An Integrated Mathematical Programming Approach and Artificial Intelligence Techniques for Optimizing Learning Elements in E-Learning Based Educational Systems

Iraj Mahdavi, Hamed Fazlollahtabar, Narges Yousefpoor

Abstract—The increasing use of an internet improved internet technologies as well as web-based applications. Also, increasing effectiveness of the e-Learning has become one of the most practically and theoretically important issues in both educational engineering and information system fields. The online training environment enables learners to undertake customized training at any time and any place. Therefore the costs and benefits in implementation is of significance. There is a competence set consisting of ideas, knowledge, information, and skills for solving a decision problem. In order to effectively acquire the needed skills in the competence set to solve the problem, finding appropriate learning sequences of the needed skills for decision makers should be taken into account. This paper concerns with identification of varied cost elements in e-learning educational system and optimization by the means of mathematical programming. Then an effective method to estimate the learning cost between any two skills by using the grey relational analysis and a radial basis function network is proposed.

Keywords—E-learning systems; Mathematical programming; Grey relational analysis (GRA); Learning sequences; Radial basis function network (RBFN)

I. INTRODUCTION

E-Learning system is an internet based service like the application system or the internet based virtual course study service [1]. This system is able to be interpreted in various ways such as “computer based, education delivery system which is provided through the Internet”, or “an educational method that is able to provide opportunities for the needed people, at the right place, with the right contents, and the right time” [2]. The e-Learning system is one of many methods of the education (the teaching and learning procedure) that allows flexible learner-centered education. It is an information system based on the World Wide Web [3]. E-Learning provides an inter-disciplinary approach to information technology and educational engineering, and an assessment of e-Learning effectiveness could also be achieved [4]. As of IT, the end user assessment, the quality of the information system, and the system’s user satisfaction could be measured. As of educational engineering, however, the learner’s academic achievement or the degrees of self-study ability could be measured. The academic achievement is an assessment of the learner’s e-Learning environment, while self-study ability is an assessment of one’s aptitude regarding his or her self-study [5]. This approach reveals the extensive and effective trends resulting from an e-Learning research [6]. Many researchers are quite divided over the various views regarding educational engineering and information systems. Many researchers are on exploratory level trying to get explanations regarding the variations of e-Learning effectiveness i.e., [7]. The tendency of educational engineering to introduce theoretical variables in order to explain e-Learning effectiveness is insufficient except for limited numbers of information systems i.e., [8]. Moreover, this approach of putting together information systems and educational engineering is rarely observed. There is a competence set consisting of ideas, knowledge, information, and skills for solving a decision problem [9]. When decision makers have acquired the needed competence set and are proficient in it, they will be comfortable and confident in making decisions [10-11]. Otherwise, they must acquire the needed competence set to solve the problem. In order to acquire a needed set of skills to cope with a decision they face, finding appropriate learning sequences of acquiring needed skills is very necessary. Skills can be roughly classified into two types: one is single skills, and the other is compound skills. A compound skill represents a collection of single skills that might be acquired by decision makers. A compound skill may facilitate the acquisition of other single skills [12-13]. For instance, the courses “marketing management”, “financial management” and “business policy” are needed single skills for obtaining a bachelor’s degree of business administration, and are denoted by $m, f$ and $b$, respectively. When both “marketing management” and “financial management” have been acquired or learnt, these two single skills form a compound skill, denoted by $m \wedge f$. Usually, whether $m$ or $f$ has been acquired or not for an undergraduate student can be determined by considering the corresponding grade that can be got at the end of a semester. It seems to be easier for us to...
learn “business policy” after the aforementioned compound skill (i.e., $m \times f$) has been already acquired in comparison with the situation when only “marketing management” or “financial management” has been acquired. It is reasonable to assume that the determination of which compound skills are useful for facilitating the acquisition of single skills can be subjectively pre-specified by users.

The total cost of a learning sequence only consisting of single skills would be above that of a learning sequence consisting of single and useful compound skills. In many methods regarding the generation of learning sequences, such as the deduction graph with an integer programming method [13], the minimum spanning table method [14] and the stage expansion method [11], the learning costs are assumed to be known. Sometimes, the cost required for learning directly from one skill to another skill is measured by time or money. But, how much money or time will be spent is too subjective to be approved for all users. It is reasonably considered that, the larger the interrelationship between two skills, the smaller is the learning cost between these two skills. In fact, the skills in the competence set are strongly interrelated [12]. Another aim of this paper is to propose a hybrid method combining a relational analysis technique, the grey relational analysis (GRA) proposed by Deng [15] and a well-known regression tool, namely a radial basis function network (RBFN) [22], to determine the learning costs between two skills. Consider the special properties of single and compound skills, the grey relational analysis (GRA) is first employed to perform relational analysis for the single skills by dealing with individual data series measured by different criteria. Subsequently, a learning cost table obtained by the GRA is used to train the RBFN. As a result, the learning cost of learning a single skill from a compound skill can be obtained by presenting an appropriate input pattern to the trained network.

