Using Innovative Technology in QFD to Improve Marketing Quality

CHIH-HUNG HSU[1], SHIH -YUAN WANG[2], LIANG-TZUNG LIN[3]
[123]Department of Industrial Engineering and Management, Hsiuping Institute of Technology,
11, Gungye Road, Dali, Taichung 41249, Taiwan, R.O.C.

Abstract: - Software design is a high value-added business, so the sequence decision of design requirements is a critical issue for the satisfaction of customer needs for improving marketing quality. On the other hand, data mining has been successfully applied in many fields. However, little research has been done in the quality function development of identifying the future sequence decision of design requirements, using data mining and grey theory. This study applied a time series-based data mining cycle and grey relational analysis, using sales questionnarie database, to identify the future sequence decision of design requirements for software designers. Certain advantages may be observed when the future sequence decision of design requirements was identified, using the data mining cycle and grey relational analysis. The future design requirement of each customer was found and satisfied in advance. The results of this study can provide an effective procedure of identifying the future design requirements to satisfy customer needs and enhance the marketing quality and competitiveness of software company in the marketplace.

Key-Words: - Data mining; Grey relational analysis; Quality function deployment; Marketing quality;

1 Introduction

Software design is a high value-added business, so the sequence decision of design requirements for software company is a critical issue for the satisfaction of customer needs. The software design includes various systems, such as large-scale on-line systems for the different businesses, basic software operating systems, application systems and others. Software satisfaction has traditionally been defined in terms of fitness for use [1]. A software product is believed to be fit for use if it performs to some level of user satisfaction, in terms of functionality and continuous operation [2]. In order to fulfill high customer satisfaction, the sequence decision of design requirements should be identified. Requirements analysis of a software system is always considered as one of very important steps in the software development procedure [3].

Quality Function Deployment (QFD) is a Japanese development and design technology. QFD has been widely applied to achieve customer needs and improve customer satisfaction in many fields. Some researchers defined QFD as follows: “This technology focuses and coordinates skills within an organization, first to design, then to manufacture and market products that customers want to purchase and will continue to purchase [4].” Some companies have claimed great success with QFD. Proponents assert that QFD has helped them reduce production costs, design time and cost; increase customer satisfaction and marketing quality [5, 6].

QFD was first introduced by Akao in 1972 at Mitsubishi’s Kobe shipyard site, and then Toyota and its suppliers developed it further for a rust prevention study [7]. After the concept of QFD was introduced in the US through auto manufacturers and parts suppliers [8], many US firms, such as AT&T, Digital Equipment, Ford, GM, Hewlett-Packard, Procter & Gamble, and Raychem, applied QFD to improving communication, product development [5, 9].

Because the voice of the customer is essential, the House of Quality (HOQ) converts each customer need (CN) into one or more quality characteristics (design requirements; DRs) in the first phase of QFD. The main goal of HOQ is to identify customer needs and weights for the product (WHATs) and then to convert these needs into design requirements (HOWs). It has a great benefit that predicting the future sequence decision of design requirements for software designers could help companies provide better products, enhance their competitiveness in marketplace, increase customer satisfaction. On the other hand, the application domain of data mining is quite broad and plausible in health insurance [10], surface roughness prediction [11], biomedical technology [12], risk prediction [13], human resource management [14], semiconductor manufacturing [15], production schedule [16], marketing [17] and others. However, little research has also been applied to the future sequence decision of design requirements in QFD using data mining cycle. This study applied a time series-based data mining cycle and grey relational analysis, using sales questionnarie database, to identify the future sequence decision of design requirements for software designers. By applying the proposed approach, the future sequence decision of design requirements can be found to enhance marketing quality and competitiveness in the software marketplace.

2 Time Series-based Data Mining and Grey Relational Analysis

Berry and Linoff defined data mining as the analysis of huge amounts of data by automatic or semi-automatic means, in order to identify significant patterns or rules [18]. One of the most important data mining techniques is time series analysis. Time series data often arise when monitoring industrial processes or tracking corporate business metrics [19].

