Mining customer knowledge for tourism new product development and customer relationship management

Shu-hsien Liao [a,*], Yin-Ju Chen [a], Min-yi Deng [b]

[a] *Department of Management Sciences and Decision Making, Tamkang University, No. 151, Yingjuan Rd., Danshuei Jen, Taipei 251, Taiwan, ROC*
[b] *Graduate School of Resource Management, National Defense University, Management College, P.O. Box 90046-17 Jon-Ho, Taipei County, Taiwan, ROC*

ARTICLE INFO

ABSTRACT

In recent years tourism has become one of the fastest growing sectors of the world economy and is widely recognized for its contribution to regional and national economic development. Tourism product design and development have become important activities in many areas/countries as a growing source of foreign and domestic earnings. On the other hand, customer relationship management is a competitive strategy that businesses need in order to stay focused on the needs of their customers and to integrate a customer-oriented approach throughout the organization. Thus, this paper uses the Apriori algorithm as a methodology for association rules and clustering analysis for data mining, which is implemented for mining customer knowledge from the case firm, Phoenix Tours International, in Taiwan. Knowledge extraction from data mining results is illustrated as knowledge patterns, rules, and knowledge maps in order to propose suggestions and solutions to the case firm for new product development and customer relationship management.

© 2009 Elsevier Ltd. All rights reserved.

1. Introduction

In recent years tourism has become one of the fastest growing sectors of the world economy and is widely recognized for its contribution to regional and national economic development. Tourism product design and development have become important activities in many areas/countries as a growing source of foreign and domestic earnings. In this regard, marketing decisions and strategic planning of tourism new product development require knowledge of factors, attributes, patterns of customer demand and market supply affecting destination choice, customer preference/capability, product characteristic, type of trips and forecast of tourism flows in the short and long term. Thus, it can be suggested that the purpose of the study of tourism new product development (NPD) is to improve the ability to estimate and/or forecast and understand travel behavior, traveler satisfaction, and tourism management (Bramwell, 1998; Witt & Witt, 1995). In addition, an important objective of tourism product demand and development analysis is to improve the understanding of public behavior towards particular customer purchases profiles and patterns. It is, therefore, interesting to know how customers select their tourism products and investigates which factors and attributes are determining their choices become important sources not only understand the demand of tourism but also investigate the seg-

mentation of possible tourism product development (Seddighi & Theocharous, 2002).

Customer relationship management (CRM) is the key competitive strategy businesses need to stay focused on the needs of the customers and to integrate a customer facing approach throughout the organization. By using information and communication technology, businesses are trying to get closer to the customer so that they can create long-term relationships in tourism industry (Sevki & Rifat, 2006). Customer relationship management refers to all business activities directed towards initiating, establishing, maintaining, and developing successful long-term relational exchanges and it is the set of methodologies and tools that help an enterprise manage customer relationships in an organized way (Lawson-Body & Limayem, 2004). As customers and businesses interact more frequently, businesses will have to leverage CRM and related technologies to capture and analyze massive amounts of customer information. Because, information and communication technology allows customer data to be collected, consolidated, manipulated, and analyzed on an unprecedented scale. However, CRM demands more than information and communication technology. The customer must become the focal point of the organization. All members of the organization must understand and support the shared values required for CRM (Piccoli, O'connor, Capaccioli, & Alvarez, 2003).

In addition, most of the parties involved in the product design and development, such as the tourism suppliers and retailers, are aware of the importance and need for tourism firms to acquire

---

and share better customer knowledge. But this is easier to say than done since customers' knowledge is concealed within the customers. It is available but not accessible, and there is little possibility of exploring the full volume of data that should be collected for its potential value. Therefore, how to effectively process and use customer data is becoming increasingly important. This calls for new techniques to help analyze, understand or even visualize the huge amounts of stored data gathered from business and scientific applications (Liao & Chen, 2004). Among the new techniques developed, data mining is the process of discovering significant customer knowledge, such as patterns, associations, changes, and significant structures from large amounts of data stored in databases (Liao, Chen, & Wu, 2008; Liao, Hsieh, & Huang, 2008). Customer knowledge extracted through data mining can be integrated with product and marketing knowledge from research and can be provided to up tourism stream suppliers as well as downstream retailers. Thus it can serve as a reference for product development, product promotion and customer relationship management.

In terms of integrating data mining approach and tourism, Wickramasinghe, Amarasiri, and Alahakoon (2004) proposed an approach that integrates traditional mathematical, data mining, and evolutionary techniques with a multi-agent system. It is implemented as a travel optimizer application for the e-tourism domain. Law, Bauer, Weber, and Tse (2006) developed a rough sets based model that can capture the essential information from business travelers. In addition, Junping, Min, and Xuyan (2008) introduced the concept of the holiday tourism information data mining, which improves a distributed sampling association rule mining algorithm: DS-ARM, define the realization process of the algorithm, test the capability of the algorithm, and use the algorithm in the analysis of the holiday traveler destination traveling behavior. However, a few research considered the integration of data mining and tourism problem on new product development and customer relationship management.

On the other hand, map display is a powerful tool with the ability to convey a large amount of information in a limited space, and it also provides an interactive tool to allow the user to interact with the underlying information (Lin, 1997). Thus, the mapping approach, which focuses on the use of IT, can be used as a tool to support new product development. Holmlund and Strandvik (1999) proposed perception configuration as a new concept, and introduced configuration maps as tools for analyzing perceptions in business relationship studies. Tülin and Russell (1998) presented market maps with a probabilistic spatial panel data model that allows the positions of products sharing the same name to be correlated across product categories. In a business setting, the combination of perceptions by two parties (such as buyers and sellers) can be represented as a perception configuration. All the perceptions are depicted on the horizontal and vertical axes of the map. This map can be used to capture both the composition and the dynamics of perception configurations, and it is generically applicable to dyadic perception studies. Daniel, Wilson, and McDonald (2003) utilized a marketing map to represent the best practice in marketing and also used the process map to understand how IT can be deployed in order to support a marketing information system. In addition, marketing map and product map are designed and implemented on business alliances and new product development (Liao, Chang, & Lee, 2008). Thus, the map mainly illustrates the links between various stages of the planning and marketing process (Liao, Chen, & Tseng, 2009; Liao & Wen, 2009). Based on this concept, this study implements a knowledge map to illustrate that new product development and customer relationship management are essentially the function that matches the customer profile and product segmentation.

Accordingly, this paper investigates the following research issues in a Taiwan tourism firm: What exactly are the customers' profiles for tourism market? Are tourism knowledge of the customers and the product itself reflected in the needs and wants of the market? Can tourism product design and planning for product mix be developed according to the knowledge of customers? Can the knowledge of customers be transformed into knowledge assets of the case firm for new product development and customer relationship management? In addition, regarding the marketing methods, the direct marketing model can also be considered to ensure that the products developed are customer-oriented after customer/product patterns and segmentations been found. Clustering analysis and the Apriori algorithm are methodologies for data mining, which is implemented to mine knowledge from customers for NPD and CRM. The knowledge extracted from data mining results is illustrated as knowledge patterns, rules, and knowledge maps in order to propose suggestions and solutions to the case firm. The rest of this paper is organized as follows. In Section 2, we present the background of the current life insurance market in Taiwan. Section 3 introduces the proposed data mining system, which includes system framework, and physical database design. Section 4 presents the data mining process, including clustering analysis, Apriori algorithm, knowledge extraction process, and result analysis for NPD and marketing. Managerial implications are presented in Section 5; and Section 6 presents a brief conclusion.