II. COSTS FACTOR IDENTIFICATION FOR IMPLEMENTATION OF AN E-LEARNING SYSTEM

Table 1. Direct learning costs comparison

<table>
<thead>
<tr>
<th>Face to face system</th>
<th>E-learning system</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Wages and salaries of trainers</td>
<td>• Development</td>
</tr>
<tr>
<td>• Payments to outside vendors</td>
<td>• Purchase and/or licensing of materials</td>
</tr>
<tr>
<td>• Facilities expenses</td>
<td>• Hardware for delivery</td>
</tr>
<tr>
<td>• Development</td>
<td></td>
</tr>
<tr>
<td>• Production and distribution of materials</td>
<td></td>
</tr>
<tr>
<td>• Travel expenses</td>
<td></td>
</tr>
<tr>
<td>• Administrative and support costs</td>
<td></td>
</tr>
</tbody>
</table>

2) Indirect Costs

Indirect costs are defined as: “compensation the wages and benefits paid to learners while they are learning as well as the
overhead costs associated with both the direct and indirect costs”. A researcher points out a relationship between direct and indirect costs and in so doing signposts the cost advantage of e-learning over classroom learning: “The available evidence suggests that the indirect, compensation costs of traditional learning are typically at least as great as the direct costs. When overhead (which also applies against the direct cost) is added in, the indirect costs of traditional learning are likely to be twice the direct costs. And that is one of the major attractions of e-learning; it holds the promise of reducing all three categories of costs, but perhaps most significantly, the indirect and opportunity costs of learning. Under some circumstances, firms have estimated the total cost of e-learning to be less than half the costs of traditional learning.

3) Opportunity Costs

Opportunity costs are business opportunities lost because employees are busy learning and not available for other responsibilities. A conservative estimate of opportunity costs is that they are equal to indirect costs. They can turn out to be dramatically higher, especially when the learner works in sales and marketing.

Because e-learning (1) is so time-effective, and (2) avoids learning-related travel, it incurs much lower opportunity costs than classroom learning.

4) Fixed Costs

The fixed cost of learning includes content development, that is, intellectual property development and licenses, instructional design, studio costs and programming costs. Occasionally, e-learning fixed costs will include extending or upgrading the network. These occasional costs should not be carried by one course but spread across a number of courses that all benefit from an improved infrastructure. Fixed costs for e-learning are significantly higher than for classroom learning reflecting the higher number and value of resources required to author an e-learning course compared with a classroom course.

5) Variable Costs

The variable or marginal costs of learning delivered in a classroom are significantly impacted by the number of learners. In contrast, the variable costs of e-learning are negligible. That is because certain cost items will remain fixed for a given course delivery regardless of the number of students, up to the delivery capacity of that item. Beyond that point, an additional cost would be incurred by adding a student to the course. For example, if an instructor books a classroom that holds 15 learners, whether one learner or 15 register, the variable cost remains the same of course, indirect costs will vary. However if 20 learners register, the cost of delivering the course doubles because two classrooms are required with an instructor in each. Fixed and Variable costs per number of students are represented in Figure 3.

So long as the overall size of a workforce remains constant, e-learning is not usually subject to stepped costs. An acquisition, on the other hand, could generate stepped indirect costs for e-learning, for example, distributed content management and server upgrades. Large virtual classes can also generate stepped costs. One instructor to 20 learners is the rule of thumb for keeping virtual classes effective and interactive. With larger classes, adding one assistant instructor for each additional 20 learners maintains effectiveness and interactivity.

