Research Article

A New Classification Analysis of Customer Requirement Information Based on Quantitative Standardization for Product Configuration

Zheng Xiao, Zude Zhou, and Buyun Sheng

School of Mechanical and Electronic Engineering, Wuhan University of Technology, No. 122, Luoshi Road, Hongshan, Wuhan, Hubei 430070, China

Correspondence should be addressed to Zheng Xiao; reallylaugh@whut.edu.cn

Received 13 October 2015; Accepted 30 December 2015

Academic Editor: David Bigaud

Copyright © 2016 Zheng Xiao et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Traditional methods used for the classification of customer requirement information are typically based on specific indicators, hierarchical structures, and data formats and involve a qualitative analysis in terms of stationary patterns. Because these methods neither consider the scalability of classification results nor do they regard subsequent application to product configuration, their classification becomes an isolated operation. However, the transformation of customer requirement information into quantifiable values would lead to a dynamic classification according to specific conditions and would enable an association with product configuration in an enterprise. This paper introduces a classification analysis based on quantitative standardization, which focuses on (i) expressing customer requirement information mathematically and (ii) classifying customer requirement information for product configuration purposes. Our classification analysis treated customer requirement information as follows: first, it was transformed into standardized values using mathematics, subsequent to which it was classified through calculating the dissimilarity with general customer requirement information related to the product family. Finally, a case study was used to demonstrate and validate the feasibility and effectiveness of the classification analysis.

1. Introduction

In the past decades, the classification of customer requirement information (CRI) has become increasingly important in the entire product development and manufacturing process. Numerous researchers have accomplished valuable work in terms of developing approaches to classification, and three of these methods, that is, those based on the Kano model, product hierarchical structure, and data format, are the most prominent.

The Kano model was introduced by Noriaki Kano who classified customer preferences according to a threshold, performance, and excitement to guide design decisions [1]. The Kano model is often associated with quality function deployment (QFD) and fuzzy mathematics, which was utilized to determine the weights or importance of customer requirements [2, 3]. The product hierarchical structure classifies CRI into different types according to modern product hierarchies,

such as function, form, extension, and price requirements [4, 5], and is a form of qualitative analysis intended to provide performance indicators. The data format divides CRI into binary, option, parameter, description, and interpretation data types in terms of common data formats on the Web [6, 7]. Compared with the product hierarchical structure and the KANO model, the classification based on the data format has already solved the problem of quantitative expression and analysis for CRI analysis. Moreover, the premise of a classification based on the data format is to collect CRI via the Web, an approach that is already adopted and is likely to be widely utilized in the future [8]. The classification of CRI using the data format facilitates not only the subsequent analysis of the information, but also the management thereof. Nevertheless, classification based on the data format has certain limitations, because this method does not take into account the number of data formats that exist on the Web or how this classification relates to the entire product configuration process.

In conclusion, traditional classification methods focus on customer requirements rather than on product configuration, such that the classification results are neither extendable nor scalable. The processes of CRI classification and product configuration can occur independently of one another. For instance, customers' preferences are able to complete transformation into product design [9, 10] without performing CRI classification, although CRI classification results are useful for determining them. Additionally, the lack of mathematical analysis of CRI means that it is difficult to use the classification results directly for product configuration purposes. Therefore, the modern approach to CRI classification has not been a simple issue of dividing the information into groups; instead, it has to consider (i) how to analyze and express CRI with mathematical methods and (ii) how to classify CRI for the purpose of product configuration.

In this paper, we introduce a classification analysis based on quantitative standardization. Section 2 describes a mathematical model to analyze CRI in terms of product families. Section 3 provides details as to how to transform CRI into quantitative standards. On the basis of Sections 2 and 3, Section 4 presents the classification method. Finally, in Section 5 a case is demonstrated to confirm the feasibility and effectiveness of the proposed method.