2. The case firm – the Phoenix Tours International

2.1. Background of the case firm

The case firm, the Phoenix Tours International, founded in 1957, is the only OTC listed companies with ISO 9001 international quality certification in Taiwan, Republic of China. Before 1980, Phoenix Tours focused on inbound tourism business. After Taiwan government opened foreign tourism market in 1981, the case firm starts to extend its services to global tourism business. In 2001 and 2002, the case firm was selected as "the best tourism firm in Asia" by Travel Trade Gazette (TTG) and this is the only firm obtained this honor in Taiwan. In the last five years (2003–2007), the case firm occupied 58% inbound and outbound tourism market share and earned 1.5 billion NT dollars business value in average.

2.2. The new product development procedure of the case firm

The case firm plays the role of middle stream on whole tourism supply chain and provides retailing service and product to the tourism market. Besides selling travel products to downstream firms and individual customers, its main task is also to design and develop new product, such as inbound and outbound travel products, new destination discovery, place marketing, customization tour, and cooperate with upper stream suppliers to produce and release new product to the market. The NPD procedure of the case firm is shown on Fig. 1.

There are several tasks for planning and operation department, which is responsible for product development and customer relationship management, including:

(1) To collect information in order to investigate possible current and future market opportunities. All information from supply chain is gathering from inside and outside sources.
(2) To study and analyze information in order to figure out market situations, future development and design/develop new product.
(3) To cooperate with partners in order to develop and maintain inbound and outbound supply chains and markets.
(4) To evaluate/test cost and market in order to confirm that new product is workable and profitable. Thus, product and
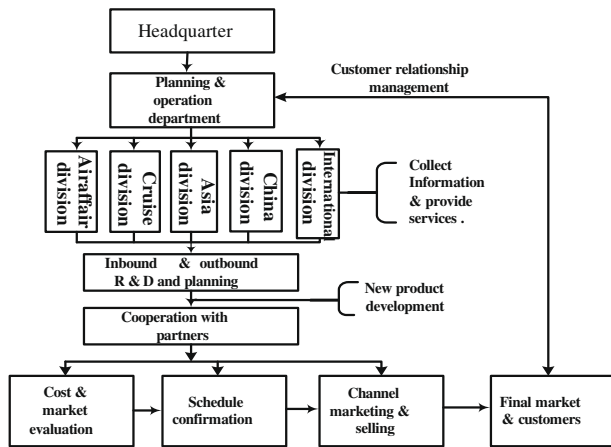
Fig. 1. The product development procedure of the case firm.

product lines are designed and developed for final products and services. In addition, product content and schedule are confirmed precisely.

(5) To release/sell final products and services through channels to market/customers in order to make profit and maintain customer relationship.

Currently, both inside and outside factors are challenging the case firm. This study considers that data mining approach may provide a more active method on enhancing its capability on product development and customer relationship management.

3. Methodology

3.1. Research framework

This paper proposes the association rules and clustering analysis for data mining to extract market knowledge of customers, products/services, and purchased data from a database. Fig. 1 presents the data mining system framework. Using a database, association rules and clustering analysis, this paper analyzes customer preference/purchase factors as well as purchases products, and then extracts relevant knowledge to explore useful information/knowledge of patterns and rules for tourism new product development and customer relationship management. Knowledge extracted from this analysis can serve as useful input for upper management and analysts of planning and operation and marketing departments (see Fig. 2).

3.2. Questionnaire design and data collection

In terms of understanding Taiwanese tourism customers' preference, purchased behavior and their market segmentation, this study designs a questionnaire contains five sections to record customers' basic data, preference data, and purchased data in order to describe relationships between customer profile and product segmentation. Tourism product/service includes inbound/outbound tourism products (group and individual product), tourism ticks booking and sales, resident arrangement/reservation, travel schedule arrangement, travel certification application, travel insurance, travel consulting, and compliant petition. All of sections and items of questionnaire are tested their content validity and expert validity through examination and modification by manager and senior staff of operational and planning department.

Between June 1 and August 31 of 2006, customer data were collected at the case firm headquarters located in Taipei, the Capital of

ROC. Due to the limitation of Personal Data Protection Act in Taiwan; the case firm is prohibited to release customer data to this research. Thus, interview with questionnaire is the only way to collect data from the case firm. By conducting interviews at the case firm and during various days of week, this research tried to collect data as randomly as possible. In addition, a total of 1400 samples were collected from the case firm in August, 2006 and 1050 of them were identified as valid. Subject data describe on the Table 1.

In Table 1, female subject number is greater that male. Mostly, subjects' age is concentrating on 20–40 years old (67/3%). 77.9% subjects' educational status is with second degree or higher degree in University. It means that subject's background is belonging to skilled workers. Single male/female shares majority marriage status. Manufacture and service industries are two main job categories of subjects, share 18% and 17.05 in samples, respectively. Finally, most subjects are live in the north area of Taiwan, where is the wealthy place in this Island.

3.3. Relational database design

The concept of the relational database was first developed in the 1970s by Codd to represent interrelated data in the form of a table. The representation of data in an interrelated table hence becomes the main characteristic of the relational database Relational databases organize data as a collection of tables in which all data relationships are represented by common values in related tables. These databases can relate data stored in one table to data in another, as long as the two tables share a common data element. The tables appear similar to flat files, but the information in more than one file can be easily extracted and combined with Structured Query Language (SQL), which is the standard data manipulation language for relational database management system. Some research articles have shown that the association rules of a relational database can provide a useful method for mining knowledge on different application areas (Thabtah, Cowling, & Hammoud, 2006). In this study, the relational database contains 20 entities, 19 relationships, and 74 attributes (Fig. 3).

3.4. Association rule – Apriori algorithm

As stated in Agrawal, Imilienski, and Swami (1993), discovering association rules is an important data-mining problem, and there has been considerable research in using association rules in the field of data-mining problems. The association rule algorithm is employed mainly to determine the relationships between items or features that occur synchronously in the database. For instance, during a trip to the shopping center, if the people who buy item X also buy item Y as well, there exists a relationship between item X and item Y. Such information is useful for decision-makers. Therefore, the main purpose of implementing the association rule algorithm is to obtain synchronous relationships by analyzing the random data and to use these relationships as reference during decision-making. The association rules are defined as follows (Wang, Chuang, Hsu, & Keh, 2004).