B) Cost Optimization for Implementation

Considering the stated costs, to make the implementation and administration of an e-learning system of education more economic an optimization is essential. In this paper this optimization is done by mathematical programming. The notations are as follows:

**Notations:**

- $C_F$: Fixed costs
- $C_V$: Variable costs
- $C_D$: Direct costs
- $C_I$: Indirect costs
- $C_O$: Opportunity costs
- $C_{Development}$: Cost of development
- $C_{Purchase}$: Cost of purchasing material
- $C_{Hardware}$: Cost of hardwares
- $T$: Time unit (Hour)
- $t'$: Period of time per month unit
- $E_1$: Initial Expense of development
- $E_2$: Expense of Purchasing materials
- $E_3$: Expense of Hardwares
- $R$: Revenue of a user per hour
- $F$: Facility payment to students
- $\phi$: User’s ability level
- $I$: Interest rate
- $X_V$: Variable cost per each user limited to B2
- $X_F$: Fixed cost per an appropriate number of user limited to B1
- $N$: Total number of users, $N=1,2,3,....N$
- $M$: Subset of users, $M=1,2,3,....S$
Number of subsets is defined in levels as follows:

\[
M = \begin{cases} 
1 & 1 \leq N \leq s \\
2 & s + 1 \leq N \leq 2s \\
3 & 2s + 1 \leq N \leq 3s \\
\vdots & \vdots \\
S & (S-1)s + 1 \leq N \leq S \times s 
\end{cases}
\]

User’s ability is needed for the educational system to provide facilities, and is defined as follows:

\[
\varphi = \begin{cases} 
0 & \text{If student is Weak} \\
1 & \text{If student is Normal} \\
2 & \text{If student is Strong} 
\end{cases}
\]

Corresponding to the stated description, the mathematical programming of our model is achieved as follows:

**Objective Function**

\[
\text{Min} Z = C_F + C_r + C_p + C_t - C_O \\
\Rightarrow \text{Min} \ Z = (M \times X_F) + (N \times X_p) + [(t' \times E_1 \times i) + (N \times X_t) + (M \times E_s)] + (F \times \varphi) - (t \times R) \\
\]

\[
S.t.
\]

\begin{align*}
1 & \leq M \leq S & (2) \\
1 & \leq N \leq S \times s & (3) \\
t' & \geq 1 & (4) \\
t & \geq 1 & (5) \\
M, N, t, t' & \in \text{integer} & (6)
\end{align*}

Equation (1) is the objective function that is the minimization of the costs. Equation (2) implies that the number of subsets is constrained between 1 and S (S depends on the number of users and the number of users in each group s). Equation (3) indicates the bounds for the number of students. Equation (4) guarantees that period of time per month unit is at least one. Equation (5) shows that the minimum time unit would be one hour. Equation (6) indicates the kind of the variables.

Regarding to the above Equations, the aim is to find the optimal values of $M, N, t, t'$, which are the decision variables. To achieve that goal, softwares such as Lingo package could be used for finding the optimal solutions. Also some sensitivity analysis might be gained for better decision making about future trends of costs.

III. DETERMINING LEARNING COSTS

Let $K$ and $f_j$ $(1 \leq p \leq K)$ denote the number of teachable single skills, which are truly needed for solving a problem by a student, and the $p$th single skill, respectively. Also, let $c(f_p, f_j)$ denote the learning cost of a student from a teacher for directly learning $f_j$ from $f_p$. It is considered that if the interrelationship that exists between two single skills, say $f_i$ and $f_j$, are much larger than another two skills, say $f_i$ and $f_k$, then it is more practical to acquire $f_j$ from $f_i$ instead of from $f_k$. In other words, $c(f_i, f_j)$ and $c(f_j, f_k)$ should reflect grades of interrelationships such that $c(f_i, f_j) < c(f_j, f_k)$. In this paper, $c(f_i, f_j)$ is interpreted to be unhelpful grades for acquiring $f_j$ from $f_i$.

In the proposed method, the learning costs between any two single skills are first obtained by the GRA.

**A) Grey relational analysis**

As mentioned above, there is an interrelationship between any two single skills. Thus, it is necessary to find the learning costs that can reflect the interrelationships. It is pertinent to treat each needed skill $f_p$ as a subsystem, and its finite output data series is evaluated by $n$ different criteria as $(f_{p_1}, f_{p_2}, \ldots, f_{p_n})$, where $n$ is the number of criteria and $f_{p_i}$ denotes the performance value for the $i$th criterion.

Hence, let $(f_{p_1}, f_{p_2}, \ldots, f_{p_n})$ denote $f_p$.