2. Background Review

2.1. CRI Structured Procedure. The increasing application of e-commerce has been transforming the acquisition method of choice from offline to online. Online CRI acquisition mainly depends on the Web, including the advantages it offers in terms of efficiency, affordability, and convenience [11]. Online CRI can be submitted in the form of XML documents, which can contain multiple kinds of data and can adapt to the dynamic development of CRI. An XML document containing CRI is referred to as a customer requirement information document (CRID). Because of the XML framework, each CRID can tag various CRI features of each customer; however, these features contain multiple data types characterized by fuzziness, concealment, and similarity such that it is difficult to correctly identify the information for use in the process of product configuration. Thus, it is necessary to transform CRI into a structured model corresponding to the product family model to enable the information to be translated into product development for manufacturing. Research on the structure model of the product family has led to the construction of general customer requirement information (GCRI) [12], which is able to abstract a series of similar CRI features whose personalized features are distinguished by specific values of cases [13]. Figure 1 illustrates the standardization procedure in which CRI is transformed into GCRI, which means that features with values in CRID categories can be transformed into the corresponding GCRI classes in a structured model.

2.2. CRI Document Model. In this model, the CRI is submitted in the form of CRID, which is a document based on a document representation model capable of enlightening the

FIGURE 1: Summary of the standardization procedure.

CRID representation. Because a document is composed of words, the word is the most widely used unit of information in document modeling [14]; namely, a document representation model can be established by using the characteristics of words and is implemented by the Vector Space Model (VSM). The VSM [15] is a vector model utilizing the extracted characteristics of words from a document in a Euclidean space [16]. A word characteristic corresponds to a separate term. If a term occurs in the document, its value in the vector is nonzero [17]. The main idea of VSM is as follows: let $D = [D_1, D_2, \dots, D_d]$ be a document set, with each document D_i ($i = 1, 2, \dots, d$) represented by a set of terms $T = [T_1, T_2, \dots, T_t]$; any T_j ($j = 1, 2, \dots, t$) corresponds to one dimension in the VSM such that D can be a $d \times t$ matrix, which means that documents can be mapped to a point in the VSM and their similarity can be calculated by distances. To date, the VSM is the most efficient and useful document representation model because it transforms the similarity between two documents into the similarity between two vectors [18].

Thus, the representation model of the CRID is a one-dimensional vector containing the values of the CRI features. If the number of CRID tags is n and the CRI features are independent of each other, any CRID D_i can be represented as a one-dimensional set $[r_{i1}, r_{i2}, \dots, r_{in}]$. On the basis of this, if the number of CRIDs is m and each CRID defines n features, there will be an $m \times n$ matrix shown as follows:

$$R = \begin{bmatrix} r_{11} & r_{12} & \cdots & r_{1n} \\ r_{21} & r_{22} & \cdots & r_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ r_{m1} & r_{m2} & \cdots & r_{mn} \end{bmatrix}. \quad (1)$$

In the matrix R , r_{ij} ($i = 1, 2, \dots, m$; $j = 1, 2, \dots, n$) is a j feature value submitted by one customer i . However, an existing problem of CRI features is that the information may occur as a combination of multiple data types such as numbers and words. This issue would have to be addressed by analyzing the similarity of these values to introduce quantitative standardization of CRI, such that r_{ij} in matrix R could be changed into a uniform fundamental unit [19].

2.3. Classification and Clustering. Classification and clustering are two major methods for information analysis, especially data in XML documents [20]. The aim of the classification is to build a classifier based on some cases with some attributes to describe the objects or one attribute to describe the group of the objects. Then, the classifier is used to

FIGURE 2: Process of CRI quantitative standardization.

predict the group attributes of new cases based on the values of other attributes [21]. The aim of clustering is to find groups of objects such that the objects in a group will be similar to one another and different from the objects in other groups. The clustering algorithm has access only to the set of features describing each object; it is not given any labels as to where each of the instances should be placed within the partition [22]. Thus, classification is supervised learning as targets are predefined, whereas clustering is generally used in an unsupervised fashion.

