International Journal of Information, Control and Computer Sciences
ISSN: 2517-9942
Vol:3, No:2, 2009

A Context-Aware Supplier Selection Model

Mohammadreza Razzazi, Maryam Bayat

***Abstract*—**Selection of the best possible set of suppliers has a significant impact on the overall profitability and success of any business. For this reason, it is usually necessary to optimize all business processes and to make use of cost-effective alternatives for additional savings. This paper proposes a new efficient context-aware supplier selection model that takes into account possible changes of the environment while significantly reducing selection costs. The proposed model is based on data clustering techniques while inspiring certain principles of online algorithms for an optimally selection of suppliers. Unlike common selection models which re-run the selection algorithm from the scratch-line for any decision-making sub-period on the whole environment, our model considers the changes only and superimposes it to the previously defined *best set of suppliers* to obtain a *new best set of suppliers*. Therefore, any re-computation of unchanged elements of the environment is avoided and selection costs are consequently reduced significantly. A numerical evaluation confirms applicability of this model and proves that it is a more optimal solution compared with common static selection models in this field.

***Keywords*—**Supplier Selection, Context-Awareness, Online Algorithms, Data Clustering.

I. INTRODUCTION

IN modern business, a key tool for success is supplier selection mechanism as it has a significant bearing on the performance and overall-profitability of the business. Qualified suppliers, at right quantities and when required, make the related business fruitful and promising. Therefore, upgrading such mechanism is a major concern in the commercial-related researches and forms the basis of an extensive research from which this paper has been extracted.

As is illustrated in figure 1, supplier selection process is usually consists of four steps [1]. The first step, Problem Definition Step, concerns decision makings which should determine the strategy of purchases e.g. the duration of new selection. Second stage of this process is called Decision Criteria Formulation. Depending on the purchasing situation, the multi-criteria nature of selecting the right suppliers may force some complexity to this decision. The best set of selection criteria is carefully determined in this phase. In the Pre-Selection of Potential Suppliers, the third step, the set of alternatives is refined with respect to the ability of satisfying a minimal threshold. Final Selection is the last phase in the supplier selection process. At this stage, ultimate suppliers are identified and orders are allocated among them while considering the system's constraints and taking into account a multitude of criteria.

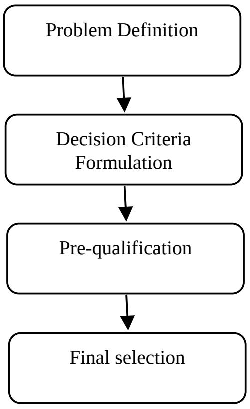

Fig. 1: General supplier selection flowchart

There are several variations of decision models and techniques which are proposed to accomplish this task. A general classification of existing techniques regarding the supplier selection process suggested in [1] is as follows:

- Single objective techniques:
    - Linear Programming Method
    - Mixed-integer Programming Method
    - Non-linear Programming Method
    - Dynamic Programming
    - Stochastic Programming
    - Decision Theory
- Multiple objective techniques:
    - Multi-objective Programming Method
    - Goal Programming Method
- Other Methods such as Neural Networks.

However, a fresh computing procedure, called ubiquitous/pervasive computing, that employs new means of automation and computation without explicit involvement of human has been emerged recently. Due to its potentials in establishing communication between human (or other objects) with computing systems, as it runs in the background of everyday life of people and tries not to be sensed by human, it is nicknamed as the calmest technology [2].

Currently, the most focused issue in ubiquitous/pervasive

M. Razzazi, PhD, is with the Department of Computer Engineering & IT, Amirkabir University of Technology (Tehran Polytechnic), Tehran, 15914, Iran (phone: +98-21-6454-2732; fax: +98-21-6649-5521; e-mail: Razzazi@ aut.ac.ir).

