

# Uncertainty in Partner Selection for Virtual Enterprises

José Crispim<sup>1</sup> and Jorge Pinho de Sousa<sup>2</sup>

<sup>1</sup> Department of Management, Escola de Economia e Gestão, Campus de Gualtar,  
University of Minho, 4710-057 Braga, Portugal

crispim@eeg.uminho.pt

<sup>2</sup> INESC Porto / Faculty of Engineering, University of Porto, Campus da FEUP,  
Rua Dr. Roberto Frias, 4200-465 Porto, Portugal

jsousa@inescporto.pt

**Abstract.** A virtual enterprise (VE) is a temporary organization that pools the core competencies of its member enterprises and exploits fast changing market opportunities. The success of such an organization is strongly dependent on its composition, and the selection of partners becomes therefore a crucial issue. This problem is particularly difficult because of the uncertainties related to information, market dynamics, customer expectations and technology speed up. In this paper we propose an integrated approach to rank alternative VE configurations in business environments with uncertainty, using an extension of the TOPSIS method for fuzzy data, improved through the use of a stochastic multiobjective tabu search meta-heuristic. Preliminary computational results clearly demonstrate the potential of this approach for practical application.

## 1 Introduction

The success of a Virtual Enterprise (VE) strongly depends on all of the participating organizations being capable of cooperating as a single unit. Therefore, an adequate selection of partners is surely very important to overcome the fragilities of this type of organization (e.g., lack of formal contracts and heterogeneity of companies), that are much related with trust [1].

Moreover problem solving with explicit consideration of uncertainty may have a very high impact on real world situations, as problems arising in practice are becoming increasingly complex and dynamic, due partially to the fast development of communications that makes the perception of changes more rapid, stochastic and difficult to forecast [2]. Stability cannot be considered a reasonable assumption any more as we do not have perfect information either in terms of the VE projects (some activities or activity features, like the processing time, or the resources capacity cannot be known with certainty) or in terms of the characteristics and behaviours of the companies that will perform those activities (e.g., market capacity entrance). Moreover, each company's performance is affected by the operations or decisions of the others, i.e., there is a propagation of uncertainty through the members of a supply chain, and in the business environment customers' needs and preferences, market forces, technologies, and even the original problem being solved may change [3].

In many situations (like in partner selection) decisions have to be made in environments with high uncertainty. Actually, in dynamic environments, the context may change at any time, making a current VE no longer viable. In such a situation a new VE composition that better fits the prevailing circumstances has to be found.

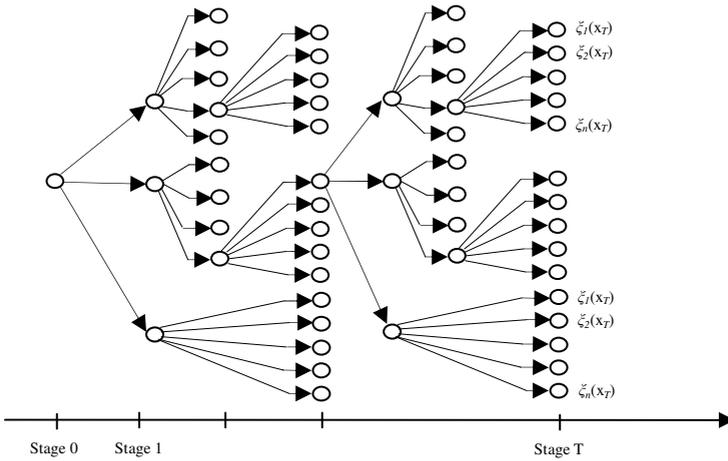
In Stochastic Combinatorial Optimisation Problems, all or part of the information about the problem is unknown, but it is possible to assume some knowledge about its probability distribution. To our best knowledge, there is in the literature no explicit reference to stochastic versions of the partner selection problem in the context of virtual enterprises. An interesting and complete survey about supplier selection that can be found in Aissaoui et al. [4] reflects this situation. Nevertheless, various models are available to select supply chain partners with uncertainty and risk (see e.g., [5-8]). Uncertainty in the configuration of collaborative networks has also been object of study by analytical or multi-agents approaches but not in terms of stochastic models (see e.g., [9] or [10]).