$I = \{i_1, i_2, \ldots, i_m\}$ denotes the item set, in which each item represents a specific literal. D stands for a set of transactions in a database in which each transaction T represents an item set such that $T \subseteq I$. That is, each item set T is a non-empty sub-item set of I. An *association rule* is an implication of the form $X \rightarrow Y$, where $X \subset I$, $Y \subset I$ and $X \cap Y = \Phi$. The rule $X \rightarrow Y$ holds in the transaction set D according to two measure standards – *support* and *confidence*. *Support* (denoted as $Sup(X, D)$) represents the rate of transactions in D containing the item set X. *Support* is employed to evaluate the statistical importance of D, and the higher its value, the more important the transaction set D is. Therefore, the rule $X \rightarrow Y$ has
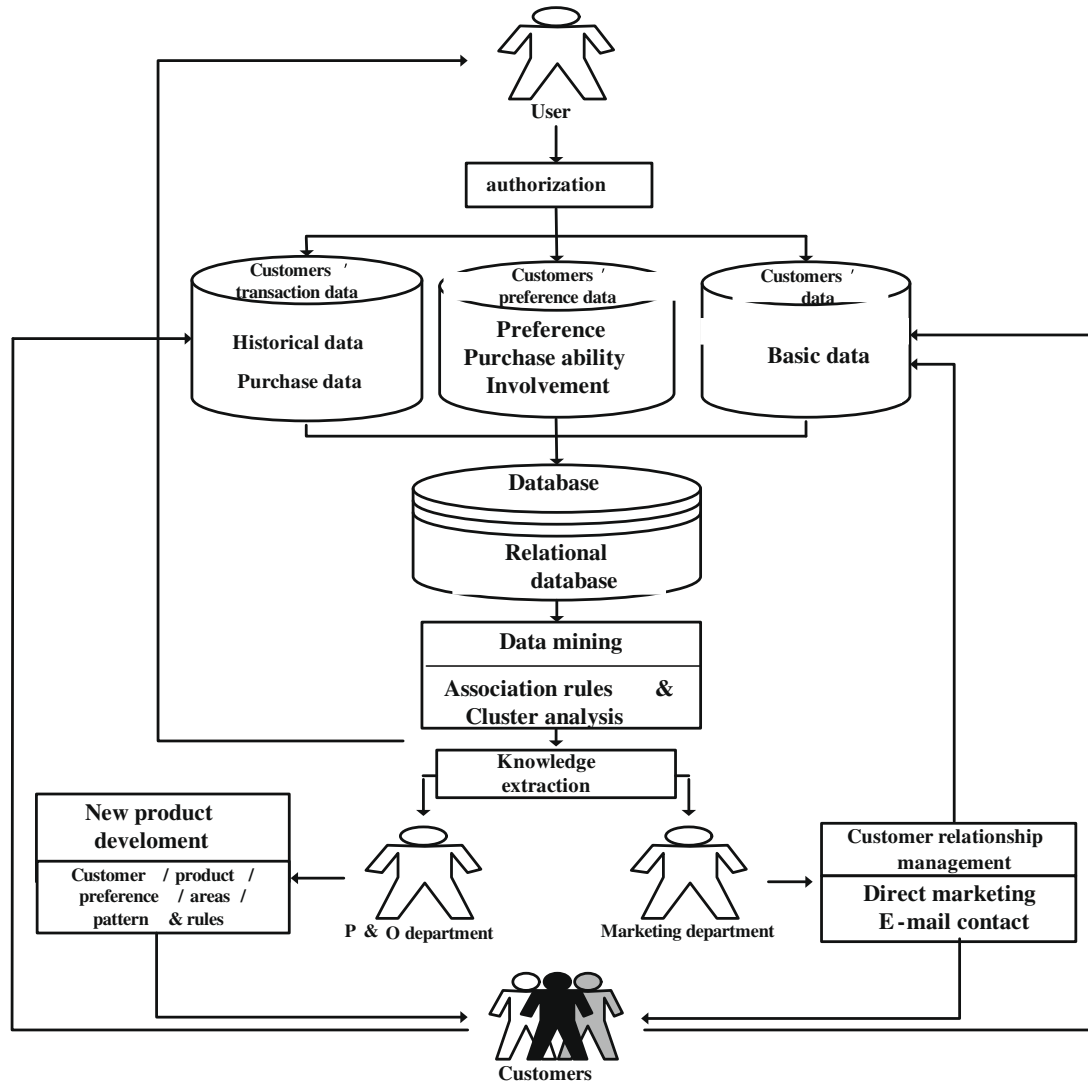
Fig. 2. System framework.

support. Sup(X ∪ Y,D) represents the rate of transactions in *D* containing *X* ∪ *Y*. Each rule *X* → *Y* also has the other measuring standard called *Confidence* (denoted as *Conf(X → Y)*), representing the rate of transactions in *D* containing both *X* and *Y*. That is, *Conf(X → Y) = Sup(X ∩ Y)/Sup(X,D)*

In this case, *Conf(X → Y)* denotes that if the transaction includes *X*, the chance that the transaction also contains *Y* is relatively high. The measure *confidence* is then employed to evaluate the level of confidence about the association rule *X* → *Y*. Given a set of transactions *D*, mining association rules is to generate all transaction rules that have certain user-specified minimum *support* (called *Minsup*) and *confidence* (called *Minconf*). According to Agrawal and Srikant (1996), mining association rules involves into two steps. The first step is to detect a large item set whose *support* is greater than *Min*sup and the second step is to generate association rules using the large item set. Such rules must satisfy the following two conditions:

1. *Sup(X ∪ Y,D)* ⩾ *Minsup*.
2. *Conf(X → Y)* ⩾ *Minconf*.

In order to reduce the possible biases incurred when using these measure standards, the simplest way to judge the standard is to use the *lift* judgment. *Lift* is defined as: *Lift = Confidence(X → Y)/ Sup(Y)* (Wang et al., 2004).

3.5. Clustering analysis

Clustering is a widely used technique, whose goal is to partition a set of patterns into disjoint and homogeneous clusters. According to Boutsinas and Gnardellis (2002), clustering algorithms have been widely studied in various fields including machine learning, neural networks and statistics. The *K*-means algorithm, hill-climbing and the density-based DBSCAN are the most popular partition-clustering methods. The goal of the *K*-means algorithm is to partition the data into *K* clusters so that the within-group sum of squares is minimized. There are many variants of the *K*-means algorithm that improve its efficiency in terms of reducing the computing time and achieving a smaller error. The Two-step can be employed to cluster the dataset into distinct groups when the initial status of these groups is unknown. The first step makes a single pass through the data, during which it compresses the raw input data into a manageable set of sub-clusters. The second step uses a hierarchical clustering method to progressively merge the sub-clusters into larger and larger clusters, without requiring another pass through the data. Hierarchical clustering has the advantage

Table 1
Subject data.

Items	Variables	Count	Percentage	Accumulated percentage
Sexuality	Male	445	42.38	42.38
	Female	605	57.62	100.0
Age (years)	Below 12	1	0.1	0.1
	13–19	3	0.29	0.3
	20–29	310	29.52	29.82
	30–39	396	37.71	67.53
	40–49	237	22.57	90.1
	50–59	89	8.48	98.57
	Above 60	14	1.33	100.0
Education	Elementary school	10	0.95	0.95
	Junior school	36	3.43	4.38
	High school	186	17.71	22.09
	College	266	25.33	47.42
	University	412	39.24	86.67
	Graduate school	140	13.33	100.0
Job category	Farmer and fisherman	15	1.43	1.43
	Manufacture	189	18	19.43
	Construction industry	94	8.95	28.38
	Service industry	179	17.05	45.43
	Information technology	103	9.81	55.24
	Finance industry	172	16.38	71.63
	Official and officer	70	6.67	78.3
	Doctor and lawyer	25	2.38	80.68
	Media	20	1.9	82.6
	Work at home	30	2.86	85.46
	Wait for employment	4	0.38	85.86
	Housekeeper	37	3.52	89.38
	Student	78	7.43	96.81
	Retired	18	1.71	98.52
	Others	16	1.52	100.0
Marriage	No	474	45.14	45.14
	Yes (without children)	139	13.24	58.38
	Yes (with children)	436	41.52	99.9
	Others	1	0.1	100.0
Living area (Taiwan)	North	710	67.62	67.62
	East	107	10.19	77.81
	Central	148	14.1	91.91
	South	78	7.43	99.34
	Remote islands	6	0.57	99.9
	Others	1	0.1	100.0

of not requiring the numbers of clusters to be selected ahead of time. The initial pre-clustering of Two-step makes hierarchical clustering fast even for large datasets (SPSS Inc., 2003).