M. Bayat is with the Department of Computer Engineering & IT, Amirkabir University of Technology (Tehran Polytechnic), Tehran, 15914, Iran (e-mail: Maryam.Bayat@ aut.ac.ir).

computing field is context-aware computing. A context-aware system is able to adapt its operations to a given context, without explicit user intervention and thus aims at increasing usability and effectiveness by taking environmental context into account [3]. As proposed by Dey and Abowd, "context" may be defined as "any information that can be used to characterize the situation of entities (i.e. either a person, place or object) that are considered relevant to the interaction between a user and an application, including the user and the application themselves" [3]. According to this definition, three distinct entities are identified concerning context, namely Places, People and Things. Each entity is characterized by certain attributes such as identity, location, status and time [3].

Regarding the pervasive computing concepts, our new idea in this work is to apply the basis of context-awareness in the supplier selection process so that it can run in the background of business and adapt the best set of suppliers to the changes of the environment.

In developing a context-aware supplier selection model, the environment may be identified as a set of possible supply candidates which may be changed in course of time. For example, suppliers frequently change their products' price in response to the governing political/social/financial conditions; new suppliers may appear and certain supplier may even lunch new trade strategies for boosting their business. Context, therefore, should also reflect these evolving-candidates' influences. To make the supplier selection process adaptable to the changes of the environment, one may suggest a re-run of the selection model every time a minor change takes place. However, most of current selection models are static in their nature and consider just a single snapshot of the environment at a time without considering changes' effects to the past and future states (i.e. they don't consider the changes in suppliers' status in the course of time). For this reason, in static selection models, a combination of changed and unchanged elements (i.e. candidate suppliers) are to be re-considered as a new set-up to make the selection process fair. This imposes additional computational cost and time to the business and, therefore, may inversely affect its overall profitability.

To develop a cost-efficient context-aware supplier selection model, online algorithms concepts and data clustering techniques has been used in this work. The proposed selection model is capable for considering all changes in the environment at each decision making step while making use of the solutions found in the past sub-periods. Therefore, the proposed model can find the new optimum set of suppliers while minimizing the overall cost of performing this process.

Within the remaining parts of this paper followings are covered: in section 2 a brief review of online algorithms and data clustering techniques are presented. The so-called context-aware supplier selection model for time-sensitive supplies is discussed in the third section. The result of evaluating the proposed selection model is demonstrated in section 4 and finally sections 5 presents some concluding remarks.

II. Background

A. Online algorithms

In one of decision-making process that is commonly known as *online decision algorithms*, a sequence of decisions can be made with minimal knowledge of future events [4]. In other words, an *online algorithm* does not rely on immediate access to the entire information concerning a given problem. Instead, the problem is revealed to the algorithm incrementally, and in response to each incomplete portion of the input, the algorithm must take an irreversible action with no access to the future input [5].

Since in our supplier selection problem, changes in the candidates' environment usually occur incrementally, the online algorithms may be regarded as a powerful means for decision making at first glance. However, applicability of the available *online algorithms* with such specific usage is not a straight forward task and there exist certain constraints in their usage in this field. For example, a pioneer work on this subject [4] considers a randomized "follow the expected leader" algorithm while the work in [6] utilizes a deterministic algorithm based on the expected gradient ascent. Another class of thoughts in this field is based on Aggregating Algorithm (AA), for solving online decision problems optimally [5]. One common feature in all versions of *online algorithms*, however, is their lengthy computational requirements due to their complicated mathematical basis which make them expensive tools for commercial applications. In other words, as our immediate goal in this work is to lunch a new context-aware supplier selection model while minimizing the computational cost, none of existing online algorithms seems to be capable of fulfilling such requirement and decreasing the selection costs. Therefore, we only inspired from the nature of these algorithms in developing our online supplier selection model.

B. Data Clustering

In all aspects of human life, one of the most important means in analyzing phenomena and objects is classifying them into categories or *clusters* in order to extract their describing features and, also, comparing them with other objects or phenomena on the basis of their similarities and dissimilarities [7].