In this work we propose the use of a multistage stochastic model that captures both the stochastic and dynamic elements of real world situations, as a way to deal with uncertainty. Unfortunately, realistic stochastic models often lead to optimisation problems impossible to solve due to their astronomical number of possible outcomes. Therefore some kind of approximation procedure has to be performed. Our approach is based on a scenario tree formulation and makes use of the flexible hybrid algorithm presented by Crispim and Sousa [11] adapted to deal with uncertainty. This algorithm combines multiobjective (tabu search) with a multiattribute (fuzzy TOPSIS) technique. The scenario reduction technique adopted aims at reducing the vast number of possible scenarios to a manageable scenarios subset, keeping the approximation as close as possible to the original.

The remainder of the paper is organized as follows. In Section 2 the problem is formulated as a scenario tree, in section 3 the method used to solve the problem is presented, in Section 4 an illustrative example is described and finally, in Section 5 some preliminary conclusions are presented.

## 2 Scenario Trees

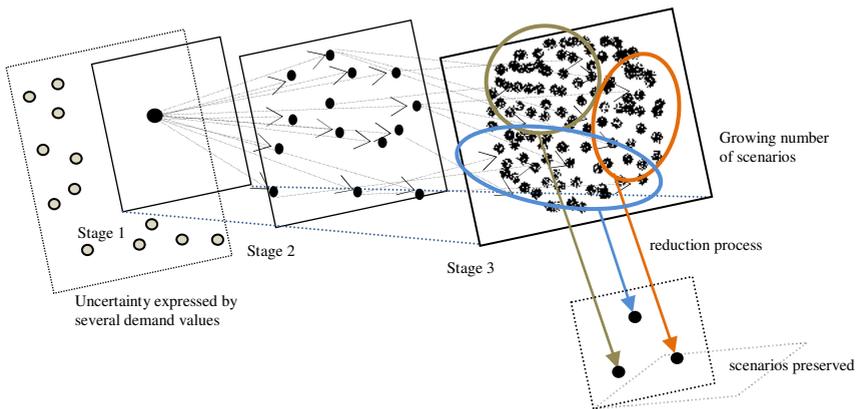
The discretization of the problem formulated as a scenario tree is a standard approach to solve multistage stochastic programs. A scenario is a path starting at the root node and ending at a node of the last period  $T$ . Along time, in different stages, several uncertain events can occur. In this way a representation is obtained for a finite number of possible realizations of the future outcomes of a given variable  $x$  (Figure 1). At each stage we have as many nodes as different realizations of  $\xi$ . The stages do not necessarily refer to time periods, but they correspond to steps in the decision process [12]. A (conditional) probability  $p_{ij} > 0$  (such that  $\sum_j p_{ij} = 1$ ) is associated to each arc of the tree. Therefore, at each stage, decisions must be made for different probability values. In scenario-based multistage stochastic programs, for feasibility reasons, one assumes that the probability distribution is discrete, and concentrated on a finite number of points, or branches. We also assume that the probability distributions of the various stages are independent of each other.



**Fig. 1.** Multistage problem formulation

It was found that uncertainties on demand quantity and timing are the two most common changes occurring in supply chain management [13]. The scenario-based approach attempts to capture uncertainty by representing it in terms of a moderate number of discrete realizations of random quantities. We assume here that the values taken by random variables  $\zeta_T$  are independent between stages.

Unfortunately, realistic models often lead to optimisation problems impossible to solve because of their size. According to Kim [14], in most large-scale stochastic programming problems, the total number of outcomes is huge and hence it is practically impossible to enumerate those outcomes. Therefore an approximation has to be done, i.e., a scenario tree is generated/aggregated/reduced (Figure 2).