Forecasting is a important tool of time series analysis. Forecasting can do for just that - if a time series has behaved a certain way in the past, the future behavior can be predicted within certain confidence limits by building models [20, 21]. One of the most important forecasting techniques is exponential smoothing analysis for time series analysis. Forecasts generated with this method are a weighted average of the past values of the variable. The weights decline for older observations. The rationale is that more recent observations are more inuential than older observations. The forecast for period t+1 calculated in period t is called Ft+1. Therefore, Ft is the forecast for period t calculated in period t-1. The forecast for period t+1 is,

$$F_{t+1} = \alpha A_t + (1-\alpha) F_t$$

which represents a weighted average of the actual alue (At) and the forecast (Ft) of the actual value (calculated at t-1). The higher the value of alpha the more weight is given to current values [22].

On the other hand, one of the most important grey theory techniques is grey relational analysis. Let the original reference sequence and comparability sequences be represented as $x_0^{(O)}(k)$ and $x_i^{(O)}(k)$, i = 1, 2, . . .,m; k = 1, 2, . . ., n, respectively.

Data preprocessing is normally required since the range and unit in one data sequence may differ from the others. Data preprocessing is also necessary when the sequence scatter range is too large, or when the directions of the target in the sequences are different. Data preprocessing is a process of transferring the original sequence to a comparable sequence. Depending on the characteristics of data sequence, there are various methodologies of data preprocessing [23, 24, 25, 26] available for the grey relational analysis.

If the target value of original sequence is infinite, then it has a characteristic of “the-larger-the-better”. The original sequence can be normalized as follows:

$$x_i^*(k) = \frac{x_i^{(O)}(k) - \min x_i^{(O)}(k)}{\max x_i^{(O)}(k) - \min x_i^{(O)}(k)}$$

when the-smaller-the-better is a characteristic of the original sequence, then the original sequence should be normalized as follows:

$$x_i^*(k) = \frac{\max x_i^{(O)}(k) - x_i^{(O)}(k)}{\max x_i^{(O)}(k) - \min x_i^{(O)}(k)}$$

However, if there is a definite target value to be achieved, then the original sequence will be normalized in the form:

$$x_i^*(k) = 1 - \frac{|x_i^{(O)}(k) - \mathrm{OB}|}{\max\{\max x_i^{(O)}(k) - \mathrm{OB}, \mathrm{OB} - \min x_i^{(O)}(k)\}}$$

Or, the original sequence can be simply normalized by the most basic methodology, i.e. let the values of original sequence are divided by the first value of the sequence:

$$x_i^*(k) = \frac{x_i^{(O)}(k)}{x_i^{(O)}(1)}$$

where $x_i^{(O)}(k)$ is the original sequence, $x_i^*(k)$ the sequence after the data preprocessing, max $x_i^{(O)}(k)$ the largest value of $x_i^{(O)}(k)$ and min $x_i^{(O)}(k)$ the smallest value of $x_i^{(O)}(k)$.

After data preprocessing is carried out, a grey relational coefficient can be calculated with the preprocessed sequences. The grey relational coefficient [23, 24, 25, 26] is defined as follows:

$$\gamma(x_0^*(k), x_i^*(k)) = \frac{\Delta_{\min} + \zeta \Delta_{\max}}{\Delta_{0i}(k) + \zeta \Delta_{\max}},$$
$$0 < \gamma(x_0^*(k), x_i^*(k)) \le 1$$

where $\Delta_{0i}(k)$ is the deviation sequence of the reference sequence $x_0^*(k)$ and the comparability sequence.

$x_i^*(k)$, i.e.

$$\Delta_{0i}(k) = |x_0^*(k) - x_i^*(k)|,$$
$$\Delta_{\max} = \max_{\forall j \in i} \max_{\forall k} |x_0^*(k) - x_j^*(k)|,$$
$$\Delta_{\min} = \min_{\forall j \in i} \min_{\forall k} |x_0^*(k) - x_j^*(k)|$$

ζ: distinguishing coefficient, $\zeta \in [0, 1]$.

The grey relational grade [23, 24, 25, 26] is a weighting-sum of the grey relational coefficient. It is defined as follows:

$$\gamma(x_0^*, x_i^*) = \sum_{k=1}^{n} \beta_k \gamma(x_0^*(k), x_i^*(k)), \quad \sum_{k=1}^{n} \beta_k = 1$$

Here, the grey relational grade $\gamma(x_0^*, x_i^*)$ represents the level of correlation between the reference sequence and the comparability sequence.