Theoretically, classification systems are either supervised or unsupervised, depending on whether they assign new inputs to one of a finite number of discrete supervised classes or unsupervised categories, respectively. In supervised classification, the mapping from a set of input data vectors ($x \in R^d$, where *d* is the input space dimensionality), to a finite set of discrete class labels, is modeled in terms of some mathematical function [7]. In unsupervised classification, called *clustering* or exploratory data analysis, no labeled data are available. The goal of clustering is to separate a finite unlabeled data set into a finite and discrete set of "natural," hidden data structures, rather than provide an accurate characterization of unobserved samples generated from the same probability distribution [8].

Clustering techniques are useful in several exploratory pattern-analysis, grouping, decision-making, and machine-learning applications. Some times there exist little prior information (e.g., statistical models) and decision-maker also has to make as few assumptions about the data as possible. In these situations clustering methodology is particularly appropriate for exploring interrelationships between data points in order to make an assessment of their structure [8].

Selecting one of different clustering techniques depends on the nature of the problem to be solved. Theoretically, these techniques may be classified as hierarchical and non-hierarchical clustering algorithms. Hierarchical techniques, such as Growing Hierarchy Self Organizing Map (GHSOM) clustering, produce a nested set of partitions and are usually employed in discovering natural structure of some phenomena. While, nonhierarchical clustering methods only partition data into a pre-specified set of clusters such as Self Organizing Map (SOM) method. In addition to the above mentioned major classes, other clustering alternatives such as fuzzy clustering, nearest neighbor clustering and evolutionary algorithms for clustering have been developed [8]. Evolutionary approaches, motivated by natural evolution, make use of evolutionary operators and a number of solutions to obtain the globally optimal partition of the data [8]. In other words, unlike other clustering methods which work with static data sets - data that neither move, nor disappear nor emerge - evolutionary clustering algorithms consider dynamic data set and try to obtain the *globally optimal* cluster set of data.

Some researchers including [10], [11], [12], and [13] have considered evolution of clusters of dynamic data sets. The most recent effort [12] proposes a novel method for online clustering of the dynamic data set, based on state space model where the measurement equation is represented by a Gaussian mixture with unknown number of components where the state equation is not explicit. They solve this problem by deriving a SEM algorithm which updates the current clustering C_t from a window of snapshots of the dynamics data set, denoted by $D_{t-1} \dots D_t$ [11].

Evolutionary clustering technique offers an attractive tool in developing a context-aware supplier selection model due to its dynamic nature. However, its theoretical complexity as well as its high computational cost, prohibits its applicability in our problem. For this reason, static clustering algorithms seem to be more appropriate to our work. From several alternatives of static clustering algorithms we have chosen GHSOM method. As the hierarchy of clusters is not important in supplier selection problem, we have used it just in layer one.

On the ground of supplier selection, every single data is regarded as a vector composed of some components. These components are the criteria chosen for selection process. For every supplier a value is to be assigned to these selection criteria by means of which the vector is capable of describing a particular supplier.

In clustering a given set of candidates, it is preferable to lift restrictions on the number of clusters to a predefined value as the number of clusters may vary because of emergence and elimination of candidates. For this reason, GHSOM is a promising clustering algorithm that eliminates the need for predefining the number of clusters.

Every added/removed data can change the quality of the cluster that it belongs to. Clusters' quality, therefore, should be monitored constantly. By predefining a quality threshold, when the set threshold is reached, a *bad cluster* has to be split into new smaller clusters with qualities below that threshold. We consider a *bad cluster* as a new map that requires a re-run of GHSOM. To measure quality of a cluster, we have adapted the sum of Weighted Euclidean Distance of data from the reference vector of their immediate cluster that has been suggested by GHSOM itself.

III. The Proposed Model

This section represents the overall framework of our approach by adapting the general supplier selection framework discussed in section 1.