**Fig. 2.** Scenario tree generation/reduction

Scenario reduction techniques aim at reducing the vast number of possible scenarios to a manageable subset (with or without a predefined cardinality). In the literature the studies along this idea basically follow one of two different perspectives:

- partition of the simulated scenarios, for example through cluster analysis (e.g., [15]);
- aggregation methods, for example by merging nodes with similar states of the stochastic parameters [16].

The objective is to make the scenario tree practically manageable, with a loss of information as small as possible. To avoid having to deal with exponentially growing scenarios we have used a reduction scheme based on cluster analysis.

The cluster simulation method adopted is similar to those introduced by Gülpinar et al. [17] or Shen and Zhang [15]. The main idea is to partition the simulated scenarios into random clusters and select one scenario in each cluster taken as representative – this scenario is known as the “centroid” (Figure 2). Therefore, the centroids<sup>1</sup> of the various clusters should be (in some sense) far away from each other. In our work, since we do not know the number of clusters in advance, we have chosen an agglomerative “hierarchical clustering” approach. In this procedure, to start the process, we assume each data point has its own cluster, and with each step of the clustering process, these clusters are combined to form larger clusters, in an iterative way. “Similarity” among each cluster’s members is measured through an Euclidian distance formula (see [18]).

These principles resulted in the design of a two phase scenario tree reduction procedure, as follows:

*Phase 1: Simulation*: randomly generate random variables (demand, production capacity and processing times) data through simulation.

*Phase 2: Clustering*: the generated data is grouped into clusters around a given number of centroids according to the hierarchical clustering scheme. Initially, we consider that the probability associated to the various clusters is the same and equals a given value -  $1/\#\_of\_clusters$ . After the clustering process, these probabilities have to be redistributed amongst the remaining scenarios (which correspond to the centroids). Therefore, we consider that the probability of each centroid is proportional to the number of elements in the respective cluster.

### 3 The Algorithm

The algorithm we have designed to cope with uncertainties in partner selection was firstly presented in Crispim and Sousa [11] and comprises two phases, namely: identifying good VE configurations, and then rank them according to a set of weights provided by the decision maker. In the first phase a multiobjective tabu search metaheuristic is used. As we take a multi-objective perspective, we are basically looking for a set of nondominated alternative solutions (the Pareto frontier).

---

<sup>1</sup> The centroid of a cluster is the average point in the multidimensional space defined by the scenarios, i.e., the cluster’s centre of gravity.

A stochastic feasible solution (i.e., a potential VE configuration) is represented by a set of companies in the network, associated to the different project activities, along with the corresponding attribute values, considered in an uncertain environment.

In other words, a solution is a subset of potential candidates (that all together form a so-called Virtual Breeding Environment) that fulfils the requirements of the project, respecting a set of constraints (associated to competencies, capacities...) and resulting in values for the different, possibly conflicting objectives.

In order to evaluate the stochastic Pareto solutions we adopted the stochastic domination concept proposed by Medaglia et al. [19] with the exception that we use distances instead of probabilities. Let  $x$  and  $y$  be a pair of feasible solutions for the partner selection problem, we say  $y$  stochastically dominates  $x$  (i.e.,  $y \preceq x$ ) if and only if the following conditions hold:

- i)  $E[f_k(y, \omega)] \geq E[f_k(x, \omega)]$  and  $D\{f_k(y, \omega) - T_k\} \leq D\{f_k(x, \omega) - T_k\}$ , for all  $k$ ;
- ii) there exists a  $k$  such that  $E[f_k(y, \omega)] > E[f_k(x, \omega)]$  and  $D\{f_k(y, \omega) - T_k\} > D\{f_k(x, \omega) - T_k\}$

where  $E[f_k(x, \omega)]$  is the expected value of the  $k^{\text{th}}$  objective and  $D\{f_k(x, \omega) - T_k\}$  is the distance between  $f_k(x, \omega)$  and the target value  $T_k$  specified by the DM. In our algorithm we propose  $T_k$  as the ideal value of the objective and, since we use fuzzy sets to express the information,  $T_k$  assumes the value of 1 in case of a benefit criterion or 0 in case of a cost criterion. Therefore, the differences to the deterministic version of the problem is in the way we evaluate each neighbourhood solution, as here a given number of samples is randomly determined to obtain the expected value of each objective, and the correspondent probability for each stochastic variable.

