is the saturation function,

$$sat\left(\frac{S}{\phi}\right) = \begin{cases} 1 \ \ if \ \frac{S}{\phi} > 1 \\ \frac{S}{\phi} \ \ if \ -1 \leq \frac{S}{\phi} \leq 1 \\ -1 \ \ if \ \frac{S}{\phi} < -1 \end{cases} \tag{22}$$

The use of saturation function instead of *sign* function is justified to avoid chattering phenomenon.

4.1 Wind Turbine Rotational Speed Command

In this part, we develop a law command to control the rotational speed. Taking into account of (8) and (10), we obtain,

$$\dot{\Omega}_{mec} = -\frac{f}{J}\Omega_{mec} + \frac{0.5\rho\pi R^5 C_p}{JG^3\lambda^3}\Omega_{mec}^2 - \frac{1}{J}C_{em} = c_1\Omega_{mec} + c_2\Omega_{mec}^2 + c_3 C_{em} \tag{23}$$

with $c_1 = -\frac{f}{J}$, $c_2 = \frac{0.5\rho\pi R^5 C_p}{JG^3\lambda^3}$, $c_3 = -\frac{1}{J}$

The tracking error is defined by, $e = \Omega_{mec} - \Omega_{mec}^d$

According to (18), the switching surface is given by,

$$S = e = \Omega_{mec} - \Omega_{mec}^d \tag{24}$$

$$\dot{S} = \dot{\Omega}_{mec} - \dot{\Omega}_{mec}^d = c_1\Omega_{mec} + c_2\Omega_{mec}^2 + c_3 C_{em} - \dot{\Omega}_{mec}^d = -\frac{k_1}{\delta_0 + (1-\delta_0)e^{-\alpha|S|^p}} sat\left(\frac{S}{\phi}\right)$$

$$C_{em} = \frac{1}{c_3}\dot{\Omega}_{mec}^d - \frac{c_1}{c_3}\Omega_{mec} - \frac{c_2}{c_3}\Omega_{mec}^2 - \frac{k_1}{c_3\left(\delta_0 + (1-\delta_0)e^{-\alpha|S|^p}\right)} sat\left(\frac{S}{\phi}\right) \tag{25}$$

$$\text{Where, } \begin{cases} u_{eq} = \dfrac{1}{c_3}\dot{\Omega}_{mec}^{d} - \dfrac{c_1}{c_3}\Omega_{mec} - \dfrac{c_2}{c_3}\Omega_{mec}^{2} \\ u_{disc} = -\dfrac{k_1}{c_3\left(\delta_0 + (1-\delta_0)e^{-\alpha|S|^p}\right)} sat\left(\dfrac{S}{\phi}\right) \end{cases} \tag{26}$$

4.2 Stator Active Power Command

In this part, we develop a law command to control the stator active power.

According to the first equation of (16), the quadrature currents on the rotor axis is given by:

$$\dot{I}_{rd} = -\frac{R_r}{\left(L_r - \dfrac{L_m^2}{L_s}\right)} I_{rd} + \omega_{sl} I_{rq} + \frac{1}{\left(L_r - \dfrac{L_m^2}{L_s}\right)} V_{rd} = c_4 I_{rd} + c_5 I_{rq} + c_6 V_{rd} \tag{27}$$

with: $c_4 = -\dfrac{R_r}{\left(L_r - \dfrac{L_m^2}{L_s}\right)}$, $c_5 = \omega_{sl}$, $c_6 = \dfrac{1}{\left(L_r - \dfrac{L_m^2}{L_s}\right)}$

The tracking error is defined by, $= I_{rd} - I_{rd}^{d}$

According to (18), the switching surface is given by,

$$S = e = I_{rd} - I_{rd}^{d} \tag{28}$$

$$\dot{S} = \dot{I}_{rd} - \dot{I}_{rd}^{d} = c_4 I_{rd} + c_5 I_{rq} + c_6 V_{rd} - \dot{I}_{rd}^{d} = -\frac{k_2}{c_3\left(\delta_0 + (1-\delta_0)e^{-\alpha|S|^p}\right)} sat\left(\frac{S}{\phi}\right)$$

$$V_{rd} = \frac{1}{c_6}\dot{I}_{rd}^{d} - \frac{c_4}{c_6} I_{rd} - \frac{c_5}{c_6} I_{rq} - \frac{k_2}{c_6 c_3\left(\delta_0 + (1-\delta_0)e^{-\alpha|S|^p}\right)} sat\left(\frac{S}{\phi}\right) \tag{29}$$

Where,

$$\begin{cases} u_{eq} = \dfrac{1}{c_6}\dot{I}_{rd}^{d} - \dfrac{c_4}{c_6} I_{rd} - \dfrac{c_5}{c_6} I_{rq} \\ u_{disc} = -\dfrac{k_2}{c_6 c_3\left(\delta_0 + (1-\delta_0)e^{-\alpha|S|^p}\right)} sat\left(\dfrac{S}{\phi}\right) \end{cases} \tag{30}$$

4.3 Stator Reactive Power Command

In this section, we calculate the law command to control the stator reactive power. According to the second equation of (16), the quadrature currents on the rotor axis is giver by:

$$\dot{I}_{rq} = -\frac{R_r}{\left(L_r - \frac{L_m^2}{L_s}\right)} I_{rq} - \omega_{sl} I_{rd} - \frac{\omega_{sl} L_m V_{sq}}{\omega_s \left(L_r L_s - L_m^2\right)} + \frac{1}{\left(L_r - \frac{L_m^2}{L_s}\right)} V_{rq}$$

$$\dot{I}_{rq} = c_7 I_{rq} + c_8 I_{rd} + c_9 + c_{10} V_{rq} \tag{31}$$

$$\text{with : } c_7 = -\frac{R_r}{\left(L_r - \frac{L_m^2}{L_s}\right)}, \quad c_8 = -\omega_{sl}, c_9 = -\frac{\omega_{sl} L_m V_{sq}}{\omega_s \left(L_r L_s - L_m^2\right)}, c_{10} = \frac{1}{\left(L_r - \frac{L_m^2}{L_s}\right)_{rq}}$$

The tracking error is defined by, $= I_{rq} - I_{rq}^d$

According to (18), the switching surface is given by,

$$S = e = I_{rq} - I_{rq}^d. \tag{32}$$

$$\dot{S} = \dot{I}_{rq} - \dot{I}_{rq}^d = c_7 I_{rq} + c_8 I_{rd} + c_9 + c_{10} V_{rq} - \dot{I}_{rq}^d$$
$$= -\frac{k_3}{c_3 \left(\delta_0 + \left(1 - \delta_0\right) e^{-\alpha |S|^p}\right)} sat\left(\frac{S}{\phi}\right) \Rightarrow$$

$$V_{rq} = \frac{1}{c_{10}} \dot{I}_{rq}^d - \frac{c_7}{c_{10}} I_{rq} - \frac{c_8}{c_{10}} I_{rd} - \frac{c_9}{c_{10}} - \frac{k_3}{c_{10} c_3 \left(\delta_0 + \left(1 - \delta_0\right) e^{-\alpha |S|^p}\right)} sat\left(\frac{S}{\phi}\right) \tag{33}$$

Where,

$$\begin{cases} u_{eq} = \frac{1}{c_{10}} \dot{I}_{rq}^d - \frac{c_7}{c_{10}} I_{rq} - \frac{c_8}{c_{10}} I_{rd} - \frac{c_9}{c_{10}} \\ u_{disc} = -\frac{k_3}{c_{10} c_3 \left(\delta_0 + \left(1 - \delta_0\right) e^{-\alpha |S|^p}\right)} sat\left(\frac{S}{\phi}\right) \end{cases} \tag{34}$$

5 Simulation Results

The control strategy was applied to a 660-kW wind turbine using Matlab/Simulink software. The simulation results show the effectiveness of the control strategy used in this study. The parameters of the wind turbine are:

$$R = 21.165\,m, G = 39, J = 28\,Kg.m^2, \rho = 1.225\,Kg/m^3, R_s = 0.0146\,\Omega, f = 0.01.$$
$$R_r = 0.0238\,\Omega, L_s = 0.0306\,H, L_r = 0.0303\,H, L_m = 0.0299\,H, P = 2$$

Nominal frequency: $50Hz$.

The exponential reaching law parameters: $\delta_0 = 0.08$, $p = 2$, $\alpha = 1000$, $k = 3$.
The optimal active power can be written with neglected losses as [11]:

$$P_s^d = P_t^{opt}$$

Form (1) and (3) we have: $P_s = k\Omega_{mec}^3$ with $k = \dfrac{\rho\pi C_p^{max} R^5}{2G^3 \lambda_{opt}^3}$

So, $I_{rq}^d = -\dfrac{2L_s}{3V_{sq}L_m}\dfrac{\rho\pi C_p^{max} R^5}{2G^3\lambda_{opt}^3}\tilde{\Omega}_{mec}$

With $\tilde{\Omega}_{mec}$ is the estimated rotational speed and Ω_{mec}^d is determined using the fuzzy maximum power point tracking MPPT proposed by S. Abdeddaim et al. [25] to maximize the output power and to adjust in real time the rotational speed of wind turbines.

on the other hand, the optimal reactive power is set to zero to ensure a unity power factor operation of the studied wind turbine: $Q_s^{ref} = 0$ [25],

Or, form (15), we have : $I_{rd}^d = \dfrac{V_{sq}}{L_m \omega_s}$

The Bees algorithm is employed to determine the control parameters k_1, k_2 and k_3, it is an optimization algorithm inspired by the natural foraging behaviour of honey bees to find the optimal solution [26]. The parameters of the Bees algorithm are: the number of scout bees is 30, the number of best selected patche is 20, the number of elite selected patches is 10, the number of recruited bees around best selected patches is 15 and the number of recruited bees around elite selected patches is 30. In the literature, we can find many objective functions as a performance criterion. In this paper, we use the integral of time multiplied by squared error (ITSE), defined by :

$$ITSE = \int_0^{t_{ss}} te^2(t)$$

Where t_{ss} is the final time.

Fig. (4) present a variable wind profile comprised between 5 m/s and 12 m/s, and it varies under the nominal value (12 *m/s*). This allows observing the operation of the wind turbine for low wind speed.

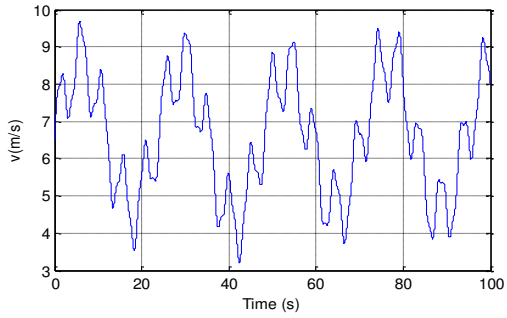


Figure 4
Wind speed

Fig. 5 shows the generator rotational speed under the proposed control method and the reference. We can see clearly that the estimated $\tilde{\Omega}_{mec}$ track perfectly the reference.

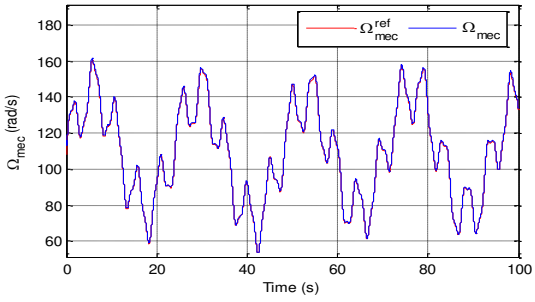


Figure 5
Generator rotational speed

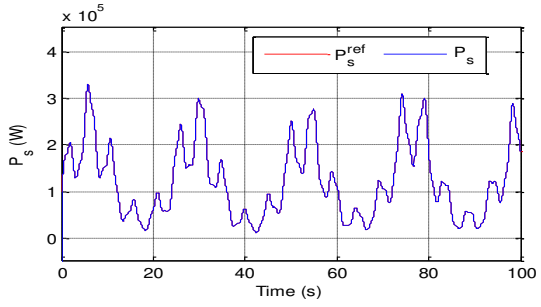


Figure 6
Active power

The active and the reactive powers are presented in Fig. 6 and Fig. 7. As we can see, the powers converge to their desired references with good dynamics.

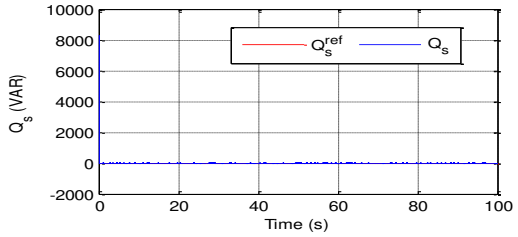


Figure 7

Reactive power

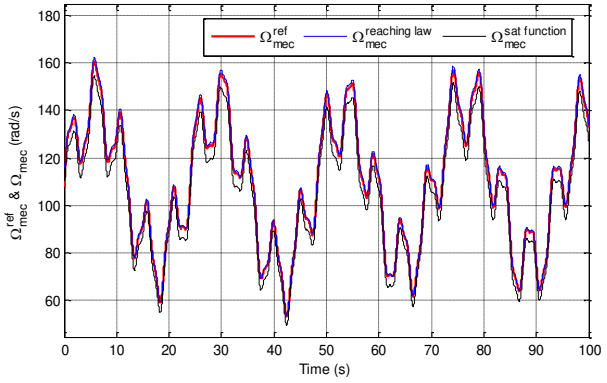


Figure 8

Generator rotational speed: comparaison between conventional sliding-mode control and sliding mode control with ERL and Bees algorithm

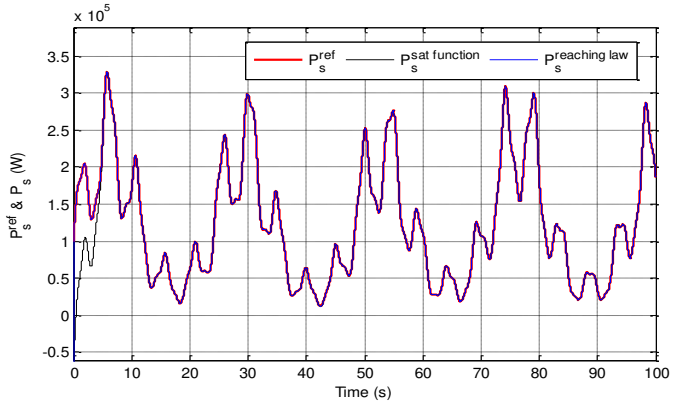


Figure 9

Active power: comparaison between conventional sliding-mode control and sliding mode control with ERL and Bees algorithm

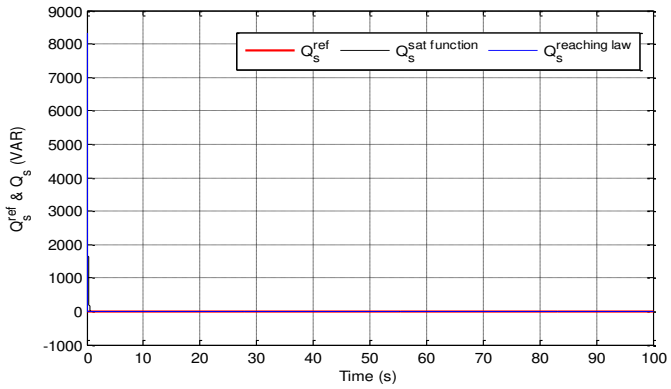


Figure 10

Reactive power: comparaison between conventional sliding-mode control and sliding mode control with ERL and Bees algorithm

Fig. 8, Fig. 9 and Fig. 10 represent a comparison between the conventional sliding-mode control and the sliding mode control with ERL and Bees algorithm. These results show the effectiveness of the proposed approach, especially the very good tracking performance of the desired trajectory.

Conclusion

In this paper, the control of a 600-*kW* wind turbine for low wind speed was presented. To achieve this objective, the nonlinear sliding mode control with exponential reaching law and Bees algorithm are used to conceive the controllers of mechanical and electrical parts. The fuzzy maximum power point tracking MPPT is using to determine the rotational speed reference. The advantage of the proposed approach, compared to the conventional method is the good reference tracking and the suppression of chatter phenomenon. Simulation results show the effectiveness of the applied control strategy.

Acknowledgement

This work was supported by the CMEP- TASSILI project under Grant 14MDU920.

References

[1] T. Burton, D. Sharpe, N. Jenkins, E. Bossanyi: Wind Energy Handbook, John Wiley & Sons, Chichester, 2001

[2] S. Abdeddaim, A. Betka: Optimal Tracking and Robust Power Control of the DFIG Wind Turbine. International Journal of Electrical Power and Energy Systems, Vol. 49, 2013, pp. 234-242

[3] R. Burkart, K. Margellos, J. Lygeros: Nonlinear Control of Wind Turbines: An Approach Based on Switched Linear Systems and Feedback Linearization. IEEE CDC-ECE, Orlando, 2011, pp. 5485-5490

[4] B. Boukhezzar, H. Siguerdidjane: Nonlinear Control with Wind Estimation of a DFIG Variable Speed Wind Turbine for Power Capture Optimization. Energy Conversion and Management, Vol. 50, 2009, pp. 885-892

[5] D. Kairous, R. Wamkeue, B. Belmadani: Advanced Control of Variable Speed Wind Energy Conversion System with DFIG. IEEE EEEIC, Prague, 2010, pp. 41-44

[6] A. Mullane, G. Lightbody, R .Yacamini: Adaptive Control of Variable Speed Wind Turbines. Rev. Energ. Ren.: Power Engineering, 2001, pp.101-110

[7] AG. Aissaoui et al. :A Fuzzy-PI Control to Extract an Optimal Power from Wind Turbine. Energy Conversion and Management, Vol. 65, 2013, pp. 688-696

[8] F. Hachicha, L. Krichen: Rotor Power Control in Doubly Fed Induction Generator Wind Turbine under Grid Faults. Energy, Vol. 44, 2012, pp. 853-861

[9] B. Boukhezzar, H. Siguerdidjane: Comparison between Linear and Nonlinear Control Strategies for Variable Speed Wind Turbine Power Capture Optimization. Control Engineering Practice, Vol. 18, 2010, pp. 1357-1368

[10] Y. Bekakra, D. Ben attous: Active and Reactive Power Control of a DFIG with MPPT for Variable Speed Wind Energy Conversion using Sliding Mode Control. World Academy of Science, Engineering and Technology, Vol. 60, 2011, pp. 12-21

[11] K. Elkhatib, A. Aitouche, R. Ghorbani, M. Bayart: Robust Nonlinear Control of Wind Energy Conversion Systems. International Journal of Electrical Power and Energy Systems, Vol. 44, 2013, pp. 202-209

[12] Boubekeur, Boukezzar. Mohamed, M'Saad. : Robust Sliding Mode Control of a DFIG Variable Speed Wind Turbine for Power Production Optimization, Mediterranean Conference on Control and Automation, Ajaccio, 2008, pp. 795-800

[13] Kairous, D. Wamkeue, R. Belmadani, B.: Sliding Mode Control of DFIG based Variable Speed WECS with Flywheel Energy Storage, International Conference on Electrical Machines (ICEM), Rome, 2010, pp. 1-6

[14] Djilali, Kairous. René, Wamkeue. : Sliding-Mode Control Approach for Direct Power Control of WECS based DFIG, International Conference on Environment and Electrical Engineering (EEEIC), Rome, 2011, pp. 1-4 (2011)

[15] C. J. Fallaha et al. :Sliding Mode Robot Control With Exponential Reaching Law. IEEE Transactions on Industrial Electronics, Vol. 58, 2011, pp. 600-610

[16] N. Abu-Tabak : Stabilité dynamique des systèmes électriques multimachines: modélisation, commande, observation et simulation. PhD thesis. École centrale de Lyon, 2008

[17] S. Le-peng, T. De-dong, W. Debiao, L. Hui: Simulation for Strategy of Maximal Wind Energy Capture of Doubly Fed Induction Generators. IEEE ICCI, Beijing, 2010, pp. 869-873

[18] Hansen O. L: Aerodynamics of Wind Turbines. ISBN 978-1-84407-438-9. Earthscan, London, Sterling VA, 2008

[19] F. K. A. Lima, E. H. Watanabe, P. Rodriguez, A. A. Luna: Simplified Model for Wind Turbine Based on Doubly Fed Induction Generator. IEEE ICEMS, Beijing, 2011, pp. 1-6

[20] S. T. Jou, S. B. Lee, Y. B. Park, K. B. Lee: Direct Power Control of a DFIG in Wind Turbines to Improve Dynamic Responses. Journal of Power Electronics, Vol. 9, 2009, pp. 781-790

[21] S. Y. Guo and L. Chen : Terminal Sliding Mode Control for Coordinated Motion of a Space Rigid Manipulator with External Disturbance, Applied Mathematics and Mechanics, Vol. 29, 2008, pp. 583-590

[22] K Széll and Korondi P: Mathematical Basis of Sliding Mode Control of an Uninterruptible Power Supply, Acta Polytechnica Hungarica Vol. 11, 2014, pp. 87-106

[23] C. Ben Regaya, A. Zaafouri, A. Chaari : A New Sliding Mode Speed Observer of Electric Motor Drive Based on Fuzzy-Logic, Acta Polytechnica Hungarica Vol. 11, 2014, pp. 219-232

[24] J. J. E. Slotine, W. LI: Applied Nonlinear Control. Prentice Hall Englewood Cliffs, ISBN 0-13-040890-5, New Jersey, 1991

[25] S. Abdeddaim and A. Betka: Optimal Tracking and Robust Power Control of the DFIG Wind Turbine, International Journal of Electrical Power and Energy Systems, Vol. 49, 2013, pp. 234–242

[26] L. Ozbakir, A. Baykasoglu, P. Tapkan: Bees Algorithm for Generalized Assignment Problem, Applied Mathematics and Computation, Vol. 215, 2010, pp. 3782-3795

Multicriteria HR Allocation Based on Hesitant Fuzzy Sets and Possibilistic Programming

Zoran Ciric, Dragan Stojic, Otilija Sedlak

Faculty of Economics Subotica, Segedinski put 9-11, 24000 Subotica, Serbia
e-mail: zoran.ciric@ef.uns.ac.rs, stojicd@ef.uns.ac.rs, otilijas@ef.uns.ac.rs

Abstract: We focus in this paper on posing and solving a problem of human resource management, namely allocating n people into m groups, such that each group satisfies minimal quality conditions. Each individual is evaluated for every quality by the expert who grades them using hesitant fuzzy numbers, thus enabling a more realistic mark giving method by avoiding artificially imposed accuracy. We compare results with crisp models of Lin and Gen and discuss the significance of fuzzy methodology.