A. Problem Definition

Similar to other selection models, our first step is Problem Definition step (see Fig. 1) in which the overall planning period for selecting appropriate suppliers is determined. This is a *planning-level decision*, and may be as long as one year or may be as short as a quarter of a year and is called "total selection period". Then the so-called "selection sub-periods" are to be defined as fractions of total selection period depending on the strategy of the firm.

B. Decision Criteria Formulation

The second step, the Decision Criteria Formulation step, consists of two categories of criteria to be considered [14]. The first category concern is the *supply profits* and is formed by those criteria having direct impact on the profitability potential of selected suppliers. The second category consists of those criteria handling the *supply risk* and includes variances of influential supply profit criteria, such as variance of price or variance of quality. The greater a variance, the more risky would be selecting that candidate. As discussed in section 1, these criteria are regarded as context in our work.

We have chosen quality, quantity, price, delivery-cost, delivery-capacity, delivery-lead-time as the *supply profit criteria*. The *quality* of supplied materials is defined as the number of defected items in a lot, and the *quantity* is the capability of the supplier to meet orders. The price criterion is the offered cost-price per unit of supplied items.

Delivery criterion is handled by the flow networking techniques, as discussed in [15]. Regarding the nature of flow networks, the unit cost of delivering the supplied material to the producer, delivery-cost, and the capacity of the flow (road) between supplier and producer, delivery-capacity, is covered. To consider the lead-time factor, i.e. the required time for delivery of goods (delivery-lead-time), another selection criterion is also included.

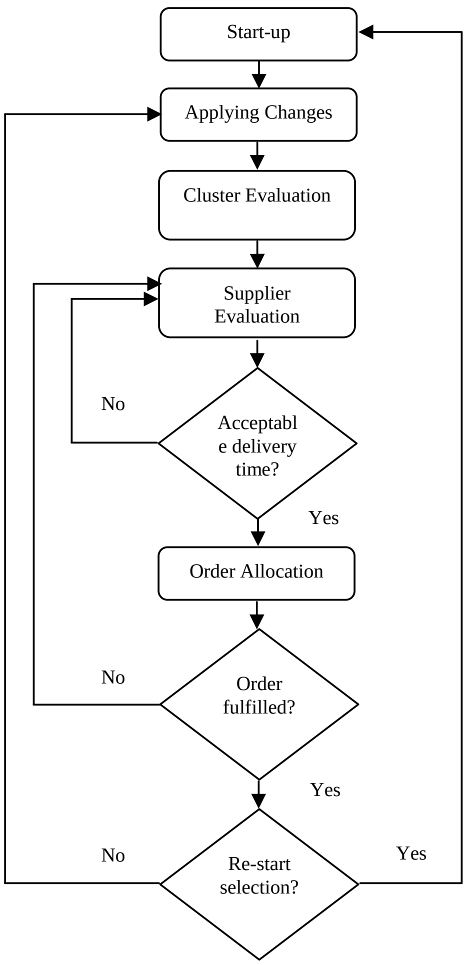

Fig. 2: The Selection flowchart

In our model, no allowance has been given to the *supply risk* criterion. However, to handle the supply risk, we employed a new strategy which is more reliable and runs in the context of selection process. Unlike a proposed approach in [14] that handles the supply risk by means of the value of some supply profit criteria at variance with its value from immediate previous sub-period, our risk management technique considers a history of criteria from T_0 (first sub-period) to present (T_{t-1}) in order to monitor the long term performance of all suppliers. This technique enables us to find out the degree of trust-ability of suppliers in the future. The technique is based on a statistical regression function and makes certain decisions in accordance with its forecasts for the subsequent sub-period.

C. Selection

Two final steps of general supplier selection framework are combined to form the Selection step in our model. This is the most important development in the proposed framework and consists of some sub-phases as shown in Fig. 2.

1) Start-up

This phase is to be run at T_0 only. At the first sub-period, we assume having no pre-knowledge about the environment. Therefore, one of common static supplier selection models must be employed. The model described in [14] seems to fit better than other methods as it clusters the candidates by means of three supply profit criteria (quality, quantity and price). This static selection model at T_0 provides a set of the best suppliers.