In the search (first) phase of the global procedure, we have introduced in the meta-heuristic algorithm the directional search concept, with the aim of generating a good approximation of the set of Pareto solutions, hopefully not neglecting any region of the search space. This concept tries to incorporate some important features into the algorithm: the ability to generate all available non-dominated solutions and be relatively easy to implement and apply.

The algorithm starts by exploring all objective functions choosing a specific objective  $f_1$ , to be improved, when  $f_1$  has not been improved for a certain (large) number of iterations. In this situation, in the next iteration, the search only makes use of objective  $f_1$ .

Since we admit the exploration of unfeasible solutions during the search (e.g., one potential VE formation may be unfeasible because of the lack of production capacity to satisfy the demand), we apply the same scheme to the constraints, i.e., in cases where the search has been performed in unfeasible regions of the solution space for too long, in the next iteration, the algorithm only accepts solutions that are feasible for the constraint with higher unfeasibility. To direct the search in such occasions we make use of two matrices, one for constraints and another for objectives. They are somehow similar to a tabu list, but they are used to force, and not to forbid, the search in a given direction..

For the second phase of the algorithm, we have designed a fuzzy TOPSIS approach that accepts different types of variables (numerical, interval, linguistic, etc.) to facilitate

the Decision Maker preferences expression. For a more detailed description of this procedure see Crispim and Sousa [11].

The algorithm steps are as follows:

*Step 0. Initialization:* Initialize the tabu list (the list of moves that is forbidden for a given number of iteration) and the ND (non-dominated) solutions list.

*Step 1. Select the current solution:* Randomly (uniformly) select a single current solution from the set of ND solutions.

*Step 2. Search the neighbourhood:* Search all possible defined moves, according to the adopted neighbourhood structure.

*Step 2.1 Directional search*

- If the objective parameter is activated, make the correspondent objective function active, otherwise, all objective functions are activated.
- If the constraint parameter is activated, only feasible solutions with respect to the activated constraint are kept.

*Step 2.2 Choose*

- Calculate the expected value of the  $k^{\text{th}}$  stochastic objective and the distance to the ideal value for each neighbour.
- Choose the non-tabu candidate solution with the best activated stochastic objective function(s) value(s) (or if it is tabu, but dominates any solution in the ND solutions list) as the best candidate solution.

*Step 3. Stochastic scheme (applied in each stage)*

Simulate a given number of scenarios for the stochastic variables (demand, processing time and production capacity) and compute the number of representative centroids and the respective occurrence probabilities.

*Step 4. Update the ND solutions list and the tabu lists*

## 4 Illustrative Example

Suppose that we have 3 manufacturing stages for a given product in which some stochastic events can occur, namely variations in demand (with an impact on the capability of firms to respect production capacity constraints and on the production costs). This example will allow us to demonstrate how the approach reacts to uncertainty influencing the objective functions and the constraints. Figure 3 shows a scenario tree in which several realizations of the uncertain demand are considered at each of three distinct stages (events). These events in which uncertainty unfolds over the planning horizon can be, for example, market research reports (important in case of fluctuating markets or in case of innovative and technological products), publicity actions, new market entrances, new competitors, etc. The number of scenarios represented in Figure 3 is only used for illustrative purposes (as the total number of scenarios of the studied example is 512).

Making use of the data of the problem example proposed by Crispim and Sousa [11], namely: a project composed by 6 activities (Table 1) and a network composed

by 100 (candidate) enterprises where 12 different activities that require 10 different resources has to be performed. These companies are characterized by: enterprise identification code (a number from 1 to 100); activity; interval time for the availability of resources; capacity; and 8 evaluation attributes (Table 2).