The typical clustering algorithm is K -means [23], which aims to partition n observations into k clusters. Each observation belongs to the cluster with the nearest mean. Since the sum of squares is the squared Euclidean distance [24], this is intuitively the “nearest” mean. K -means clustering is able to compute fast and compatible with massive data. However, due to its unsupervised fashion, there must be an issue of how to choose k and centroids. Subsequent research on semisupervised clustering [25] is to remedy this defect because guiding a clustering algorithm is very efficient for improving its quality.

In practice, if the centroids and k can be defined by some technique indicators, a classifier will be built based on dissimilarity calculation among objects. In the next sections, aiming at CRI, the construction of this classifier will be introduced.

3. Quantitative Standardization of CRI

The CRI features contain multitype data values in a CRID, including nominal and scaled variables. The nominal variables are binary and multiple, and the scaled variables are measured. The process of CRI quantitative standardization is shown in Figure 2 and has the purpose of realizing quantification for the assignment of uniform fundamental units of CRI feature values.

3.1. Quantitative Analysis. The nominal binary variables have the values of 0 and 1, where 0 or 1 means a CRI feature value does not exist or exists, respectively. The scaled variables are real values with fundamental units. The nominal multiple variables are different from the former two, which may not correspond to real values with multiple states. Thus, there must be a quantitative analysis for variable assignment.

The differences among the various features and the differences between the feature states are accounted for by proposing methods based on fuzzy mathematics for the purpose of linguistic representation for the quantitative analysis.

(i) Quantitative Analysis Based on Fuzzy Mathematics. Let U be a discussion universe; if mapping $\mu_{\underline{A}} : U \rightarrow [0, 1]$, $u \mapsto \mu_{\underline{A}}(u) \in [0, 1]$ defines a fuzzy set $\underline{A}$ in U , $\mu_{\underline{A}}(u)$ will be named

the membership function of the fuzzy set $\underline{A}$ [26], which can be expressed as

$$\underline{A} = (\mu_{\underline{A}}(u_1), \mu_{\underline{A}}(u_2), \dots, \mu_{\underline{A}}(u_n)). \quad (2)$$

The defuzzification calculation formula of $\underline{A}$ is

$$A = \sqrt{\mu_{\underline{A}}(u_1)^2 + \mu_{\underline{A}}(u_2)^2 + \dots + \mu_{\underline{A}}(u_n)^2}. \quad (3)$$

The membership function is also represented as a fuzzy distribution, of which the trapezoidal distribution is commonly used. The trapezoidal distribution $\underline{M}$ is described by four parameters $\underline{M} = (l, m_1, m_2, n)$, whose fuzzy membership function can be expressed as

$$\mu_{\underline{M}}(u) = \begin{cases} \frac{u-l}{m_1-l} & l \leq u \leq m_1 \\ 1 & m_1 \leq u \leq m_2 \\ \frac{u-n}{m_2-n} & m_2 \leq u \leq n \\ 0 & \text{others.} \end{cases} \quad (4)$$

According to Chen's model [27], the defuzzification calculation formula of $\underline{M}$ is

$$M = \frac{l + m_1 + m_2 + n}{4}. \quad (5)$$

(ii) Quantitative Analysis Based on Linguistic Representation. The linguistic representation is achieved by constructing a fuzzy function with a linguistic evaluation scale rule in Table 1. In terms of the number of nominal multiple variable states, the number of linguistic terms is determined so that a linguistic set $\{N, VL, L, RL, LL, M, LH, RH, H, VH, P\}$ can be defined. Finally, the linguistic set is transformed into fuzzy functions by linguistic representation [28, 29].

3.2. Standard Transform. The similarity of quantitative CRI with different fundamental units is measured by using a standard transform formula:

$$r_{ij} = \frac{r_{ij} - \bar{r}_j}{\sqrt{(1/(m-1)) \sum_{i=1}^m (r_{ij} - \bar{r}_j)^2}} \in [-1, 1], \quad (6)$$

$$r_{ij}^* = \frac{r_{ij} - \min\{r_{ij}\}}{\max\{r_{ij}\} - \min\{r_{ij}\}} \in [0, 1].$$

This standard transform formula can eliminate the influence of fundamental units and unify CRI standardized values in the range of $[0, 1]$.