2) Applying changes

At the beginning of the subsequent sub-period (T_{t-1}), we have a changed set of candidates. Any potential supplying offer may undergo changes (e.g. changes in price and/or quality), some supplier may seize activity, and fresh candidate may appear at this time. Therefore, it is essential to consider theses changes and accommodate them within new appropriate clusters. As pointed out in section 2.2 changing data may lead to changes in the quality of clusters. Clusters with a quality below the specified threshold should be split (Fig. 3). By the end of this phase a new state of environment forms.

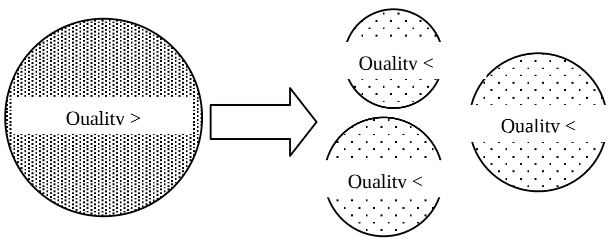

Fig. 3: splitting a bad quality cluster

3) Cluster Evaluation

Having a changed environment of candidates, candidates should then be evaluated and scored subsequently. To safeguard the ultimate goal of reducing computational cost, we perform the evaluation process at a concise cluster level. As mentioned before, a group of data that forms a cluster has commonly a reference vector representing that particular cluster. These indices in our model act as abstractions of data and may be evaluated instead of actual clusters. The evaluation process is based on two advanced criteria belonging to the *supply risk* category. The first criterion concerns historical performance analysis of a given cluster in which a sequence of the cluster performance in previous sub-periods are plotted by means of statistical regression methods in order to enable forecasting cluster's performance in the subsequent sub-period. In mathematical sense it may be written as:

$$\nabla = reg\{\zeta[(cls)_i, T_0], \ldots, \zeta[(cls)_i, T_{t-1}]\} \qquad (1)$$

Where:

$(cls)_i$ = i^{th} cluster.

reg = regression function

$\zeta\ [(cls)_i, T_j]$ = performance of i^{th} cluster at T_j sub-period

∇ = performance trend of i^{th} cluster at $[T_0 ... T_{t-1}]$

And the forecasted performance $(\dot{\zeta}_i)$ at T_t may be computed using:

$$\dot{\zeta}_i = \nabla\left[(cls)_i, T_t\right] \tag{2}$$

$$(\Gamma_P)_i = \dot{\zeta}_i / \sum_i \dot{\zeta}_i \tag{3}$$

Where: $(\Gamma_P)_i$ stands for the computed score of the i^{th} cluster performance

The second criterion for evaluating a cluster has been based on a definition of the *quality* as the sum of Weighted Euclidean Distance of data from their reference vector. As the cluster reference vector represents actual data in our model, the denser the data points, the better estimates can be obtained. Therefore, the smaller the sum of Weighted Euclidean Distance of data the better cluster.

$$Q(cls)_i = \sum_j d_j \tag{4}$$

Where:

d_j = weighted Euclidean distance of j^{th} data in i^{th} cluster from its reference vector

$Q(cls)_i =$ Quality of i^{th} cluster

And to compute the quality score of i^{th} cluster $(\Gamma_Q)_i$ we may write:

$$(\Gamma_Q)_i = 1 - Q(cls)_i \tag{5}$$

Integrating the resulted scores calculated by (3) and (5) provides a sound basis for evaluating clusters and assigning an overall score (τ_i) to each one.

$$\tau_i = \alpha \cdot (\Gamma_P)_i + \beta \cdot (\Gamma_Q)_i \tag{6}$$

In (6), α and β are arbitrary coefficients taking into account the desire weight of each component in that equation. The greater α the more weight is given to the historical performance and the greater β the denser cluster would result. This should be noticed that α and β vary between 0 and 1 so as their sum always equals 1.