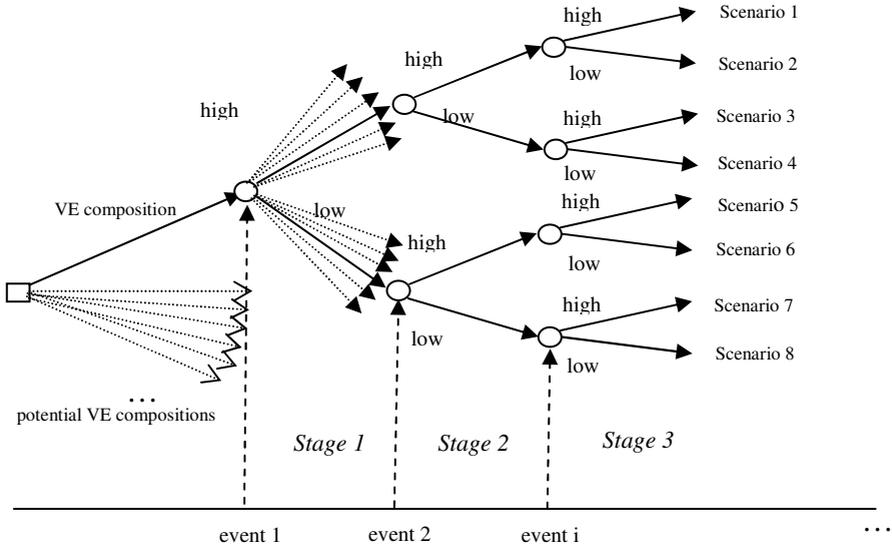


Fig. 3. Scenario tree

Table 1. Project data

Project						
Activities (code)	Resources	Precedent activities	Duration	Earliest start time	Latest start time	Quantity of resources
A	7	-	36	0	106	400
B	8	-	62	0	97	604
C	3	-	67	0	122	528
D	5	A	16	36	122	275
E	4	B	25	62	122	368
F	8	C,E,D	43	87	165	304

Table 2. Description of attributes

attributes (objectives)	c1	c2	c3	c4	c5	c6	c7	c8
type	linguistic	numerical	interval	interval	linguistic	numerical	numerical	linguistic
max (+) / min (-)	+	+	-	-	+	+	-	+
cardinality (for linguistic)	7	-	-	-	3	-	-	7
weight (%)	20	23	2	7	19	13	14	2

*Impact of demand uncertainty on the constraints.* As the final demand at each stage is unknown *a priori*, we will not be sure about each firm’s capability to produce the required quantity. Therefore, for each company, we have calculated its probability of being capable of satisfying the required demand. For example, if the demand at a given stage is normally distributed with  $N(\mu=2000, \sigma=200)$ , the probability that a company with a production capacity of 2561 units is able to satisfy the demand is  $\Phi\left(\frac{2561-2000}{200}\right) \approx 0,997$ . For decision purposes, we have assumed that a solution is feasible if the probability of respecting the capacity constraints is higher or equal to 0,8 for all companies involved in the three stages considered. In a practical situation the DM would be able to define their own rules to distinguish between feasibility and unfeasibility.

*Impact of demand uncertainty on the objective functions.* At each stage we obtain several realizations (in our case several centroids for high demand and several centroids for low demand) through the clustering procedure presented above (Table 3). Once an acceptable clustering is found, it is necessary to represent each cluster by a single point to be used in the scenario tree. If the centre of the cluster does not correspond to any obtained point (e.g., in the case where the cluster of points is quite sparse), the centroid should be the closest point to the cluster centre. The probability of each centroid is proportional to the number of elements in the respective cluster. This process starts with the generation of a sample of size 100 from the normal distribution [20] based on the Central Limit Theorem (CLT). Moreover, we assume that the probability distributions at the three stages are independent.