4. Research results

4.1. New product development

4.1.1. Travel area – inbound travel (pattern A)
4.1.1.1. Inbound travel association analysis. Four association rules with lift value ranging from 1.39 to 2.87 were obtained with minimum rule support of 14% and minimum rule confidence of 42% using the variables of the number of respondents describing different inbound tourism products (Fig. 4 and Table 2).

Research results show that main factors considered by customers on inbound travel including security, travel agency reputation, travel style, and acceptable price. In these regards, the case firm can figure out how to maintain its reputation, select safety travel destination and place, design vivid travel style (such as remote island tour, natural tour, and hot spring and SPA tour), and package with acceptable travel fee are considerations for inbound travel NPD.

4.1.1.2. Inbound travel cluster analysis. Cluster analysis was then performed to compare and verify the obtained rules. Five clusters
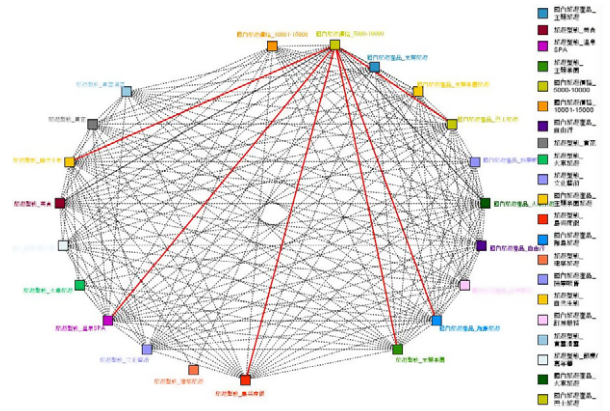

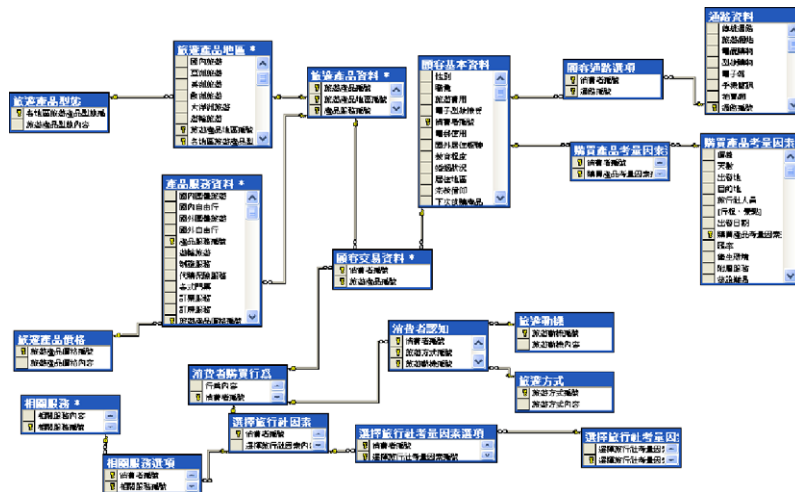**Fig. 4.** Association diagram of inbound travel new product development.

Fig. 3. Physical relational database.

Table 2
Association rules for inbound travel – pattern A (*Minsup* = 10%; *Minconf* = 55%).

Rule	Lift	Support		Confidence	Consequent	Antecedent
R_{A1}	1.66	10.5	60.9	Preference factor = Security	Preference factor = Reputation of agency	Travel style = Remote island holiday	Price = NT 5000–10,000 dollars
R_{A2}	1.59	12.5	58	Factors = Security	Preference factor = Reputation of agency	Travel style = Natural tour	Price = NT 5000–10000 dollars
R_{A3}	1.56	11.5	57	Preference factor = Security	Preference factor = Reputation of agency	Travel style = Hot spring and SPA tour	Price = NT 5000–10,000 dollars
R_{A4}	1.52	12.7	55.6	Preference factor = Security	Inbound travel product = Bus tour	Travel style = Remote island holiday	Price = NT 5000–10,000 dollars
R_{A5}	1.38	11.1	65	Travel style = Natural tour	Preference factor = Reputation of agency	Preference factor = Security	Price = NT 5000–10,000 dollars
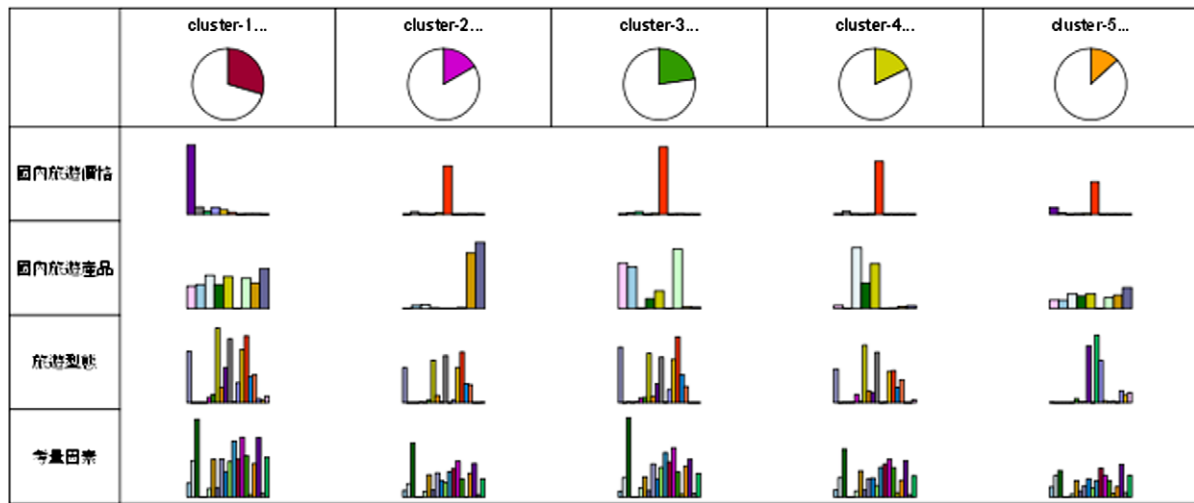
Fig. 5. Clusters of pattern A.

Table 3
Clusters of pattern A.

Cluster	CA₁	CA₂	CA₃	CA₄	CA₅
Percentage	29%	16%	22%	17%	13%
Product category of inbound travel	Remote island holiday (D3) (17.22%)	Remote island easy tour (D3) (50.8%)	Easy tour (D6) (17.22%)	Bus tour (D1) (43.72%)	Remote island tour (D3) (20.25%)
Preference factors	Price (B1) (12.67%)	Price (B1) (15.49%)	Price (B1) (16.85%)	Price (B1) (12.9%)	Environmental hygiene (B7) (11.77%)
Travel style	Remote island tour (A2) (16.19%)	Natural tour (A14) (19.53%)	Natural tour (A14) (18.64%)	Remote island easy tour (A2) (20.74%)	Train and bus service (A4) (32.75%)
Accepted price (NT dollars)	10,001–15,000 (C3) (74.98%)	5000–10,000 (C1) (92.89%)	5000–10,000 (C1) (95.16%)	5000–10,000 (C1) (94.82%)	5000–10,000 (C1) (78.51%)

were obtained from the Two-step analysis, including remote island tour, easy tour, train/bus tour, environmental hygiene, and price (customer preference factors). Clustering analysis results show that the case firm can design a new product by considering environmental hygiene situations of remote islands and possible natural style, train/bus tour, and hot spring and SPA product with accepted fee (NT 5000–10,000 dollars) for a travel product (Fig. 5 and Table 3).