TABLE 1: Linguistic evaluation scale rule.

	2	3	4	5	5	6	7	9
N			Y	Y		Y	Y	Y
VL								Y
L		Y		Y	Y	Y	Y	Y
RL			Y		Y	Y		Y
LL							Y	
M	Y	Y		Y	Y		Y	Y
LH							Y	
RH			Y		Y	Y		Y
H	Y	Y		Y	Y	Y	Y	Y
VH								Y
P			Y	Y		Y	Y	Y

4. Classification Analysis of CRI

As mentioned in the CRI standardization procedure, CRI features with values in CRIDs can be transformed into corresponding GCRI with cases in a structured model. If the GCRI is regarded as the centroid, the CRI features can be classified by their dissimilarity with GCRI cases. Because it is possible to change the selection of GCRI flexibly according to specific conditions, such as technical abilities or information variation, the classification can achieve diversification.

The classification analysis of CRI performs the following five steps:

- (A) If there are n CRIDs, in accordance with the GCRI feature set $\{F_1, F_2, \dots, F_m\}$, there will be m selected CRI features to constitute an $n \times m$ matrix X :

$$X = \begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1m} \\ x_{21} & x_{22} & \cdots & x_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ x_{n1} & x_{n2} & \cdots & x_{nm} \end{bmatrix}. \quad (7)$$

- (B) The dissimilarities among CRIDs are calculated by the Euclidean distance, which is a matrix D :

$$D = \begin{bmatrix} 0 & & & & \\ d_{21} & 0 & & & \\ d_{31} & d_{32} & 0 & & \\ \vdots & \vdots & \vdots & \ddots & \\ d_{n1} & d_{n2} & \cdots & d_{nm} & 0 \end{bmatrix}. \quad (8)$$

- (C) In (8), if $d_{ij} = 0$, CRID i and CRID j will be merged such that the dimensionality of matrix D is reduced. An $h \times m$ ($h \leq n$) matrix is shown as Y :

$$Y = \begin{bmatrix} y_{11} & y_{12} & \cdots & y_{1m} \\ y_{21} & y_{22} & \cdots & y_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ y_{h1} & y_{h2} & \cdots & y_{hm} \end{bmatrix}. \quad (9)$$

TABLE 2: Quantitative analysis of scaled feature variables.

CRI scaled features	Real values	Standardized values
CPU	2-core	0
	4-core	1
Frequency	1.0 G	0
	1.2 G	0.4
	1.5 G	1
Memory RAM	512 M	0
	1536 M	1
Dimension	4.3 IN	0
	4.5 IN	0.5
	4.7 IN	1
Screen resolution	800 × 480	0
	854 × 480	0.19
	960 × 540	1

TABLE 3: Quantitative analysis of nominal binary feature variables.

CRI scaled features	States of values		Standardized values	
Battery replacement	Replaceable	Irreplaceable	1	0
Message to enterprise	Not null	Null	1	0

- (D) In the GCRI feature set $\{F_1, F_2, \dots, F_m\}$ ($m \in \forall N$), because each feature may have no fewer than one case, the dissimilarity measurement formula between F_k and y_k ($k = 1, 2, \dots, m$) is

$$D^* = \min [|\{y_k\} - \{F_k\}|]. \quad (10)$$

- (E) In (10), if $D^* \neq 0$, the feature value will be marked, and its corresponding CRID will be submitted and recorded. The classification results depend on the number of $D^* \neq 0$ in the CRID.

5. Case Study

5.1. CRI Quantitative Standardization for a Smart Phone. The CRI scaled features for a smart phone have real values with fundamental units, some of which are chosen for the demonstration of quantitative standardization. The data is initially subjected to the process of quantitative analysis and standard transformation described in Section 3 and the resulting standardized values are listed in Table 2. Similarly, the standardized values of CRI nominal binary features are shown in Table 3.