4) Supplier Evaluation and Order Allocation

At this step ultimate suppliers are selected on the basis of assigned scores and orders allocations are made amongst them. To perform this task, candidates in the first best clusters are to be evaluated by means of delivery-capacity, delivery-cost and delivery-lead-time criteria.

This is noteworthy that separating the delivery-related criteria and *supply profit* criteria allows similar candidate to change there geographical location without any impact on their evaluation. We call this type of candidate as “mobile suppliers”.

In this evaluation, a higher delivery-capacity $(D_{cap})_j$ has a positive impact on the final supplier score and a higher delivery-cost $(D_{\cos t})_j$ has a negative impact. In addition to delivery-capacity and delivery-cost a third criterion in this category, delivery-lead-time, should also be considered. Those candidates with lengthy delivery-lead-time beyond a desired threshold are not finalized and, therefore, may be omitted. In mathematical formulation it may be written as:

$$(\Gamma_S)_j = \left[\frac{(D_{cap})_j}{\sum_j (D_{cap})_j}\right] + \left[1 - \frac{(D_{\cos t})_j}{\sum_j (D_{\cos t})_j}\right] if: (D_{time})_j \le \varepsilon \tag{7}$$

Where:

$(\Gamma_S)_j$ Stands for the computed score of the j^{th} supplier from i^{th} cluster

$(D_{time})_j$ Denotes delivery lead time of the j^{th} supplier from i^{th} cluster

$\varepsilon =$ a threshold defined by the firm

With finalized scores, suppliers are selected decently from a list of sorted scores. In the process of *orders allocation* to these selected suppliers, the minimum value of either delivery-capacity or quantity criteria (N_j) are used.

$$(\Lambda_S)_j = \min\left[(D_{cap})_j, N_j\right] \tag{8}$$

Where:

$(\Lambda_S)_j$ Stands for allocated order to j^{th} supplier from i^{th} cluster

If the overall capacity of supplying candidates of the first best cluster is short of demand, the subsequent best cluster is considered to meet the remaining portion of the demand and its candidates are evaluated with the same procedure. And so on.

IV. Evaluation

To evaluate the proposed context-aware supplier selection model, a set of 100 supplying candidates for a given industry were considered. The “total selection period” was assumed to be in order of six months and each “selection sub-period” lasts after exactly one month leading to six sub-periods: T_0, T_1… T_5.

Then a mixed-integer programming model – that is believed to be a similar approach - was considered as a stand-comparison for assessing effectiveness of the proposed model from three different points of view including number of selected suppliers, total supply costs and the computational costs.

The number of selected suppliers is important as a greater number of selected suppliers causes more difficulties in managing them and imposes more additional costs to the business. As shown in Fig. 4, our context-aware supplier selection model is capable to meet the orders with fewer numbers of suppliers.

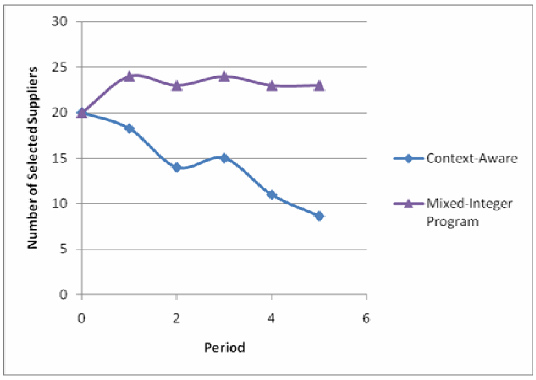

Fig. 4: Number of selected suppliers per period

The total supply cost that should be paid by a business is a function of offered prices, delivery costs and switching costs[1]. In mathematical terms it may be written as:

Total supply cost = (price × allocated order) + (delivery costs × allocated order) + switching costs

This criterion's evaluation results are presented in Fig. 5 and Fig. 6.

