Expert Systems with Applications 36 (2009) 4185–4191

Contents lists available at ScienceDirect

Expert Systems with Applications

journal homepage: www.elsevier.com/locate/eswa

Data mining to improve industrial standards and enhance production and marketing: An empirical study in apparel industry

Chih-Hung Hsu *

Department of Industrial Engineering and Management, Hsiuping Institute of Technology, Taichung 41249, Taiwan, ROC

ARTICLE INFO

Keywords:
Data mining
Cluster analysis
Industrial standards
Production management and marketing
Apparel industry

ABSTRACT

Apparel production is a high value-added industry in the global textile manufacturing chain. Standard size charts are crucial industrial standards for high-tech apparel industries to maintain competitive advantages in knowledge economy era. However, these industries suffering from production management and marketing often find it hard to obtain the accurate standard size charts. In addition to conventional experience approaches, there is an urgent need to develop effective mechanism to find the industrial standards that are the most suitable to their own industries. This study aims to fill the gap by developing a data mining framework based on two-stage cluster approach to generate useful patterns and rules for standard size charts. The results can provide high-tech apparel industries with industrial standards. An empirical study was conducted in an apparel industry to support their manufacturing decision for production management and marketing with various customers' needs. The results demonstrated the practical viability of this approach. Moreover, since the anthropometric database must be repeatedly updated, standard size charts may also be continuously renewed via application of the proposed data mining framework. By applying the proposed framework for solving industrial problems, these industrial standards will remain continually beneficial for both production planning and reducing inventory costs, while facilitating production management and marketing.

1. Introduction

Apparel manufacturing produces products with the highest added value in the global textile manufacturing chain. Standard size charts provide very crucial industrial standards, and play an important role in high-tech apparel industries (Burns & Bryant, 2000; Jongsuk & Jasper, 1993). For large-scale production, effective production management and marketing are very important factors for apparel manufacturers seeking to reduce cost and increase marketplace competitiveness. Apparel manufacturers seek to produce the best designs that meet standard size charts for fitting customers' body types and needs (Regan, Kincade, & Sheldon, 1998). Furthermore, standard size charts can correctly predict numbers of items and ratio of sizes to be produced, resulting in accurate inventory control and production planning for facilitating production management and marketing (Chung & Wang, 2006; Dai, 2004). Due to the lack of up-to-date standard size charts, many manufacturers cannot develop their own size charts for production; as a result, very often the overdue size charts do not fit the customers' body types, so consumers are forced to choose suitable apparel by trial and error, resulting in enormous inconvenience, not to mention wasted time and money (Burns & Bryant, 2000; Hsu & Jing, 1999). Owing to current variations in body type, thus, the developments of standard size charts that accurately conform to the body types of people are crucial for improving production management and marketing (Gupta & Gangadhar, 2004; Laing, Holland, Wilson, & Niven, 1999). Thus, an issue of importance is to have the current standard size charts of the customers' body types, in order to predict the production and marketplace demand of different sizes of apparel for the apparel industries.

Standard size charts originated from the experienced tailors in the late 18th century. Tailors measured the body dimensions of each customer, and then drew and cut patterns. After many original patterns had been accumulated, tailors gradually developed patterns into a system for storing apparel, which could be utilized to make clothes for people with similar body types. For the conventional approach of establishing standard size charts, Emanuel, Alexander, Churchill, and Truett (1959) developed a set of procedures to formulate standard size charts for all body types. According to this approach, people were first classified by body weight into four shape categories. Within the four shape categories, they were subdivided into two body height: tall and short. Thus, eight size groups were classified, and each category had similar body height and weight. The sizing classifications of other countries were also similar, and the classification was based on two or three

* Fax: +886 4 24961114.
E-mail address: chhsu@mail.hit.edu.tw

0957-4174/$ - see front matter © 2008 Elsevier Ltd. All rights reserved.
doi:10.1016/j.eswa.2008.04.009

sizing variables (Winks, 1997). In related studies, Tryfos (1986) suggested an integer programming approach to optimize the number of sizes in order to maximize expected sales. Chen (1998) focused on women to measure anthropometric data for developing standard size charts. McCulloch, Paal, and Ashdown (1998) proposed a non-linear optimization technique to derive a set of standard size charts from anthropometric data. Chang and Shen (1999) applied decision analysis to develop standard size charts for Taiwanese women. Laing et al. (1999) used statistical analysis to develop standard size charts for protective apparel used by New Zealand firemen. Moon and Nam (2003) measured the anthropometric data of Korean women to classify lower trunk figure types. Gupta and Gangadhar (2004) used a statistical method to develop standard size charts for young Indian women, and Chung and Wang (2006) applied statistical analysis to establish standard size charts for Taiwanese students. Human body types can be distinguished by taking various approaches. As an alternative to catering to the individual consumer in classifying apparel sizes, apparel manufacturers limit their production to a few standard sizes. Consumers are offered more limited choices, but manufacturers can avoid production and inventory problems. Therefore, it would be helpful to develop standard size charts, which have the fewest number of sizes to fit the largest number of body types, for the majority of consumers (Chung & Wang, 2006; McCulloch et al., 1998).