The CRI nominal multiple features for a smart phone have no fewer than two states, for example, depth and color. Because these states are not expressed by quantitative values, they have to be transformed by fuzzy mathematics and linguistic representation to obtain quantitative values, which are subsequently changed into standardized values. The process of quantitative standardization is presented in Tables 4 and 5.

In Table 4, three states of depth are transformed into a trapezoidal distribution by the linguistic evaluation scale rule

TABLE 4: Quantitative analysis of nominal multiple feature variables: depth.

Feature (depth)	Trapezoidal distribution	Quantitative values	Standardized values
Ultrathin	(0.0, 0.0, 0.9, 0.9)	0.45	0
Thin	(0.9, 0.9, 1.0, 1.0)	0.95	0.78
Ordinary	(1.0, 1.0, 1.2, 1.2)	1.10	1

TABLE 5: Quantitative analysis of nominal multiple feature variables: color.

Feature (color)	Fuzzy set vector	Quantitative values	Standardized values
Black	(0.00, 0.00, 0.00)	0.00	0
White	(1.00, 1.00, 1.00)	1.73	1
Grey	(0.50, 0.50, 0.50)	0.87	0.5
Light grey	(0.83, 0.83, 0.83)	1.44	0.83
Dark grey	(0.66, 0.66, 0.66)	1.14	0.66
Dim grey	(0.41, 0.41, 0.41)	0.71	0.41
Red	(1.00, 0.00, 0.00)	1.00	0.58

TABLE 6: GCRI feature set and cases.

Feature set	Standardized values
Frequency	0 0.4 1
Color	0 1
Depth	0 0.78 1
Dimension	0 0.5 1
Screen resolution	0 0.19
Message to enterprise	1 0

and fuzzy membership distribution. Then, the quantitative values of defuzzification are calculated by Chen's model. After performing a standard transformation, the standardized values of the three depth states are 0, 0.78, and 1, respectively.

Table 5 contains the results of the transformation of seven states of color into a fuzzy set vector by a color universe based on RGB [30]. The quantitative values and the standardized values are calculated by a defuzzification formula of a fuzzy set vector.

5.2. CRI Classification Analysis for a Smart Phone. If the GCRI feature set and their cases (set of standardized values) are selected as indicated in Table 6, these cases will produce the quantitative standardized feature values of 30 CRIDs listed in Table 7, which constitutes a 30×6 matrix. In accordance with the dissimilarity measurement, the dimension reduction matrix is a 22×6 matrix.

Finally, the values in Table 7 were processed in terms of the dissimilarity measurement formula for selected GCRI

FIGURE 3: Dissimilar distribution of CRIDs.

FIGURE 4: Product configuration with GCRI and product family.

cases and their corresponding feature values, and the results are listed in Table 8. The dissimilarity distribution is shown in Figure 3.

According to results of dissimilarity with GCRI cases, those matching CRIDs are 1, 5, 8, 10, 14, 15, 17, 19, 20, 21, 22, 23, 24, 26, 29, which means that they are compatible with product family model. Their product configuration scheme can be generated. Contrarily, the remaining CRIDs and their

FIGURE 5: Global analysis diagram.

TABLE 7: Quantitative standardized feature values for smart phone.

CRID	Frequency	Color	Depth	Dimension	SR	Message
1	0.4	0	0	0.5	0.19	0
2	0.4	1	0.78	0.5	0.19	1
3	0.4	0.87	0.78	0.5	0.19	1
4	0.4	0	0.78	0.5	0.19	1
5	0.4	0	0.78	0.5	0.19	0
6	0.4	0.5	0.78	0.5	0.19	1
7	0.4	0.5	0.78	0.5	0.19	0
8	0.4	1	0.78	0.5	0.19	0
9	0.4	1	0.78	0.5	0.19	1
10	0.4	1	0.78	0.5	0.19	0
11	0	1	0.78	1	0	1
12	0	1	0.78	1	0	1
13	0	0.87	0.78	1	0	1
14	0	0	0.78	1	0	0
15	0	0	0.78	1	0	0
16	0	0	0	1	0	1
17	0	0	0	1	0.19	0
18	0	0	0	1	0.19	1
19	0	0	0	1	0.19	0
20	0	0	0	0	1	0
21	1	0	0	0	1	0
22	1	0	0	0	1	0
23	1	0	0	0	1	0
24	1	0	0	0	1	0
25	1	0.5	0	0	1	0
26	1	1	0	0	1	0
27	1	0.5	0	1	1	1
28	1	1	0	1	1	1
29	1	1	0	1	1	0
30	1	0.5	0	1	1	0