Keywords: hesitant fuzzy sets; multicriteria optimization; possibilistic programming

1 Introduction

Just as in many other areas, project management makes use of mathematical modeling in cases such as setting strategic goals, formulation of strategies, selection of human resources and realization of the chosen strategy and control. Criteria and restrictions of alternatives are also encompassed in the space of uncertainty and indeterminacy.

Resource Allocation Problem (RAP) is the process of allocating resources among the various projects or business units for maximization of profit or minimization of cost [10]. The process of the RAP is focused on finding an optimal allocation of limited resources to a certain number of tasks while controlling for the given resource constraint. Resource may be any entity used to accomplish a goal, e.g. a person, an asset, material, etc. This paper deals with human resource allocation. We start from a set of individuals who were tested for certain qualities by the experts. The novelty lies in the fact that marks given by the experts need not to be exact numbers, but qualitative values, such as: absolutely true, highly true, moderately true, etc. Furthermore, if experts do have certain hesitations it will be allowed to choose more than one 'mark', thus avoiding to be artificially precise. Methodology which allows us to model indeterminacy is fuzzy sets theory, which is particularly well designed for dealing with non-probabilistic uncertainties. Following the work of Saaty et al. [20] we observe marketing sector vacancies.

We expand their work in two directions: expanding the grading system and, indirectly, the coefficients of model constraints which become fuzzy. On the other hand, we allow each candidate to be placed at the position best suited for his qualities, which is not necessarily the position he/she would apply for in the first place.

Chapter two focuses on literature review, both of human resources allocation and hesitant fuzzy sets. Chapter three presents the data and variables. We pose the model and perform an in-depth analysis and construction of constraints' coefficients. LINDO software is used for solving binary variables problems. Results are presented and differences between results of similar studies are pointed out. Finally, the last chapter concludes and the guidelines for future research are given.

2 Literature Review

We begin with the review of papers which inspired us the most, dealing with human resources allocation. Due to the considerable amount of subjectivity involved in hiring people, it is helpful to have prioritization techniques to deal with intangibles and make the process more objective, Saaty [18] and Saaty et al. [19]. They propose Analytic Hierarchy Process (AHP), and linear programming to rate and derive the best combination of people assigned to jobs [24] such as Rezaei et al. [15].

Filho et al. [6] uses the Constraint Satisfaction Problem approach to solve human resource allocation problems in cooperative health services. They proposed a new tool for planning human resources utilization in hospital plants, i.e. health units. They used simulations for measuring the performance of the proposed heuristics. Another contribution of this work was the mathematical modeling of the unary, multiple, numeric and implicit constraints.

Ballesteros-Pérez et al. [2] described a new application to key sociometrical concepts for the selection of workers within organizations in area of project development, in order to assure the achieving of the highest possible overall work efficiency from the viewpoint of social interaction.

Silva and Costa [21] focused on software development projects and presented a methodology based on dynamic programming for assigning human resources to the observed projects [24]. The methodology takes into account the complexity of each project and the existing capabilities of staff and the skills required for the project. A simulation is used to demonstrate the decision model.

The concept of Hesitant Fuzzy Sets (HFS) is a new one, dating back to Torra [23]. HFS are a generalization of intuitionistic fuzzy sets of Atanassov and Gargov [1] who allowed each fuzzy membership function to have a degree of indeterminacy.

Torra defines a HFS through a function that returns a set of membership values for each element in the domain. The applications of HFS followed various directions, mostly in generalization of the existing results dealing with (ordinary) fuzzy sets. Rodríguez et al. [17] focused on group decision making (GDM) and hesitant situations that could arise from multiple opinions. They have shown how to generate comparative linguistic expressions by using a context-free grammar. Rodríguez et al. [16] extended context-free grammar to that defined in Rodríguez et al. [17], and further expanded the topics in GDM in terms of expressing and modeling doubt between different linguistic terms, requiring richer expressions to express the experts' knowledge more accurately.

On the other line of reasoning, but also tackling the decision making processes, Xia and Xu [25] presented an extensive study on hesitant fuzzy information aggregation techniques and their application in decision making with anonymity. They have developed certain hesitant fuzzy operational rules based on the interconnection between the hesitant fuzzy set and the intuitionistic fuzzy set, as well as a series of aggregation operators for dealing with various situations, and risk. Portik and Pokoradi have applied the fuzzy rule based risk assessment and the authors use it in the decision making process [14].

3 Methodology and Model Explanation

In this chapter we shall present in detail the proposed methodology and provide the readers with the model for human resources allocation to different positions in marketing sector. When evaluating candidates, HR management is confronted with a significant number of tradeoffs over a diverse range of criteria.

Multiple-objective optimization problems have gained momentum in researchers' community with various backgrounds since early 1960. In the realm of multiobjective optimization problem, multiobjective functions are optimized simultaneously. Therefore, there does not necessarily exist a solution that is best with respect to all objectives due to the incommensurability and confliction among objectives. A solution may be the best in one objective but underperforming in another. The so-called Pareto optimal solutions are a set of solutions for the multiple-objective case which cannot simply be compared with each other, i.e. no improvement is possible in any objective function without sacrificing at least one of the other objective functions.

Fuzzy mathematical programming (FMP) was developed in order to adequately solve optimization problems which poses non stochastic indeterminacies. Inuiguchi [7] classifies FMP in three categories based on the types of indeterminacies: 1) FMP with ambiguities, also known as flexible programming, developed by Bellman and Zadeh [3], allows fuzzy preferences for the decision maker. 2) FMP with vagueness deals with fuzzy coefficients in constraints and

goals. Dubois and Prade [4] were the first researchers who investigated systems of linear equations with fuzzy coefficients and suggested possible applications of fuzzy mathematical programming. Dubois [5] observed fuzzy coefficients as possibilistic distributions of coefficient values, hence the name possibilistic programming for the second type of FMP. Following the early work of Dubois, we shall apply possibility theory in our linear model. Finally, third type, robust programming combines fuzzy coefficients and ambiguous preferences of the decision makers. Negoita et al. [12] were the first to formulate such a linear programming problem.

Managing and modeling of uncertainty by different forms of information, used by experts to provide their preferences, can be done in various ways, including utility vectors, fuzzy preference relations, linguistic variables, interval values, multiplicative preference relations, hesitant fuzzy sets.

Hesitant fuzzy sets are a generalization of 'ordinary' fuzzy sets and will be used in possibilistic programming. For every element x from the domain X, membership function of a HFS A is given by $h(x) = A_x$, where A_x is any (finite) subset of [0,1]. Suppose that the experts deciding on candidate's communication skills should choose a grade from the set of linguistic descriptions, L = {none, very low, low, medium, high, very high, absolute}. Of course, if unsure of the exact grade, an expert has a possibility to opt for two or more grades[1]. Grades can be naturally ordered, and modeled by fuzzy sets, as shown in Figure 1.

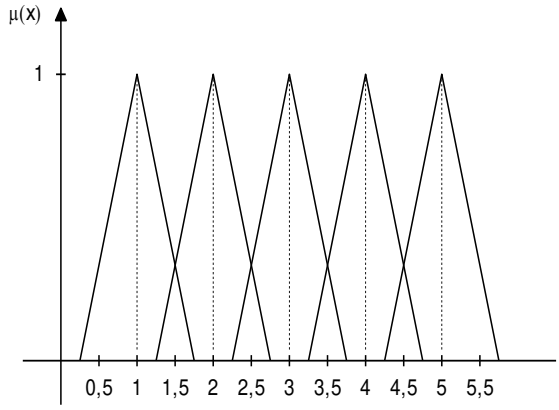


Figure 1

Choosing more than one grade would result in 'merging' the graphs of chosen grades into one convex fuzzy set. It is done by means of envelope which is a linguistic interval whose limits are minimal and maximal linguistic term, as shown in Figure 2.

[1] Choosing all options is equivalent to total absence of knowledge.

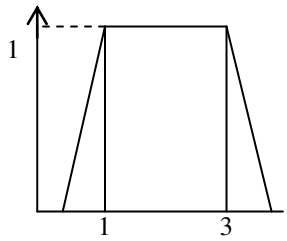


Figure 2

We observe only symmetrical marks, bounded with linear functions, though different methods of forming asymmetrical functions exist and can be found in Inuiguchi et al. [8]. We also refer an interested reader to Klir and Yuan [9] for any comprehensive and in-depth reading on fuzzy sets and fuzzy logic, such as Nagy et al. ([11]) and Takacs [22].

3.1 The Model

Let us observe a company 'Comp' consisting of S sectors (Manufacturing, Marketing, R&D, etc.) Each sector consists of up to P positions (e.g. Marketing sector embodies marketing vice president, marketing assistant, customer service representative, shipping clerk, etc.), i.e. $P = maxP_i$, where P_i is the number of positions in sector *i, i=1,2,…S*. Each position employs up to *K* employees, where K being $\max K_{ij}$, and K_{ij} is the number of employees on the position j of the sector *i* (e.g. there are two positions for marketing VP, 3 for marketing assistant, 6 for customer service representative, 2 for shipping clerk, etc.)

There is a total of *N* candidates, each being tested for C characteristics, regardless of the position he/she apply for. The program chooses based on the test results the position one is suited best for, and that doesn't necessarily mean is the position one applied for. Characteristics include experience, organizational skills, dependability, computer knowledge, contacts, formal education, etc. These characteristics are measured by both numerical values and experts' opinions in terms of hesitant fuzzy sets, and their systematization with graphical expressions is given below:

Experience is divided into three broad intensity groups: high (corresponds to 15+ years of experience), medium (6-15 years), and low (up to 5 years). For each broad group experts are asked to grade the applicant in terms how satisfactory (in terms of Figures 1 and 2) his actual achievements are, being in a certain intensity group. Similar reasoning goes for following characteristics: quality contacts and leadership, which are also divided into high, medium, and low, with further sub-divisions by the experts (Figure 3). Organizational and computer skills, as well as dependability are solely divided into broad groups, without additional experts'

opinion. Finally, education is divided into PhD, MSc, BSc, and secondary, while the experts give opinion on the quality of the school attended, thus allowing for slight differences within the same educational level.

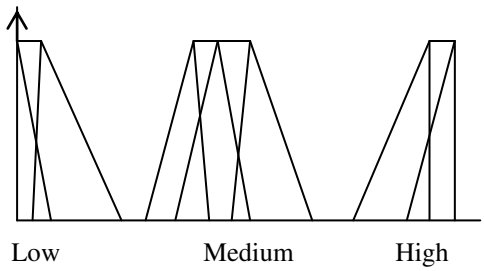


Figure 3

Let us denote by x_{ijk} a binary variable which has a value 1 if the candidate k, $k=1...N$ is chosen for the position i, $i=1...P$ in sector j, $j=1...S$, otherwise, x_{ijk} has a value 0. Natural constraints imposed on x's are following:

$$\sum_i^P \sum_j^S \sum_k^N x_{ijk} = 1, \text{for every } k = 1..N. \tag{1}$$

One candidate can be employed at exactly one position.

$$\sum_i^P \sum_j^S \sum_k^N x_{ijk} \leq N \tag{2}$$

Total number of candidates exceeds number of positions.

Each sector is choosing candidates in order to maximize it's manpower:

$$\sum_i^P \sum_j^S \sum_k^N c_{ijk} x_{ijk} \tag{3}$$

Coefficients c_{ijk} in the linear expression (3) are fuzzy numbers of triangular, or trapezoidal shape and are obtained through summation of grades for each position.

Analytical and graphical expression of a trapezoidal number $A = (a, b, \alpha, \beta)$ is given by following membership function and in Figure 4:

$$\mu_A(x) = \begin{cases} 1 - \dfrac{a-x}{\alpha}, a-\alpha \leq x \leq \alpha \\ 1 \qquad\qquad\quad , a \leq x \leq b, \\ 1 - \dfrac{x-b}{\beta}, b \leq x \leq b+\beta \end{cases}$$

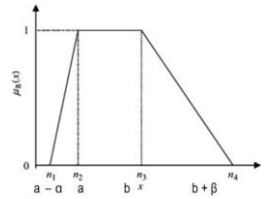


Figure 4

The linear combination of fuzzy trapezoidal numbers, which is present as a constraint expression as well as a goal function, is again a number of the same shape. If we denote by (a_{ijk}, b_{ijk}, α_{ijk}, β_{ijk}) coefficient c_{ijk}, the total manpower TM, defined at (3) is given by:

$$\left(\sum_{i,j,k} a_{ijk}x_{ijk}, \sum_{i,j,k} b_{ijk}x_{ijk}, \sum_{i,j,k} \alpha_{ijk}x_{ijk}, \sum_{i,j,k} \beta_{ijk}x_{ijk}\right) = \left(A(x), B(x), \alpha(x), \beta(x)\right)$$

The possibilistic linear function is not uniquely determined, therefore its maximization, as well as constraints of the form $\left(A(x), B(x), \alpha(x), \beta(x)\right) \geq M$ may not be completely meaningful. For overcoming this problem we introduce a so-called 'fuzzy inequality', $\rho(A,B)$, where A and B are two fuzzy sets. If the possiblistic distribution μ_A of a possibilistic variable α, the measurements of possibility and necessity of an event that α belongs to the set B ([5]):

$$Poss_A(B) = sup_x min(\mu_A(x), \mu_B(x)) \quad (4)$$

$$Nes_A(B) = inf_x max(1 - \mu_A(x), \mu_B(x)), \quad (5)$$

where μ_B is a membership function of a set B. $Poss_A(B)$ estimates the level up to which it is possible that possibilistic variable α belongs to the set B. on the other side, the expression $Nes_A(B)$ measures the level of necessity up to which is certain that α belongs to the set B. Let α be a possibilistic variable and $B = (-\infty, g]$, a crisp set. One could easily obtain the following: $Pos(\alpha \leq g) = sup\{ \mu_A(x) \mid x \leq g\}$ and $Nes(\alpha \leq g) = 1 - sup\{ \mu_A(x) \mid x > g\}$, which is exhibited in Figure 5. Similarly, for $B = [g, +\infty)$ we obtain following equations presented in Figure 6:

$Pos(\alpha \geq g) = sup\{ \mu_A(x) \mid x \geq g\}$ and $Nes(\alpha \geq g) = 1 - sup\{ \mu_A(x) \mid x < g\}$.

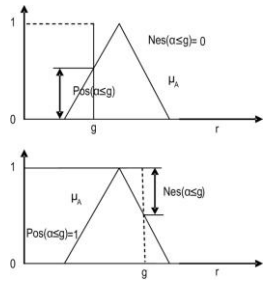


Figure 5

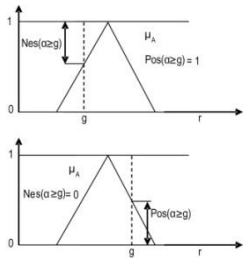


Figure 6

The constraints on the number of people in each position are given by:

$\sum_{i,j} x_{ijk} = p_{ij}$, where i is the number of sectors, $i=1...S$ and j is the number of positions in each sector, $j=1...P$.

Maximization of the total human potential of the company is done by maximizing the right spread of the possibilistic distribution, allowing for the best possible marks to be incorporated in the optimal team:

$$\text{Maximize } \sum_{i,j,k} \left(\beta_{ijk} + b_{ijk}\right) x_{ijk} \tag{6}$$

One of the possible restrictions is a funding restriction. Each employee has suggested his expected and minimal wage, w_e and w_{min} respectively. We would like not to break funds F with high necessity, of e.g. $h = 0.8$ or $h = 0.9$. One easily concludes from Figure 5 that this constraint is equivalent to:

$$\frac{\sum x_i w_{ei} - F}{\sum x_i \left(w_{ei} - w_{\min i}\right)} \leq 1 - h \tag{7}$$

Other restrictions are possible, namely if we want to impose a minimal mark value in the certain job group, e.g. average computer skills among marketing assistants should be at least M. Since we observe all the marks as possibilistic distribution functions, we can pose this requirement differently, namely, we could demand that the minimal average computer skill is necessarily greater than M, where we use high necessity measure of e.g. $h = 0.8$ or $h = 0.9$, as presented in Figure 6. The set of restrictions is given by:

$$\frac{M - \sum a_i x_i}{\sum a_i x_i} \leq 1 - h \tag{8}$$

4 Results and Discussion

Numerical example is provided in this section. Data from 30 applicants were collected, and only the marketing sector has been covered. Data can be found in Appendix 1[2]. In addition, applicants were asked to suggest their salary in terms of expected and minimum value. Values of salaries are given in Appendix 2. This group was tested for education levels, contacts, experience, leadership skills, organizational, computer secretarial, and communicational skills, as well as dependability. Not all skills are necessary for each position, namely: VP Marketing (up to two people) are tested for experience, contacts, education and leadership. Marketing assistants (MA, up to three) for organization, computer skills and education, Customer service representatives (CSR, up to four) for education, computer, secretarial and communicational skills. Finally, Shipping clerks (SC, up to five) were tested for dependability, organizational and computer skills. Objective function was to maximize the total manpower of future employees.

Results without limited funds are presented in Table 1. Results include restrictions imposed on VPs to have an average score of at least 20. For MAs for having an average mark on education and organization at least 8, on CSRs to have average communicational skills at least 2.5.

Table 1

Optimal choice of applicants, necessity coefficient 0.8. Unlimited funds

Position	Ordinal number of the applicant	Costs
VP	13, 23	185-190
MA	7, 12, 22	285-295
CSR	4, 9, 26, 30	267-280
SC	8, 15, 16, 21, 29	192-207

Total human capital equals 426.5. Total costs up to 972.

However, if the funds are decreased to 600 (in thousands of US dollars per annum) we observe the results in Table 2:

Table 2

Optimal choice of applicants, necessity coefficient 0.8.Funds limited to 600

Position	Ordinal number of the applicant	Costs
VP	7, 12	195-200
MA	4, 8, 16	100-108
CSR	6, 11, 15, 25	138-150
SC	9, 14, 20, 27, 29	121-138

[2]Applicants are ordered from A.A. through B.D., and in tables in current chapter we used ordinal notation: A.A. is 1st,...B.D. is 30th.

Total human capital is decreased to 355.4. Total costs up to 596. Major decrease in human capital is due to choosing applicants for CSR and SC positions with lower education levels, as well as low contacts and leadership skills. Even though these characteristics influence the total score, they are almost irrelevant for the positions in question, and chosen applicants show excellent results in characteristics needed for those positions (e.g. high computer and communicational skills).Further decrease in funding is presented in Table 3a and 3b, together with the change in desired necessity levels, from 0.8 to 0.95.

Table 3a

Further decrease in funds, F = 500, h = 0.8

Position	Ordinal number of the applicant	Costs
VP	7, 12	195-200
MA	4, 8, 16	100-108
CSR	9, 15, 25, 29	97-106
SC	2, 14, 18, 24, 27	76-88

Human capital is slightly decreased to 329.2, while funds are at most 502, though crossing the limit by 2000 is unlikely, since necessity level is 0.8. For extreme risk averse investors, Table 3b exhibits the same funds, with necessity level 0.95.

Table 3b

Further decrease in funds, F = 500, h = 0.95

Position	Ordinal number of the applicant	Costs
VP	7, 30	187-195
MA	4, 8, 16	100-108
CSR	9, 15, 25, 29	97-106
SC	2, 14, 18, 24, 27	76-88

HC for this constellation equals 326.5. Maximal amount of money needed is almost certainly less than 497. Huge savings are possible if focusing on applicants with lower levels of formal education, yet skillful.

Conclusion

We tried to present a possible application of hesitant fuzzy sets in the field of human resources allocation. Certain qualities of the applicants cannot be measured precisely, and an expert's opinion is needed, which as well may not be precise, or the expert might have doubts or hesitate to give a final mark. Hesitant fuzzy sets are designed to perfectly model these situations. We suggested possible objective functions, as well as an array of possible constraints. There were no restrictions on number of positions one applicant could apply. We tested the model on 30 applicants to Marketing sector of a company wanting to expand to 14 new positions. Results show that insignificant loss in total human potential, measured by sum of marks in all observed characteristics, can lead to substantial savings.

Possible limitation might be the omission of interaction among cooperative individuals. When choosing a team of people it is relevant to measure a level of synergy within a team and test whether a team contributes more (or less) than the sum of independent performances of each team member. Surely it is possible to employ other mathematical methods in order to model synergy levels. However, the focus of this paper was on presenting a new method in decision making and mark giving, but this interesting topic will be fully investigated in our future work.