Data mining has been successfully applied in many domains, such as health insurance (Chas, Ho, Cho, Lee, & Ji, 2001), biomedicine (Maddour & Elloumi, 2002), human resource management (Min & Emam, 2003), semiconductor manufacturing (Chien, Hsiao, & Wang, 2004), production schedule (Sha & Liu, 2005), knowledge management (Hou & Yang, 2006), education (Chang, 2007) and course planning (Hsia, Shie, & Chen, 2008). However, there is a lack of research in developing industrial standards by using the data mining approach. This study aims to develop a data mining framework for industrial standards to explore useful patterns and rules from anthropometric data. By applying the proposed framework, body types can be classified. These industrial standards can then be developed to facilitate apparel production and marketing. Thus, production management and marketing will be enhanced with these standard size charts. An empirical study for industrial standards in one of the largest apparel company in Taiwan is studied to demonstrate the validity of this approach.

2. Data mining and cluster analysis

The analysis of large amounts of data by automatic or semi-automatic means, in order to uncover significant patterns and rules, is the definition of data mining (Berry & Linoff, 2000). Among the data mining techniques, cluster analysis helps in classification of data. Cluster analysis seeks to maximize between-group variances and minimize within-group variances, including both hierarchical and non-hierarchical methods (Berson, Smith, & Thearling, 2001).

Agglomerative hierarchical algorithms are commonly used with hierarchical methods, to calculate the distance between observations; the proximity of the two closest observations determines which such pairs will form a cluster. This procedure ends when all of the observations have been joined into a single cluster. Crucial to this application, is Ward's minimum variance in the employment of an agglomerative hierarchical algorithm. The prime determination of combination is based on attaining the smallest increase in global within-group variance. Without doubt, for non-hierarchical algorithms, the *K*-means algorithm is a highly popular method for choosing clusters as well as the total of clusters. There is an alignment of clusters according to the closest centroid to a particular observation, until all observations have been assigned a cluster-based on the closest centroid. This method ensures that cluster variance will be minimized (Giudici, 2003; Hair, Anderson, Tatham, & Black, 1997).

Recent clustering studies have advocated synthesizing the hierarchical and non-hierarchical methods (Kuo, Ho, & Hu, 2002). To derive a cluster-based data mining approach, this study combines Ward's minimum variance method with the *K*-means algorithm. For a cluster-based data mining approach, this study proposed a two-stage operation for mining patterns and rules concerning anthropometric data in order to develop standard size charts that may facilitate apparel production management and marketing.

3. An empirical study

This study constructed a data mining framework to explore the anthropometric data for developing industrial standards. This framework includes four major steps: problem definition, data preparation, data mining, and evaluation and application as shown in Fig. 1. The previous steps can be served as the baseline reference for the next step.

3.1. Problem definition

Since there is a dearth of complete and up-to-date size charts for adult Taiwanese females, a anthropometric database was constructed based on 52 anthropometric variables, measured in each of 986 females in one of the largest apparel company in Taiwan, according to the definition of the garment construction and anthropometric surveys of the (ISO-8559); this resulted in 51,272 pieces of anthropometric data. This study aims to analyze large amounts of data, via data mining framework. In this way, systematic patterns and rules will be uncovered within the mass of anthropometric data, based on adult female lower body type classification so that industrial standards can be developed for the mutual benefit of consumers and manufacturers.

3.2. Data preparation

The data was processed, and analyzed, in order to boost the efficiency and ensure the validity of the results, data was first verified before the mining, and all abnormal data were not used (Pyle, 1999). Accordingly, out of 986 subjects measured, subjects whose heights and weights deviated by $\pm 3\sigma$ from the mean (Moon, & Nam, 2003), were regarded as exhibiting abnormal data and consequently were excluded from the sample. The 30 subjects of abnormal data were deleted. In total, 956 subjects were used for further analysis.

Since only a number of the 52 anthropometric variables matched the requirements for the development of standard size charts, in coordination with international standards (ISO/TR10652), as well as the judgment of domain experts, this study identified 16 key anthropometric variables. These 16 anthropometric variables included nine linear measurements and seven girth measurements.

Even 16 anthropometric variables, in developing standard size charts, would prove overly complex, so only the most important factors were first examined. According to Kaiser–Meyer–Olkin's measure of sampling adequacy (0.86) and Bartlett's test ($p < 0.01$), the 16 dimensions were all determined to be suitable for factor analysis which provided the eigenvalues of these 16 anthropometric variables. In regard to Kaiser's eigenvalue criterion, two factors were chosen with eigenvalues above 1. Anthropometric variables that exhibited factor loadings above 0.5, were clustered inside Factors 1 and 2 as shown in Table 1. The major anthropometric variables