features must be divided and marked by the dissimilarity. In view of the overall situation, the classification analysis derives what CRIDs and which CRI can meet product configuration. From the microscopic view, the classification analysis also derives which GCRI feature is easy to be challenged or

ignored. Thus, in terms of those challenged GCRI features, the product family model would be considered to make appropriate adjustments for better product configuration. Furthermore, because the GCRI is able to be renovated, it is possible to achieve a flexibly classification if it is the centroid.

5.3. Further Discussion. Production in enterprises is largely oriented towards CRI. Because original CRI is hardly utilized for modeling product family, GCRI is widely adopted in order to transform CRI into a structure model. This structured model composed of GCRI refers to product family model. Based on the GCRI and the product family, the modular product configuration is able to select appropriate modules and manufacture desirable products. Take the case of a smart phone; Figure 4 presents product configuration with GCRI and its product family. Thus, using GCRI to distinguish CRI can effectively assist in product configuration.

In CRI classification analysis for the smart phone, the dissimilar distribution indicates what CRIDs can meet product configuration. The global analysis diagram shown in Figure 5 affords a proportion of dissimilarity with a GCRI set in all CRIDs and that with GCRI cases in each CRID. The enterprise can determine the processing level of each CRID according to the proportion of dissimilarity between GCRI cases, whereas CRIDs with a higher proportion often require more complex treatment.

Likewise, the CRI classification analysis discovers which GCRI feature is easy to be challenged or ignored. The feature analysis diagram shown in Figure 6 exposes the statistics of features that differ from those in the GCRI cases. Popular CRI features, such as the color feature, are clearly shown. Moreover, enterprises are able to select those elements they wish to include in the GCRI feature set by either adding or removing features presented in this analysis result. These results are expected to be useful for product family renewal.

6. Conclusions

This paper proposes a classification approach for realizing CRI quantitative analysis aimed at supporting product family adaptation in enterprises. The CRI classification analysis not only considers the scalability of classification results but also regards subsequent application to product configuration.

TABLE 8: Dissimilarity results.

	CRID	$D^*_{\text{Frequency}}$	D^*_{Color}	D^*_{Depth}	$D^*_{\text{Dimension}}$	D^*_{SR}	D^*_{Message}
1	1	0	0	0	0	0	0
2	2 & 9	0	0	0	0	0	1
3	3	0	0.13	0	0	0	1
4	4	0	0	0	0	0	1
5	5	0	0	0	0	0	0
6	6	0	0.5	0	0	0	1
7	7	0	0.5	0	0	0	0
8	8 & 10	0	0	0	0	0	0
9	11 & 12	0	0	0	0	0	1
10	13	0	0.13	0	0	0	1
11	14 & 15	0	0	0	0	0	0
12	16	0	0	0	0	0	1
13	17 & 19	0	0	0	0	0	0
14	18	0	0	0	0	0	1
15	20	0	0	0	0	0	0
16	21 & 22 & 23 & 24	0	0	0	0	0	0
17	25	0	0.5	0	0	0	0
18	26	0	0	0	0	0	0
19	27	0	0.5	0	0	0	1
20	28	0	0	0	0	0	1
21	29	0	0	0	0	0	0
22	30	0	0.5	0	0	0	0

FIGURE 6: Feature analysis diagram.