Appendix 1: List of 30 applicants and their grades in various characteristics

	API	A.A.	A.B.	A.C.	A.D.	A.E	A.F.	A.G.	A.H.	A.I.	A.J.	A.K.	A.L.	A.M.	A.N.	A.O.	A.P.	A.Q.	A.R.	A.S.	A.T.	A.U.	A.V.	A.W.	A.X.	A.Y.	A.Z.	B.A.	B.B.	B.C.	B.D.
VP Marketing	Experience	2	1,5	2,3	2,0	5,2	5,0	5,4	3	4,4	4,2	3,2	5,2	7,2	2,4	3,1	3,2	2,4	3,2	2,4	4,4	5,3	4,3	6,6	3,2	2,3	4,4	3,2	4,3	3,4	5,5
		2	1,4	2,2	2,0	5,1	5,0	5,2	3	4,2	4,1	3,1	5,1	7,1	2,3	3,0	3,1	2,2	3,1	2,2	4,3	5,2	4,2	6,4	3,1	2,2	4,3	3,1	4,2	3,3	5,4
		2	1,3	2,1	2,0	5,0	5,0	5,0	3	4	4	3	5,1	7,1	2,2	2,9	3	2	3	2	4,3	5,1	4,2	6,2	3,1	2,1	4,2	3,1	4,2	3,2	5,4
		2	1,2	2	2,0	4,9	5,0	4,8	3	3,8	3,9	2,9	5,0	7,0	2,1	2,8	2,9	1,8	2,9	1,8	4,2	5	4,1	6	3	2	4,1	3	4,1	3,1	5,3
	Contact	2,5	2,3	3,5	3,5	5,2	5,2	6,2	3,4	4,3	4,8	3,9	5,9	6,8	3,3	3,6	3,8	2,7	3,3	3,0	4,4	5,5	5,3	6,2	3,2	3,3	5,2	3,5	4,7	3,8	5,6
		2,4	2,2	3,4	3,4	5,0	5,1	6,0	3,2	4,2	4,6	3,8	5,8	6,7	3,2	3,5	3,7	2,6	3,2	3,0	4,3	5,5	5,2	6,1	3,2	3,2	5,1	3,5	4,6	3,7	5,5
		2,4	2,2	3,4	3,4	4,8	5,1	5,8	3,0	4,1	4,4	3,8	5,7	6,6	3,2	3,5	3,6	2,6	3,2	3,0	4,2	5,5	5,2	6,1	3,2	3,2	5,1	3,5	4,5	3,7	5,4
		2,3	2,1	3,3	3,3	4,6	5,0	5,6	2,8	4,0	4,2	3,7	5,6	6,5	3,1	3,4	3,5	2,5	3,1	3,0	4,1	5,5	5,1	6,0	3,2	3,1	5,0	3,5	4,4	3,6	5,3
	Education	2,1	1,8	4,0	3,7	6,2	5,7	7,2	2,0	3,8	5,8	3,5	6,0	7,5	3,5	2,1	2,0	1,9	1,7	3,3	3,2	5,2	5,5	7,3	1,5	3,0	5,0	3,2	5,0	2,1	4,1
		2,0	1,7	4,0	3,6	6,1	5,5	7,1	2,0	3,7	5,6	3,5	6,0	7,5	3,4	2,0	2,0	1,8	1,6	3,2	3,1	5,1	5,5	7,2	1,4	3,0	5,0	3,1	5,0	2,0	4,0
		2,0	1,6	4,0	3,6	6,0	5,3	7,1	2,0	3,6	5,4	3,5	6,0	7,5	3,4	1,9	2,0	1,7	1,5	3,2	3,0	5,1	5,5	7,2	1,3	3,0	5,0	3,0	5,0	2,0	4,0
		1,9	1,5	4,0	3,5	5,9	5,1	7,0	2,0	3,5	5,2	3,5	6,0	7,5	3,3	1,8	2,0	1,6	1,4	3,1	2,9	5,0	5,5	7,1	1,2	3,0	5,0	2,9	5,0	1,9	3,9
	Leadership	2,8	2,8	4,1	4,3	5,2	5,3	6,5	3,7	4,2	5,1	4,3	6,3	6,6	3,7	3,8	4,2	2,8	3,3	3,2	4,3	5,6	5,7	6,0	3,3	3,8	5,5	3,6	4,8	4,0	5,6
		2,7	2,7	4	4,2	5,1	5,2	6,5	3,5	4,1	5	4,2	6,2	6,4	3,5	3,7	4	2,7	3,2	3,1	4,2	5,5	5,5	6	3,22	3,7	5,4	3,4	4,7	4	5,5
		2,6	2,6	4	4,1	5,1	5,1	6,5	3,3	4,1	5	4,1	6,2	6,2	3,3	3,7	4	2,6	3,2	3,1	4,1	5,4	5,3	6	3,2	3,6	5,4	3,2	4,7	4	5,5
		2,5	2,5	4	4,0	5,0	5,0	6,5	3,1	4	4,9	4	6,1	6,0	3,1	3,6	3,8	2,5	3,1	3	4	5,3	5,1	6	3,1	3,5	5,3	3	4,6	4	5,4
Marketing assistant	Organization	1,3	2	3,1	4,1	1,2	2,3	4,0	3	1,4	2,2	3,1	4,7	2,0	2,3	3,3	4,3	1,2	2,1	1	2,4	3,3	4,4	1,2	2,1	3	4,2	1,5	2,2	3,4	4,2
		1,2	2	3	4,0	1,1	2,2	4,0	3	1,2	2,1	3	4,6	2,0	2,2	3,2	4,2	1,1	2	1	2,3	3,2	4,2	1,1	2	3	4,1	1,5	2,1	3,3	4,1
		1,1	2	2,9	4,0	1,1	2,2	4,0	3	1	2,1	3	4,6	2,0	2,2	3,1	4,1	1,1	2	1	2,2	3,1	4	1,1	2	3	4,1	1,5	2,1	3,2	4,1
		1	2	2,8	3,9	1,0	2,1	4,0	3	0,8	2	2,9	4,5	2,0	2,1	3,0	4	1	1,9	1	2,1	3	3,8	1	1,9	3	4	1,5	2	3,1	4
	Computer	1,1	2,1	2,9	4,1	0,9	2,1	3,1	5	1,1	2	3,2	4,5	1,0	2,2	3,2	4,2	1,3	2,1	1,2	2,3	3,1	4,4	1,1	2,1	3,2	4	1,4	2,3	3,2	4,1
		1	2	2,8	4,0	0,8	2,0	3,0	5	1	2	3,1	4,5	1,0	2,1	3,1	4,1	1,2	2	1,1	2,2	3	4,3	1	2	3,1	4	1,3	2,2	3,1	4
		1	2	2,7	4,0	0,8	2,0	3,0	5	1	2	3,1	4,5	1,0	2,1	3,1	4,1	1,1	2	1,1	2,1	3	4,2	1	2	3	4	1,2	2,2	3,1	4
		0,9	1,9	2,6	3,9	0,7	1,9	2,9	5	0,9	2	3	4,5	1,0	2,0	3,0	4	1	1,9	1	2	2,9	4,1	0,9	1,9	2,9	4	1,1	2,1	3	3,9
	Education	2,1	1,8	4	3,7	6,2	5,7	7,2	2	3,8	5,8	3,5	6,0	7,5	3,5	2,1	2	1,9	1,7	3,3	3,2	5,2	5,5	7,3	1,5	3	5	3,2	5	2,1	4,1
		2	1,7	4	3,6	6,1	5,5	7,1	2	3,7	5,6	3,5	6	7,5	3,4	2	2	1,8	1,6	3,2	3,1	5,1	5,5	7,2	1,4	3	5	3,1	5	2	4
		2	1,6	4	3,6	6	5,3	7,1	2	3,6	5,4	3,5	6	7,5	3,4	1,9	2	1,7	1,5	3,2	3	5,1	5,5	7,2	1,3	3	5	3	5	2	4
		1,9	1,5	4	3,5	5,9	5,1	7	2	3,5	5,2	3,5	6	7,5	3,3	1,8	2	1,6	1,4	3,1	2,9	5	5,5	7,1	1,2	3	5	2,9	5	1,9	3,9
Customer service representative	Computer	1,1	2,1	2,9	4,1	0,9	2,1	3,1	5	1,1	2	3,2	4,5	1	2,2	3,2	4,2	1,3	2,1	1,2	2,3	3,1	4,4	1,1	2,1	3,2	4	1,4	2,3	3,2	4,1
		1	2	2,8	4	0,8	2	3	5	1	2	3,1	4,5	1	2,1	3,1	4,1	1,2	2	1,1	2,2	3	4,3	1	2	3,1	4	1,3	2,2	3,1	4
		1	2	2,7	4	0,8	2	3	5	1	2	3,1	4,5	1	2,1	3,1	4,1	1,1	2	1,1	2,1	3	4,2	1	2	3	4	1,2	2,2	3,1	4
		0,9	1,9	2,6	3,9	0,7	1,9	2,9	5	0,9	2	3	4,5	1	2	3	4	1	1,9	1	2	2,9	4,1	0,9	1,9	2,9	4	1,1	2,1	3	3,9
	Education	2,1	1,8	4	3,7	6,2	5,7	7,2	2	3,8	5,8	3,5	6	7,5	3,5	2,1	2	1,8	1,6	3,2	3,1	5,1	5,5	7,3	1,5	3	5	3,2	5	2,1	4,1
		2	1,7	4	3,6	6,1	5,5	7,1	2	3,7	5,6	3,5	6	7,5	3,4	2	2	1,8	1,6	3,2	3,1	5,1	5,5	7,2	1,4	3	5	3,1	5	2	4
		2	1,6	4	3,6	6	5,3	7,1	2	3,6	5,4	3,5	6	7,5	3,4	1,9	2	1,7	1,5	3,2	3	5,1	5,5	7,2	1,3	3	5	3	5	2	4
		1,9	1,5	4	3,5	5,9	5,1	7	2	3,5	5,2	3,5	6	7,5	3,3	1,8	2	1,6	1,4	3,1	2,9	5	5,5	7,1	1,2	3	5	2,9	5	1,9	3,9
	Secretarial	1,5	2,0	3,3	4,0	2,8	3,4	4,8	3,3	6,3	3,3	3,3	5,1	3,5	2,7	2,9	3,5	1,5	2,0	1,8	2,6	3,9	4,8	3,2	1,9	3,1	4,4	2,0	3,2	2,9	4,1
		1,4	2,0	3,2	4,0	2,6	3,3	4,7	3,2	6,2	3,2	3,2	5,0	3,4	2,6	2,8	3,5	1,4	2,0	1,7	2,5	3,8	4,7	3,1	1,8	3,0	4,3	2,0	3,1	2,8	4,0
		1,4	2,0	3,2	4,0	2,4	3,2	4,5	3,2	6,1	3,1	3,1	5,0	3,2	2,6	2,7	3,5	1,4	2,0	1,6	2,5	3,7	4,7	3,1	1,8	3,0	4,2	2,0	3,1	2,7	4,0
		1,3	2,0	3,1	4,0	2,2	3,1	4,4	3,1	6,0	3,0	3,0	4,9	3,1	2,5	2,6	3,5	1,3	2,0	1,5	2,4	3,6	4,6	3,0	1,7	2,9	4,1	2,0	3,0	2,6	3,9
	Communicatio	1,3	2,0	2,9	3,8	1,7	2,7	3,8	4,0	6,0	2,4	3,2	4,7	2,4	2,3	3,2	4,1	1,5	2,3	1,3	2,7	3,5	4,4	2,1	2,3	3,0	4,1	1,7	2,6	3,3	4,4
		1,2	2	2,8	3,7	1,6	2,6	3,7	4	6	2,3	3,1	4,6	2,3	2,2	3,1	4	1,4	2,2	1,2	2,6	3,4	4,3	2	2,2	3	4	1,6	2,5	3,2	4,3
		1,2	2	2,7	3,6	1,5	2,4	3,6	4	6	2,3	3,1	4,5	2,2	2,1	3	4	1,4	2,1	1,2	2,4	3,3	4,2	2	2,1	3	4	1,4	2,5	3,2	4,1
		1,1	2	2,6	3,5	1,4	2,3	3,5	4	6	2,2	3	4,4	2,1	2	2,9	3,9	1,3	2	1,1	2,3	3,2	4,1	1,9	2	3	3,9	1,3	2,4	3,1	4
Shipping clerk	Dependability	1,55	1,8	2,6	3,05	3,05	3,55	4,25	4	2,75	3,1	3,2	4,85	4,1	2,3	3,15	3,7	1,85	2,65	1,8	3,35	4,2	4,35	3,85	2,65	2,75	4,2	2,3	3,3	3,3	4,8
		1,5	1,7	2,5	3	3	3,5	4,2	4	2,7	3	3,1	4,8	4	2,2	3,1	3,6	1,8	2,6	1,7	3,3	4,1	4,3	3,8	2,6	2,7	4,1	2,2	3,2	3,2	4,7
		1,5	1,6	2,4	3	3	3,4	4,2	4	2,6	3	3	4,6	4	2,1	3,1	3,5	1,6	2,5	1,5	3,3	4	4,2	3,8	2,6	2,5	4	2,1	3,1	3,1	4,5
		1,4	1,5	2,3	2,9	2,9	3,3	4,1	4	2,5	2,9	2,9	4,5	3,9	2	3	3,4	1,5	2,4	1,4	3,2	3,9	4,1	3,7	2,5	2,4	3,9	2	3	3	4,4
	Organization	1,3	2	3,1	4,1	1,2	2,3	4	3	14	2,2	3,1	4,7	2	2,3	3,3	4,3	1,2	2,1	1	2,4	3,3	4,4	1,2	2,1	3	4,2	1,5	2,2	3,4	4,2
		1,2	2	3	4	1,1	2,2	4	3	1,2	2,1	3	4,6	2	2,2	3,2	4,2	1,1	2	1	2,3	3,2	4,2	1,1	2	3	4,1	1,5	2,1	3,3	4,1
		1,1	2	2,9	4	1,1	2,2	4	3	1	2,1	3	4,6	2	2,2	3,1	4,1	1,1	2	1	2,2	3,1	4	1,1	2	3	4,1	1,5	2,1	3,2	4,1
		1	2	2,8	3,9	1	2,1	4	3	0,8	2	2,9	4,5	2	2,1	3	4	1	1,9	1	2,1	3	3,8	1	1,9	3	4	1,5	2	3,1	4
	Computer	1,1	2,1	2,9	4,1	0,9	2,1	3,1	5	1,1	2	3,2	4,5	1	2,2	3,2	4,2	1,3	2,1	1,2	2,3	3,1	4,4	1,1	2,1	3,2	4	1,4	2,3	3,2	4,1
		1	2	2,8	4	0,8	2	3	5	1	2	3,1	4,5	1	2,1	3,1	4,1	1,2	2	1,1	2,2	3	4,3	1	2	3,1	4	1,3	2,2	3,1	4
		1	2	2,7	4	0,8	2	3	5	1	2	3,1	4,5	1	2,1	3,1	4,1	1,1	2	1,1	2,1	3	4,2	1	2	3	4	1,2	2,2	3,1	4
		0,9	1,9	2,6	3,9	0,7	1,9	2,9	5	0,9	2	3	4,5	1	2	3	4	1	1,9	1	2	2,9	4,1	0,9	1,9	2,9	4	1,1	2,1	3	3,9

Appendix 2: Expected and minimal salaries per annum (in thousands USD), per applicant

applicant	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30
exp	15	15	41	55	60	65	100	20	40	60	42	100	100	18	21	43	15	15	17	40	100	95	90	16	22	90	17	55	23	95
min	13	13	38	53	55	60	95	18	36	54	38	100	97	16	20	39	13	14	15	35	96	90	88	13	20	86	15	52	19	92

References

[1] Atanassov, K.; Gargov, G: Interval Valued Intuitionistic Fuzzy Sets, Fuzzy Sets and Systems, Vol. 3, No. 3, 1989, pp. 343-349

[2] Ballesteros-Pérez, P.; González-Cruz, C.; Fernández-Diego, M: Human Resource Allocation Management in Multiple Projects Using Sociometric Techniques, International Journal of Project Management, Vol. 30, No. 8, 2012, pp. 901-913

[3] Bellman, R. E.; Zadeh, L. A: Decision-Making in a Fuzzy Environment. Management Sci. Vol. 17, 1970, pp. 141-164

[4] Dubois, D.; Prade, H: Systems of Linear Fuzzy Constraints, Fuzzy Sets and Systems, Vol. 3, 1980, pp. 37-48

[5] Dubois, D: Linear Programming with Fuzzy Data, in: J. C. Bezdek (Ed.), Analysis of Fuzzy Information 3: Applications in Engineering and Science, CRC Press, Boca Raton, FL, 1987, pp. 241-263

[6] Filho, C. F. F. C.; Rocha, D. A. R.; Costa, M. G. F.; Pereira, W. C. A: Using Constraint Satisfaction Problem Approach to Solve Human Resource Allocation Problems in Cooperative Health Services, Expert Systems with Applications, Vol. 39, No. 1, 2012, pp. 385-394

[7] Inuiguchi, M.; Sakawa, M.; Kume, Y: The Usefulness of Possibilistic Programming in Production Planning Problems, Int. J. Prod. Economics, Vol. 33, 1994, pp. 45-52

[8] Inuiguchi, M.; Tanino, T.; Sakawa, M: Membership Function Elicitation in Possibilistic Programming Problems, Fuzzy Sets and Systems, Vol. 111, No. 1, 2000, pp. 29-45

[9] Klir, G.; Yuan: Fuzzy Sets and Fuzzy Logic: Theory and Applications, Prentice Hall PTR, 2005

[10] Lin, C-M; Gen, M: Multi-Criteria Human Resource Allocation for Solvingmultistage Combinatorial Optimization Problems Using Multiobjective Hybrid Genetic Algorithm, Expert Systems with Applications, Vol. 34, 2008, pp. 2480-2490

[11] Nagy, K.; Diveki, Sz.; Odry, P.; Sokola, M.; Vujicic, M: A Stochastic Approach to Fuzzy Control, Acta Polytechnica Hungarica, Journal of Applied Sciences, Vol. 9, No. 6, 2012, pp. 29-48

[12] Negoita, C. V.; Minoiu, S.; Stan, E: On Considering Imprecision in Dynamic Linear Programming. Economic Comput. Economic Cybernet. Stud. Res, Vol. 3, 1976, pp. 83-95

[13] Osman, M. S.; Abo-Sinna, M. A.; Mousa, A. A: An Effectivegenetic Algorithm Approach to Multiobjective Resource Allocationproblems, Applied Mathematics and Computation, Vol. 163, 2005, pp. 755-768

[14] Portik, T.; Pokorádi, L: The Summarized Weighted Mean of Maxima Defuzzification and Its Application at the End of the Risk Assessment Process, Acta Polytechnica Hungarica, Vol. 11, No. 3, 2014, pp. 167-180

[15] Rezaei, P.; Rezaie, K.; Nazari-Shirkouhi, S.; Jamalizadeh Tajabadi, M. R: Application of Fuzzy Multi-Criteria Decision Making Analysis for Evaluating and Selecting the Best Location for Construction of Underground Dam, Acta Polytechnica Hungarica, Vol. 10, No. 7, 2013, pp. 187-205

[16] Rodríguez, R. M.; Martínez, L.; Herrera, F: A Group Decision Making Model Dealing with Comparative Linguistic Expressions Based on Hesitant Fuzzy Linguistic Term Sets, Information Sciences, Vol. 241, 2013, pp. 28-42

[17] Rodríguez, R. M.; Martínez, L.; Herrera, F: Hesitant Fuzzy Linguistic Term Sets for Decision Making, IEEE Transactions on Fuzzy Systems, Vol. 20, No. 1, 2012, pp. 109-119

[18] Saaty, T. L: Decision Making with Dependence and Feedback: the Analytic Hierarchy Process, RWS Publications, 4922 Ellsworth Avenue, Pittsburgh PA 15213, 1996

[19] Saaty, T. L.; Vargas, L. G.; Dellmann, K: The Allocation of Intangible Resources: the Analytical Hierarchy Process and Linear Programming, Socio-Economic Planning Sciences, Vol. 37, 2003, pp. 169-184

[20] Saaty, T. L.; Peniwatib, K.; Shanga, J. S: The Analytic Hierarchy Process and Human Resource Allocation: Halfthe Story, Mathematical and Computer Modelling, Vol. 46, 2007, pp. 1041-1053

[21] Silva, L. C.; Costa, A. P. C. S: Decision Model for Allocating Human Resources in Information System Projects, International Journal of Project Management, Vol. 31, No. 1, 2013, pp. 100-108

[22] Takács, M: Multilevel Fuzzy Approach to the Risk and Disaster Management, Acta Polytechnica Hungarica, Journal of Applied Sciences, Vol. 7, No. 4, 2010, pp. 91-102

[23] Torra, V: Hesitant Fuzzy Sets, International Journal of Intelligent Systems, Vol. 25, 2010, pp. 529-539

[24] Zemova, B.; Talasova, J: Fuzzy Sets in HR Management, Acta Polytechnica Hungarica, Journal of Applied Sciences, Vol. 8, No. 3, 2011, pp. 113-124

[25] Xia, M.; Xu, Z: Hesitant Fuzzy Information Aggregation in Decision Making, International Journal of Approximate Reasoning, Vol. 52, 2011, pp. 395-407

A New Wind Turbine Concept: Design and Implementation

Bedri Kekezoğlu, Muğdeşem Tanrıöven, Ali Erduman

Yildiz Technical University, Electrical and Electronics Faculity, Electrical Engineering Department, Istanbul, Turkey
bkekez@yildiz.edu.tr, tanriov@yildiz.edu.tr, aerduman@yildiz.edu.tr

Abstract: Electrical energy demand has been continuously increasing. Depleting fossil fuel reserves, environmental concerns, and insufficiency of conventional generation techniques in meeting growing demand, renewable energy use has been widely adopted in the world. When considering the application of renewable energy sources in the world, it can be seen that wind energy is mostly preferred over other renewable energy sources. In this study, a new prototype wind energy conversion system suitable for urban use is designed and manufactured. The proposed design is modular and has flexible structure. In the new design, an outer gear ring attached to turbine blades is used. In the design stage, both the number of blades and the number of outer gear rings are varied to analyze their effect on turbine performance. Performance analysis of the prototype wind turbine is completed under real life conditions and results are given. As a result of this study, it is shown that increasing the gearwheels and blade numbers caused turbine output power increases. The most efficient structure identified during the field analysis is a three gearwheel with six blade system.