Given one reference sequence, say $f_{p_i}$ $(1 \leq p \leq K)$, and some comparative sequences, say $f_i(1 \leq i \leq K)$, we can easily obtain the grey relation between $f_{p_i}$ and $f_i$ by viewing $f_{p_i}$ as a desired goal [20]. Formally, given the reference sequence $f_{p_i}$ and the comparative sequences $f_i$ with the normalized form, the grey relational coefficient (GRC) $\xi(f_{p_i}, f_{i_j})$ between $f_{p_i}$ and $f_{i_j}$ $(1 \leq j \leq n)$ is able to be computed as follows [19,21]:

\[
\xi(f_{i_j}, f_{p_i}) = \frac{\Delta_{\text{min}} + \rho \Delta_{\text{max}}}{\Delta_{\text{ij}} + \rho \Delta_{\text{max}}},
\]

where $\rho$ is the discriminative coefficient $(0 \leq \rho \leq 1)$, and usually $\rho = 0.5$ [19]. It should be noted that the appropriate value of $\rho$ is dependent on requirements of individual applications. Moreover, $\Delta_{\text{min}} = \min_{i,j} |f_{p_i} - f_{i_j}|$, $1 \leq i \leq k$, $1 \leq j \leq n$ (8) $\Delta_{\text{max}} = \max_{i,j} |f_{p_i} - f_{i_j}|$, $1 \leq i \leq k$, $1 \leq j \leq n$ (9) $\Delta_{\text{ij}} = |f_{p_i} - f_{i_j}|$ (10) where $\cdot |$ denotes the absolute value. Clearly, $\xi(f_{i_j}, f_{p_i})$ is between zero and one. Then, the grey relational grade (GRG) denoted by $\gamma(f_i, f_{p_i})$ can be computed as follows:
\[
\gamma(f_i, f_p) = \frac{1}{n} \sum_{j=1}^{n} \xi(f_{ij}, f_{pj}), \quad (11)
\]

\[0 \leq \gamma(f_i, f_p) \leq 1 \text{ thus holds. Let } c(f_o, f_j) \text{ denote the learning cost for directly learning } f_j \text{ from } f_i.\]

\[c(f_i, f_p) = 1 - \gamma(f_i, f_p), \quad 1 \leq i, p \leq K. \quad (12)\]

It can be seen that, the larger the relationship that exists between two skills, the smaller is learning cost between these two skills. When \( c(f_o, f_j) \) has been determined by the GRA, an initial learning cost table like Table 2 can be generated, in which it can be seen that there are \( K \) single skills. In Table 2, only \( K^2 - K \) learning costs are taken into account. That is, it is not necessary to consider \( c(f_o, f_j) \) since it is impossible to learn a single skill, say \( f_p \), from \( f_o \). It should be noted that, if a learning cost exists between any two skills in a competence set, then the expansion of this competence set can be categorized as a cyclic expansion problem [12].

<table>
<thead>
<tr>
<th>single skill</th>
<th>( f_1 )</th>
<th>( f_2 )</th>
<th>( f_3 )</th>
<th>( \ldots )</th>
<th>( f_{k-1} )</th>
<th>( f_k )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( f_1 )</td>
<td>-</td>
<td>( c(f_1, f_2) )</td>
<td>( c(f_1, f_3) )</td>
<td>( \ldots )</td>
<td>( c(f_1, f_{k-1}) )</td>
<td>( c(f_1, f_k) )</td>
</tr>
<tr>
<td>( f_2 )</td>
<td>( c(f_2, f_1) )</td>
<td>-</td>
<td>( c(f_2, f_3) )</td>
<td>( \ldots )</td>
<td>( c(f_2, f_{k-1}) )</td>
<td>( c(f_2, f_k) )</td>
</tr>
<tr>
<td>( f_3 )</td>
<td>( c(f_3, f_1) )</td>
<td>( c(f_3, f_2) )</td>
<td>-</td>
<td>( \ldots )</td>
<td>( c(f_3, f_{k-1}) )</td>
<td>( c(f_3, f_k) )</td>
</tr>
<tr>
<td>( \ldots )</td>
<td>( \ldots )</td>
<td>( \ldots )</td>
<td>( \ldots )</td>
<td>( \ldots )</td>
<td>( \ldots )</td>
<td></td>
</tr>
<tr>
<td>( f_{k-1} )</td>
<td>( c(f_{k-1}, f_1) )</td>
<td>( c(f_{k-1}, f_2) )</td>
<td>( c(f_{k-1}, f_3) )</td>
<td>( \ldots )</td>
<td>-</td>
<td>( c(f_{k-1}, f_k) )</td>
</tr>
<tr>
<td>( f_k )</td>
<td>( c(f_k, f_1) )</td>
<td>( c(f_k, f_2) )</td>
<td>( c(f_k, f_3) )</td>
<td>( \ldots )</td>
<td>( c(f_k, f_{k-1}) )</td>
<td>-</td>
</tr>
</tbody>
</table>

