Determining Optimal Cost Drivers in Activity-Based Costing by Using Artificial Neural Networks

Mansooreh Jalayeria, Tahereh Aliheidari Biosk*

a Graduated in Engineering of Agricultural Economics, and Expert in Export Development Bank of Iran, Tehran, Iran.
b Department of Economic Science and Research Branch, Islamic Azad University, Yazd, Iran.

Abstract

Despite in attention to the application of Artificial Neural Networks (ANNs) in Activity-Based Costing (ABC) until now, employing ANNs in this field may have numerous advantages. Considering the capability of ANNs for optimizing both linear and nonlinear problems, they are good candidates for solving two problems of optimal cost driver determination in ABC. As for the first one is related to the Cost Driver Optimization (CDO) and for the second one i.e., Cost Estimation Relationship (CER) problem. Hence, in this paper, a new procedure for nonlinear cost allocation and determining the optimal cost driver employing ANNs is presented. The proposed procedure is implemented in a large bank in Iran. The results of the implementation of the proposed procedure have been compared with the results of using conventional method from 1999 to 2008. In addition, the durability of the selected optimal cost drivers within this period has been discussed.

© 2014 Published by RRAMT France Ltd.

Keywords: Activity-Based Costing, Artificial Neural Networks, Banking, Cost Driver, EDBI.

1. Introduction

All the methods used for identifying, exploration and determination of the products and services costs are called Costing (Babad et al, 1993). In this sense, Activity-Based Costing (ABC) technique has been used extensively for the past decades because of its ability to overcome the problems of the traditional costing systems (Bjørnenak et al, 2002).

Under finding of Bhimani and Pigott (1992), Cooper and Kaplan (1992), Gordon and Sivester (1999), Innes (1990), Innes and Mitchell (1995), Kennedy and Affleck-Graves (2001), and Turney and Anderson (1989), Many potential advantages of the ABC systems include: (1) helping to determine non value-added activities, (2) improving the ability of managers to make decisions related to pricing, production, and investment through the provision of more accurate product and process costs, and (3) provision of the required conditions in order to improve the market share and better control and also monitoring over the costs.

Despite the mentioned advantages of this technique, it has some defects and shortcomings such as implementation complexity, lack of flexibility, and uncertainty of the selected Cost Drivers (C.Ds) as the optimal. With a view to the fact that C.Ds are factors which often have cause and effect relationships with costs, in case they

* Corresponding Author name.
are not selected appropriately, ABC technique does not function accurately and this may cause distortion in decision making (Barfield et al, 2004).

In the recent years, so many efforts have been made to reduce these shortcomings which have led to the development of a method called Time Driven Activity-Based Costing (TDABC) by Kaplan and Anderson in 2004. However, some of the ABC’s shortcomings have been eliminated in this method (e.g., Everaert and Buggeman, 2008a, 2008b; Kaplan and Anderson, 2004, 2007) but, selecting the optimal C.D.s as an important task is still of great importance (Alinezhad et al, 2013; Ratnatunga, 2012; Everaert et al, 2008b).

Previous studies imply that conventional ABC has two significant problems concerned with determination of the optimal C.D.s. One of them addresses the Cost Driver Optimization (CDO) problem due to the lack of the absolute criteria to select the optimal C.D.s. Whereas Horngren et al. (1997) described that cost functions often show nonlinear behaviour in practice and the other problem i.e. Cost Estimation Relationship (CER), is that ABC assumes the graph of total costs versus a set of C.Ds forms a straight line within the relevant range. In other side, several researchers have made efforts to overcome the problem of nonlinearity function estimation. They have showed that Artificial Neural Networks (ANNs) can model not only linear but nonlinear functions in a more appropriate and convenient manner in comparison with parametric methods such as regression (Bode, 1998; Creese, 1995; Garza, 1995, Smith, 1997; Cavalieri, 2004; Şenyiğit, 2013).