Keywords: renewable energy resources; wind energy; wind turbine prototype; urban use

1 Introduction

Wind energy outshines all other renewable energy resources due to the recent technological improvements. Electrical energy generation from wind power has increased rapidly and due to the increased interest many studies on efficient wind turbine design have been performed.

There are several studies about improvement of wind turbine performance in literature. Ameku et all. designed a 3 kW wind turbine prototype focusing on blade design [1]. Kosasih and Tondelli analyzed a low power turbine that has a conical structure to speed up air flow through the wind turbine in the laboratory environment [2]. Chong et al. presented a vertical axis wind turbine for high buildings [3]. A. Ali et al. developed a new vertical axis wind turbine. In the

study, the power generation potential of the system was analyzed and turbine blades were tested for different configuration and wind speed values [4]. Park et al. developed a 3 MW offshore wind turbine [5]. Solero et al. improved a wind turbine which is used for cold weather conditions [6]. Nagai et al. presented a performance of 3 kW wind turbine with a variable slope control system [7].

Although wind power plants contribute to electrical energy generation, they are built far away from consumption centers due to several factors such as, large area needed (not suitable for urban use), shadow effect and noise problems. Therefore, a new wind turbine design for the cities, which has low wind speed profile and high-population-density considerations, becomes a critical need.

Mainly, wind turbines that are used in urban areas, and produce power at low wind speeds, should be visually compatible with the environment and should work with low noise. In determining mounted places of wind turbines used in urban areas, turbulence should be minimized by taking into account the surrounding structures. Despite these challenges, there remains a growing interest for roof mounted wind turbines [8, 9, 10].

In this study, a new mini wind turbine concept suitable for urban use and low wind speeds is demonstrated. The mini wind turbine concept has a modular structure and can be optimized for different conditions and wind sites. In this paper, performance analyses of the mini wind turbine concept are performed and the results are summarized.

The paper is organized as follows: General information and mathematical background of this study is given in Section II. The design parameters, details of proposed wind turbine and obtained results are presented in Section III. Analysis of the results and future studies are given in Section IV. Section V presents the conclusion.

2 Wind Energy Conversion Systems

Wind energy conversation systems (WECS) has become the most useful renewable energy source among all others. World Wind Energy Association (WWEA) data shows that installed wind power is 254 GW as of mid-year 2012. Also, it is expected that installed wind power will be 273 GW by the end of 2012 [11]. European Union generated %6,3 of its energy using 93,6 GW from installed wind power in the year 2011 [12].

As known, kinetic energy of wind is converted to electrical energy by using wind turbines. Mechanical power production of a wind turbine is given in Equation 1 [13].

$$P_m = \frac{1}{2} \cdot \rho . A . v^3 . C_p \quad (W) \tag{1}$$

where, P_m, turbine mechanical power output, ρ, air density, A, turbine swept area, v, wind speed and C_p, turbine power coefficient. Power coefficient varies with blade tip speed ratio (λ) and blade angle (ϑ). Power coefficient of a wind turbine could be calculated by using Equation 2 and Equation 3, respectively.

$$\mathrm{C_p}(\lambda, \vartheta) = c_1 \left(c_2 \frac{1}{\beta} - c_3 \vartheta - c_4 \vartheta^x - c_5 \right) e^{-c_6 \frac{1}{\beta}} \tag{2}$$

$$\frac{1}{\beta} = \frac{1}{\lambda + 0.08\vartheta} - \frac{0.035}{1 + \vartheta^3} \tag{3}$$

Coefficients used in Eq. 2 are empirical constants. Blade tip speed ratio could be obtained by Equation 4.

$$\lambda = R\varpi / v \tag{4}$$

where, R, blade diameter and ω, angular velocity of turbine blade.

3 Mini Wind Turbine Concept

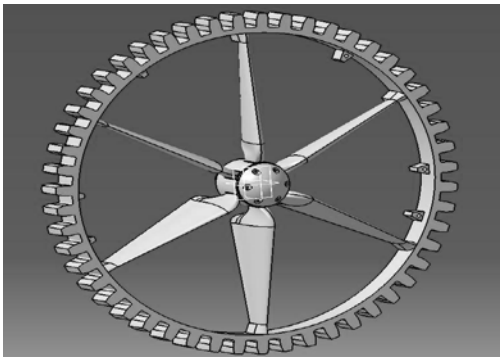


Figure 1
Design of mini wind turbine concept

In this study, a new wind turbine system is presented. The proposed mini wind turbine could be used in low altitudes with low wind speed and can be optimized for various conditions thanks to its modular structure. In the new turbine, the

turbine blades are located in the gearwheels. Thus, the swept area can be adjusted according to the requirements. The designed system is shown in Figure 1.

The realized wind turbine prototype is designed to obtain maximum output power with the smallest blades possible. The blades could be coupled to the outer gearwheel. Thereby, surface area could be increased to achieve higher output power. The outer gearwheel is flexible for variable blade numbers. This feature provides optimum system designs for different wind sites.

3.1 Design Studies

All components of new mini wind energy conversion system are individually designed. Blades, outer gearwheel and other components of the system are respectively designed and realized. Following sub-sections explains design details.

3.1.1 Blade Design

The introduced mini wind turbine concept is envisaged to produce high power for low wind speeds. Therefore, the blade structure is designed for low wind speed conditions with high efficiency. The blade type S833 which was developed by National Renewable Energy Laboratory (NREL) is used for the prototype system. This blade profile is placed in a thick blade category with a 18% blade thickness ratio. Also, this type of blade is especially used up to 1-3 m blade diameters.

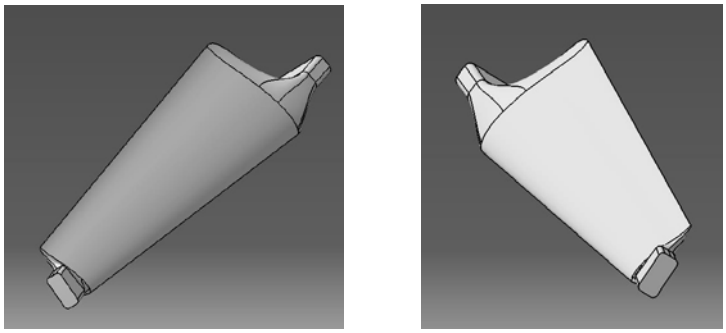
The blade length of the first prototype intended to be 30 cm. Other quantities of turbine blades are determined parametrically based on the blade length. Calculated aerodynamic parameters of the mini wind turbine are given in Table 1.

Table 1
Aerodynamic Parameters of Turbine Blade

Parameter	Value
Angle of Attack	10^{O}
Torsional Angle	2^{O}
Taper Ratio	0,414

The aim is that developed mini wind turbine blades are positioned within a gearwheel. For this purpose, additional junction points are added on the upper ends of the wings to provide a gearwheel connection.

As known, interconnected gearwheels have a different direction of rotation. Because of this reason, mini wind turbine blades must act in different directions to work with multiple gearwheels. This situation was taken into account in the design studies and two different blade designs were developed for clockwise and counterclockwise rotations. Design studies are shown in Figure 2 for both blade shapes.

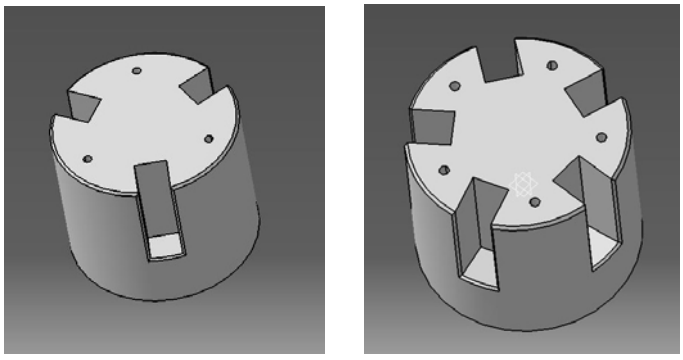


(a) (b)

Figure 2
(a) Clockwise rotating blade, (b) Counter-clockwise rotating blade

3.1.2 Hub Design

The mini wind turbine is intended to be used with different blade numbers and can be optimized for different wind potentials. For this purpose, four different hubs suitable for 3, 4, 5 and 6 blades utilization, are designed for prototype studies. Grooves are located on the designed hub to facilitate the blade connection. Hubs developed for different blade numbers are, shown in Figure 3.



(a) (b)

Figure 3
(a) Hub design for three blade connection, (b) Hub design for five blade connection

In the design phase, the most important challenge was the connection between hub and blade. In order to keep the turbine blades stable with no oscillation, a gradient is given in the grooves to prevent forward thrusting of the mounted blades. Blades are locked into the hub by using a hub cover which is shown in Figure 4.

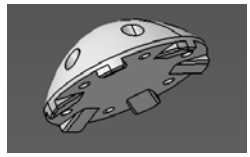


Figure 4
Hub cover for six blade connection

3.1.3 Gearwheel Design

The main purpose of the new wind turbine system is to have a modular structure, which provides a way to achieve increased surface area. Herein wind turbine blades are positioned within a gearwheel. Gearwheels are designed so as to include 3, 4, 5 or 6 of the same length blades. Technical details of gearwheels are given in Table 2.

Table 2
Aerodynamic Parameters of Turbine Blade

Parameter	Value
Module (m)	10
Tooth number (z)	71
Internal diameter (d_t)	710 mm
Tooth root diameter (d_f)	486.68 mm
Tooth top diameter (d_a)	530 mm
Step	31.416 mm

The designed gearwheel is given in Figure 5.

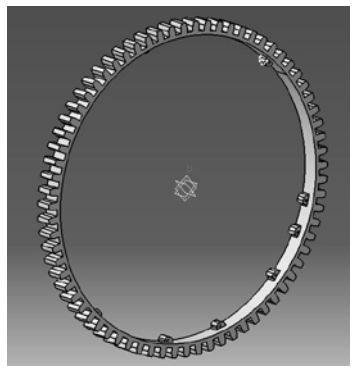


Figure 5
Gearwheel used on mini wind turbine

One may think that the proposed gear system could cause additional friction on the baled system. This is minimized, firstly, because the gear wheels are made of polyethylene and manufactured to obtain minimum friction which results in the friction on the gear wheels having a relatively negative effect on the power generation. As known, wind turbine power generation is proportional to the blades swept area, and, secondly the swept area of the wind turbine can be optimized according to the wind profile of the site. Therefore, the friction losses can be ignored in comparison to power generation gain, because of the gear wheel additions.

3.2 Prototype Studies

In the prototype study, the purpose is to manufacture light weight wind turbine components. For this purpose, the blades of the turbine is manufactured with a fiberglass reinforced composite material. Because of its high strength and thermal conductivity, fiberglass is minimally affected by environmental variations. Most commercial wind turbines, today are also manufactured with fiberglass reinforced composite materials.

For gears, polyethylene materials are used since these materials are not affected by variable ambient conditions. Additionally, polyethylene material is useful for reducing the noise problem. The gearwheel that surrounds the blades is produced by using CNC machines. Aluminum materials are preferred for fabrication of the turbine hub and its cover. The reason for this is the strength and light weight of aluminum. Thus, blades of the turbine could easily bear these components. Also, these components are manufactured by CNC machines.

A DC generator that has 650 W nominal power is used in this prototype wind turbine. Technical parameters of the generator are given in Table 3. Output power of the generator with respect to rotational speed is presented in Figure 6.

Table 3
Technical parameters of the generator

Parameter	Value
Generator Type	Brushless DC
Nominal Spin	700 rpm
Nominal Power	650 W
Output Voltage	24 V
Output Current	25,8 A

As seen in Figure 6, output power of the generator increases rapidly over 200 rpm.

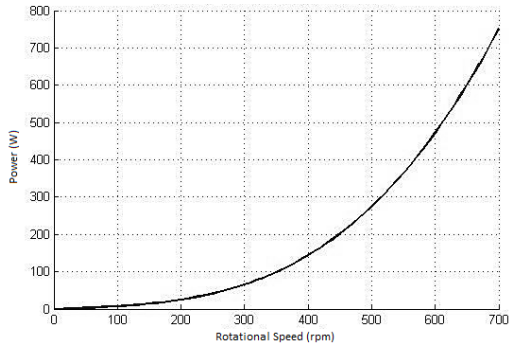


Figure 6
Generator characteristics

Components which are used to design the prototype wind turbine are given in Figure 7.

Figure 7
Components of the prototype wind turbine

The main advantage of the proposed system is its modular structure so that various blade numbers and gearwheel configurations can be used, thus allowing different working conditions to be analyzed. The platform of the prototype system is made of aluminum profiles. Three gearwheels could be used on this platform for synchronous applications. The three gearwheel with six blade system configuration is given in Figure 8.

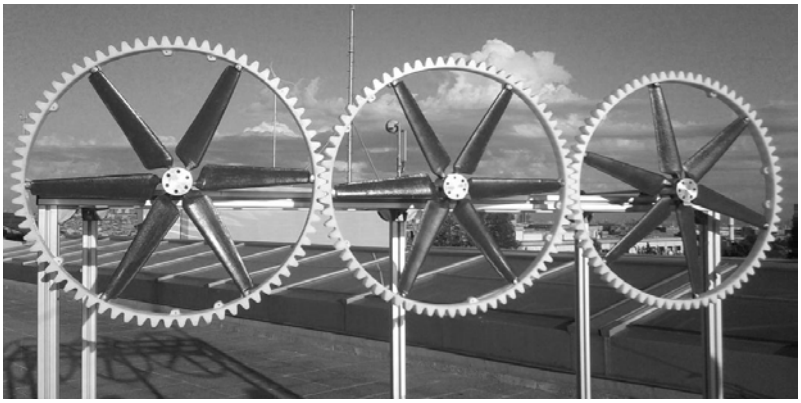


Figure 8
Three gearwheels with six blades system

3.3 Prototype Performance Analysis

After manufacturing the proposed prototype, performance values were obtained under actual operating conditions. Characteristic curves of the mini turbine system were obtained by analyzing the measured data. The variations of rotational speed versus wind speed data for three gearwheels with six blades system is presented in Fig. 9.

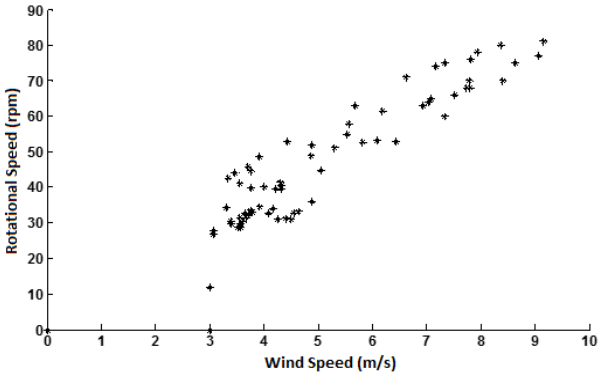


Figure 9
Data distribution for three gearwheels with six blades system

All the measured system data obtained, from the rotor speed versus the wind speed variation of the mini wind turbine study, is shown in Figure 10. Herein, the performance curves are given a six blade structure with a three-gearwheel system. As stated, the mini wind turbine prototype was designed to be able to be used with different gearwheel and blade numbers.

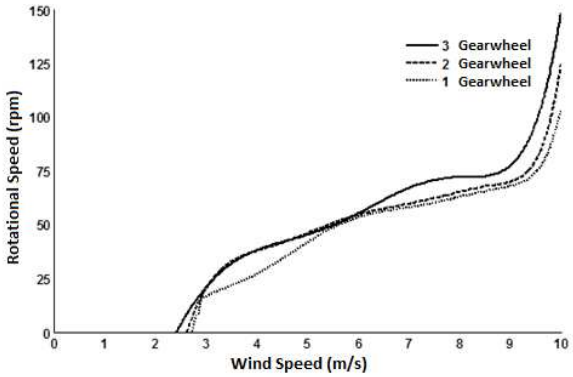


Figure 10
Performance curves of six blade structure

The study results presents that the number of blades and gearwheels directly effects the turbine performance. The friction of gearwheels used in the system is minimized by the manufacturing technique. The measurement results show that turbine performance increases with increasing gearwheel numbers. The effect of gearwheel number for high wind speeds can be clearly seen in Figure 10. Similarly, increased blade numbers also increases the turbine performance.

The power curve of the proposed system is given in Figure 11, where the 600 W generator is driven by a gearbox with a 1:5 conversion rate. As seen in Fig. 11, the power generation begins at approximately 4 m/s wind speed using a six blade, three gearwheel structure. Power generation significantly increases at higher wind speeds and reaches the nominal output power after 9 m/s. Based on the site requirements, the mini wind turbine can be optimized by changing dimensions or increasing the gearwheel number, in other words increasing turbine surface area.

As seen in Figure 11, when the rotational speed of the generator uses an additional gear wheel, more power can be obtained with the proposed system. As an example, when a gear wheel with 1:5 ratio is used, nominal output power of 600 W can be obtained at a 10 m/s wind speed. It can be seen that generated power using the proposed wind turbine with additional gear wheel is comparable to nominal power outputs of commercial wind turbines at 10-12 m/s wind speed.

Nowadays, wind turbines cut-in speed levels are 3-4 m/s. The proposed system does start power generation at a 4 m/s wind speed, but it works at lower efficiency rates below a 7 m/s wind speed. It is expected that, a more efficient system design can be achieved by selecting lighter materials for the whole system and increasing the number of gear wheels used in the system.

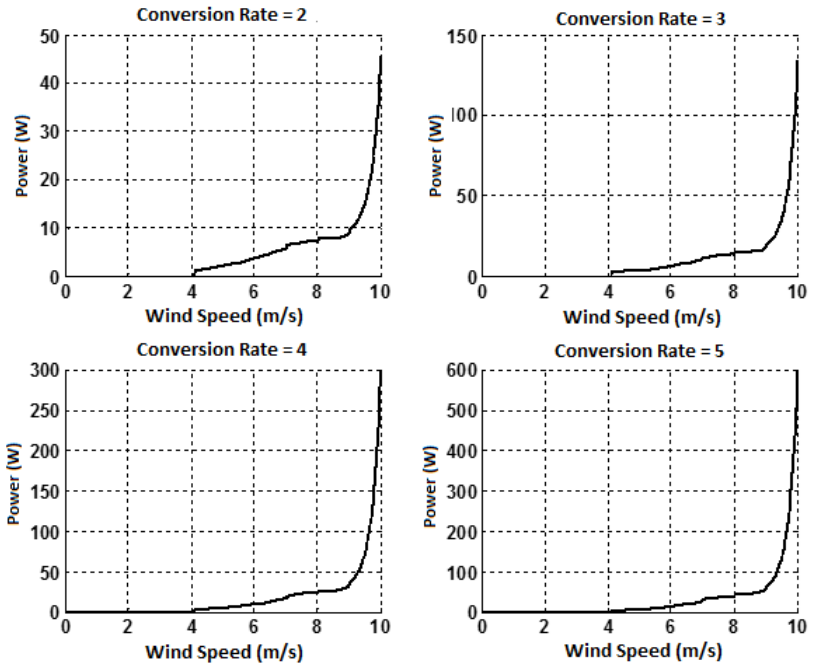


Figure 11
Power curves of mini wind turbine with 1:5 gearbox conversion rate

4 Future Studies

The analysis results show that the developed system is heavier than that of commercial wind turbines. Thus, if the mini wind turbine can be produced with lighter components, it will be possible to generate power at lower wind speeds.

Since the generator used in the proposed system should have a high torque response at lower wind speeds, a generator design should be completed to improvebetter system performance.

In addition, an electromechanical system that can track the wind direction variation should be included in the proposed system in order to maximize wind utilization.

Conclusions

In this paper, a new urban wind turbine concept that can be mounted on roofs, is presented. With the developed urban wind turbine system, wind energy potential can be used more efficiently in supplying city demands.

In this study, electrical energy generation is realized based on a gearwheel coupled to turbine blades. The number of blades can be varied depending on wind characteristics. The developed turbine is designed to be used with three, four, five or six blades.

The produced prototype system consists of a gearwheel fixed on the blades. The number of gearwheels can be also varied. Thus, surface area of the turbine can be enlarged to harness more wind turbine.

All components of the mini wind power wind turbine are manufactured based on modeling and design studies. Analysis results show that increasing both blade and gearwheel numbers increases the turbine mechanical power output. The maximum performance is obtained when a six blade three gearwheel structure is used for the given wind speed profile.

Due to its modular structure, the proposed design can be adopted for different buildings.

Acknowledgement

This work was supported by Yildiz Technical University Scientific Research Projects Coordination Department. Project Number: 29-04-02-KAP01.