At the technical level, considering that CRI feature values consist of multiple data types, such as numbers and words, a quantitative analysis based on fuzzy mathematics and linguistic representation is presented. This analysis is capable of revealing not only the differences between the CRI features of a product, but also the differences among the states in each CRI feature, thereby avoiding shortcomings such as incomplete expression of states and meaningless assignment of features. Furthermore, the dissimilarity among CRI feature values is measured by utilizing a standard transformation for eliminating the influence of different fundamental units. An association between the classification analysis and product configuration is achieved by using a flexible classification

based on the fact that the selective GCRI is regarded as the centroid. Therefore, the determination of the classification results is no longer an isolated operation; instead, it derives which CRI can meet product configuration and which GCRI feature can assist in improving product family.

In engineering practice, classification analysis enables CRI to be quantified recognized information, which will be compatible with other management systems in enterprises such as ERP or PDM. It is helpful for the final product rapid development and intelligent configuration. Meanwhile, using GCRI to discover the specific features, enterprises can determine market positioning of their future product and predict the corresponding product family model. Although the proposed approach is demonstrated by analyzing selected features of smart phones to verify the feasibility and effectiveness, it can be extended to other consumer electronics.

As future work, we will consider the study of a new framework that enables CRI classification analysis to deal with Big Data.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

Acknowledgments

This research is funded and supported by the Natural Science Foundation of Hubei Province, China (2015CFA115), and Science and Technology Support Program of Hubei Province,

China (2015BAA058). The authors express their most sincere appreciation to Professor Quan LIU, who provided an experimental platform together with valuable advice. The authors are also grateful for assistance received from Professor Qingsong AI and Professor Ping LOU regarding revising the paper.