Comparing results of clustering analysis and association rules, accepted price for customer inbound travel is NT 5000–10,000 dollars with remote island travel product. In addition, comparing Tables 2 and 3, associations of remote island tour, environmental hygiene, and train/bus tour for a package product with NT 5000–10,000 fee is a new rule for possible NPD. Thus, the case firm can cross-examine the results of association rules and clusters in order

to find and check anomalous or unknown pattern or rule. By doing so, the case firm can see a clear picture on customer profile and product/service segmentation.

After that, this study integrates analysis results of two data mining approach illustrating a knowledge map in terms of describing complete pattern and rule knowledge of inbound travel analysis (Fig. 6). Fig. 6 shows a knowledge map for a new product development for customers with considering different factors/variables. In the knowledge map, there are totally four different cells, which correspond to diverse factors/variables of customer profile and product segmentation. After specific association analysis, certain decision variables will be combined and collected in the form of a pattern. Different patterns will develop different rules. As shown in Fig. 6, the knowledge map contains ten rules available for further NPD considerations.
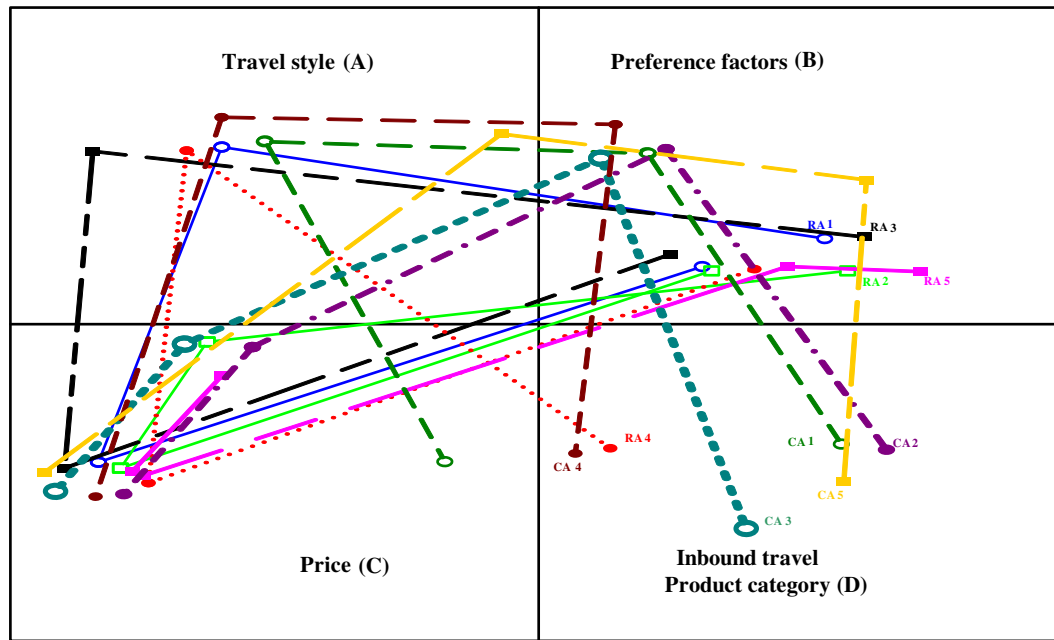
Fig. 6. Knowledge map of inbound travel new product development.

Thus, in the process of new product development, it is suggested to develop at least ten inbound travel products. Thus, these travel products should be the most preferable considerations according to the NPD knowledge map. In addition, similar types of products may be assorted for others in terms of product line extensions. For instance, for travel products, different types such as travel style, customer preference factors, product category, and price are possible product mix acceptable by different customers. Therefore, the diversity of product lines and mixes should be considered as an important issue during the process of new product development.

4.1.2. Travel area – outbound travel – Asia (pattern B)
4.1.2.1. Outbound travel association analysis.
In terms of Asia area, with lift value 1.62 and mini confidence 59 for Apriori algorithm analysis, this study finds a association pattern with three association rules, such as Thai, Buddha temple tour, theme park holiday, acceptable product price for NT 10,000–20,000 dollars, security, and environmental hygiene for preference factors (Table 4).

According to Table 4, main factors considered by customers on Asia area travel include security, environmental hygiene, been travel country, travel style, and acceptable price. In these regards, the case firm can figure out how to develop its regular Asia travel customers, select safety travel destination and place, care about environmental hygiene, design depth and carefree travel style (such as Buddha temple tour and theme park holiday), and package with acceptable travel fee are considerations for outbound travel NPD.

4.1.2.2. Outbound travel cluster analysis – Asia area.
Three clusters were obtained from clustering analysis, including Japan, Thai, and security and hygiene (customer preference factors). Clustering analysis results show that the case firm can design a new product by considering environmental hygiene and security situations of Buddha temple tour, gourmet tour to Japan or Thai with accepted fee (NT 10,000–20,000 dollars) for a package product (Table 5). In practice, Thai travel is a highly price competitive tourism product in Taiwan. In this regard, pricing is always chaos on the market and quality as well as reputation is usually denounced by customers. Thus, the case firm should maintain its reputation and travel quality on Thai travel for NPD consideration.

Comparing results of clustering analysis and association rules, accepted price for customer Asia area travel is NT 10,000–20,000 dollars with Thai travel and security consideration are identical analysis results. In addition, considering three clusters of pattern B, associations of Japan and hop spring tour, Thai and great mountain tour, and Thai and shopping tour with reputation and security consideration for NT 20,000–50,000 fee are new rules for possible NPD.

In Taiwan, mostly, travel disputes are happened in low price product. In contrast with Japan travel, Thai travel belongs to low price travel destination. Therefore, works on quality and reputation could help the case firm to gain customer loyalty. On the other hand, in this study, Japan and Thai are found for two favor travel countries with some customer profile and product segmentation. This means that the case firm can develop new product aiming

Table 4
Association rules for Asia area – pattern B (*Minsup* = 5%; *Minconf* = 55%).

Rule	Lift	Support	Confidence	Consequent	Antecedent			
R_{B1}	1.62	5.8	59	Preference factor = Security (B10)	Travel style = Buddha temple tour (A4)	Travel style = Theme park holiday (A3)	Price = 10,001–20,000 (C2)	Been travel Asia country = Thai (D4)
R_{B2}	1.57	5.2	57.4	Preference factor = Security (B10)	Travel style = Buddha temple tour (A4)	Price = 30,001–40,000 (C4)	Been travel Asia country = Japan (D2)	Been travel Asia country = Thai (D4)
R_{B3}	1.54	6	55.6	Preference factor = Hygiene (B7)	Travel style = Gourmet tour (A1)	Price = 30,001–40,000 (C4)	Price = 10,001–20,000 (C2)	Been travel Asia country = Japan (D2)

Table 5
Clusters of pattern B.

Cluster	CB$_1$	CB$_2$	CB$_3$
Percentage	48%	30%	20%
Been travel Asia country	Japan (D2) (18.32%)	Thai (D4) (17.06%)	Thai (D4) (19.64%)
Preference factors	Security (B10) (11.39%)	Reputation (B11)(16.09%)	Security (B10) (13.7%)
Travel style	Hot spring tour (A6) (14.78%)	Great mountain tour (A7) (19.37%)	Shopping (A2) (15.35%)
Accepted price (NT dollars)	20,000–30,000 (C3) (48.11%)	40,000–50,000 (C5) (32.47%)	10,000–20,000 (C2) (87.24%)

these two countries' segmentation by providing differential product and service.