B) A Neural Network-based Approach

1) Introduction to RBFN

RBFN have been widely used in function approximations [22-23-24]. RBFN is a structure with locally tuned and overlapping receptive fields [22]. The structure of an RBFN with \( n \) receptive field units, \( s \) hidden units and a single output is depicted in Figure 2.
In an RBFN, each hidden node has its own radial basis function, such as a Gaussian function [25]:

\[ R_i(x) = \exp \left( -\frac{\|x - u_i\|^2}{2\sigma_i^2} \right), \quad i = 1, 2, \ldots, s - 1, \quad (13) \]

where \( R_i(x) \) is the radial basis function in the \( i \)th receptive field unit, \( \sigma_i \) is the spread width, and \( u_i \) is the center of that unit and is a vector with the same dimension as \( x \). In particular, there are no connection weights between the input layer and the hidden layer. Since the \( s \)th hidden unit is a bias node, let \( R_s(x) \) be equal to 1. Furthermore, the function of the output node, which performs the weighted sum associated with each receptive field, is similar to the defuzzification in the fuzzy inference [26].

Using back-propagation update rules, a RBFN can be trained by adjusting the aforementioned parameter specifications including \( u_i = (u_{i1}, u_{i2}, u_{i3}, \ldots, u_{in}) \), \( \sigma_i \), and the connection weights between the hidden layer and the output layer, where \( 1 \leq i \leq s - 1 \) and \( 1 \leq j \leq n \). When the \( t \)th training epoch is performed, \( w_i \), \( u_{ij} \), and \( \sigma_i \) can be determined by the update rules as follows:

\[ w_i(t) = w_i(t-1) + \Delta w_i(t-1), \quad (14) \]
\[ u_{ij}(t) = u_{ij}(t-1) + \Delta u_{ij}(t-1), \quad (15) \]
\[ \sigma_i = \sigma_i(t-1) + \Delta \sigma_i(t-1), \quad (16) \]

Where,

\[ \Delta w_i(t-1) = \eta_{R1}(d(t-1) - u(t-1))R_i(x)(t-1) + \alpha_{R1}\Delta w_i(t-2), \quad (17) \]
\[ \Delta u_{ij}(t-1) = \eta_{R2}(d(t-1) - u(t-1))w_i(t-1)R_i(x)(t-1) \frac{x(t-1) - u_{ij}(t-1)}{\sigma_i^2} + \alpha_{R2}\Delta u_{ij}(t-2), \quad (18) \]
\[ \Delta \sigma_i(t-1) = \eta_{R3}(d(t-1) - u(t-1))w_i(t-1)R_i(x)(t-1) \frac{\|x(t-1) - u(t-1)\|^2}{\sigma_i^2} + \alpha_{R3}\Delta \sigma_i(t-2). \quad (19) \]

In the update rules, the learning rates (i.e., \( \eta_{R1}, \eta_{R2}, \eta_{R3} \)) and the momentum parameters (i.e., \( \alpha_{R1}, \alpha_{R2}, \alpha_{R3} \)) in the individual update rules must be pre-specified for training an RBFN. After \( t_k \) training data have been presented to the RBFN, the training procedure is terminated. Usually, we can stop the training when the mean squared error \( E_{\text{ave}} \), shown as (13), reaches below the pre-given tolerant error:

\[ E_{\text{ave}} = \frac{1}{2N} \sum_{j=1}^{N} (d_j - o_j)^2, \quad (20) \]

where \( d_j \) and \( o_j \) are the actual output and the desired output of the \( j \)th input training data, respectively, and \( N \) is the number of training data. It should be noted that, since each radial basis function and each connection weight between the hidden layer and output layer are equal to a multidimensional composite membership function of the premise part and a consequent part of a fuzzy rule, respectively, an RBFN is functionally equivalent to a zero-order Sugeno fuzzy inference system under some conditions [27].
2) Estimating learning costs
For a decision problem $E$, in order to keep the learning costs between any single skills fixed by using the GRA [28], it is not appropriate to compute the learning costs between a single and a pre-given compound skill at the same time by using the GRA. That is, the learning cost between two single skills should be the same for each user. For instance, as shown in Table 1, $e(f_i, f_j)$ in the case of only considering $f_1, f_2, \ldots, f_K$ may be not equal to that in the case of considering $f_1, f_2, \ldots, f_K$ and a compound skill, say $f_1 \wedge f_2$. Also, the determination of which compound skills are useful for facilitating the acquisition of single skills could be dependent on users [29]. For this, the estimation of the cost of learning a single skill from a compound skill is obtained by using a trained RBFN. But, the acquisition of estimated cost of learning a compound skill from a single skill is not meaningful.