**Table 3.** Centroids of the stochastic demand

	Stage 1		Stage 2		Stage 3	
	N(2000,200)		N(2500,250)		N(1500,150)	
high demand	demand value	probability	demand value	probability	demand value	probability
	2146	0,37	2245	0,26	1367	0,15
	2356	0,12	2834	0,18	1828	0,10
	1906	0,36	2537	0,56	1612	0,31
	1721	0,15			1487	0,34
				1233	0,10	
low demand	N(1000,100)		N(1500,150)		N(500,50)	
	873	0,20	1698	0,15	436	0,23
	1046	0,26	1409	0,35	545	0,29
	973	0,37	1847	0,8	494	0,48
	1128	0,17	1550	0,32		
		1240	0,10			

In terms of production costs we are assuming the VBE companies benefit from a quantity discount depending on the demand level that follows a discrete uniform distribution.

The algorithm found 18 non-dominated solutions for the project, with capacity and time windows constraints. It should be noted that for activity D, only 5 companies have an adequate production capacity, this leading to the formation of a consortium between some of them. Table 4 shows the ranking and composition of the alternative configurations. These coalitions are those that prove to be more “robust” to face the demand uncertainty with impact on the objective function and on the constraints.

**Table 4.** Results

Rank	VE					Project activities					
		$\bar{d}_i^+$	$\bar{d}_i^-$	$\bar{R}_i$	A	B	C	D	E	F	
1	5	307.975	152.146	0.0470763	35	71	14	31	47	71	
2	10	308.717	137.358	0.0425979	21	44	20	101	72	44	
3	4	308.682	134.007	0.0416065	21	28	14	90	95	28	
4	2	309.287	132.176	0.0409843	35	6	14	26	72	6	
5	14	309.216	127.847	0.0397040	21	97	14	26	72	44	
6	3	309.221	127.776	0.0396821	21	97	14	26	72	97	
7	18	309.221	127.776	0.0396821	21	97	14	26	72	97	
8	6	309.640	118.609	0.0368921	74	44	20	90	57	44	
9	11	309.822	115.400	0.0359096	74	44	20	101	72	44	
10	8	309.967	115.299	0.0358630	7	44	20	26	72	44	
11	9	309.762	114.749	0.0357210	46	44	20	101	72	44	
12	7	310.010	110.672	0.0344689	74	44	20	101	39	44	
13	16	310.243	106.676	0.0332415	74	97	14	26	72	94	
14	17	310.320	104.907	0.0327005	74	97	14	26	72	6	
15	12	310.320	104.613	0.0326120	74	97	14	26	72	44	
16	15	310.320	104.613	0.0326120	74	97	14	26	72	44	
17	13	310.260	103.886	0.0323986	46	97	14	26	72	44	
18	1	310.459	975.056	0.0304506	46	97	14	101	39	97	

Note: The company n° 101 consists in one consortium formed by 7 individual companies (company n°s: 9, 13, 43, 49, 55, 68, 79)

## 5 Conclusions

The approach developed in this work creates a quite general and flexible research framework that can be used to analyse numerous partner selection scenarios. With this framework the decision maker is able to easily change the objectives and constraints in order to obtain a satisfactory solution, and allowing the use of a mix of variable types to express his/her preferences. All types of evaluation criteria can be used and, in a real situation, the decision maker should use criteria that are in accordance with the available (or obtainable) information.

The proposed algorithm includes an innovative multiobjective directional stochastic tabu search metaheuristic. The flexibility provided by this approach becomes even more important if we think in the uncertainty propagation within the network and/or in the specificity of the virtual environment. This efficient quantitative tool should be able to provide an adequate, useful support, in simulating and assessing different alternative solutions for the uncertain business environments associated to VE formation or re-organization.