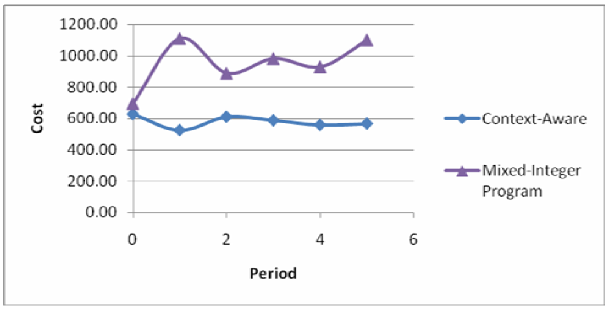

Fig. 5: Total supply cost per period

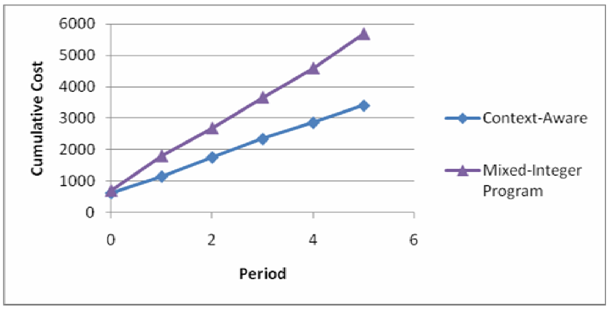

Fig. 6: Cumulative cost per period

To assess computational costs of running the models on a PC with an Intel Pentium 1.80GHz processor and 512 MB of RAM, as an identical computing tool for both models, an exactly the same set of candidates and demands were considered and the results of analysis are shown in Fig. 7 and Fig. 8.

[1] In literature, switching cost is defined as the investment in training and technology when selecting a new supplier.

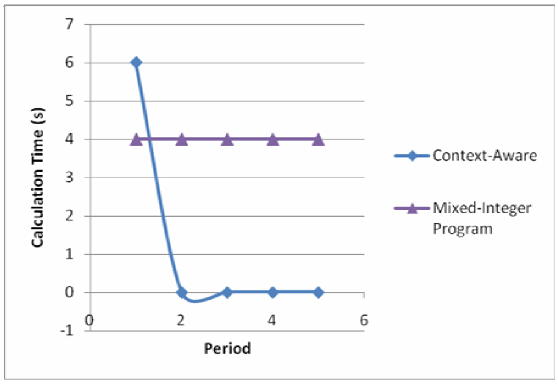

Fig. 7: Computional time per period

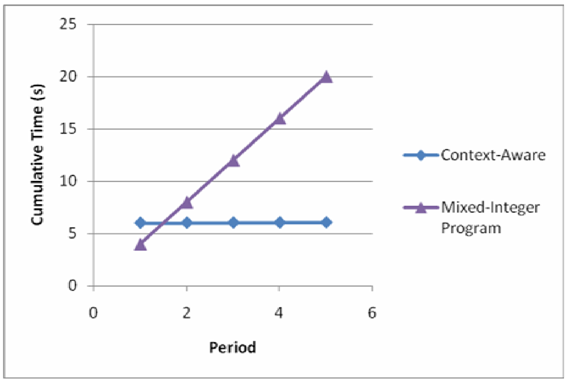

Fig. 8: Cumulative computational time per period

V. Conclusion

Selecting the best possible suppliers is commonly regarded as a crucial issue due to its bearings on the total profitability and success of any business. Optimization of this business-process and reducing its potential inflating costs, therefore, is an important field of studies in the modern commerce. The emergence of new computing procedures such as pervasive computing, introduces new means for performing processes of a business more efficiently. In this area, context-aware computing develops systems which are able to adapt themselves to changes of the environment without explicit intervention of human. Hence, by applying the concept of context-awareness in the supplier selection process, it may be possible to significantly improve profitability of a business. To formulate a *context-aware supplier selection model* on the basis of the current static selection models may force a business to re-run the selection processes all over again on the whole set of candidates to find the new best set of suppliers regardless of the severity of the environment changes, which is usually minor. This increases computational cost that may be avoided by considering changes only and making use of the previously selected set of suppliers for subsequent sub-period.