For solving the optimization problem with nonlinear and complex nature, some of researchers have suggested the heuristic algorithms such as ANNs and Genetic Algorithm (GA) (Babad et al, 1993; Levitan et al, 1996, Kim et al, 2003).

Consequently, employing ANNs is a good way for solving the CER and the CDO problems in ABC system considering the capability of ANNs for modeling the nonlinear problems and optimizing.

Hence, in this novel work, applying ANN approach, a procedure for nonlinear cost allocation and determination the optimal C.D.s is presented. In addition, employing the proposed procedure, the optimal C.D sets within a ten-year period for a large bank in Iran’ ABC system have been determined annually and these results have been compared with the results of using conventional method for this ABC system.

The rest of this paper is organized as follows: The next section considers some of the literature and a procedure for the optimal C.D.s determination employing ANN approach is explained in section 3. In section 4, ABC in the bank is described and Section 5 is devoted to the results of using the proposed procedure for in the bank’ ABC system. Ultimately, the results of the optimal C.D.s determination within a ten-year period for the bank’ ABC system are presented in section 6. Conclusions and final remarks are given in section 7.

2. Review of literature

Cost driver is a measurable and logic variable which can be used to determine the amount of the resource consumption by each activity and the activity consumption by cost objects. Actually, what is important in every costing method is the selection of the C.Ds. Since in contrast to the traditional costing methods, multiple C.Ds are applied in two stages in Activity Based Costing (Lievens, 2003). In the first stage, organizational activities are determined and the indirect costs are allocated to activity centers. This stage of the cost allocation is carried out using “resource cost drivers” which represent the activity’s usage of indirect cost and in the second stage, the costs in each activity centers are assigned to “cost objects” using “activity cost drivers”. Therefore, the activity cost consumption for each product or service is determined by activity C.Ds through the second stage (Schniederjans, 1997).

CDO problem as a critical issue in ABC systems, have been considered by several studies which conducted with various methods. In 1993, Blanchandar used the Greedy Algorithms to optimize the C.D determination process. One year later, Dekin used the logical relationships for this purpose and in 1997, Schniederjans and Garvin applied the AHP method for the optimized selection of the C.D (Schniederjans, 1997).

Kocakulah and Iekmann (2001) at South America Indiana University made a comparison between business loan profitability of a Bank using ABC and traditional costing method. It should be noted that in this study, C.D determination method is of a great importance has been examined using the T-test correlation regression coefficient.
3. Modeling with ANNs

ANNs mimic the ability of the biological neural systems in a computerized way by resorting to the learning mechanism as the basis of human behavior (Cui et al, 2008). ANNs are information-processing systems that generate output values based on specific logic. The ANNs 'learns' the governing relationships in the input and output data by modifying the weights between its nodes. In essence, a trained Neural Network (NN) model can be viewed as a function that maps input vectors onto output vectors.

ANNs can be applied to problems with non-linear nature or with too complex algorithmic solution to be found. Their ability to perform complex decision-making tasks without prior programming tasks makes ANNs more attractive and powerful than parametric approaches especially for nonlinear problems.

ANNs have at least two potential strengths over the more traditional model fitting technique such as parametric methods and multiple regression analysis. First, ANNs can detect and extract nonlinear relationships and interactions among inputs and outputs. Second, they are capable of estimating a function without requiring a mathematical description of how the output functionally depends on the inputs and other irritating assumptions of parametric methods.

The architecture of a NN model usually consists of three parts: an input layer, hidden layers and an output layer. The information contained in the input layer is mapped to the output layer through the hidden layers. Each neuron can receive its input only from the lower layer and send its output to the neurons only on the higher layer.

3.1. The procedure of NN modelling and optimizing

Determination of the optimal C.Ds using ANNs, is performed in two steps as follows: The first step is constructing the various NN models of cost function which vary in their inputs (the relevant C.Ds) and the second step, is the selection of the optimal C.Ds which are related to the most fitted model based on the performance indicators. For this purpose, a procedure consisting of these two steps is presented according to Fig. 1.