4.2. Customer relationship management

4.2.1. Travel service

4.2.1.1. Travel service association analysis (pattern C). In service industry, service is the core of a product. Customer perceives value through service (satisfaction) and pays the cost for a product. A firm's value proposition is the set of benefits or values it promises to deliver to customers to satisfy their needs and wants. Thus, how to provide good service becomes the way to maintain and develop customer relationships. There are four association rules were found with minimum rule support of 3% and minimum rule confidence of 42% (Table 6). The pattern C contains preference service, job category, age, marriage status, annual travel expense, education, religious, and living area for association rules.

Table 6 shows that main factors considered by customers on travel service include free process/delivery of paper work and travel consulting on the customer profile of male/female, college degree, single or married couple with children, age between 20 and 39, and living the north area of Taiwan. It means that paper work convenience and expertise consulting on middle-class customer are the critical factors for perceiving service preference. In these regards, the case firm can figure out how to provide free process/delivery and consulting services to this customer profile for possible CRM.

4.2.1.2. Travel service cluster analysis. Two clusters have been found by clustering analysis, including free process/delivery paper work, and travel digital map (Table 7). Cluster 1 shows that male who works in service industry, married with children, needs free process/delivery service for paper work. This means that service which save time and space will be convenient to this customer profile. On the other hand, single female students prefer having travel digital map for travel service. With lower annual travel expense than cluster 1, easy travel or pack tour may become the travel style of cluster 2 customers. Thus, travel digital map service becomes another convenient preference service to this customer profile.

Table 7
Clusters of pattern C.

Cluster	CH$_1$	CH$_2$
Percentage	60%	40%
Preference service	Free delivery (I1) (24.91%)	Travel digital map (I3) (18.37%)
Marriage status	Married with children (G3) (60.55%)	Single (G1) (97.34%)
Religious	Buddhism (E4) (28.91%)	None (E7) (57.71%)
Living area	North area (F1) (65.54%)	North area (F1) (87.78%)
Age	30–39 (B4) (53.01%)	20–29 (B3) (83.59%)
Sexuality	Male (A1) (51.3%)	Female (A2) (67.72%)
Education	University (C5) (36.39%)	University (C5) (45.9%)
Annual travel expense (NT dollars)	20,000–40,000 (H3) (33.68%)	10,000–20,000 (H2) (31.25%)
Job category	Service business (D4) (18.64%)	Student (D13) (23.49%)

From Tables 6 and 7, job and education factor influencing customer service are new found on clustering analysis. Cluster 1 and 2 are new pattern for customer service preference. Thus, free paper work process/delivery service, travel consulting, and travel digital map are important service items for CRM considerations on these customer profiles.

Fig. 7 shows a knowledge map of customer service for customers with different pattern and rules. In the knowledge map, there are totally nine different cells, which correspond to different factors of customer profile and preference service. As shown in Fig. 7, the knowledge map contains six rules available for further CRM considerations.

4.2.2. Direct marketing

4.2.2.1. Travel web site usage association analysis (pattern D). Direct marketing is direct communications with targeted customers to obtain immediate response and give it interactive feedback. Recent technological advances, including the widespread use of the Internet, have created what some have called a New Economy. A few would disagree that the Internet and other powerful new connect-
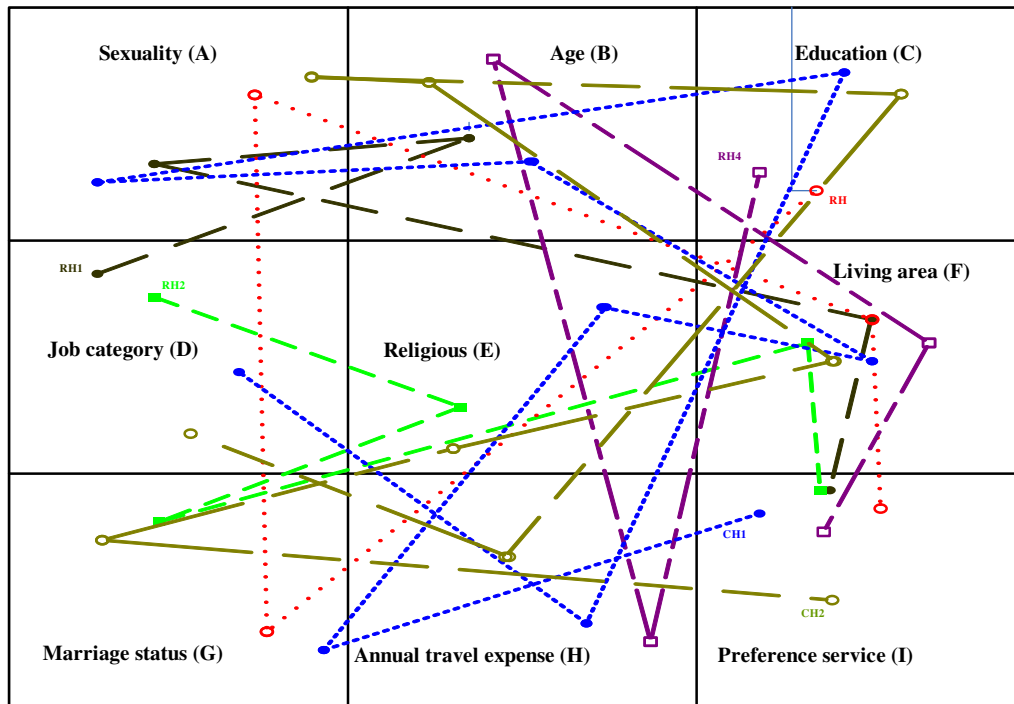
Table 6
Association rules for travel service – pattern C (*Minsup* = 3%; *Minconf* = 30%).

Rule	Lift	Support	Confidence	Consequent	Antecedent		
R$_{H1}$	1.5	3.1	32.8	Preference service = Free delivery(I1)	Job category = Manufacture worker (D2)	Age = 30–39 (B4)	Sexuality = Male (A1)	Living area = North area (F1)
R$_{H2}$	1.48	3.2	32.3	Preference service = Free delivery (I1)	Job category = Manufacture worker (D2)	Religious = None (E7)	Marriage status = Single (G1)	Living area = North area (F1)
R$_{H3}$	1.46	3.8	31.8	Preference service = Consulting (I2)	Education = College(C4)	Marriage = Yes with children (G3)	Sexuality = Female (A2)	Living area = North area (F1)
R$_{H4}$	1.38	3.1	30	Preference service = Consulting (I2)	Education = College (C4)	Annual travel expense = 20,000–40,000 (H3)	Age = 20–29 (B3)	Living area = North area (F1)

Fig. 7. Knowledge map of customer service.

ing technologies are having a dramatic impact on business and customers. On this web trend, e-business model for tourism industry becomes a necessary development not only contact with customers but also develop new strategies and practices better suited to new environment. Travel web site is a window, which communicates between sellers and buyers, and becomes one of major tourism product/service channels and offers a method for direct marketing. There are two association rules were found with minimum rule support of 7% and minimum rule confidence of 50% (Table 8). The pattern D contains communication tool, Internet service, age, marriage status, and sexuality for association rules.