References

[1] Ameku, K., Nagai, B. M. and Roy, J. N., Design of A 3 KW Wind Turbine Generator with Thin Airfoil Blades, Experimental Thermal and Fluid Science, 2008, 32:1723-1730

[2] Kosasih, B. and Tondelli, A., Experimental Study of Shrouded Micro-Wind Turbine, Procedia Engineering, 2012, 49:92-98

[3] Chong, W. T., Pan, K. C., Poh, S. C., Fazlizan, A., Oon, C. S., Badarudin, A. and Nik-Ghazali, N., Performance Investigation of a Power Augmented Vertical Axis Wind Turbine for Urban High-Rise Application, Renewable Energy, 2013, 51:388-397

[4] Ali, A., Goldeb, S., Alama, F. and Moria, H., Experimental and Computational Study of a Micro Vertical Axis Wind Turbine, Procedia Engineering, 2012, 49:254-262

[5] Park, J., Kim, J., Shin, Y., Lee, J. and Park, J., 3 MW Class Offshore Wind Turbine Development, Current Applied Physics, 2010, 10:307-310

[6] Solero L., Caricchi F., Crescimbini F. and Falchetta M., Direct-Drive Wind Generator Pilot Plant for Stand-Alone Units in Extremely Cold Climates, Int J REE, 2001, 3(2):326-329

[7] B. M. Nagai, B. M., Ameku K. and Roy J. N., Performance of A 3 kW Wind Turbine Generator with Variable Pitch Control System, Applied Energy, 2009, 86:1774-1782

[8] L. Ledo, P. B. Kosasih and P. Cooper, Roof Mounting Site Analysis For Micro-Wind Turbines, Renewable Energy, 2011, 36:1379-1391

[9] Y. F. Wang and M S. Zhan., 3- Dimentional CFD Simulation and Analysis on Performance of a Micro-Wind Turbine Resembling Lotus In Shape, Energy ang Buildings, 2013, 65:66-74

[10] L. Santoli, A. Albo, D. A. Garcia, D. Bruschi and F. Cuma, A Preliminary Energy and Enviromental Assessment of a Micro Wind Turbine Prototype In Natural Protected Areas, Sustainable Energy Technologies and Assessments, 2014, 8:42-56

[11] WWEA, World Wind Energy Association. 2012 Half year report. WWEA; 2012

[12] EWEA, The European Wind Energy Association. A report by The European Wind Energy Association 2011. Belgium: EWEA; 2011

[13] Gilbert, M. M., Renewable and Efficient Electric Power Systems, 2004, John Wiley & Sons, USA

Effects of Increasing the Power of Retail Chains on Competitive Position of Wholesalers

Stipe Lovreta[1], Jelena Končar[2], Ljiljana Stanković[3]

[1] Faculty of Economics, University of Belgrade, Kamenicka 6, 11000 Belgrade, Serbia; slovreta@ekof.bg.ac.rs

[2] Faculty of Economics Subotica, University of Novi Sad, Segedinski put 9, 24000 Subotica, Serbia; jkoncar@ef.uns.ac.rs

[3] Ministry of Trade, Tourism and Telecommunications, Government of the Republic of Serbia, Nemanjina 22-26, 11000 Belgrade, Serbia; ljiljana.stankovic@mtt.gov.rs

Abstract: Highly noticeable processes of retail market concentration have marked the end of the last and the beginning of this century. The intensive development of retail chains and high-level of increasing of their power caused the shifting of powers, changes of positions and the establishment of new relationships between the members of marketing channels. Retailing revolution which makes substantial changes to former marketing channels is under way. A lack of understanding these changes is especially noticeable on the developing markets where large retail chains assume key roles in creating and managing the channels. Strengthening of their position is more than evident. The position of wholesalers in these trends, i.e. need for their intermediating role between producers and retailers, becomes especially interesting. The situation is alarming, primarily, in the field of food and other fast moving consumer goods, where the changes manifest themselves to a high extent and with special content. Wholesalers are facing a less significant role with growing retailers, which are constantly making more numerous and more complex demands of them. In such conditions, many wholesalers disappear from the market. However, wholesalers still have significant possibilities and advantages of cooperation with large retail chains, which can strengthen their position and ensure survival in the marketing channels. This study confirms it, by proving that increasing the power of retailers brings a higher level of satisfaction for wholesalers and that they achieve a higher level of satisfaction in doing business with large rather than with small and medium-sized retailers.

Keywords: marketing channels; retail chains; wholesalers; power

1 Introduction

Modern markets, both mature and consolidating ones, are characterized by intensive trends, which immensely changes positions and relationships between members in marketing channels. Due to the power of the changes those trends can surely be considered as a retailing revolution. Despite the fact that individual markets have their own specificities, all market economies are characterized by certain common processes. Parallel to the increase of retail market concentration and increase of market share of the individual retail chains, the power of retailers is constantly increasing compared to the power of the other members of marketing channels. Retail chains become gatekeepers for approaching consumers [6], and simultaneously, they become larger and more powerful, individually. Also, retail as a sector, both in national and global markets, becomes more powerful. As John Dawson [3] [4], one of most eminent theorists in this field, emphasizes in his works, retail is transforming from a reactive to a highly proactive economic sector. Retailers become considerably larger than all the other members of marketing channels, even more so than the producers themselves. In accordance with this, due to their enormous buying power, large market shares, and sophisticated managements, giant retailers are considered power retailers and category killers, which clearly shows their power and dominant position [18].

With an-ever-growing market share, there arises the opportunity for large retailers with large retail stores to demonstrate their buying power to producers and to their suppliers in general. In such conditions, the power imbalance between members within marketing channels creates the opportunity for more powerful members to act in an opportunistic manner and to impose force [1], which eventually can endanger cooperation and efficient functioning of marketing channels. The complex situation, produced by change in balance between the competitive power of retailers and wholesalers, is especially evident. Wholesalers find themselves in undesirable situations which can be handled only by a small number of them.

Primarily, less numbers of ever growing retailers, in all the markets, have begun to control approaches to consumers. On the other hand, they become more demanding towards their suppliers, in order to rationalize buying and logistic activities, but also to concentrate their buying power, retailers look for fewer larger suppliers which can cooperate with them [10]. They are intensively striving to establish control and power both in horizontal and vertical dimensions of marketing channels [20], thus becoming the primary initiator of the vertical integrations and cooperation processes.

Also, it is not in the wholesalers' favor, as Desai states [5], that all market players have a high-level of understanding that one more level in marketing channels, except when intermediary brings special contribution and lower costs, doubtlessly increases inefficiency and decreases profitability. In that sense, no one wants to have worse business results, especially retailers with significant market and

negotiating power. These strong market players are gradually taking upon themselves distributive tasks and numerous functions, formerly traditionally performed by wholesalers and producers. In this situation, the most important question is which one of the players retain the authority of performing a sufficient number of functions, so that its membership and intermediating role in a certain channel is justified [18].

However, if they keep their position in marketing channels, as well as all the other suppliers, wholesalers will also face the fact that retail chains mercilessly use their increased negotiating power, based primarily on accumulated market power, to shift negotiating positions with suppliers in their favor and purchase products at the lowest prices possible. The negotiating power of retailers and their constant pressure to decrease transactional prices create for the opposite side a constant need to increase efficiency and effectiveness, but unfortunately, also often seriously endanger the survival of the suppliers themselves. In these kinds of situations, sellers usually participate in supply to the extent when their profits become negative, while the dominant buyer maximizes its profit [7]. However, many of them withdraw too late. Wholesalers are especially exhausted by these situations, from which many of them never recover.

On the other hand, it is important to emphasize that by offering lower prices to its consumers, primarily based on achieved lower prices in purchases and increase of efficiency, and by ensuring a higher level of services in comparison to their competitors, large retail chains can continue to spill market share over into their favor. In this way they, also, contribute to a decrease in the number of small and medium-sized, classical, retailers which not only represent the wholesalers, but also the producers, and traditionally close a less demanding marketing channel, in which they can achieve significantly higher differences in prices and compensate that which was lost in competition with retail giants. In relation to that, small and medium-sized traders are undoubtedly very important for every market. For sure, the support policy connected with them, should be focused on: a) the enhancement of competitiveness, and b) innovation development [9]. At the same time, the most important areas for them to develop, are those such as cost management, trade and marketing, production, technical development and finances [12]. And, there is no time to lose.

In any case, in accordance with Maruyama [15], the modern changes in structure of the wholesale sector and in the position of wholesalers in marketing channels can be seen as a consequence of three most important factors:

- changes in retail structure which have directly influenced the changes in wholesale
- intensive development of informational technologies in distribution (“points of sale” system, electronic data exchange system and alike); and

- application of new strategies of distribution, such as supply chain management and strategic alliance between producers and distributors with implementation of efficient consumer response systems, logistics reforms and changes in supply management.

It is obvious that wholesale sector, parallel to strengthening of large retail chains and disappearance of small and independent retailers from the market, in highly complex conditions, fights for its own survival and undergoes a strong transformation. Vulnerable wholesalers are, with the purpose to preserve and increase their role in marketing channels, enforced to apply new strategies and to make significant adjustments to the new situation. Their survival under modern conditions requires permanent dynamics so that, in accordance with the level of a given market development, a set of wholesale`s roles substantially differs in transition, developing markets, industrialized ones and those with post-industrial market characteristics [19].

The changes have enforced the wholesalers to undergo significant transformations and to find different strategies for strengthening position in modern marketing channels. Many of them have made radical changes in their marketing behavior, and that as a precondition for survival within the market. Specialization and focus on the services they provide have become their basic advantages, and they keep on strengthening their position by finding new innovative ways to implement these [17]. Many market players have become aware of the fact that the flow of goods through the existing channels with more wholesale levels can be the least expensive way for marketing goods and services [8]. Today, along with other usual wholesale services, institutional wholesalers provide support, even to the largest retailers, and primarily they provide information on supply and demand trends on the market. They are aware of different situations. One of the most important is that a large informational gap between producers and retailers requires the intermediation by wholesalers so as to bridge the gap, making the marketing channels longer, while, opposite to this, a small informational gap, due to the development of product brands and presence of big retailers, makes distributional channels shorter and leaves little space for intermediation [21].

In accordance with this, wholesalers search for their place in marketing channels. Maybe the most impressive example for “outsourcing” the management of certain categories of products is when strong retailers give this activity to the leading suppliers in a given field. A practice known as “category captainship” means that wholesalers, category captains, usually conduct analyses on category and provide retailers with a detailed plan which includes recommendations related to which brands should be included in category, how to define prices for each product individually, how much space to allocate to each brand and where to position each brand on the shelves [14].

And, in addition to this, long-term relationships between partners in marketing channels are becoming a more important part of a long-term strategy of any

participant, which is particularly significant for wholesalers due to the specificity of their intermediating role [2]. Wholesalers, more than any other member of marketing channels in the current conditions, should be oriented to the development of long-term relationships, in accordance with their marketing orientation [11]. However, a precondition for the establishment of such cooperation is acceptance of the power imbalance. This is the first step towards successful building of mutual relationships between wholesalers and large retailers [10]. The best example of a very strong wholesale role in modern conditions is market of Japan in which large international retailers enter with a common desire to implement in this local market their original global distribution system, trade customs and business models. They are, primarily, interested in developing direct relationships with producers, i.e. with their suppliers, excluding the wholesalers, in order to avoid uncertainty and higher costs, which an additional intermediary can produce. However, when it comes to the Japanese market, they are relatively often enforced to change their patterns of conduct and use Japanese wholesalers [16]. Nishimura [16] states that, faced with specificities and complexity of distribution in Japan, many foreign retailers change their usual strategies, which Carrefour did soon after it entered the market, by deciding to rely on local wholesalers for goods originating from Japan.

Generally speaking, wholesalers in all markets do not favour the development of large retail chains and enforcement of their power. One can often hear the wholesalers, accustomed to cooperating with small, i.e. traditional retailers, criticizing the modern, huge, retailers and accusing them for the failure of a large part of institutional wholesalers. However, many proactive wholesale intermediary firms within the developed markets, and more often in the market of Serbia, find a way to turn the current situation to their advantage and use all the benefits that cooperation with large retailers provides. They are starting to observe strong retailers not as a danger, but as a huge opportunity to develop business.

2 Objective and Hypotheses of the Research

The subject of this research is the insufficiently known and rarely analyzed effects which modern changes in the sphere of retail cause to the competitive position of wholesalers. The aim is to evaluate the consequences of increasing the power of retailers on wholesalers, which requires thorough and comprehensive observation of positive and negative influences on the position and real interests of wholesalers. In order to enlighten the mentioned influences, we have conducted the researches which should confirm or disprove the defined hypotheses:

H1: *Increasing the power of retailers brings a higher level of satisfaction for wholesalers in their business activities, and it is in the interest of large, but not of small wholesalers.*

H2: *Wholesalers, in market conditions characterized by the presence of strong retail chains, express higher level of satisfaction in doing business with large retailers than with small and medium-sized ones.*

A detailed survey on wholesalers' satisfaction in Serbia, in accordance with the defined elements of a business relationship, enlightens the level of wholesalers' satisfaction with doing business with other members in marketing channels in modern conditions characterized by increasing the power of retailers, with special emphasis on cooperation with large and with small and medium-sized retailers individually. The whole picture of that cooperation is completed by the survey of the satisfaction index for individual elements of business relationship of wholesalers with large and with small and medium-sized retailers. Apart from that, the concrete wholesalers' responses also contribute to the evaluation of the existing level of their satisfaction in doing business with large and with small and medium-sized retailers.

The results of this research point to the real impact of changes in relationships between members in marketing channels in the market conditions characterized by constant increase of the retail power. Thereby, the results especially emphasize the existence of positive impact those changes make on the performance of the activities in the trade sector of the economy. The results should redirect uncritical opinions away from accusing the more and more efficient modern retailers of the unfavourable competitive position of wholesalers, which, again, due to their insufficient proactive approach, in terms of adjusting to new competitive conditions on the market, lose their positions and disappear from the market.

3 Research Methodology

With a view to researching the effects of demonstration of large retailers power on wholesalers in the marketing channels, comprehensive studies have been conducted on the market of the Republic of Serbia. In accordance with the defined research methodology, the obtained results have been analyzed and the effects of increasing and usage of retailers' power on the position of wholesalers have been evaluated. The obtained results were basis for giving answers to research hypotheses.

Instead of the (hardly feasible) possibility to precisely measure the power of large retailers and give a precise evaluation of positive and negative effects of its use on wholesalers, attitudes, i.e. opinions of a certain number of wholesalers were taken into consideration. Surveys, with satisfying scope for the current conditions on the Serbian market, presented a relatively good picture on the significance and effects of increasing the power of retailers on the competitive position of wholesalers in the marketing channels.

In order to perceive a new position of wholesalers and their relationship with other members in the marketing channels, satisfaction of wholesalers with doing business with other members in the marketing channels was determined by calculating satisfaction indices for each individual relationship. To test the defined hypotheses, it is essential to conduct deeper analysis by segmenting wholesalers and retailers into large and small and medium-sized ones, where the criterion of size was a number of employees (up to 200 employees for small and medium-sized and 200 employees for large ones).

Satisfaction indices of wholesalers were calculated for the following individual business relationships (*ij*): 1. wholesalers (all) with producers (1a); 2. wholesalers (all) with large retailers (1b); 3. wholesalers (all) with small and medium-sized retailers (1c); 4. large wholesalers with producers (2a); 5. large wholesalers with large retailers (2b); 6. large wholesalers with small and medium-sized retailers (2c); 7. small and medium-sized wholesalers with producers (3a); 8. small and medium-sized wholesalers with large retailers (3b); 9. small and medium-sized wholesalers with small and medium-sized retailers (3c).

With the purpose of calculating the satisfaction indices of wholesalers for the individual business relationship, nine key criteria (k) (elements of a business relationship) were defined. These criteria were a basis for evaluating cooperation between the members and for each of the criteria, the survey respondents defined:

- relevance for cooperation - points from 1 to 10; and
- level of satisfaction with the existing cooperation (satisfaction rate) - points from 1 to 5.

The following table (Table 1) shows the elements of a business relationship which were used in the questionnaire:

Table 1
Elements of a business relationship - criteria (k)

(k)	**Elements of a business relationship – criteria (k)**
1	Prices, rebates and additional payments
2	Terms of payment (due to date) and regularity of payments
3	Sales potential
4	Range of assortment
5	Activities and costs to be borne by a concrete member
6	Cooperation in terms of promotional and other marketing activities
7	Data exchange and electronic communication
8	Level of trust
9	Potential for development and possibility of a long-term partnership

Apart from this, significance of single elements of a business relationship was a subject to analysis as well, meaning that satisfaction index for all the elements of a business relationship was calculated individually for relationships of wholesalers

with large retailers and with small and medium-sized retailers. Within the obtained data processing, the first step was normalization of satisfaction rate ($V_{ij,k}$) obtained for each element of a business relationship (k) individually, as follows:

$$X_{ij,\,k} = 100/4\ (V_{ij,k} - 1\ ),$$

(i=1, 2 or 3; j=a, b or c; k=1, 2,..., 9)

(in case that a wholesaler rated sales potential of a retailer with 5, then a value obtained by normalization would be 100, with 4 – 75, with 3 – 50, with 2 – 25 and with 1 – 0).

Satisfaction index $X_{ij,}$ which measures a relationship (*ij)*, was calculated as a weighted average of the obtained normalized rates $X_{ij,k}$ by each criteria. Ponder was obtained rate of relevance for cooperation of certain criteria, i.e. rated elements of a business relationship (from 1 to 10). Calculating satisfaction index for each individual relationship was performed according to the formula for weighted average:

$$X_{i\,j} = \frac{\sum w_{ij,k} X_{ij,k}}{\sum w_{ij\,.k}} \quad ,$$

where $W_{ij,k}$ stands for average ponders of a k-criteria of an individual relationship (*ij*).

The following table (Table 2) illustrates the procedure for calculating satisfaction index:

Table 2

Illustration of procedure for calculating satisfaction index

Elements of a business relationship – criteria (k)	**Satisfaction rate by criteria**	**Normalization of satisfaction rate**	**Relevance for cooperation by criteria**	**Result**
1	3	50	8	400
2	3	50	8	400
3	3	50	10	500
4	3	50	10	500
5	3	50	8	400
6	4	75	9	675
7	2	25	7	175
8	3	50	9	450
9	3	50	10	500
Total			79	4000

Satisfaction index in the presented case values 4000/79=50.63. The whole procedure provides the value of satisfaction index, which ranges from 0 to 100, and the same methodology is also applied for calculating satisfaction index by the elements of a business relationship individually (by criteria individually) for the relationship of wholesalers with large retailers and with small and medium-sized retailers.

The research was conducted on a stratified and purposely (non-random) chosen sample. The sampling included 15 wholesalers who generate most of their turnover through selling to the retailers within the market of the Republic of Serbia. In the chosen sample, 4 wholesalers with more than 200 employees (large wholesalers) and 11 wholesalers with less than 200 employees (small and medium-sized wholesalers) were interviewed. On average, they co-operate with 290 suppliers and 3620 customers, where the average number of customers among the retailers is 3240. All the wholesalers focus on food and other fast moving consumer goods.

In defining the sample, the researchers took into consideration adequate geographic distribution, size of undertakings measured by the number of employees and turnover, as well as a group of products with the largest share in turnover of a selected wholesaler. The sampling included 60% of turnover of food and other fast moving consumer goods which is generated in the wholesaler-retailer relation within the Serbian market. The basic method used for obtaining the data from wholesalers was a direct personal interview (face-to-face) which was guided by a pre-defined questionnaire, during the first quarter of 2013. The data collected within the conducted survey research were processed and analyzed using statistical program SPSS 15.0.

4 Research Results

In the questionnaire they filled out within the conducted survey research, the wholesalers, primarily, determined the relevance for cooperation of each offered element of business relationship with producers and, separately, with retailers. After that, they rated the level of current satisfaction within the existing business relationships with producers, and, separately, with large and with small and medium-sized retailers, based on the same elements.

According to the pre-defined methodology, and on the basis of the obtained rates, the average satisfaction indices of wholesalers in doing business with different partners in marketing channels were obtained. The results are presented in the following table (Table 3):

Table 3
Average satisfaction indices of wholesalers in doing business with other members in marketing channels (sample size: 15 wholesalers)

Satisfaction of → With ↓	**Wholesalers**	**large wholesalers**	**Small and medium-sized wholesalers**
Producers	67.23	77.78	63.26
large retailers	62.99	**70.54**	58.98
Small and medium-sized retailers	54.27	57.77	**52.75**

Judging by the obtained results, in the current conditions of reached, relatively significant, level of concentration within the market of the Republic of Serbia, wholesalers, especially the large ones, express the highest level of satisfaction with doing business with producers, which indicates that there is one completely new position of wholesalers as intermediary and their current significant role for the producers. Apart from this, large wholesalers doubtlessly express significantly higher level of satisfaction with doing business with all other members in marketing channels, whether it concerns producers, or large retailers or small and medium-sized retailers, in comparison to small and medium-sized wholesalers.

Also, it is noticeable that wholesalers express significantly higher level of satisfaction with doing business with large retailers than with small and medium-sized ones. It is interesting enough that even small and medium-sized wholesalers achieve higher level of satisfaction with their relationship with large retailers than with small and medium-sized ones which were their traditional partners until now.

Again, there is a large difference in the satisfaction indices when observing the relationship between large wholesalers and large retailers, on one side, and the relationship between small and medium-sized wholesalers and small and medium-sized retailers, on the other side. The satisfaction index is significantly higher for the relationship between large wholesalers and large retailers.

The research results also gave a more detailed picture of business relationship of wholesalers with retailers per individual elements, which is presented in Table 4.

This table also shows the relevance of certain elements for wholesalers in terms of their business relationship with retailers, as well as satisfaction indices of wholesalers with large and with small and medium-sized retailers for those elements, with ranking. It is noticeable that the most relevant element for wholesalers, in their relationship with retailers, is “prices, rebates and additional payments”, which includes sale prices and rebates given to retailers, as well as compensations to be paid additionally to retailers such as payments for “enlisting the products” and alike.