![Fig.1. The procedure of NN modeling.](image)

Phase1. Selection of NN model topologies: There are about 31 different NN topologies, which are being employed in researches at present. The most common networks are: Multilayer Perceptrons (MLP) also called Multilayer Feed forward Networks, Adaptive Resonance Theory Models (ART), Recurrent Associative Networks (RAN), and Self-Organizing Maps (SOM)( Krycha et al, 1999).
Each NN type and training algorithm combination is selected for different situations, depending on the networks’ purpose of usage. For instance, the Learning Vector Quantization (LVQ) network is particularly suitable for classification problems, while MLP trained with BP networks are suitable as black-box models of systems which the underlying relations are poorly known or nonlinear.

In this phase, the topology of the models and their training algorithms considering the nature of modeling problem are chosen. The other decision making addresses the definition of the input variables (e.g., C.Ds) and the output variables (e.g., total cost) in order to assign them to the network of each model as inputs and outputs, respectively.

Also, the way of presenting data to the network have to be determined. One way to do this is to present all available process variables as network inputs, and then let the network modify itself during training so that the connection of any insignificant variables becomes weak. Another approach is to be more selective, and introduce as inputs only those variables that are surely affecting on the process outputs. The first approach is called the “global network” while the second is termed the “focused network” (Wilcox et al, 1998).

Phase 2. Collection and preparation of training data set: In the Data collection stage, it is necessary to ensure the sufficiency and integrity of the data used to train and test the network, whereas the network performance can be directly influenced by the presented data. It is not simply possible to say how many data sets are required, because this depends on the nature of the process modeling problem and the cost of providing data.

The prepared data set is categorized randomly into two sets: Training data set and test data set. Generally, the training data should cover the whole of data variation range, unless there is a good reason to resort to stratification or data blocking. In order to enhance the model fitness, several tasks for preparing data may be performed such as: (1) data integrity check, (2) extreme data removal, (3) data scaling, and (4) data coding.

Phase 3. Constructing and fitting the NN models: The constructing of the NN model architecture includes determining few features:
- The number of layers;
- The number of neurons in each layer;
- Each layer’s transfer function;
- How the layers are connected to each other.

So the performance of a NN model depends on the fitness of the network features. For instance, too few neurons may result in under fitting, but too many neurons may yield over fitting, which means that all the training data fit well, but the NN model performance for test data is low.

The optimal configuration of each network is selected according to the training process results. So, various architectures for each network are proposed with diverse features. During the training process, representative examples of inputs and their corresponding outputs are presented to the models. Each NN model that its network trained and learned the governing relationships in the data set by modifying its weights and biases, called the fitted model. Ultimately, the optimal configuration for each network based on the user-specified error function is chosen by trial and error.

There are various training algorithms to fit NN models. The most popular algorithm in optimization and estimation applications is the standard back propagation (BP) (Nascimento et al., 2000). This algorithm is a widely used iterative optimization technique that locates the minimum of a function expressed as

$$E = \frac{1}{2} \sum_{m} (y_{dm} - y_{m})^2$$  \hspace{1cm} (1)$$

Where $y_{dm}$ is the target value of the output layer, and $y_{m}$ is the ratiocinated value of the output layer. Based on BP algorithm, during the training process, the deviation between the network output and the desired output at each presentation is computed as an error. This error, in quadratic form, was then fed back (back propagated) to the network and used for modifying the weights by a gradient descent method. Such that the error decreases with each interaction and the neural model gets closer and closer to generate the desired output.
The training process carries on while one of three user-specified conditions is met at least. These conditions consist of (1) exceeding the maximum number of epochs, (2) meeting the performance goal, and (3) decreasing of the gradient descent rate to the less than the allowable limit.