## References

1. Camarinha-Matos, L.M., Afsarmanesh, H.: A framework for virtual organization creation in a breeding environment. *Annual Reviews in Control* 31, 119–135 (2007)
2. Bianchi, L., Dorigo, M., Gambardella, L.M., Gutjahr, W.J.: Metaheuristics in stochastic combinatorial optimization: a survey. Technical Report IDSIA-08-06. Dalle Molle Institute for Artificial Intelligence, Manno, Switzerland (2006)

3. Molokken-Ostfold, K., Jorgensen, M.: A comparison of software project overruns - flexible versus sequential development models. *IEEE Transactions on Software Engineering* 31, 754–766 (2005)
4. Aissaoui, N., Haouari, M., Hassini, E.: Supplier selection and order lot sizing modeling: A review. *Computers & Operations Research* 34, 3516–3540 (2007)
5. Goetschalckx, M., Vidal, C.J., Dogan, K.: Modeling and design of global logistics systems: A review of integrated strategic and tactical models and design algorithms. *European Journal of Operational Research* 143, 1–18 (2002)
6. Paulraj, A., Chen, I.J.: Strategic supply management: theory and practice. *International Journal of Integrated Supply Management* 1, 457–477 (2005)
7. Ding, H., Benyoucef, L., Xie, X.: A simulation-based multi-objective genetic algorithm approach for networked enterprises optimization. *Engineering Applications of Artificial Intelligence* 19, 609–623 (2006)
8. Wu, D., Olson, D.L.: Supply chain risk, simulation, and vendor selection. *International Journal of Production Economics* 114, 646–655 (2008)
9. Ivanov, D., Kaeschel, J., Sokolov, B., Arkhipov, A.: A Conceptual Framework for Modeling Complex Adaptation of Collaborative Networks. In: Camarinha-Matos, L.M., Afsarmanesh, H., Ollus, M. (eds.) *Network-Centric Collaboration and Supporting Frameworks*, pp. 15–22. Springer, Boston (2006)
10. Tolkacheva, V., Ivanov, D., Arkhipov, A.: Assessment of Collaborative Networks Structural Stability. In: Camarinha-Matos, L.M., Afsarmanesh, H., Novais, P., Analide, C. (eds.) *Establishing The Foundation of Collaborative Networks*, pp. 75–82. Springer, Boston (2007)
11. Crispim, J.A., Sousa, J.P.: Partner selection in virtual enterprises. *International Journal of Production Research* (2008) (in press),  
<http://dx.doi.org/10.1080/00207540802425369>
12. Dupačová, J.: Applications of stochastic programming: Achievements and questions. *European Journal of Operational Research* 140, 281–290 (2002)
13. Das, S.K., Abdel-Malek, L.: Modeling the flexibility of order quantities and lead-times in supply chains. *International Journal of Production Economics* 85, 171–181 (2003)
14. Kim, J.: Event tree based sampling. *Computers & Operations Research* 33, 1184–1199 (2006)
15. Shen, R., Zhang, S.: Robust portfolio selection based on a multi-stage scenario tree. *European Journal of Operational Research* 191, 864–887 (2008)
16. Blomvall, J., Shapiro, A.: Solving multistage asset investment problems by the sample average approximation method. *Mathematical Programming* 108, 571–595 (2006)
17. Gülpinar, N., Rustem, B., Settergren, R.: Simulation and optimization approaches to scenario tree generation. *Journal of Economic Dynamics and Control* 28, 1291–1315 (2004)
18. Balopoulos, V., Hatzimichailidis, A.G., Papadopoulos, B.K.: Distance and similarity measures for fuzzy operators. *Information Sciences* 177, 2336–2348 (2007)
19. Medaglia, A.L., Graves, S.B., Ringuest, J.L.: A multiobjective evolutionary approach for linearly constrained project selection under uncertainty. *European Journal of Operational Research* 179, 869–894 (2007)
20. Tavares, L.V., Oliveira, R.C., Themido, I.H., Correia, F.N.: *Investigação Operacional*. McGraw-Hill, Lisboa (1996)