Table 4
Satisfaction indices of wholesalers per the elements of a business relationship of wholesalers with large and small and medium-sized retailers (sample size: 15 wholesalers)

Rank	Elements of a business relationship of wholesalers with retailers	Relevance of the element	Large retailers		Small and medium-sized retailers	
			Satisfaction index	Rank	Satisfaction index	Rank
1	**Prices, rebates and additional payments**	9.47	56.69	**8**	73.94	**1**
2	**Terms of payment (due to date) and regularity of payments**	9.27	61.15	**5**	59.53	**2**
3	**Sales potential of retailers**	8.47	74.61	**2**	56.89	**4**
4./ 5.	**Range of assortment of retailer**	7.67	79.13	**1**	58.04	**3**
4./ 5.	**Activities and costs to be borne by a concrete member**	7.67	58.70	**7**	55.00	**5**
6	**Level of trust in retailer**	7.40	66.89	**3**	52.25	**6**
7	**Cooperation in terms of promotional and other marketing activities**	7.33	62.27	**4**	49.32	**7**
8	**Potential for development and possibility of a long-term partnership**	7.13	59.58	**6**	48.60	**8**
9	**Data exchange and electronic communication**	5.60	47.92	**9**	34.82	**9**
AVERAGE SATISFACTION INDEX			**62.99**		**54.27**	

It is followed by “terms of payment (due to date)” and “regularity of payments”. At the same time, those two elements offer to wholesalers in their relationship with small and medium-sized retailers the highest level of satisfaction, rank 1 and 2, respectively. When it comes to large retailers, the satisfaction with these two elements is at a lower rank, while the highest rank, with the highest level of satisfaction, is related to the element “range of assortment of retailer,” which is followed by “sales potential of retailer”. However, it is highly significant that only for the first element, which is related to prices, rebates and additional payments is the satisfaction index higher for small and medium-sized retailers than for large ones, while for all other elements the significant advantage, in terms of the obtained satisfaction indices, is on large retailers’ side. This finally leads to decisively higher average satisfaction indices for the relationship of wholesalers with large retailers.

However, the concrete responses that wholesalers gave to the precise questions in connection to their attitudes about doing business with large and with small and medium-sized retailers make the picture of the current situation whole (Table 5):

Table 5
Concrete wholesalers' responds to the questions related to doing business with large and with small an medium-sized retailers (sample size: 15 wholesalers)

Questions	Responses	
Do you prefer doing business with large, modern, retailers or with small and medium-sized, traditional, retailers (who do you rather choose for a partner)?	With large, modern, retailers	With small and medium-sized, traditional, retailers
	40.00%	60.00%
Do you think that development of retail chains is in your and interest of other wholesalers?	I think	I do not think
	27.00%	73.00%

Although the level of satisfaction in doing business with large, modern retailers is significantly higher, only 40% of wholesalers responded that they rather choose for their partners large, modern, retailers. Apart from this, only 27% think that development of large retail chains is in their interest and in the interest of other wholesalers, while 73% of them believe the opposite.

5 Testing the Hypotheses

Researches have confirmed the first hypothesis that "*increasing the power of retailers brings a higher level of satisfaction for wholesalers in their business activities, and it is in the interest of large, but not of small wholesalers*". Average satisfaction indices have shown a very high-level of satisfaction of wholesalers with their business relationship with producers, which confirms that, in conditions of ever growing retail market concentration, producers find the wholesaler's role very important, as well as that wholesalers are relatively satisfied with the way producers treat them. The t-test assesses that the level of satisfaction of wholesalers with their business relationship is significantly higher (t=2.868 (prob. 0.021)), so that we can accept the first hypothesis.

Also, when observing their satisfaction, on one hand, with large, and on the other hand, with small and medium-sized retailers, we notice a considerably higher level of satisfaction of large, but also of small and medium-sized, wholesalers with relationship with large retailers, in comparison to relationship with small and medium-sized retailers. However, these two groups of wholesalers do not show statistically significant differences in satisfaction with relationships with large retailers (t = 1.706 (prob. 0.126)), while the level of satisfaction with the cooperation with small retailers is significantly higher in large wholesalers (t = 2.704 (prob. 0.027)).

Level of satisfaction of large wholesalers is in all cases, i.e. in their relationship with producers, with large retailers, but also with small retailers, significantly higher than the satisfaction that small and medium-sized wholesalers express for the same relationships. Generally, the lowest level of satisfaction is expressed by small and medium-sized wholesalers in terms of doing business with small and medium-sized retailers.

Researches have also confirmed the second hypothesis that "*wholesalers, in market conditions characterized by the presence of strong retail chains, express a higher level of satisfaction in doing business with large retailers than with small and medium-sized ones".* Average satisfaction index of wholesalers with their relationship with large retailers is on a significantly higher level than their average satisfaction index with the relationship with small and medium-sized retailers. This difference in the higher level of satisfaction of wholesalers business with large retailers is statistically significant ($t = 2.273$ (ver. 0.053)), so that we can accept the second hypothesis.

Also, if we observe individual elements of a business relationship, we will notice that satisfaction indices for almost all the elements are on a significantly higher level for relationship of a wholesaler with large retailers. The only element which has the satisfaction level higher for the relationship of a wholesaler with small and medium-sized retailers is the element, which is currently the most important element of business relationship and it is related to prices, rebates and additional payments. When it comes to this element, the great advantage is on the small and medium-sized retailers' side.

However, concrete responses of wholesalers and an insufficient percentage of those wholesalers which rather choose large retailers for their partners and consider that the development of large retail chains is in their interest, indicate a very difficult situation in wholesale sector. Those wholesalers which are successful importers or distributors of relatively strong brands, or those specialized in specific product groups, generally prefer cooperation with large, modern retailers and consider the development of retail chains to be in their interest. It seems that they are more willing to rise to the new challenges. Opposite to this, the wholesalers with a wide range of assortment fear more of the future that surrounds the development of the retail giants and growth of their power.

Although the cooperation with large retail chains brings a higher level of satisfaction to wholesalers, it is obvious that the insufficient percentage of wholesalers within the observed market of Serbia has transformed adequately and increased its efficiency and effectiveness in terms of adapting to the modern distribution demands which have been dictated by large retail chains. Judging by the obtained responses, it is evident that many of them identify opportunities and advantages of doing business with large retail chains, but only few of them do believe they can use it for their own development and long-term survival in marketing channels.

Conclusions

It is evident that increasing the power of retailers makes numerous effects on the position of institutional wholesalers in marketing channels and sets new demands for this type of intermediary. In that sense, many effects can be considered as negative ones. However, the development of large retailers also brings numerous positive impacts on position of institutional wholesalers in marketing channels. This is even confirmed by the Serbian market, the market, which is, one could say, in the initial phases of an intensive transformation.

At any rate, position of wholesalers in marketing channels is significantly changing with the development of the retail market. Today, the existence of wholesalers in marketing channels depends on the level of efficiency and effectiveness they manage to reach in carrying out their distributive activities. If they want to participate in distribution, they still must offer services of superior quality and also undergo transformations and changes in accordance with the demands and intensive changes imposed by their partners. Wholesalers will be able to avoid the danger of elimination by focusing on constant adjustments to the changes, and by flexibility and innovation.

Wholesalers must put more intensive efforts in finding and developing new areas of work which will represent their specific advantages. If they want to avoid complete isolation from the marketing channels, in conditions when strong retail chains endeavor to cut out wholesalers as intermediaries, they must, primarily, impose their role to the producers themselves, by providing them assistance and benefits in different segments. In accordance with Katsikeas's [13] statements, the power of wholesalers towards producers is exactly in their informational power, referent power, i.e. in a position to be a leader of numerous activities, their professional knowledge and other legitimate sources of power. In that sense, today, in their efforts to strengthen their role in the conditions of modern distribution, wholesalers turn out to be a sort of marketing and informational centers, highly important for the functioning of a channel as a whole and its results. The one which possesses relevant information on supply and demand can benefit from a direct inclusion in distributive process. And, this is the key characteristic of wholesale: to converse information into a profit.

Need for intermediation for producers, for large retailers, and certainly for small and medium-sized retailers, the one which would perform wholesale functions exists both on the market of Serbia and on the markets of the most developed countries. However, the question is if there are adequate partners on wholesalers` side which can guarantee mutual success. Wholesale intermediaries which are able to meet the challenges imposed by a modern, highly concentrated retail market, will have a higher level of satisfaction in these conditions and will use numerous advantages offered by cooperation with modern, more efficient and effective retail chains. Still, it is hard to enter into the circle of winners, and even harder to last there.

References

[1] Belaya, V., Hanf J. H. (2009) „The Two Sides of Power in Business-to-Business Relationships: Implications for Supply Chain Management“, *The Marketing Review*, Vol. 9, No. 4: 361-381

[2] Black, S. G. (2010) „Relationalism: A Vintage But Sound Concept in Distribution Channel Relationships“, *Atlantic Economic Journal,* No. 38: 245-246

[3] Dawson, J. (2004) „New Cultures, New Strategies, New Formats and New Relationships in European Retailing: Some Implications for Asia*", Journal of Global Marketing*, Vol. 18, No. 1-2: 73-97

[4] Dawson, J. A., Findlay, M. A., Sparks, L. (2008) „*The Retailing Reader"*, Routlage, USA and Canada

[5] Desai, P., Koenigsberg, O., Purohit, D. (2004) „Strategic Decentralization and Channel Coordination“, *Quantitative Marketing and Economics*, No. 2: 5-22

[6] Dobson, P., Waterson, M., Davies, S. (2003) „The Patterns and Implications of Increasing Concentration in European Food Retailing“, *Journal of Agricultural Economics*, Vol. 54, No. 1: 111-125

[7] Ertek, G., Griffin, M. P. (2002) „Supplier – and Buyer – Driven Channels in a Two-Stage Supply Chain“, *II E Transactions*, Vol. 34, No. 2: 691-700

[8] Flath, D., Nariu, T. (2008) „The Complexity of Wholesale Distribution Channels in Japan“, *The Japanese Economy*, Vol. 35, No. 2: 68-86

[9] Hadžić, M., Pavlović, P. (2012) „Serbian Small and Medium-sized Enterprises in Times of Crisis“, *Acta Polytechnica Hungarica,* Vol. 9, No. 3: 45-64

[10] Hingley, K. M. (2005) „Power Imbalance in UK Agri-Food Supply Channels: Learning to Live with the Supermarkets“, *Journal of Marketing Management,* No: 21: 63-88

[11] Human, G., Peter, N. (2010) „Relationship and Innovation Orientation in a Business-to.Business Context“, *South African Journal of Business Management,* No. 41 (4): 59-70

[12] Kadocsa, G., Francsovics, A. (2011) „Macro and Micro Economic Factors of Small Enterprise Competitiveness”, *Acta Polytechnica Hungarica*, Vol. 8, No. 1: 23-40

[13] Katsikeas, C., Goode, M., Katsikea, K. (2000) „Sources of Power in International Marketing Channels“, *Journal of Marketing Management*, No. 16: 185-202

[14] Kurtuluş, M., Toktay, L. B. (2011) „Category Captainship vs. Retailer Category Management under Limited Retail Shelf Space“, *Production and Operations Management*, Vol. 20, No. 1: 47-56

[15] Maruyama, M. (2004) „Japanese Distribution Channels“, *The Japanese Economy*, Vol. 32, No. 3: 27-48

[16] Nishimura, J. (2004) „The Linkage of Trades in Terms of Wholesale Business Formats in Japanese Distribution Systems“, *Journal of Global Marketing*, Vol. 18, No. ½: 167-186

[17] Rosenbloom, B. (2007) „The Wholesaler's Role in the Marketing Channel: Disintermediation vs. Reintermediation”, *Int. Rev. of Retail Distribution and Consumer Research*, Vol. 17, No. 4: 327-339

[18] Rossenbloom, B. (2013) „*Marketing Channels - A Management View",* South-Western, New York

[19] Samli, A. C. (2007) „American Wholesaler: Quo Vadis?“, *Journal of Marketing Channels*, Vol. 14, No. 3: 93-101

[20] Thomas, R. A., Wilkinson, J. T. (2011) “The De-Devolution of Marketing – Is America's Marketing Model Fighting Hard Enough to Keep Up?”, *Marketing Management*, spring: 19-25

[21] Torii, A., Nariu, T. (2004) „On the Length of Wholesale Marketing Channels in Japan“, *The Japanese Economy*, Vol. 32, No. 3: 5-26

HR Management at Subsidiaries of Multinational Companies in CEE in Light of Two Surveys of Empirical Research in 2008-2009 and 2012-2013

József Poór[1], Allen D. Engle, Sr[2], Ildikó Éva Kovács[3], Agnes Slavic[4], Geoffrey Wood[5], Katalin Szabó[6], Marzena Stor[7], Kinga Kerekes[8], Zsuzsa Karoliny[9], Ruth Alas[10], Krisztina Némethy[11]

[1]J. Selye University, Bratislavská cesta 3322, 945 01 Komárno, Slovakia, Szent István University, Páter Károly utca 1, 2100 Gödöllő, Hungary, poor.jozsef@gtk.szie.hu [2] Eastern Kentucky University, 521 Lancaster Ave, KY 40475 Richmond, US, allen.engle@eku.edu [3] Szent István University, Páter Károly utca 1, 2100 Gödöllő, Hungary, kovacs.ildiko.eva@gtk.szie.hu [4] University of Novi Sad, Segedinski put 9-11, 24000 Subotica, Serbia, slavica@ef.uns.ac.rs [5] Warwick Business School The University of Warwick, Coventry, CV4 7AL Warwick, UK, geoffrey.wood@wbs.ac.uk [6] Szent István University, Páter Károly utca 1, 2100 Gödöllő, Hungary, szabo.katalin@gtk.szie.hu [7] Wroclaw University of Economics, Komandorska 118/120, 53-345 Wroclaw, Poland, marzena.stor@wp.pl [8]Babeş-Bolyai University, Strada Mihail Kogălniceanu 1, 400084 Cluj-Napoca, Romania [9] University of Pécs, Szántó Kovács János u. 1/B, H-7633 Pécs, Hungary, karoliny@ktk.pte.hu [10] Estonian Business School, Lauteri 3, 10114 Tallinn, Estonia, ruth.alas@ebs.ee [11] Óbuda University, Bécsi út 96/b, H-1034 Budapest, Hungary, nemethy.krisztina@rh.uni-obuda.hu

Abstract: There is only limited literature on Human Resource Management (HRM) activities in Multinational Companies (hereafter MNC) in the Central and Eastern European region. The research presented below is part of an ongoing research project developed by the Central and Eastern European International Research Team (CEEIRT), a consortium of researchers from universities across the Central and Eastern European (CEE) region with the aim of describing the ongoing patterns of HR practices, policies, practices and pursued by regional MNCs. Statisticians and descriptive results are present from surveys carried out in 2008-2009 and again in 2012-2013.

Keywords: HR practices; multinational companies; Central-Eastern Europe

1 Introduction

According to UNCTAD's 2014 Report, the volume of FDI has fallen globally since the 1990s. A record value of US $1,492 bn was recorded in 2000, whereas the total global FDI volume dropped to US $735 bn only two years later. This same report identifies the declining number of cross-border acquisitions as being among the most important trends behind this shrinking level of FDI. According to the UNCTAD 2014 Report, the volume of FDI increased again between 2004 and 2008. Another record value of US $1.401 bn was reached in 2006, but as the result of the global financial crisis of 2008 this volume has been drastically reduced. This decline is in contrast to trends reflected in other macroeconomic indicators such as GDP, foreign trade balance or employment rates, which once again started to rise globally (UNCTAD, 2014). According to the UNCTAD forecast, FDI 2014 will remain at close to the 2012 level. Issues such as the structural weaknesses of the financial system, political instability or an increasingly unfavourable macroeconomic environment could result in a further decline in investor confidence and as a consequence may trigger further declines in FDI. Whereas previously the most developed countries had a large share of the regional FDI, recently the trend has changed and the emerging and developing economies have a 50% share of the total FDI (UNCTAD 2010). FDI inflow into the CEE region between 1990 and 2012 amounts to 1.040 bn. $ (830 bn. €) (UNCTAD 2014). In this region, it will be seen that for most of the 2000s, FDI increased, but fell back after the 2008 crisis. (Figure 1)

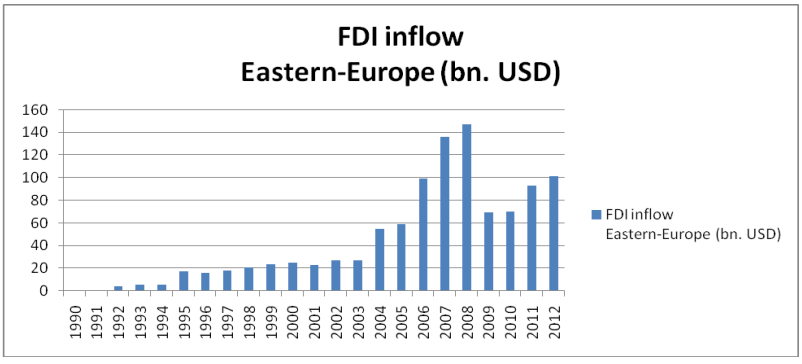


Figure 1

FDI inflow Eastern-Europe (1990-2012)

Source: UNCTAD (2014): World Investment Report, 2013. United Nations, Geneva

Previous research about the region has suggested that the inflow of FDI could lead to increasing differences in employment and wage levels (Golejewska, 2002; Jenkins, 2009 and Falusné, 2000), and this could exacerbate social and economic polarization. Antalóczy and Sass (2005) found significant correlations between employment, average wages and the size of the foreign capital, but that these values were also associated with significant levels of economic inequality. Bohle

et al. (2007) found that inward FDI is associated with stronger levels of high-tech exports, which contributes to a tendency for focal firms to upgrading their capabilities, so as to be more valuable, rare, difficult to imitate and non-substitutable. Allen Aldred (2011) notes, that the subsidiaries capabilities increase their autonomy in relation to the headquarters in their attempts to gain an enlarged role. These complex macro relationships are a topic of ongoing interest for academics and policy-makers alike.

1.1 The Research Model

The project, which had established a basis for a longer term and its most recent part of the project (2011-2013), is a combination of descriptive and deep analytical study based on the following research model. (Figure 2)

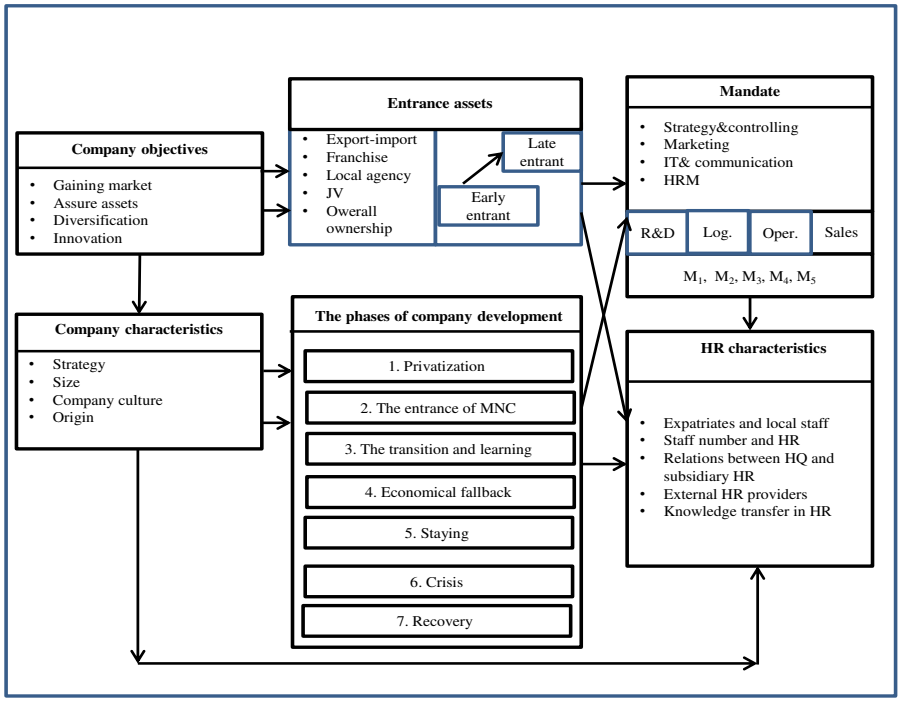


Figure 2

Research Model: HR and its Context in Multinational Subsidiaries

Source: Primary research by the authors

This model combines several influencing and contextual factors evolved from a variety of disciplines (e.g. international management, international human resource management and HRM): Objectives of the firms: MNCs enter to foreign markets for traditional reasons (market acquisition, securing resources, and diversification), but lately they are also seeking better economies of scale and a

more rational allocation of expenditures, via shifting processes and activities to lower cost nations (Dowling et al., 2013). These realignment activities necessitate coordinating and refining HR practices, such as altering methods for expatriate compensation, reviewing training policies, and realigning the transfer of managers across borders (Fisher et al., 1993, Francesco and Gold, 1997; Venaik et al., 2005 and Brewster, 2006). Earlier research suggests that formal labor regulation does not have a significant relationship to FDI level (see Wood et al. 2014).This means that although firms may enter countries because of cheaper labor, they feel either that they can circumnavigate rules or that they can compensate for variations in regulation through a pattern of HR policies.

Phases of company development: Major stages of company development and HR implications in Central and Eastern Europe during 1988-2013 can be classified into the following seven categories (Poór et al., 2010):

- Stage One and Two – Privatization/ firm establishment: This period is characterized by the transition from centralized state control and state socialism to private ownership of the means of production in a wide range of industries (Koenings, 2001 and Pavlinek, 2002)
- Stage Three and Four – Restructuring and development: Later the new owners shift interest more toward economic rationalization and the divestiture of unprofitable units and functions. The key role of HR was to contribute to institutional changes and transformation (Claessens and Djankov, 2002).
- Stage Five – Consolidation and renewal: The majority of companies in the competitive sector of CEE economies has finished the structure changing after privatization.
- Stage Six and Seven - Crisis and Recovery: This period encompasses the global economic crisis that erupted in 2008 and the slow recovery since, of particular interest is (Kotler and Caslione, 2011) the impact of this stage on new or modified HR practices (Balázs and Veress, 2009) and executive expectations for the future (Fodor et al., 2011).