Phase 4. Model validation: In this phase, all of fitted models are validated. The validation data set is used to ensure that there is no over fitting in the final result. In order to validate the models, a data set is selected randomly from training data. When a significant over fitting has occurred, the error of validation data starts to increase and the training process is stopped.

Phase 5. Selection of performance indicators: To find out the most reliable model, several performance indicators can be used. The performance of the NN models based on their reliability is evaluated by a regression analysis between the predicted values by the NN models and the actual values. The indicators used with aim of performance evaluation for both training and test data sets are the root mean square error (RMSE) and correlation coefficient (Hosoz et al, 2007). The root mean square error is calculated by:

$$RMSE = \sqrt{\frac{1}{m} \sum_{i=1}^{m} (y_i - \hat{y}_i)^2}$$

(2)

Where \( \hat{y}_i \) is the predicted value by the NN model, \( y_i \) is actual value and \( m \) is the number of points in the data set. The absolute fraction of variance, a statistical criterion that can be applied to multiple regression analysis, is calculated by:

$$R^2 = 1 - \frac{\sum_{i=1}^{m} (y_i - \hat{y}_i)^2}{\sum_{i=1}^{m} (y_i - \bar{y})^2}$$

(3)

The absolute fraction of variance ranges between zero and one. Ideally, \( R^2 \) should be close to one, whereas a poor fit results in a value near zero.

Phase 6. Model performance evaluation: By termination of process training and validation, the networks are ready for prediction. So, the input vectors from the separate test data are introduced to the trained network and the responses of the network, i.e., the predicted output, are compared with the actual ones using the performance indicators. The most reliable model for the total cost function is chosen in respect to the evaluation and training results. Hence, the C.Ds related to the chosen model is the optimal C.Ds.

4. Activity-Based Costing in the bank

One of the most significant capabilities of the finance institutions such as Banks is the possession of rich and huge databases. Taking advantage of this privilege ensures the observance of the rule of benefit to cost excess in ABC method. Hence, the Activity Based Costing for the bank has been selected as the case of this study. The proposed model for banking service costing as a pilot is illustrated in Fig.2.
According to Fig. 2, the allocation of the costs from resources to services is carried out in two levels. In level one, total indirect cost via two stages is allocated to first level cost objects by C.Ds and in level two; service cost at each branch is calculated with the same structure, individually. First level incorporates; 12 groups of resources (e.g., Depreciation cost, Hardware and software costs, and etc), seven groups of activities (e.g., Credit activity, Legal affairs, and etc) and 29 cost objects (e.g., Branches of bank) and second level consists of 12 groups of resources, seven groups of activities, and 19 cost objects (e.g., Letter of credit).

As mentioned earlier in Section 1, determining the optimal C.Ds is one of the most significant factors for the successful implementation of ABC. A conventional way for selecting appropriate C.Ds from a larger set of primal candidate C.Ds is using human judgments, supported occasionally by parametric methods or an analysis using simple accounting techniques (Hosoz et al, 2007; Rupp et al 1995).

Considering disagreement over the optimal C.Ds for the Hardware and software costs, it was selected as a case. The proposed C.Ds relevant to the Hardware and software cost are as follows: Total of records ($r_T$), Total of entries ($T_e$), Total of cash records ($T_{cr}$), Total of non cash records ($T_{nr}$), Total of entry of current account ($T_{ea}$). “$T_i$” is defined as the summation of “$T_{cr}$” and “$T_{nr}$” in order that as a hint can incorporate some side information into the learning process in addition to the normal information(Sukthomya et al, 2005). Each one of C.Ds or a combination of them may be defined as input for NN models.

5. The NN modeling of the bank `ABC system

Export Development Bank of Iran uses the Oracle database for the storage of daily transaction data. The critical steps for the model implementation of Activity-Based Costing are derivation, transfer, transformation, and data integrating.

In this study, the clustering analysis has been used to eliminate the noises of data which is caused due to the diversity of branches or some tasks such as closing of accounts in the last month of year.