In this work, by adapting data clustering techniques and inspiring from online algorithms, we have proposed a fresh supplier selection model that adapts the best set of suppliers with the changes of the environment while minimizes selection costs. Numerical evaluations confirm applicability of the proposed procedure and prove that, compared with static selection models, more optimal solution may be obtained while some costs savings are secured.

ACKNOWLEDGEMENTS

The authors would like to acknowledge Iran Telecommunication Research Center (ITRC) for supporting this work.

REFERENCES

[1] N. Aissaoui, M. Haouari, and E. Hassini., "Supplier Selection and Order Lot Sizing Modeling: A Review", *Computers & Operations Research*, vol. 34, pp. 3516-3540, 2007.

[2] M. Weiser, and J.S. Brown, "The Coming Age of Calm Technology", Xeroc PARC, 1996.

[3] M. Baldauf, S. Dustdar, and F. Rosenberg, "A Survey on Context-Aware Systems", *International Journal of Ad Hoc Ubiquitous Computing*, vol. 2, no. 4, pp. 263-277, 2007.

[4] A. Kalai, and S. Vempala, "Efficient Algorithms for Online Decision Problems", *Journal of Computer and System Sciences*, vol. 71, pp. 291-307, 2005.

[5] T. Levina, Y. Levin, J. McGill, and M. Nediak, "Linear Programming with Online Learning", *Operation Research Letters*, vol. 35, pp. 612-618, 2007.

[6] V. Zinkevich, "Online Convex Programming and Generalized Infinitesimal Gradient Ascent", in *Proc. 20th International Conference on Machine Learning*, 2003, pp. 928-936.

[7] R. Xu, and D. Wunsch, "Survey of Clustering Algorithms", *IEEE Transactions on Neural Networks*, vol. 16, no. 3, pp. 645-678, 2005.

[8] A.K. Jain, M.N. Murty, and P.J. Flynn, "Data Clustering: A Review", *ACM Computing Surveys*, vol. 31, no. 3, pp. 264-323, 1999.

[9] M. Dittenbach, D. Merkl, and A. Rauber, "The Growing Hierarchical Self-Organizing Map", in *Proc. International Joint Conference on Neural Networks*, Como, Italy, 2000, pp. 15-19.

[10] J. Beringer, and E. Hüllermeier, "Online Clustering of Parallel Data Streams", *Data & Knowledge Engineering*, vol. 58, pp. 180-204, 2006.

[11] D. Chakrabarti, R. Kumar, and A. Tomkins, "Evolutionary Clustering", in *Proc. 12th ACM SIGKDD international conference on Knowledge discovery and data mining*, Philadelphia, PA, USA, 2006, pp. 554 - 560.

[12] Y. Huang, S. Liu, and Y. Wang, "Online Detecting and Tracking of the Evolution of User Communities", in *Proc. of the Third International Conference on Natural Computation*, 2007, pp. 681-685.

[13] Y. Wang, S. Liu, J. Feng, "Mining Naturally Smooth Evolution of Clusters from Dynamic Data", in *Proc. of SIAM Conf. on Data Mining*, Minneapolis, Minnesota, 2007, pp. 125-134.

[14] G.H. Hong, S.C. Park, D.S. Jang, and H.M. Rho, "An Effective Supplier Selection Method for Constructing a Competitive Supply-Relationship", *Expert Systems with Applications*, vol. 28, pp. 629-639, 2005.

[15] F. Fayazbakhsh, and M. Razzazi, "Coordination of a Multi-Commodity Supply Chain with Multiple Members using Flow Networks", in *Proc. Second International Conference on Digital Society*, Sainte Luce, Martinique, 2008, pp. 25-30.