Thus, this study proposed a time series-based data mining cycle and grey relational analysis, in order to mine the future sequence decision of design requirements in QFD for satisfying customer needs.

3 The Data Mining Procedure

This study uses data mining cycle and grey relational analysis to identify the future design requirements with each respective step closely involved. The data mining cycle involves a series of activities, from defining the problem to evaluating and applying the results. The previous steps can be served as the baseline reference for the next step, and the steps for identifying the future design requirements in QFD for software design are described below.

3.1 Defining the problem for data mining

Owing to unknown weights of future customer needs for the future sequence decision of design requirements, a large questionnaire database was created for a professional software design company, based on many sales questionnaries, measured in each of the 2500 customer needs about first impressions, according to four period questionnaries; this resulted in a huge amount of data.

The goal of this study was to explore and analyze a huge amount of data, by employing a time series-based data mining cycle and grey relational analysis in QFD, so as to identify the weights within customer questionnaries in each period. Based on these the weights of customer needs, the future sequence decision of design requirements may be discovered and the results can be encouraged and beneficial for sofeware designers and company.

3.2 Data preparation and analysis

The data was processed, and analyzed, in order to enhance the efficiency and ensure the accuracy of the results [27]. Before mining the data, it had to be checked and processed, with all abnormal or missing data being separated out. As a result, of the 10,000 questionnaries, 44, which had missing or abnormal data, were deleted; this left a total of 9956 valid questionnaries regarding the customer needs of first impressions. There are six customer needs and five deign requirements for each questionnaire, as shown in Table 1 and Table 2. To simplify the space, each customer need is denoted as CN, and the notation of DR is used for each design requirement.

Table 1. Definitions of customer needs

	Customer Needs
CN1	Looks good
CN2	Easy to learn initially
CN3	Easy to use after learned
CN4	Easy to install
CN5	Well packaged
CN6	Good price

Table 2. Definitions of deign requirements

	Deign requirements
DR1	Preliminary screen designs
DR2	Interaction time for common operations
DR3	Estimated time for novice to learn
DR4	Estimated time for novice to inatall
DR5	Packaging layout

3.3 Data mining by time series analysis

The weight for each customer need is evaluated by a 1–10 scale, where a customer need with a higher value is more important. The weights of four period for each customer need are periodically computed in Table 3.

Table 3. Weights of customer needs

	Period 1	Period 2	Period 3	Period 4
CN1	5.1	5.9	6.5	6.7
CN2	4.8	6.3	6.5	7.2
CN3	5.2	5.5	6.4	7.8
CN4	5.4	4.8	5.2	5.4
CN5	3.2	3.4	3.3	3.8
CN6	7.1	7.5	7.8	8.3

On the other hand, it is essential for software company to reflect customer needs by corporate language and then fulfil those deign requirements to satisfy customer needs. When customer needs are translated by HOWs, the software company has to check the relationship between WHATs and HOWs.

QFD represent the respective strong (with a weight of 9), moderate (with a weight of 3), and weak relationship (with a weight of 1), while the blank is zero. Taking period 1 as an example, the matrix relationship between customer needs and deign requirements was shown in Table 4.

Table 4. The HOQ of period 1

	Period weight	DR1	DR2	DR3	DR4	DR5
CN1	5.1	9				1
CN2	4.8	1		3		
CN3	5.2	3	1	9	3	
CN4	5.4	3	9	1	1	
CN5	3.2		1	3	9	9
CN6	7.1					3

Through checking the relationship between WHATs and HOWs, the matrix relationship between customer needs and deign requirements were determined. Subsequently, data mining was undertaken, using a time series-based data mining cycle, to mine the weights and determine the trend of each customer need for the next period.

According to the data mining cycle, the weights of customer needs in the next period (period 5) would be estimated as shown in Table 5. As shown, the weight of CN1 in the period 5 (Predicted) is 6.6; thus, these predicted weights of customer needs were chosen for the next processing by using grey relational analysis.