Mandate: During the analysis, following Delany (1998) and White-Poynter (1984), we classified the participants into five groups based on how much of the value chain is covered by the range of activities of the local subsidiary.

1) "Mandate 1": "Mandate 1": This is a business in which "front end" value chain activities – some subassemblies, manufacturing, distribution and retailing - are accomplished for the local trading area. The business is a small-scale replica of the parent.

2) "Mandate 2": This is a business producing a designated set of component parts for a multi-country or the global market. Operational activities are limited to packaging, bulk breaking, some final processing and warehousing, distributing.

3) "Mandate 3": This is a business that does not have control over the entire value chain of a business unit but has activities in a number of stages of the value chain. This might be the preparation of manufacturing activities or a regional logistics brief (responsibility).

4) "Mandate 4": This is a business that develops and markets a certain product chain for global markets. Products, markets and basic technologies are similar to the parent company, but the exchange of information between the subsidiary and the parent are rare.

5) "Mandate 5": This is a business that has the freedom and resources to develop lines of business for either local, multi-country or the global market. The subsidiary is allowed unconstrained access to global markets and freedom to pursue new business opportunities.

Nature and time of market entry: Data relevant for international comparison show that the majority of international companies chose to solve the problem of gaining majority control by acquisitions or implementing a greenfield developments in the region (UNCTAD 2012). The great migration of MNCs into Hungary took place in the 1990s – in contrast with the neighbouring Slovakia where this process largely occurred between 2002 and 2007. Many of the large multinational companies now present in Hungary have been operating here continuously for nearly two decades. However, the actors of some industries (e.g. automotive suppliers) have been known to move operations in and out of nations within the CEE region and outside of the region very quickly. If the situation is not favourable, these companies walk away promptly. However, the decision that the primary actors in these companies make to stay or leave also depends largely on whether their main buyers stay in-country or depart. In support of the role of "cheap manufacturing and service provider" the number of Hungarian subsidiaries capable of operating with shorter delivery times increased during the recent financial crisis.

Origin: The origin here means that the institutional arrangements and purposes of a MNE are influenced by associations and beliefs related to the firm's country of origin (Kotler-Keller, 2006: 893). According to Yan (2003), one of the key researchers in this topic area (this group would include Hofstede, 1991; Rosenzweig and Nohira, 1994 and Jackson and Artola, 1997) management practices and structures are heavily influenced by national cultures. In contrast, others (Pauly and Reich, 1997) believe that the path and form of technology development is fundamentally impacted by these national differences. Companies of different origins to choose various paths to internationalization, and it should also be stressed that companies of American, Japanese or European origin have built up their present organization structure and hierarchy in a variety of ways (Dowling et al., 2013). Pudelko-Hatzing (2007) stress that human resources management practices of the subsidiary level of an MNC resemble the practices in the home country more so than practices of the local firms (country of origin

effect). However, these authors emphasize too, that national cultures and institutional environment can limit the transfer of HRM practices. Besides that, the subsidiary level HRM practice may follow the practice of the country that sets the HRM standards for what is perceived as a global best practice, too (the so-called dominance effect).

1.2. Research Participants - Samples

The data collection was made through both face-to-face interviews and online completion of the standardized questionnaire, which was based on the model shown in Figure 2. The number of valid responses by participating countries is shown in Table 1. Participants were solicited from MNE firms known to the local, national researchers, based on previous contacts, ongoing consulting activities and area or national listings. (Table 1)

Table 1

Number of valid responses from participating countries in the two CEEIRT research phases (2008-2009 and 2011-2013)

Country	Number of valid responses	
	2008-2009	2011-2013
Estonia	45	0
Croatia	11	0
Poland	88	53
Hungary	75	118
Romania	17	34
Serbia	20	19
Slovakia	23	30
Total	279	254

Employment Scope: The first survey (2008-2009) covered seven countries and included results from 279 subsidiaries. These firms employed almost 300.000 people. We can state that the companies in this survey period are split equally between large (51.9%) and small enterprises (48.1%) based on the number of their employees. In our second survey (2012-2013), we analyzed the data of five countries. A total of 254 foreign owned subsidiaries took part in the research, employing total 245 thousand people. Some 48% of the respondents had more than 250 employees.

Mandate: We also asked how wide a portion of the value chain the local subsidiary has control of, in other words, what is the „mandate" of the local organization (Delany 1998 and White-Poynter 1984). Some 60% participating subsidiaries have a typical subsidiary mandate, such as purchasing, production/operation and sales/distribution in 2008-2009 survey. This figure is a little bit higher in the 2012-2013 survey period.

MNE Origin: The subsidiaries participating in the total sample of the research (2008-2009) came from 33 different countries. More than 74% of them came from four countries: Germany and Austria (47.8%), and the USA and Canada (27%). The remaining subsidiaries came from Western and Southern European EU countries (11.3%) as well as from outside of Europe (13.9%). The data on the origin of the 254 subsidiaries participating in the survey in 2012-2013 show that they were from 34 countries, 67% of them have their parent company in Germany (23.8%), USA (18.9%), France (9.4%), Austria (6.7%) or Switzerland 4.3%). Among the respondents 6.3% is headquartered in emerging or transitional countries. In both surveys MNC companies originates mainly from Europe, especially from the European Union (EU).

Time of market entry: The foreign owners of more than half (52.7%) of the subsidiaries examined in survey 2008-2009 were established before 1995 in Hungary. Almost one quarter of the companies arrived in Hungary between 1996 and 2000 (24.3%) and the remaining firms (23.0%) set up operations in the new millennium. In the other samples most of the respondents established subsidiaries after 2000, especially in Romania and in Serbia where this happened for 75.0% and 81.8% of respondents respectively, In the total sample of the 2012-2013 survey nearly a third of the subsidiaries were established before 1990 (31.7%) and almost a quarter (22.7%) between 1990 and 2000. Almost half (45.7%) of subsidiaries were established after 2000. Patterns of regional arrival and resettlement by MNEs are at once complex and dynamic.

Nature of market entry: About 44% of the foreign owners of the companies participating in the survey 2008-2009 came to Hungary creating greenfield investments and around 56% of them obtained majority control from extant Hungarian companies during privatization and the following period of acquisition. In the Croatian and Serbian samples (81.8% and 75% respectively) the primary mode of entry was through acquisition as was the case, to a lesser extent, in Estonia (63.6%). While in the Polish, Romanian and Slovakian samples more multinational companies established subsidiaries via greenfield investments. In the 2012-2013 survey results 51% of the respondents were established via a green-field project and 49% were established through acquisition. Some 45.7% of responding firms carry out traditional production, 16.5% are in trading and 35.4% provides services.

Strategic orientation: A significant number of the respondents (39.7%) in the total sample of 2008-2009 indicated that they were seeking growth development during the period examined. Over 35% of the companies surveyed characterized their strategy as stability. The fact that 23.2% – nearly a quarter – of the respondents chose the redundancies option indicates a slow recovery from the crisis. By the time of the 2012-2013 survey the main strategic orientation is growth/market expansion for the majority of the respondents (59.1%), while the priority for 36.9% of respondents is to maintain stability. Only 3.2% of the respondents indicated a reduction of capacities. This is a significant change when compared to

the 2008-2009 survey, when 23.2% of respondents considered reduction in force and rationalization of capacity. In comparison with survey in 2008-2009, few of MNC subsidiaries in 2012-2013 were struggling with issues of cut-back or downsizing.

1.3 Key Indicators of HR Activity

In this section we give an overview of the following HR characteristics: (1) number and workload of the HR staff, and (2) the main indicators representing the importance, results, and efficiency characteristics of the HR activity (e.g. labor cost – total cost ratio, a relative weight of the training budget).

1.3.1. Number, Workload and HR Staff

HR effectiveness ratio: In the total sample of 2008-2009 the average number of employees per HR professional was under 100 employees for 73.6% of the organizations and the average number of employees per HR professional of 200 or less characterized 96.6% of surveyed organizations. The average headcount serviced by one HR employee is 64 across the responding 279 companies in the 2008-2009 sample in contrast to 69 across the responding 254 companies in the 2012-2013 sample. This finding is consistent with other observations by the authors that the average headcount/HR staff ratio is declining. Some sources explain this with the increasing importance of the HR function (Balázs–Veress, 2009 and Pudlowski, 2009). The average number of employees served by an HR professional is far below the traditional firm ratio. It is well known from management theory and practical experience that it is not reasonable to maintain a separate HR apparatus under a certain number of employees (e.g. 80-100 persons) within an organization. However, the actual ratio also depends on the industry and the composition of the workforce. (Table 2)

Table 2

Number and workload of HR staff (HR effectiveness ratio)

Years	Total staff	HR staff	Staff/HR staff
2009-2010	292 697	4 605	64
2012-2013	247.000	3.550	69

The average size of the HR department: The average number of HR professionals was between one and five position holders in the total sample 2008-2009, and in sample 2012-2013 too. In the Hungarian and Polish samples the average number of HR professionals ranged between one and ten incumbents. Respondents in five out of the seven countries indicated no HR staff; ranging from some 5% (in Hungary) up to 10% (in Croatia) of the subsidiaries. This includes 45.1% of respondents having only 1-4 people in HR. Some 18.9% of respondents reported 5-10 HR staff, and 14.4%of respondents reported 10-20 HR people, while 14.3 % of respondents operate an HR department that is larger than 20 people. In some

cases there is no formally appointed/identified HR leader in the organization. In these cases the management or HR issues are performed by the head of the organization or the CFO. The average number of HR professionals was between one and five HR position holder in each sample. This figure is in line with regional data of Cranet global HR survey (Cranet, 2011).

1.3.2. The Main Indicators

Training and Development spending: Literature describing HRM operations often considers the relative weight of the training budget (compared to the entire annual labor cost) as an important indicator of modern and effective HR activity. It is also seen as a sign of the relative commitment of a firm to its people (although, of course, it is possible that a high expenditure in basic hiring training may simply be a product of high staff turnover rates).The average annual training and development spending – as a percentage of the annual labor costs – is less than 1% for 26.0% of the respondents, between 1 and 3% for 42.4% of respondents and 31.6% of respondents spend more than 3% of the labor costs on training and development programs in the 2012-2013 survey. The global average of this indicator calculated using the formerly mentioned Cranet (2006 and 2011) international comparative HR database was 3.36%, the Eastern European index was 3.15% and the Hungarian statistic is 3.54%. (Table 3)

Table 3
Training cost ratio %

	2009-2010	2012-2013	Cranet global	Cranet CEE
% of average training budget	3,1	3,4	3,54	3,15

Source: Primary research by the authors and Cranet (2006 and 2011). International Executive Report 2005, Cranet Survey on Comparative Human Resource Management.Cranfield:Cranet-Cranfield University

Relationship of the local and central HR organization: There are several models to describe how the corporate center and local HR units share accountabilities. Taylor et al. (1966) describe the relationship between the subsidiaries and the parent company with the following three basic systems of relations:

In the exportive system of relations, HR systems developed in the parent company are adopted without changes.

In the adaptive system of relations, local subsidiaries adapt the HR systems adopted from the parent company according to their local needs. In the integrative system of relations, all good and applicable solutions are attempted to be spread and implemented in all units of the company regardless of the origin of the HR system.

Lawler (2006) concluded from his research conducted among American subsidiaries operating in Asia and Europe that the most dominant deciding factor in the adoption and adaptation of HR systems is the size of local companies. The

logic is understandable: which solution should be applied in a certain case? Lawler states that the system to be implemented depends on the sum of the impacts of internal and external factors that form and influence the organization. In certain cases the national culture of the host country and the legal, regulatory environment are considered influencing factors.

We found several different function sharing practices among the companies examined in the 2008-2009 survey:

The typical solution that was implemented by nearly half of the respondents in the total sample was that the HR department of the company's headquarters lays down general guidelines and provides a standard framework for the work of HR departments of the subsidiaries and requires information and reporting from them. While 20% of the companies' headquarters performed the auditor's role.

In addition, in the case of almost 36% of the companies the headquarters was responsible for providing resources and advice when requested.

Around 15% of the respondents stated that the headquarters provided the subsidiary with a detailed HR model, policies, procedures, and rules.

On the other hand, about 15% of the HR departments of the responding subsidiaries reported getting hands-offs treatment, almost complete freedom from the headquarters resulting in what they considered a decentralized HR activity.

Finally, in almost 6% of companies the headquarters provided central control.

A majority of the respondents in the 2012-2013 survey (47.5%) indicated that the central HR function issues general guidelines and frameworks and the subsidiary HR staff implement their local practices within these frameworks. Some 32.6% of the respondents have detailed HR models and policies issued by the headquarters, and 19.8% of respondents reported that the headquarters provided them with full authority, performs some audit functions, and expects regular reports. This survey finding underpins the findings in other research regarding the division of HR accountabilities locally between line management and HR specialists. According to Cranet (2006, 2011) some HR decisions are typically the prevue of line management, whereas other decision areas (namely HRIS and labor management) are more influenced by the local HR staff.

- International assignees: As presented in Perlmutter's (1969) seminal paper, multinational companies following a typology of four personnel strategies have different priorities in their selection and recruitment policies. The company can follow an ethnocentric, polycentric, regiocentric or geocentric selection mechanism. In the ethnocentric orientation, key positions of the local company are held by professionals from the parent company. In polycentric companies, local key positions are held by locals but their promotion to higher positions is very limited. In companies following the regiocentric selection mechanism, locals can hold key positions not only in

the subsidiary but also in the center coordinating the management of the region. In companies following the geocentric selection mechanism, locals can obtain positions even in the top management of the company.

- Usually two types of long-term emissaries are distinguished – employees arriving from abroad from the parent company or from a third country (who are also confusingly often called expatriates) and a second set of employees from the local CEE regional subsidiary appointed for a long-term deputation abroad at the parent company or subsidiaries operating in other countries. In the majority of the results reported from 2008-2009 most foreign expatriates are managers. Exceptions are found in the Estonian sample, where foreign expatriates are more likely to be non-managers (58.6%). In the Romanian and Slovakian samples the proportions of foreign expatriates in managerial and non-managerial roles are almost equally divided. It is important to note that companies appear to be sending an increasing number of employees abroad for a short term "international assignments" (Dowling et al., 2013), for a variety of limited projects, and therefore did not meet the traditional definition of a long term assignment associated with the traditional definition of an expatriate.
- Among the total 247,000 employees working at the 254 respondents, 1000 expatriates are listed in employment estimates, merely 0.4% of the employee reported in the survey of 2012-2013.These assignments are typically in management positions. At 60% of the respondents no expatriates at all were reported at the time of the research. The number of local employees on international assignment is 800, and they were spread across 27% of the responding subsidiaries. Only 33% of these assignees worked in a managerial position.
- Knowledge Management in HR: In terms of developing personal competencies amongst the HR staff, the most important channels or methods of development is reported to be local training and development, and learning assignments at the mother company's headquarters. This result is consistent for both samples. It is interesting to note that responses indicate the extent or degree of knowledge transfer follows a similar pattern a) inside the subsidiary HR department, b) between the local HR department and other local business units, and c) between the subsidiary HR group and the headquarters HR unit.

1.4 Analysis

With the help of correlation analyses and cross-tab analyses, this paper presents the factors which have a statistically significant role in influencing the characteristics of the HR variables presented above and the nature of the relationships in cases found to be significant.

The correlations between seven elements of the firm and contextual elements in our model (mandate, size, origin, strategic orientation, nature, time of entry into the market and firm characteristics) and five HR characteristics (number of HR staff in the subsidiary, number of expatriates, the role of the HR department of the parent, the use of external HR service providers and knowledge transfer in the field of HR) were tested with the 'Chi-squared' test. The strength of the correlation was described by means of the Cramer V index. (If the strength of the correlation established between two variables under investigation is around 0, we can speak of 'independence'; the relationship is 'weak' if below 0.3, 'medium' between 0.3 and 0.69, and 'close' at/over 0.7. (Table 4, 5 and 6).

Table 4

Detailed analysis of hypothesis explanations of two consecutive researches

Fact.inflHRM	Hypothesis explanations of two consecutive researches	
	2008-2009	2012-2013
H1 mandate[1]	According to the Chi-squared test the mandate affects the number of HR employees and expatriates (expats) but does not affect the relationship between the local HR and HQ HR. The Chi-squared test shows a significant connection between the mandate and the employment of external HR service providers in the cases of recruitment, selection and training/development. According to the cross table analysis companies which are also engaged in research and development apply external service in the highest proportion – not counting selection. Independently from the mandate the use of external service is the most typical for training/development (75%) and recruitment (62%) Between the mandate and the knowledge flow from the parent company to the subsidiary there is just a 10% significance level connection. (Chi-squared test, ANOVA.) While in the case of the local training and education both tests show a significant connection (on 5% level) with the mandate variable. Overall we accept the hypothesis.	According to the Chi-squared test the mandate partly affects the number of HR employees (in the case of production and sales) but it does not affect the number of expatriates and the relationship between the local HR and the HQ HR. Between the employment of external service providers and the different mandate types the Chi-squared test shows significant connections just in a few times. According to the cross table analysis (similarly to the previous research) research and development companies apply external service in the highest proportion and the growth in that happens there in the most cases – not counting the selection. Generally, the use of external service is now the most typical for the field of the training/development of the staff (65%), recruitment (65%) and Occupational Safety and Health (OSH)[4] (66%).No statistically significant connection can be shown between the mandate and the knowledge flow from the parent company to the subsidiary. However, in one case (production/service) there is a significant

1 In the two researches the variable affects of the mandate can be examined differently. While in the research of 2008-2009 the different value chain role combinations were different attributes of a variable (These were Sales/marketing, Sales/marketing/production/operation, Sales/marketing /production/ operation/research/development), and from these one could be marked, in the research of 2012-2013 the companies could mark the applying mandates one by one (Sales, marketing, production, etc.) This means that each mandate category was a separate variable. So the connections with other variables (e. g. number of HR-employees, expats) inside a single category could be examined just in reference to the given yes/no questions.

4 This field was examined only in the research of 2012-2013.

		connection with the local training and education. (According to the Chi-squared test, ANOVA).We party accept the hypothesis.
H2 origin	The effect of the origin cannot be examined by Chi-squared test in the most cases. (The conditions of the test are not met). According to the analysis of contingency tables at Eastern European and Japanese/Korean companies the number of expats is higher than in the case of the other managing cultures. Eastern European and Japanese/South Korean companies reported on a much higher growth (double to fivefold) in the case of employment of external service providers in several fields. (The biggest difference appears in the cases of recruiting, selection and HR planning. The significance of the knowledge flow from the parent company to the subsidiary is in the Eastern European companies the greatest while it is the smallest in the Japanese/Korean ones. The other origin categories behave very similarly so they are put together and ANOVA shows on 10% level a significant connection between the two variables. For Eastern Europeans there is a much greater importance of the local training and education than in the case of the other managing cultures, however there is no significant connection between the two variables. (According to ANOVA) On the relationship between the local HR and HQ HR[5] the Chi-squared test shows a significant connection in three cases – not counting centralization. Decentralization is much more typical for companies of Eastern European origin (38%) and of German origin (40%) while for Anglo-Saxon companies the provision of general guidelines (75%) and detailed HR model (49%).	The effect of the origin cannot be examined by Chi-squared test (because in no cases of the variables are the conditions of the test met.) All in all, (compared with the previous research) independently from the category of origin the proportion of those firms which do not have a HR officer is considerably higher (5% vs 54%) and much higher is the number of those where there are expats (57% vs 93%). According to the cross table analysis the firms of Eastern European origin show a significantly different picture in the case of certain examined variables than the other firms from different managing cultures. Among these is the proportion of those firms the highest (69%) where there are no HR-employees and they are characterized by having no expats or just a low number of them. In reference to the relationship between the local HR and the HQ HR[6] decentralization is typical for them (47%). There was not a single firm for which centralization was typical. In the use of external HR providers no tendency can be experienced in the case of origin. However, the Eastern European firms use external providers in much less proportion in the most fields than in the previous research. In the case of the local HR training no significant connection can be shown between the variables either. (ANOVA) Between the origin and the HR knowledge transfer variables there is a weak but significant relation which is the least typical for Asian and Latin-South European firms and the most for Anglo-Saxon ones.We partly accept the hypothesis (because the Eastern European firms behave differently from the other ones).
H3 date of the establish	Chi-squared test shows a connection on a 10% significance level between the date of the establishment and the number of HR employees. A high number of HR employees (more than 5) is the most characteristic for the firms which were	The criteria of the Chi-squared test are not met in the most cases. On the basis of the cross table analysis the firms established before 2010 show a very similar picture in regard to the most variables while the very new firms

5 The relationship between the local and central HRs could be examined in the two researches differently. Look at footnote number 2.