References

- [1] Q. L. Xu, R. J. Jiao, X. Yang, M. G. Helander, H. M. Khalid, and O. Anders, "Customer requirement analysis based on an analytical Kano model," in *Proceedings of the IEEE International Conference on Industrial Engineering and Engineering Management*, pp. 1287–1291, Singapore, December 2007.
- [2] L. Xie and Z. Li, "A customer requirements rating method based on fuzzy kano model," in *Future Control and Automation*, vol. 173 of *Lecture Notes in Electrical Engineering*, pp. 141–146, Springer, Berlin, Germany, 2012.
- [3] M. A. Shafia and S. Abdollahzadeh, "Integrating fuzzy kano and fuzzy TOPSIS for classification of functional requirements in national standardization system," *Arabian Journal for Science and Engineering*, vol. 39, no. 8, pp. 6555–6565, 2014.
- [4] Z. Lingyun, *Research on the Key Technology of Customer Requirements Processing of Design for Mass Customization*, Hefei University of Technology, Hefei, China, 2008.
- [5] S. Li, Y. Liu, J. Wang, and H. Zeng, "An intelligent interactive approach for assembly process planning based on hierarchical classification of parts," *The International Journal of Advanced Manufacturing Technology*, vol. 70, no. 9–12, pp. 1903–1914, 2014.
- [6] Y. G. Jing, B. Dan, S. Peng, and L. F. Guo, "Intelligent mapping of semi-structured customer needs for web-based product customization," *Advanced Materials Research*, vol. 201, pp. 1496–1499, 2011.
- [7] Z. Xiao, Q. Liu, and Q. S. Ai, "Acquisition and analysis for customer requirement information," *Advanced Materials Research*, vol. 139–141, pp. 1468–1471, 2010.
- [8] China Internet Network Information Center (CNNIC), "Research Report of Chinese Online Shopping Market," 2013, <http://www.cnnic.cn/>.
- [9] Y. Wang and M. M. Tseng, "Integrating comprehensive customer requirements into product design," *CIRP Annals—Manufacturing Technology*, vol. 60, no. 1, pp. 175–178, 2011.
- [10] Y. Wang and M. M. Tseng, "Identifying emerging customer requirements in an early design stage by applying bayes factor-based sequential analysis," *IEEE Transactions on Engineering Management*, vol. 61, no. 1, pp. 129–137, 2014.
- [11] S. Negash, T. Ryan, and M. Igbaria, "Quality and effectiveness in web-based customer support systems," *Information & Management*, vol. 40, no. 8, pp. 757–768, 2003.
- [12] Y. Tan and G. Wei, *Theory and Method for Modular Product Intelligent Configuration*, Southwest Jiaotong University Press, 2011.
- [13] J. Wang, *Research on Identification and Description Approaches to the Information about Customer Needs for Mass Customization*, Chongqing University, Chongqing, China, 2009.
- [14] Y. Wang, *Information Extraction from Clinical Notes*, University of Sydney, Sydney, Australia, 2009.
- [15] G. Salton, A. Wong, and C. S. Yang, "A vector space model for automatic indexing," *Communications of the ACM*, vol. 18, no. 11, pp. 613–620, 1975.
- [16] J. Han, M. Kamber, and J. Pei, *Data Mining: Concepts and Techniques*, Morgan Kaufmann, San Francisco, Calif, USA, 2006.
- [17] L. Jing, M. K. Ng, and J. Z. Huang, "Knowledge-based vector space model for text clustering," *Knowledge and Information Systems*, vol. 25, no. 1, pp. 35–55, 2010.
- [18] A. Strehl, J. Ghosh, and R. Mooney, "Impact of similarity measures on web-page clustering," in *Proceedings of the Workshop on Artificial Intelligence for Web Search (AAAI '00)*, pp. 58–64, July 2000.
- [19] J. Xie and C. Liu, *Methods and Applications of Fuzzy Mathematics*, Huazhong University of Science and Technology Press, Wuhan, China, 2006.
- [20] X. Bi, X. Zhao, G. Wang, Z. Zhang, and S. Chen, "Distributed learning over massive XML documents in ELM feature space," *Mathematical Problems in Engineering*, vol. 2015, Article ID 923097, 13 pages, 2015.
- [21] A. S. Fathima, D. Manimegalai, and N. Hundewale, "A review of data mining classification techniques applied for diagnosis and prognosis of the arbovirus-dengue," *International Journal of Computer Science Issues*, vol. 8, no. 6, p. 322, 2011.
- [22] K. Wagstaff, C. Cardie, S. Rogers, and S. Schrödl, "Constrained k-means clustering with background knowledge," in *Proceedings of the 18th International Conference on Machine Learning (ICML '01)*, vol. 1, pp. 577–584, Williamstown, Mass, USA, June 2001.
- [23] T. Kanungo, D. M. Mount, N. S. Netanyahu, C. D. Piatko, R. Silverman, and A. Y. Wu, "An efficient k-means clustering algorithm: analysis and implementation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 24, no. 7, pp. 881–892, 2002.
- [24] Wikipedia, Euclidean Distance, 2015, https://en.wikipedia.org/wiki/Euclidean_distance.
- [25] Y. Chen, M. Rege, M. Dong, and J. Hua, "Non-negative matrix factorization for semi-supervised data clustering," *Knowledge and Information Systems*, vol. 17, no. 3, pp. 355–379, 2008.
- [26] L. A. Zadeh, *Fuzzy Sets and Fuzzy Information Granulation Theory*, Beijing Normal University Publishing Group, Beijing, China, 2005.
- [27] S.-M. Chen, "Fuzzy group decision making for evaluating the rate of aggregative risk in software development," *Fuzzy Sets and Systems*, vol. 118, no. 1, pp. 75–88, 2001.
- [28] F. Herrera and L. Martínez, "A 2-tuple fuzzy linguistic representation model for computing with words," *IEEE Transactions on Fuzzy Systems*, vol. 8, no. 6, pp. 746–752, 2000.
- [29] F. Herrera and L. Martínez, "A model based on linguistic 2-tuples for dealing with multigranular hierarchical linguistic contexts in multi-expert decision-making," *IEEE Transactions on Systems, Man, and Cybernetics Part B: Cybernetics*, vol. 31, no. 2, pp. 227–234, 2001.
- [30] Z. Zhou, Z. Xiao, Q. Liu, and Q. Ai, "An analytical approach to customer requirement information processing," *Enterprise Information Systems*, vol. 7, no. 4, pp. 543–557, 2013.