Table 5. Predicted weights for the customer needs

	Period 1	Period 2	Period 3	Period 4	Predicted
CN1	5.1	5.9	6.5	6.7	6.6
CN2	4.8	6.3	6.5	7.2	6.9
CN3	5.2	5.5	6.4	7.8	7.3
CN4	5.4	4.8	5.2	5.4	5.3
CN5	3.2	3.4	3.3	3.8	3.7
CN6	7.1	7.5	7.8	8.3	8.1

3.4 Grey relational analysis

Through determining the future weights of customer needs, the HOQ between design requirements and future weights of customer needs were shown in Table 6.

Table 6. The HOQ between DRs and CRs

	Period weight	DR1	DR2	DR3	DR4	DR5
CN1	6.6	9				1
CN2	6.9	1		3		
CN3	7.3	3	1	9	3	
CN4	5.3	3	9	1	1	
CN5	3.7		1	3	9	9
CN6	8.1					3

Owing to the target value of original sequence is infinite, this study has a characteristic of "the-larger-the-better". Furthermore, the original sequence would be normalized as shown in Table 7.

Table 7. The normalization of original sequence

DR1	DR2	DR3	DR4	DR5
1	0	0	0	0.11
0.33	0	1	0	0
0.33	0.11	1	0.33	0
0.33	1	0.11	0.11	0
0	0.11	0.33	1	1
0	0	0	0	1

After data preprocessing was carried out, a grey relational coefficient would be calculated with the preprocessed sequences. The grey relational coefficient was defined as shown in Table 8.

Table 8. The grey relational coefficient

DR1	DR2	DR3	DR4	DR5
0	1	1	1	0.89
0.67	1	0	1	1
0.67	0.89	0	0.67	1
0.67	0	0.89	0.89	1
1	0.89	0.67	0	0
1	1	1	1	0

The grey relational grade is a weighting-sum of the grey relational coefficient. It was computed as shown in Table 9. Furthermore, the sequence of grey relational grade for the future sequence decision of design requirements was shown in Table 10.

Table 9. The grey relational grade

DR1	DR2	DR3	DR4	DR5
1	0.5	0.5	0.5	0.53
0.60	0.5	1	0.5	0.5
0.60	0.53	1	0.60	0.5
0.60	1	0.53	0.53	0.5
0.5	0.53	0.60	1	1
0.5	0.5	0.5	0.5	1

Table 10. The sequence of grey relational grade

	Grey relational grade	Sequence
DR1	0.6385	3
DR2	0.5786	4
DR3	0.7013	1
DR4	0.5723	5
DR5	0.6609	2

3.5 Evaluation and Application of Results

The future sequence decision of each design requirement to satisfy customer future needs was analysed in Table 10. According to the future sequence decision, DR3 and DR5 should be closely noticed since they have front sequence and could become the most important design requirement to satisfy customer needs in the future. By contrast, DR4 has a back sequence and could become the least important design requirement in the future. Different sequence of design requirement should be considered differently for software designers. The future sequence decision of design requirements can be designed and planned to satisfy customer future needs in advance.

In addition, the data mining cycle emphasizes the dataset information by repeating interaction activities. Since customer needs can change rapidly, the questionnarie database of customer needs must be updated continually; therefore, the time series-based data mining cycle and grey relational analysis, proposed in this study, will continually update the database and continually identify the future customer needs for the future sequence decision of design requirements for software designers. These revised deign requirements will exactly satisfy with customer needs, allowing software designers access to the latest customer needs, thus facilitating advanced software design.

4 Conclusion

The application domain of data mining is quite broad. However, little research has also been applied to the future sequence decision of design requirements in QFD using data mining cycle. This study applied a time series-based data mining cycle and grey relational analysis, using sales questionnarie database, to identify the future sequence decision of design requirements for software designers. Certain advantages may be observed when the future sequence decision of design requirements were identified, using the proposed approach. The future design requirement of each customer was found and satisfied in advance. The results of this study can provide an effective procedure of identifying the future design requirements to satisfy customer needs and facilitate the marketing quality in the marketplace for software company.

Acknowledgments

The authors thank peer reviews of the independent qualified referees for their comments.