6 The relationship between the local and central HRs could be examined in the two researches differently. Look at footnote number 2.

ment	established before 1995. In the cases of expats and the connection between local HR and the HQ HR[7] there is no significant connection, however, according the cross table analysis the firms established after 1995 are characterized by a higher number of expats while those which were established after 2000 are characterized by decentralization and less by provision of general guidelines and a detailed HR model. In the use of external HR service providers the Chi-squared test cannot carried out in all cases. On a 10% significance level there are connections in two cases (recruitment and labour relations) but according to the cross table no definite tendencies can be shown in the connection of the variables. In the case of the knowledge transfer from the parent company to the subsidiary there is no significant connection, but its importance is much higher in reference to the firms which were established after 2000. In the local training there is no significant difference between the firms in regard to the date of the establishment either. We partly accept the hypothesis.	(established after 2010) differ from those from many aspects. They employ less HR-officers, they do not have expats and in regard to the local HR and the HQ HR[8] they are characterized in a higher proportion by centralization while it is a lot more typical for them not to use external service providers. (We have to take into consideration that it is about just 7 firms). ANOVA does not show a significant connection between the knowledge flow from the parent company to the subsidiary and the local HR trainings. However, the subsidiaries established after 2010 are more characterized by the knowledge transfer than the others. (According to the averages.)Overall we reject the hypothesis.
H4 the way of establish ment	According to the Chi-squared test there is a significant connection between the way of establishment and the number of HR officers. Newly established firms are characterized by a lower number (1 to 5 men) of HR employees (54%) while in the case of firms gained by acquisition this number is mostly higher than 5. (59%) In the cases of the other variables (expat number, relationship between the local HR and the HQ HR[9], use of external service providers and knowledge flow from the parent company to the subsidiary) there is no significant difference between the two ways of establishment. The Chi-squared test does not show any significant connection in any cases. With the help of the cross table analysis it can be observed that the newly established firms are a slightly more characterized by the growth of the use of external service providers, while for the firms gained by acquisition is the decrease of the use of external service providers more	According to the Chi-squared test there is a weak but significant connection between the way of establishment and the number of HR officers. Newly established firms are characterized by a lower number (1 to 5 men) of HR employees (52%) while in the case of firms gained by acquisition this number is mostly higher than 5.(47%) In the expat number and the relation of the local HR and the HQ HR[10] the firms do not show significant differences in regard to the way of establishment. Between the way of establishment and the use of external service providers there is no significant connection in any fields either, however, with the help of the cross table analysis it can be observed that the newly established firms are more characterized by the growth of the use of external service providers – not counting occupational safety – in contrast with the firms gained by acquisition which are in most cases characterized by the

7 The relationship between the local and central HRs could be examined in the two researches differently.

8 The relationship between the local and central HRs could be examined in the two researches differently.

9 The relationship between the local and central HRs could be examined in the two researches differently.

10 The relationship between the local and central HRs could be examined in the two researches differently. Look at footnote number 2.

	typical. Independently from the way of establishment nearly half of the firms are mostly or totally characterized by the knowledge transfer from the HQ HR to the local HR. (48% vs 51%) Overall we reject the hypothesis.	decrease of the use of external service providers. Independently from the way of establishment nearly half of the firms are mostly or totally characterized by the knowledge transfer from the HQ HR to the local HR. (50% vs 48) Overall we reject the hypothesis.
H5: strategy	According to the Chi-squared test there is no significant connection between the number of expats and the relation of the local HR and the HQ HR.[11] The cross table analysis shows that in the case of the latter decentralization is slightly more typical for the firms which decide for growth while those firms which decide for workforce reduction (WFR) are more characterized by a detailed HR model and centralization. According to the Chi-squared test in the field of the use of external HR service providers there is a connection with strategy in many cases (HR planning, recruitment, compensation and labour relations). According to the cross table analysis in regard with WFR strategy the response proportion of the growth in the use of external service providers is significantly less in the most fields, while the decrease response is significantly higher than in the others. Labour relations are an exception where the growth and overall the use of external service providers are significantly higher than in the others. According to the Chi-squared test there is a significant connection between the knowledge transfer from the HQ HR to the local HR and the strategy. (In the case of the growth strategy it is more significant than in the others). While there is no significant relation with the role of the local training and development. We partly accept the hypothesis.	In examination of relations between the variables we first of all relied on the cross table examinations, because the data do not allow the statistical hypothesis test. (There are only eight firms with downsizing strategy and in comparison with the previous research their number decreased significantly.) It can be ascertained that those subsidiaries which follow downsizing and regression there are not HR officers in any cases and the proportion of expats is considerably higher. For those firms who choose growth in the relation between the local HR and the HQ HR [12] decentralization slightly more typical while those which choose downsizing are more characterized by general guidelines and centralization. Those firms which choose growth strategy use external providers almost in every field in a higher proportion than the others (not counting labour relations). Compared with other strategies the proportion of those who report on growth in the use of external providers is significant. (2-3 fold) Between the strategy and the knowledge transfer from the HQ HR to the local HR there is no significant connection. (ANOVA). With the help of the cross table analysis it can be observed that the knowledge transfer has slightly more significance in the case of the growth strategy, while it is the local training in the case of downsizing. We partly accept the hypothesis.

11 The relationship between the local and central HRs could be examined in the two researches differently.

12 The relationship between the local and central HRs could be examined in the two researches differently. Look at footnote number 2.

size (number of employees)	According to the Chi-squared test there is a significant, middle-strength connection between the size and the number of HR-staff. The bigger the firm is the more HR officers it has. In the case of the number of expats there is a weak but significant relation with a similar tendency. Between the relation of the HQ HR and the local HR[13] and the use of external service providers there is a significant connection in two cases (recruitment, training and development) but there is no definite tendency according the cross table analysis. In the case of the knowledge transfer from the HQ HR to the local HR there is a connection on 10% significance level (the bigger the firm is, the more important it is regarded). According to the cross table analysis in the case of the local training and development there is no significant relation. It is the most significant in medium-sized firms (250-1000 men). We partly accept the hypothesis.	According to the Chi-squared test there is a significant relation between the size of the firm (on the basis of the number of employees) and the number of HR officers (middle-strength connection) and the number of expats (weak connection). The bigger the firm is, the more HR officers and expats it has. With the relation between the local HR and HQ HR[14] there is just a connection on 10% significance level. On the basis of the cross table analysis the big firms are more characterized by decentralization, while the small ones and middle-sized ones by centralization. There is a significant connection between the knowledge transfer from the parent company to the subsidiaries and in the case of the local training as well. (Chi-squared test, ANOVA). These are more typical for the firms with more than 1000 employees. Overall we accept the hypothesis.
H7: Size of the HR department influenced by the characteristics of the firms	According to the Chi-squared test there is a significant relation between the size of the HR department and the size of the firm (middle-strength relation) and between the number of expats (weak relation). The bigger the size of the firm is, and the higher the number of the expats is, the more HR officers it has. In the case of inpatriates the Chi-squared test cannot be applied, however, on the basis of the cross table analysis there is a definite tendency: a higher number of inpatriates goes together with a higher number of HR officers. Regarding the annual training and development budget there is no significant connection, but the cross table analysis shows that labour costs lower than 1% is more typical for those firms which do not have HR officers or there are just 1-5 men, than for those where the number of HR officer is higher. In the case of the personnel cost the Chi-squared test cannot be carried out and no definite tendency can be observed with the help of the cross table analysis.We partly accept the hypothesis.	As we could observe it in the previous point, the Chi-squared test shows a middle-strength connection between the size of the HR department and the size of the company. In the case of the other variables it cannot be applied, however, the cross table analysis shows a definite tendency: the higher the number of the HR officers is, the more characteristic is for the company a high number of inpatriates/expatriates. In the case of the annual training and development budget, those firms which have more than 20 HR officers are out of line with the others, the training budget is above 1% in much higher proportion, while the labour costs are below 10%, unlike in the others. We partly accept the hypothesis.

13 The relationship between the local and central HRs could be examined in the two researches differently.

14 The relationship between the local and central HRs could be examined in the two researches differently. Look at footnote number 2.

Table 5

Analysis of HR influencing factors (2008-2009)

No	Factors	1. Mandate		2. Origin		3. Fondation time		4. Fondation way		5. Strategic orientation		6. subsidiary (No. of employees)	
		p (Pearson Khi^2 v.	Camer V v. Éta	p (Pearson Khi^2 v.	Camer V v. Éta	p (Pearson Khi^2 v.	Camer V v. Éta	p (Pearson Khi^2 v.	Camer V v. Éta	p (Pearson Khi^2	Camer V v. Éta	p (Pearso n Khi^2	Camer V v. Éta
1	Number of HR staff employed by subsidiaries	0,047**				0,077*	C=0,152	0,017**	C=0,204			0,000***	C=0,52(
2	Number of expats	0,003***										0,000***	C=0,27
3	Role of HR department at HQ												
	Local HR is free to make decisions			0,093*								0,007*	C=0,19
	Supply of resources and counselling to local HR if necessary	0,008***		0,033**									
	General guidelines and framework to local HR			0,016**								0,008*	C=0,19(
	Detailed HR model. policies. and control to local HR			0,02**									
4	4. Involvement of external HR service providers												
	Planning of human resources									0,028**			
	Recruitment	0,003***				0,072*				0,012**		0,052*	
	Selection	0,003***											
	Training-development	0,037**										0,004*	
	Compensation-Remuneration									0,001***			
	Labor relations					0,071*	C=0,150			0,007***			
	HRIS												
5	Knowledge transfer from HQ to local HR	0,068** 0,071**	É=0,140	0,059*	É=0,163					0,036**		0,059*	
6	Role of local training in the development of HR competences	0,025** 0,023**	É=0,1										
		***=1% (significance level)						**=5%		*=10%			

Table 6

Analysis of HR influencing factors (2011-2013)

No	Factors	1. Mandate		2. Origin		3. Fondation time		4. Fondation way		5. Strategic orientation		6. Size of subsidiary (No. of employees)	
		p (Pearson Khi^2 v. ANOVA)	Camer V v. Éta	p (Pearson Khi^2 v. ANOVA)	Camer V v. Éta	p (Pearson Khi^2 v. ANOVA)	Camer V v. Éta	p (Pearson Khi^2 v. ANOVA)	Camer V v. Éta	p (Pearson Khi^2 v.	Camer V v. Éta	p (Pearson Khi^2 v.	Camer V v. Éta
1	Number of HR staff employed by subsidiaries							0,048**	C=0,181			0,000***	C=0,578
2	Number of expats											0,000***	C=0,288
3	Role of HR department at HQ												
	Local HR is free to make decisions												
	Supply of resources and counselling to local HR if necessary												
	General guidelines and framework to local HR											0,092*	C=0,153
	Detailed HR model. policies. and control to local HR												
4	4. Involvement of external HR service providers												
	Planning of human resources												
	Recruitment											0,007***	C=0,191
	Selection											0,008***	C=0,19
	Training-development												
	Compensation-Remuneration												
	Labor relations												
	HRIS											0,018**	C=0,179
5	Knowledge transfer from HQ to local HR			0,038**	E=0,219							0,055* 0,032**	C=0,18 E=0,171
6	Role of local training in the development of HR competences											0,005*** 0,001***	C=0,215 E=0,237
		***=1% (minden ésszerű szignifikancia szint)						**=5%		*=10%			

Table 7

Hypotheses analysis (2008-2013)

H	Explanations	2008-2009 (Yes/ No)			2012-2013 (Yes/ No)		
		True	Partially true	Not true	True	Partially true	Not true
H1	The mandate of the subsidiaries has an impact on HR variables	Y				X	
H2	The company's origin has an impact on HR variables			N		X	
H3	The time of market entry of the MNC has an impact on HR activities.			N	Y		
H4	The nature of the subsidiary's market entry has impact on HR variables.			N		X	

H5	The strategic orientation of companies has an impact on the projected characteristics of HR activities.			N	Y		
H6	The size of the subsidiaries in terms of employee numbers has an impact on the various features of HR variables.		X		Y		
H7	The size of the employee numbers, expat numbers and revenue is has an impact HR department.		X			X	

Our hypotheses (H1-H7) and our review of the findings has already given us the opportunity to examine the connection between certain HR variables and the factors thought to influence their application. What is interesting is that although government regulation has reduced since 2009, more firms see the strategic orientation of the firm as impacting on HRM practices. This could be because there is more room for firms to experiment with new types of HRM, or simply owing to a more difficult operating environment resulting in firms being willing to take HRM more seriously. Subsidiary size also appears to be having an increasing impact on HRM practice. This could reflect the extent to which subsidiaries are gaining an increasing degree of autonomy. An interesting question would be the relative lifespan of the subsidiary as this contextual factor relates to the level of autonomy of HR managers. The hard times of 2008-2009 appear – at least for the time being – to be over. Whereas there is by no means a return to the pre-2008 days when over three fourths of MNC subsidiaries describe themselves as growing, most respondents reported they were once again expanding. Nonetheless, HR departments still remain smaller than their counterparts in the West – HRM is still evolving in the region.

Conclusions

We carried out our empirical survey for a second time within the Central and Eastern European region. In this study we presented the frequencies of the answers from the most important aspects of the survey. The results suggest that the companies adopt their HR model according to the external environment and internal business performance. The HR function appears to now be required to become more business (results) oriented, and at the same time line management is taking on more operational HR activities. The presence of HR business partners and outsourcing the non-strategic HR specialist tasks has increasingly become a prevailing operating model at the subsidiaries of MNCs in the Central-Eastern European region.

Our analysis of the pattern of the HR practices of MNC's in CEE region shows at best mixed support for the original two (efficiency, effectiveness) assumptions. While the practices of MNCs have contributed to considerable changes in HRM practices of the CEE region, yet a significant degree of diversity characterizes the region even when accounting for MNE similarities in size, strategy and subsidiary mandate. Seeking out patterns of HR resource allocation across this dynamic region continues to be a topic in need of further study.

Limitations and Future Plans

The research team readily acknowledges that these two research data sets do not constitute a replication. The sampling processes of the two iterations of a standardized instrument have a different mix of national samples for 2008-09 and 2012-13. Even so, the ability to generally track large samples of MNC activities across time, nations within the region, as well as across contextual characteristics while concentrating on a standard set of HR practices, policies, relationships and roles appears to have both an academic and practical public policy value.

The authors intend to develop the research in three different directions:

We intend to examine similarities and differences in HR practice at organizations operating in contexts of other types of capitalism (e.g. Mediterranean Market Economies, Coordinated Market Economies (CMEs), Liberal Market Economies (LMEs). An interesting question is the extent to which MNC country of origin is impacting on HR practice in the region. Based on previous research results (see Pudelko-Harzing 200); Allen-Aldred 2011; Whitcher-Chau 2012) we would predict that the pattern of human resource practices applied by MNCs in CMEs would differ from the configuration of practices in the LME region or in the CEE region.

In later stages of our ongoing research we will try to identify any patterns in type of industry, size of firm or national origin within a broader type of capitalism context. A more detailed investigation into the patterns by which MNCs set up various forms of subsidiary value chains within the region (and hence set up or modified mandates) and then went on to modify the nature and location of those operations is an interesting potential area of study. For an interesting recent conceptual review of the issues of MNC subsidiary roles and activities see Schmid, Dzedek and Lehrer (2014).

Furthermore, the relationship between the dynamics of MNC subsidiary mandate changes and MNC executive perceptions of local, area, national and CEE regional HR competencies and capabilities is a promising area of research study. To what degree do issues of the reputation of national and regional human capital impact decisions on the deployment and redeployment of MNC value chain activities and resources?

References

[1] Allen, M. M. C. Aldred, M. L. (2011) Varieties of Capitalism, Governance, and High-Tech Export Performance – A Fuzy Set Analysis of the New EU Member States. Employee Relations, Vol. 33 (4): pp. 334-355

[2] Antalóczy, K. Sass, M. (2005) Foreign Investments and Regional Location and Economic Effects (Economic Review), Vol. 52, (5): pp. 494-520

[3] Balázs, Gy and Veress, R. (2009) HR Challenges and Solutions in Crisis Time. Budapest: Hewitt Co.

[4] Bohle, D., Radice, H. and Shields, S. (2007) Introduction. Competition and Change. Vol. 11 (2): pp. 81-87

[5] Brewster, C. (2006) Comparing HRM Policies and Practices across Geographical Borders. In: Stahl, G. Bjorkman, I. (eds.) Handbook of Research in International HRM. Cheltenham: Edward Elgar Publication

[6] Claessens, S. and Djankov, S. (2002) Privatization in Eastern Europe. Journal of Public Economics, March: 307-324

[7] Cranet (2006) International Executive Report 2005, Cranet Survey on Comparative HRM. Cranfield: Cranet-Cranfield University

[8] Cranet (2011) International Executive Report 2005, Cranet Survey on Comparative HRM. Cranfield: Cranet-Cranfield University

[9] Delany, E. (1998) Strategic Development of Multinational Subsidiaries in Ireland. In: Birkinshaw, J. and Hood, N. (eds.) Multinational corporate evolution and subsidiary development.New-York: St Martin's Press

[10] Dowling, P. J., Festing, M. and Engle, A. (2013) International Human Resource Management, 6th Ed. London: Thomson-Cengage

[11] Falusné Szikra, K. (2000) Foreign Direct Invesments – Inland Jobs. Economic Review, Vol. 47, (6): pp. 446-458

[12] Fisher, C. D. ,Schoenfeld, L. F. and Shaw, J. B. (1993) Human Resource management. Boston: Houghton-Mifflin

[13] Fodor P., Kiss T. and Poór, J. (2011) Focus on the Impact of the Economic and Financial Crisis on the Human Resource Function – Four Eastern European Countries in the Light of Empirical Research in 2009. Acta Polytechnica Hungarica Journal of Applied Sciences, (1): pp. 81-104

[14] Francesco, A. M. and Gold, B. A. (1997) International Organizational Behavior. Upper Saddle River. Prentice Hall

[15] Golejewska, A. (2002) Foreign Direct Investment and its Employment Effects. Polish Manufacturing during Transition. Sopot: University Gdansk

[16] Hofstede, G. (1991) Cultures and Organizations. Software of the Mind. New York: McGraw-Hill

[17] Jackson, S. H. and Artola, M. (1997) Ethical Beliefs and Management Behavior: B Cross-cultural Comparison. Journal of Business Ethics, Vol. 16, (11): pp. 1163-73

[18] Jenkins, R. (2006) Globalization, FDI and Employment in Vietnam. Transnational Corporations, (1): pp.116-142

[19] Koenings, J. (2001) The Effect of Foreign Direct Investment on Domestic Firms. Economics in Transition, 3: 619-633

[20] Kotler, P. and Keller, K. L. (2006) Marketing Management. Upper Saddle River, N. J.: Pearson Prentice Hall

[21] Kotler, P. and Caslione, J. A. (2011) Kaotika, Budapest

[22] Lawler, J. J. (2006) Cultural and Institutional Determinants of HR Systems in International Affiliates of American Multinational Corporations. Institute of Labor and Industrial Relations, Illinois: University of Illinois

[23] Pauly, L. W. Reich, S. (1997) National Structures and Multinational Corporate Behavior: International Organization, Vol.51, (1): pp. 1-30

[24] Pavlinek, P. (2002) The Role of Foreign Direct Investment in the Privatisation and Reconstructing of the Czech Motor Industry. Post&Communist Economies, 3: 359-379

[25] Perlmutter, H. V. (1969) The Tortuous Evolution of the Multinational Corporation. Columbia Journal of World Business, Vol. 4, (1): pp.9-18

[26] Poór, J., Engle, A. and Gross, A.: Human Resource Management Practices of Large Multinational Firms in Hungary 1988-2005. Acta Oeconomica, 2010, Vol. 60 (4) pp. 427-460

[27] Pudelko, M. and Harzing, A.-W. (2007) Country of Origin, Localization or Dominance Effect? An Empirical Investigation of HRM Practices in Foreign Subsidiaries. HRM. Vol. 46 (4): pp. 535-559

[28] Pudlowski, E. M. (2009): Managing HR Cost in a Declining Economic Environment. Benefits Quarterly; (4): pp. 37-43

[29] Rosenzweig, P. M. and Nohira, N. (1994). Influence on Human Resource Management in Multinational Corporations, Journal of International Business Studies, Vol. 25, (2): pp. 229-251

[30] Schmid, S., Dzedek and Lehrer, M. (2014) From Rocking the Boat to Wagging the Dog: A Literature Review of Subsidiary Initiative Research and Integrative Framework. Journal of International Management, 20, (4): pp. 201-218

[31] Taylor, S., Beechler, S., Napier, N. (1996) Toward an Integrative Model of Strategic International HRM. Academy of Management Review, Vol. 21, (4):pp. 959-985

[32] UNCTAD (2010) World Investment Report, 2009, United Nations, Geneva

[33] UNCTAD (2012) World Investment Report, 2011 United Nations, Geneva

[34] UNCTAD (2014) World Investment Report, 2013, United Nations, Geneva

[35] Venaik, S., Midgley, D. and Devinney, T. (2005) Dual Paths to Performance: The Impact of Global Pressures on MNC Subsidiary Conduct and Performance. Journal of International Business Studies, Vol. 36, (6): pp. 655-675

[36] White, R. and Poynter, T. (1984) Strategies for Foreign-owned Subsidiaries in Canada. Business Quarterly, Vol 49, (2): pp. 59-69

[37] Witcher, B. J. and Chau, V. S. (2012) Varieties of Capitalism and Strategic Management: Managing Performance in Multinaltionals after the Global Financial Crisis. British Journal of Management. 23: S58-S73

[38] Wood, G., Cheah, J., Mazouz, K, Yin, S. (2014) "Foreign Direct Investment from Emerging Markets to Africa - the HRM Context", HRM (US/Wiley), 53 (1): pp. 179-201

[39] Yan, Y. (2003) A Comparative Study of Human Resource Management Practices in International Joint Ventures: the Impact of National Origin. International Journal of HRM, Vol. 14, (4): pp. 487-510

[40] Zsuzsanna K. Szabo, Emilia Herman (2014) Productive Entrepreneurship in the EU and Its Barriers in Transition Economies: A Cluster Analysis, Acta Polytechnica Hungarica, Vol. 11, No. 6, 2014