Table 8 shows that main factors considered by customers for direct marketing include telephone for communication, e-mail and MSN for main Internet service on the customer profile of married couple with children, and age between 40 and 49 years old. It means that MSN and e-mail are major Internet service for this customer profile to receive and deliver message on the Internet. Thus, the case firm can figure out how to communicate with targeted customers to obtain immediate response and give it interactive feedback using MSN and e-mail to customers for possible CRM.

4.2.2.2. Direct marketing cluster analysis. Results show that two clusters are come from clustering analysis, including yes or no MSN users (Table 9). Cluster 1 customer profile shows that single female customer, age between 20 and 29 years old, who has university degree using MSN service and obtains travel information from the Internet. This means that this customer profile is like to

search information on the Internet. MSN and e-mail are main ways for communicating and delivering message to this customer profile. On the other hand, married female with children not using MSN with university degree, age between 30 and 39 years old, and obtains travel information from friends or relatives. With higher annual travel expense than cluster 1, mature female customer has better purchase capability with less Internet service usage than female customers of Cluster 1. Thus, there are two possible patterns of customers for direct marketing.

Source of travel information is new clue on clustering analysis results. Although cluster 1 customers spend less annual travel expense than cluster 2 customers, cluster 1 customers are the targeted segmentation for the case firm to implement more active way to communicate message through Internet. On the other hand, customers those who are not MSN users, and rely on travel information from their friends and relatives using e-mail and spending more annual travel expense are another important segmentation on direct marketing. Thus, paper catalog, electronic catalog via e-mail is a possible way to communicate with them.

5. Discussion

5.1. In the regard of current market strategy

This study mines six patterns of inbound and outbound travel product considering customers' preference factors, and purchased

Table 8
Association rules for direct marketing – pattern D (*Minsup* = 7%; *Minconf* = 50%).

Rule	Lift	Support	Confidence	Consequent	Antecedent			
R_{J1}	1.35	7.5	53.4	Communication tool = Telephone (I2)	Age = 40–49 (B5)	Marriage status = Married with children (D3)	Internet service = MSN (J1)	Internet service = e-mail (M1)
R_{J2}	1.35	8.3	52.6	Communication tool = Telephone (I2)	Marriage status = Married with children (D3)	Internet service = MSN (J1)	Sexuality = Female (A2)	Internet service = e-mail (M1)

Table 9
Clusters of pattern D.

Cluster	CJ$_1$	CJ$_2$
Percentage	44%	56%
MSN user	Yes (J2) (51.74%)	No (J1) (54.16%)
Sexuality	Female (A2) (60.02%)	Female (A2) (51.19%)
Age	20–29 (B3) (82.37%)	30–39 (B4) (57.19%)
Education	University (C5) (52.44%)	University (C5) (44.71%)
Marriage status	Single (D1) (95.13%)	Married with children (D3) (50.77%)
Religious	None (E7) (62.8%)	Buddhism (E4) (34.38%)
Living area	North area (F1) (89.25%)	North area (F1) (60.27%)
Annual travel expense	10,000–20,000 (G2) (32.79%)	20,000–40,000 (G3) (44.06%)
Next time purchase intension	Outbound-group travel (H3) (43%)	Outbound-individual travel (H4) (34.86%)
How long to collect information before traveling	Before one month (L4) (36.27%)	Before one month (L4) (21.85%)
E-mail user	Yes (M1) (92.81%)	Yes (M1) (95.25%)
Communication tool	Telephone (I2) (32.02%)	Telephone (I2) (45.37%)
Source of travel information	Internet (K5) (22.58%)	Friends and relatives (K2) (16.86%)

product for investigating current strategies of new product development. In Fig. 7, different association rules show that inbound travel contains several customer preference factors with low price. However, security, environmental hygiene, and reputation are still concerned by inbound travel customers, and more efforts of product design on days of travel, and temperature consideration could improve customer preference for inbound traveling. On the other hand, outbound travel product including America travel, Europe travel, cruise travel, Asia travel, and Oceania travel are diversity in customer preferences. With higher price than inbound travel, higher requirements of preference factors are demanded by customers. Europe travel, cruise travel, Oceania travel, and America travel are belonging to high price travel product. The case firm can design and develop outbound travel product according to different product segmentations. In addition, in Fig. 8, schedule is the critical preference factor considered by customers both on inbound and outbound travel product. This means that schedule arrangement plays a critical role not only on pricing but also on customer satisfaction. Thus, the case firm should figure out travel schedule planning for the current NPD strategy.

5.2. In the regard of future market strategy

This study mines a pattern of future market for the case firm. This pattern is a potential market in Europe. In Fig. 9, association rule shows that based on security, environmental hygiene, temperature, reputation, and with low price, there is a new market segmentation considered by outbound customers. This segmentation

of customers is less concern on travel schedule and destination of country; however, more concern on safety and product quality is their considerations. To check out the customer profile, they are single male/female, university degree, age between 30 and 39 years old, working in service industry, using MSN and e-mail, with annual travel expense NT 10,000–20,000 dollars, and lining the North area in Taiwan. In terms of developing young/single individual customer, the case firm can design and develop new Europe travel product and maintain its customer relationships with electronic ways.

5.3. In the regard of customer value and satisfaction

Human needs are states of felt deprivation. For example, physical needs for food, clothing, shelter, and safety. Individual needs seek for knowledge, esteem, and self-expression. These needs were not created by producers or marketers; they are a basic part of the human makeup. Human wants are the form human needs take as they are shaped by culture and individual personality. They are shaped by one's society and are described in terms of objects that will satisfy needs. Therefore, producers or marketers can seek what object of customers' wants in order to satisfy their needs. In this regard, how to find the object of customers' wants becomes a critical task for businesses to generate customer value and satisfaction. Tourism is an object of some customers' wants due to different group/individual needs. Thus, this paper investigates what factors/attributes are the customers' preference needs and wants for travel products/services by extracting specific knowledge pattern
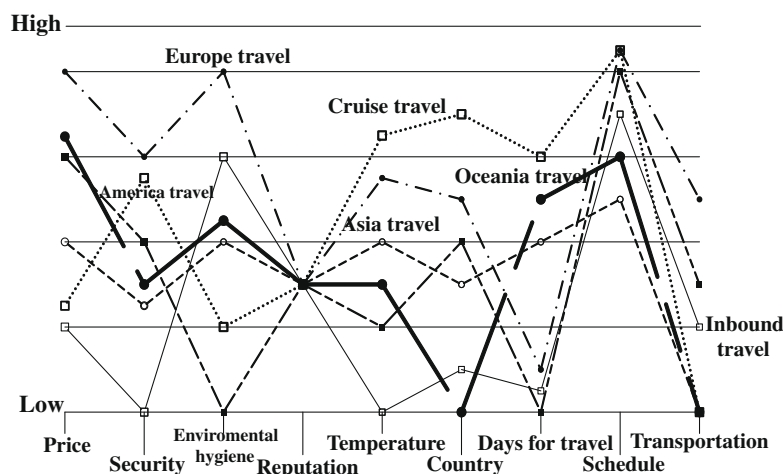

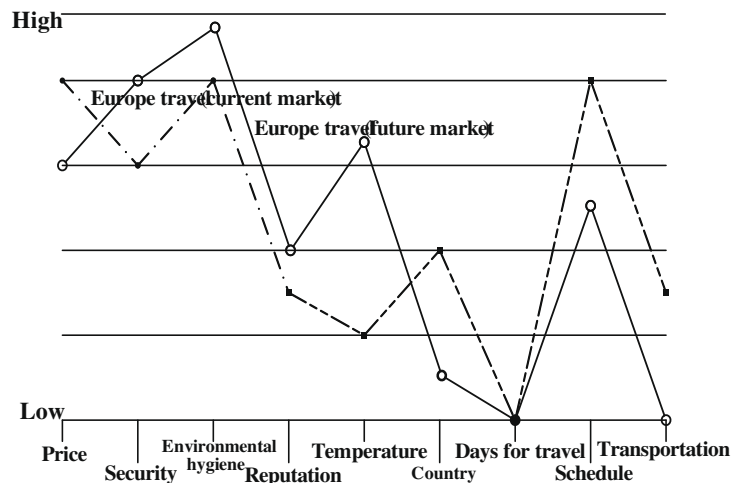**Fig. 8.** Current strategic map for the case firm.