From the learning cost table obtained by the GRA, $K^2 - K$ input-output pairs are used as the training patterns for a RBFN (i.e., $N = K^2 - K$). As illustrated in Figure 3, if the input training data is $(f_1, \ldots, f_n, e(f_1), f_{p1}, \ldots, f_{pm}, e(f_p))$, then its desired output is $e(f_p)$, where $e(f_i)$ and $e(f_p)$ are the synthetic performance values of $f_i$ and $f_p$, respectively. $e(f_p)$ is computed by the simple additive weighting method as follows:

$$e(f_p) = \sum_{i=1}^{n} w_i f_{p_i}, \quad (21)$$

where, $\sum_{i=1}^{n} w_i = 1$, \quad (22)

where $w_i$ is the relative weight of the $i^{th}$ criterion. Although there are many methods that can be used to assess the relative weights [30], they are not the focus of this paper. The reason for considering $e(f_p)$ is that $e(f_p)$ indicates the relative degree of importance among $K$ single skills. The network is iteratively trained until a termination condition is reached. $e(s_u, f_p)$ of acquiring $f_p$ from $s_u$, where $s_u$ is a compound skill, can be obtained from the output of the trained network by feeding $(s_u, f_p, e(s_u), f_{p1}, \ldots, f_{pm}, e(f_p))$ to the trained network. Without losing generality, if $s_u = f_i \wedge f_{i+1} \ldots \wedge f_{i+d-1} \wedge f_{i+d}$ (2 \leq i + d \leq K), then

$$s_u = f_i \times f_{(i+1)_h} \times \ldots \times f_{(i+d-1)_h} \times f_{(i+d)_h} , \quad (2 \leq h \leq n)$$

and $f_p$ does not belong to $\{ f_i, f_{i+1}, \ldots, f_{i+d}, f_{(i+d)+1} \}$. Of course, the performance of RBFN could be influenced by the number of hidden nodes. However, how many hidden nodes are necessary is generally not known.

IV. DISCUSSION AND CONCLUSIONS
The initial purpose of this paper is the cost optimization in e-learning system of education. To do that, different cost factors have been discussed. The elements of cost are substantial in implementing e-learning systems; therefore identifying them and trying to minimize them lead to advantages in enforcement of educational organizations. The approach which has applied in this paper for cost optimization is mathematical programming. By mathematical programming, the optimal values of decision variables would be achieved that are helpful tools in decision making for now and future of educational organizations.

The main aim of this paper is to propose a novel method to estimate the costs of learning of a student from one skill to the other skill by a teacher. Furthermore, each skill has its synthetic performance value by evaluating on different criteria. The GRA is first employed to derive the learning cost between two single skills. Then, a learning cost table is generated. The reason for not considering the compound skills at the same time is to keep the cost of learning one single skill from the other single skill fixed, and to consider the variability of the compound skills. In comparison with the single skills, which compound skills is more useful are more dependent on the subjective thinking and perception of users. For the future study, subsequently, a RBFN is trained by using the learning cost table to realize the interrelationship between two single skills and the corresponding learning cost. By presenting an input–output pair consisting of one compound skill and one single skill to the trained network, the corresponding cost of acquiring the single skill from the compound skill can be obtained. The advantage of the proposed method is to provide a reasonable way to estimate the learning costs by measuring the grade of the relationship between any two single skills instead of using money or time.

ACKNOWLEDGMENTS
This research has been supported by Mazandaran Telecommunication Research Center.

REFERENCES
Iraj Mahdavi is vice president of graduate studies and research at Mazandaran University of Science and technology. He has published more than seventy papers in international journals and conferences. His interests are mathematical modeling, information technology and fuzzy logic.

Hamed Fazlollahtabar is Research assistant at Mazandaran University of Science and technology. He has published more than twenty papers in international journals and conferences. His interests are mathematical optimization, economic evaluations, statistical analysis, electronic systems and artificial intelligence.

Narges Yousefpour is studying at Mazandaran University of Science and technology. She has published some papers in international conferences. Her interests are operations research and electronic learning.