The clustering data consisted of 348 data vectors for the year 2007 were gathered from the Oracle database. As, each vector consisted of five data elements related to the values of the proposed C.Ds and one data element relevant to their corresponding Hardware and software cost. The 348 data vectors were randomly split into 300 training data vectors and 48 test data vectors.

Five models for the mentioned cost function with different inputs were proposed (Further information is presented in Table 1). The topologies of five NN models were selected MLP and their networks were trained with
BP algorithm based on Levenberg-Marquardt (Rumelhart et al, 2013). In order to choose the optimal configuration for each network, a number of different network configurations, consisting of one to three hidden layers and different number of neurons in hidden layers with various transfer functions were considered.

The training process was run under the MATLAB environment so a minimum of the user-specification error function i.e., RMSE of 1e-001 was reached while the number of epochs is less than 1000 and grad rate is more than 1e-010.

The optimal configuration of each training network was selected based on the degree of model fitness which is demonstrated by the value of error function. As Fig 3 indicates when the number of neurons in first and second hidden layers were equal seven and ten numbers, respectively, it was the most of the training RMSE of network for model C.

It is noticeable that the collected data vectors were customized for each model by selecting some data element considering the related C.Ds. Table 1 shows the features of the configuration for networks of five NN models and the result of training process relevant to models are demonstrated in Table 2.

![Fig 3. Effect of number of neurons in first and second hidden layers on training RMSE for model C](image)

Table 1. The features of the configuration for networks of five proposed models

<table>
<thead>
<tr>
<th>Models</th>
<th>Selected C.Ds</th>
<th>MLP structure</th>
<th>Transfer function</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model A</td>
<td>$T_r$</td>
<td>1-4-6-1</td>
<td>Tansig-logsig- purline</td>
</tr>
<tr>
<td>Model B</td>
<td>$T_r$</td>
<td>1-4-1</td>
<td>logsig- purline</td>
</tr>
<tr>
<td>Model C</td>
<td>$T_r, T_{cr}$</td>
<td>2-7-10-1</td>
<td>purline-tansig- purline</td>
</tr>
<tr>
<td>Model D</td>
<td>$T_r, T_{nr}$</td>
<td>2-8-8-1</td>
<td>purline-logsig- purline</td>
</tr>
<tr>
<td>Model E</td>
<td>$T_r, T_{ea}$</td>
<td>2-8-9-1</td>
<td>purline-logsig- purline</td>
</tr>
</tbody>
</table>

So, in order to find the most reliable model, all 48 test data vectors were used for model evaluations. The test data vectors were customized and presented to each model. The results of the model evaluation using performance indicators for five models are provided in Table 2.
Table 2. The results of model evaluation using performance indicators for five models

<table>
<thead>
<tr>
<th>Models</th>
<th>RMSE Training</th>
<th>RMSE Testing</th>
<th>$R^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model A</td>
<td>0.0826</td>
<td>0.0868</td>
<td>0.2794</td>
</tr>
<tr>
<td>Model B</td>
<td>0.0799</td>
<td>0.0803</td>
<td>0.2988</td>
</tr>
<tr>
<td>Model C</td>
<td>0.0478</td>
<td>0.0572</td>
<td>0.3438</td>
</tr>
<tr>
<td>Model D</td>
<td>0.0789</td>
<td>0.0824</td>
<td>0.3068</td>
</tr>
<tr>
<td>Model E</td>
<td>0.0646</td>
<td>0.0717</td>
<td>0.3112</td>
</tr>
</tbody>
</table>

According to the results of Table 2, “$T_c , T_{cr}$” is chosen as the optimal C.Ds for year 2007 because the relevant model i.e., the model C is the most reliable one, considering its minimum value for RMSE (both training and testing) and also its maximum value for $R^2$.

6. Supplement Study

For the completion purpose, employing the proposed procedure, the optimal C.Ds for Hardware and software costs from 1999 to 2008 have been determined annually as the results is presented in Table 3. Note that used data for determining the optimal C.Ds for each year is cumulative data of last years.