Fig. 9. Future strategic map for the case firm.

and rules from customers. By doing so, this article demonstrates a knowledge extraction approach in order to examine that knowledge of the customers and the product itself reflects the needs of the market.

5.4. In the regard of new business model

On the other hand, once customer knowledge been extracted from market, different marketing strategies and tactics can implement on the case firm. For example, direct marketing and selling might be an option to the case firm. Could paper/electronic catalog marketing become one kind of direct marketing approach for marketers or salespersons on case firm when their customer profile and product segmentation of customers and needs of product are targeted? Thus, this means that case firm can design paper/electronic catalog according to customers' knowledge to sell and promote its products and service both on channel and via Internet. In addition, electronic catalogs may become another kind of mobile marketing and selling method for the case firm to mine customer knowledge regarding specific groups of customers who are accepting catalogs and coupons from their mobile phones, PDA, and computers. Thus, in this study, mining customer knowledge for direct marketing is an example of implementing a database marketing approach for increased connectivity and use of information technology, and creating internal data mining capability for analysis and support of NPD and CRM.

6. Conclusion

Customers' needs and wants are a sensitive and complicated, if a firm can understand them and make efforts to fulfill their wants and provide friendly service then the customer will be more supportive to the enterprise. During the process of developing from the product concept to the actual product itself, the customer can only passively receive new information, and can only select from the products that are currently on sale in the market. No matter which type of product, the customer cannot individually come up with a product concept and then develop it. Furthermore, buying what is available on the market does not mean that customers are satisfied with the current product, because the customers' preferences and experiences were not considered in developing the product so they can only accept the product as it is. As a result, tourism firm has responsibility to develop products that fulfill the customers' needs and wants, as this will increase the tourism firm's competitiveness and it is an essential criterion to earning

higher profits. This paper proposes Apriori algorithm as a methodology of association rule and clustering analysis for data mining, which is implemented for mining customer knowledge from the case firm. Knowledge extraction from data mining results is illustrated as knowledge patterns, rules, and knowledge maps in order to propose suggestions and solutions to the case firm for NPD and CRM.

Acknowledgement

References

Agrawal, R., & Srikant, R. (1996). Fast algorithms for mining association rules. In *Proceedings of the 20th international conference on very large database* (pp. 487–499).

Agrawal, R., Imilienski, T., & Swami, A. (1993). Mining association rules between sets of items in large database. In *Proceedings of the 1993 ACM SIGMOD international conference on management of data* (pp. 207–216).

Boutsinas, B., & Gnardellis, T. (2002). On distributing the clustering process. *Pattern Recognition Letters, 23*(10), 999–1008.

Bramwell, B. (1998). User satisfaction and product development in urban tourism. *Tourism Management, 19*(1), 35–47.

Daniel, E., Wilson, H., & McDonald, M. (2003). Towards a map of marketing information systems: An inductive study. *European Journal of Marketing, 37*(5), 821–847.

Holmlund, M., & Strandvik, T. (1999). Perception configurations in business relationships. *Management Decision, 37*(9), 686–696.

Junping, D., Min, Z., & Xuyan, T. (2008). The realization of distributed sampling association rule mining algorithm in tourism. In *Proceedings of the world congress on intelligent control and automation (WCICA)* (pp. 183–187).

Law, R., Bauer, T., Weber, K., & Tse, T. (2006). Towards a rough classification of business travelers. *Lecture Notes in Computer Science, 4093*, 135–142.

Lawson-Body, A., & Limayem, M. (2004). The impact of customer relationship management on customer loyalty: The moderating role of web site characteristics. *Journal of Computer-Mediated Communication, 9*(4)<http://www.ascusc.org/jcmc/vol9/issue4/lawson_body.html>. Retrieved 20.11.04.

Liao, S. H., Chang, W. J., & Lee, C. C. (2008). Mining marketing maps for business alliances. *Expert Systems with Applications, 35*(3), 50–62.

Liao, S. H., & Chen, Y. J. (2004). Mining customer knowledge for electronic catalog marketing. *Expert Systems with Applications, 27*(2), 521–532.

Liao, S. H., Chen, Y. N., & Tseng, Y. Y. (2009). Mining demand chain knowledge of life insurance market for new product development. *Expert Systems with Applications, 36*(2), 9422–9437.

Liao, S. H., Chen, C. M., & Wu, C. H. (2008). Mining customer knowledge for product line and brand extension in retailing. *Expert Systems with Applications, 35*(3), 1763–1776.

Liao, S. H., Hsieh, C. L., & Huang, S. P. (2008). Mining product maps for new product development. *Expert Systems with Applications, 34*(1), 50–62.

Liao, S. H., & Wen, C. H. (2009). Mining demand chain knowledge for new product development and marketing. *IEEE Transactions on Systems, Man, and Cybernetics – Part C: Applications and Reviews, 39*(2), 223–227.

Lin, X. (1997). Map displays for information retrieval. *Journal of the American Society for Information Science, 48*(1), 40–54.

Piccoli, G., O'connor, P., Capaccioli, C., & Alvarez, R. (2003). Customer relationship management a driver for change in the structure of the US lodging industry. *Cornell Hotel and Restaurant Administration Quarterly, 61*, 61–73.

Seddighi, H. R., & Theocharous, A. L. (2002). AA model of tourism destination choice: A theoretical and empirical analysis. *Tourism Management, 23*(5), 475–487.

Sevki, Ö., & Rifat, I. (2006). Customer relationship management in small–medium enterprises: The case of Turkish tourism industry. *Tourism Management, 27*(6), 1356–1363.

SPSS Inc. (2003). *Clementine user's guide*. Chicago: SPSS Inc.

Thabtah, F., Cowling, P., & Hammoud, S. (2006). Improving rule sorting, predictive accuracy and training time in associative classification. *Expert Systems with Applications, 31*(2), 414–426.

Tülin, E., & Russell, W. (1998). Econometric modeling of competition: A multi-category choice-based mapping approach. *Journal of Econometrics, 89*(1–2), 159–175.

Wang, Y. F., Chuang, Y. L., Hsu, M. H., & Keh, H. C. (2004). A personalized recommender system for the cosmetic business. *Expert Systems with Applications, 26*(3), 427–434.

Wickramasinghe, L. K., Amarasiri, R., & Alahakoon, L. D. (2004). A hybrid intelligent multi-agent system for e-business. *Computational Intelligence, 20*(4), 603–623.

Witt, S. F., & Witt, C. A. (1995). Forecasting tourism demand: A review of empirical research. *International Journal of Forecasting, 11*(3), 447–475.