References:

[1] Dunn RH, Software quality assurance: A management perspective. *Quality Progress*, 21(7), 1988, pp.52-56.

[2] O'Brien DH, Software quality starts with the customer. *Quality Progress*, 30(6), 1991, pp.22-24.

[3] Liu XF, A quantitative approach for assessing the priorities of software quality requirements. *The Journal of Systems and Software*, 42, 1998, pp.105-113.

[4] Hauser JR and Clausing D, The house of quality. *Harvard Business Review*, 66(3), 1988, pp.63-73.

[5] Griffin A, Evaluating QFD's use in US firms as a process for developing products. *Journal of Product Innovation Management*, 9, 1992, pp.171-187.

[6] Wheelwright SC and Clark KB, *Revolutionizing product development: Quantum leaps in speed, efficiency, and quality*. New York: Free Press, 1992.

[7] Wasserman GS, On how to prioritise quality characteristics during the QFD planning process. *IIE Transaction*, 25(3), 1993, pp.59-65.

[8] Sullivan LP, Quality function deployment (QFD): the beginning, the end, and the problem in between. A collection of presentations and QFD case studies. American Supplier Institute,

1987.

[9] Ansari A and Modaress B, Quality function deployment: the role of suppliers. *International Journal of Purchase Management*, 30(4), 1994, pp.28-35.

[10] Chas YM, Ho SH, Cho KW, Lee DH and Ji SH, Data mining approach to policy analysis in a health insurance domain, *International Journal of Medical Informatics*, 62, 2001, pp.103-111.

[11] Feng CX and Wang X, Development of empirical models for surface roughness prediction in finish turning, *International Journal of Advanced Manufacturing Technology*, 20, 2002, pp.348-356.

[12] Maddour M and Elloumi M, A data mining approach based on machine learning techniques to classify biological sequences, *Knowledge-Based Systems*, 15, 2002, pp.217-223.

[13] Becerra-Fernandez, I., Zanakis SH and alczak S, Knowledge discovery techniques for predicting country investment risk. *Computer and Industrial Engineering*, 43, 2002, pp.787-800.

[14] Min H and Emam A, Developing the profiles of truck DRivers for their successful recruiDRent and retention: a data mining approach, *International Journal of Physical Distribution & Logistics Management*, 33, 2003, pp.149-162.

[15] Chien CF, Hsiao A and Wang I, Constructing semiconductor manufacturing performance indexes and applying data mining for manufacturing data analysis, *Journal of the Chinese Institute of Industrial Engineers*, 21, 2004, pp.313-327.

[16] Sha DY and Liu CH, Using data mining for due date assignment in a dynamic job shop environment, *International Journal of Advanced Manufacturing Technology*, 25, 2005, pp.1164-1174.

[17] Wong KW, Zhou S, Yang Q and Yeung MS, Mining customer value: from association rules to direct marketing, *Data Mining and Knowledge Discovery*, 11, 2005, pp.57-79.

[18] Berry M and Linoff G, *Data Mining Techniques: for Marketing, Sales, and Customer Support*, Wiley, New York, 1997.

[19] Giudici P, A*pplied Data Mining: Statistical Methods for Business and Industry*, Wiley, England, 2003.

[20] Chatfield C, *The analysis of time series – an introduction*. 5th. Chapman and Hall, London, UK, 1996.

[21] Box GEP, Jenkins GM, and Reinsel GC, *Time series analysis – Forecasting and control*. 3rd ed. Prentice Hall, Englewood Cliffs, NJ, USA, 1994.

[22] Harvey HC, *Time series models*. Halstead Press, New York, NY, USA, 1981.

[23] Deng JL, Grey System. 1, 1989, pp.1-24.

[24] Deng JL, *A course on grey system theory*, Huazhong University of Science and Technology Press, Wuhan, 1990.

[25] Deng JL, *The essential methods of grey systems*, Huazhong University of Science and Technology Press, Wuhan, 1992.

[26] Fung CP, Manufacturing process optimization for wear property of fiber-reinforced polybutylene terephthalate composites with grey relational analysis. *WEAR*, 203, 2003, pp.298-306.

[27] Pyle D, *Data Preparation for Data Mining*, Morgan Kaufmann, California, 1999.