According to the results, “$T_c$” is selected as the yearly optimal C.D set for 1999 and 2000, “$T_r , T_{nr}$” for 2001 to 2004, and “$T_r , T_{cr}$” for 2005 to 2008, however, the yearly conventional selected C.Ds during this period was “$T_r$.”

The results of Table 3 imply the instability of the optimal C.Ds within a ten-year period. Investigation of the total indirect costs revealed that in 2001 and 2005, the total indirect costs have increased considerably due to an increase in Human resource costs of 20% and in Branch establishing cost of 25%, respectively.

Table 3. The results of determination of optimal C.Ds from 1999 to 2008 for the bank’ ABC system

<table>
<thead>
<tr>
<th>Year</th>
<th>1999</th>
<th>2000</th>
<th>2001</th>
<th>2002</th>
<th>2003</th>
<th>2004</th>
<th>2005</th>
<th>2006</th>
<th>2007</th>
<th>2008</th>
</tr>
</thead>
<tbody>
<tr>
<td># of data vectors</td>
<td>324</td>
<td>648</td>
<td>972</td>
<td>1296</td>
<td>1608</td>
<td>1920</td>
<td>2232</td>
<td>2496</td>
<td>2760</td>
<td>3000</td>
</tr>
<tr>
<td>Selected C.Ds</td>
<td>$T_c$</td>
<td>$T_c$</td>
<td>$T_r, T_{nr}$</td>
<td>$T_r, T_{nr}$</td>
<td>$T_r, T_{nr}$</td>
<td>$T_r, T_{cr}$</td>
<td>$T_r, T_{cr}$</td>
<td>$T_r, T_{cr}$</td>
<td>$T_r, T_{cr}$</td>
<td></td>
</tr>
<tr>
<td>Optimal NN model</td>
<td>B</td>
<td>B</td>
<td>D</td>
<td>D</td>
<td>D</td>
<td>D</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
</tr>
</tbody>
</table>

As a major portion of total indirect costs is Hardware and software cost, the proper selection of optimal C.Ds can have a considerable effect on cost price of services and consequently on calculated interest of services.

For instance, selecting the $T_r$ as the optimal C.Ds related to year 2008, results in introducing the long-term loans again real estate as the most profitable service. However, if the proposed optimal C.Ds i.e., “$T_r, T_{cr}$” set be selected as the optimal C.Ds for this year, the short-term loans again real estate.

7. Conclusions and Final Remarks

In this paper, a procedure applying NN approach for determining the optimal C.Ds was proposed and through a case i.e., the bank’ ABC system was implemented. For this aim, the model of ABC in the bank was explained and Hardware and software cost group was selected as representative of cost groups. Based on expert judgments, five NN models for Hardware and software cost function which differ in their inputs were developed.

Considering the results of reliability evaluation for these five NN models, the most fitted model and consequently, the optimal C.Ds for Hardware and software cost were selected. However, an absolute criterion for performance evaluation of optimal C.Ds selection ways is not presented yet, but learning algorithms such as ANNs can be more reliable than conventional methods which rely on human judgments.
For supplement study, the annually optimal C.Ds for Hardware and software cost using the proposed procedure was determined from 1999 to 2008. These results demonstrate in 2001 and 2005 due to a considerably increase in total indirect cost, the annually optimal C.D sets have been substituted.

Therefore, the significant issue that rarely has been noticed in the using of the optimal C.Ds is investigating the suitability of present optimal C.Ds when the total indirect cost has increased. Our analysis tends to suggest that in situations concerning the considerable increase in the total indirect cost, the suitability of present optimal C.Ds be investigated.

According to the application scope of the ANNs algorithm and on the other hand, need for the improvement of conventional ABC methods, and also the commitment for determining the optimal C.D sets at firms and organizations may lead to opportunities for further research and development.

References


