

# **APPLICATION OF DATA MINING TECHNIQUES IN ON-LINE TRAVEL INDUSTRY: A CASE STUDY FROM THAILAND**

By

Pongsak Hoontrakul  
Senior Research Fellow  
Sasin of Chulalongkorn University, Bangkok,  
Thailand  
Phayathai Road  
Bangkok  
e-mail: pongsak@hoontrakul.com

and

Sunil Sahadev<sup>1</sup>  
Lecturer  
School of Management, University of Sheffield  
9 Mappin Street, Sheffield: S1 4DT, UK  
Phone: +44 (0) 1142223434  
Fax: +44 (0) 114 222 3348  
e- mail: s.sahadev@shef.ac.uk

---

<sup>1</sup> Corresponding author

# **APPLICATION OF DATA MINING TECHNIQUES IN ON-LINE TRAVEL INDUSTRY: A CASE STUDY FROM THAILAND**

## **Abstract**

**Purpose:**The paper presents the process of segmenting an organization's customers through a process of K-mean clustering and subsequent expert judgment.

**Methodology/Approach:** The paper adopts a case study methodology

**Findings:**The case study illustrates the use of data mining techniques in segmenting and profiling of customers by a travel intermediary in Thailand.

**Research limitations/implications:** Having adopted a case study methodology, the paper limits its inferences to a single organization.

**Practical implications:** the case study suggests a practical methodology that can be used in similar instances. The methodology incorporates both a qualitative phase, followed by a quantitative phase and can easily be applied in similar situations.

**Key words:** K-means clustering, e-intermediary, segmentation, travel and tourism industry, Thailand.

**Introduction:**

Data mining techniques like clustering, decision trees etc. have been widely used for successfully segmenting and targeting customers across various industries. Data mining provides an effective approach to discover and understand patterns in customer behavior thereby helping the decision maker to better group customers. The on-line travel intermediary sector has witnessed a surge in patronage over the past three years. It is estimated that more than 50% of all travel bookings happen on-line in the USA and Europe. The migration towards on-line travel intermediation continues unabated and large online intermediaries like expedia.com, Travelocity.com etc. are expanding fast to cater to the emerging demand. However, being a highly competitive industry, there is a mushrooming of small intermediaries concentrating on niche segments and offering products at a cheaper rate. The competition between the intermediaries and suppliers and the service providers in the on-line travel and tourism booking industry is also quite well known. It is in this context that segmentation and targeting of on-line enquires become very critical for success. All enquiries do not get converted in a transaction since customers have a choice of looking at different sources for competitive rates and offers. It is therefore important for intermediaries to understand what types of enquiries get converted so as to target their offer to increase the enquiry conversion rate. The case study examines a small on-line intermediary in Bangkok, Thailand which maintains a website through which customers can book rooms in hotels in Thailand. Over the past three years of its operation, the intermediary has found that its enquiry conversion rate can be increased through better design of its products so that different customer segments are offered what they are mostly interested in. The following sections describe the basic segmentation approaches adopted for comprehending online enquiry patterns, business model of the firm and the process followed for adopting the data mining technique

**Segmentation studies in On-line marketing literature:**

The on-line marketing strategy literature has of late witnessed a number of studies which consider segmentation related issues (eg. Jainszeewski, 1998; Chen and Cooper 2001; Moe and Fader, 2004a; Moe and Fader, 2004b). Most of these studies attempt or propose

to segment customers in terms of their search behavior. Further, most of these studies look at the search behavior in the context of B2C sites that sell consumer goods like Groceries. Tourism being a information good where, at the time of purchase, the customer is only provided with information about the product and also, since tourism is typically a high value item it can be safely surmised that click stream based segmentation alone will not be able to completely describe customer behavior. Further, the purchase decision for typical tourism product entails a greater amount of risk for the customers. This is because, an a priori comprehensive assessment of the quality is impossible unless you have already experienced the product. Hence, a potential consumer can be expected to indulge in more extensive and intensive search behavior. The process between enquiry and conversion is therefore quite critical and complex when compared to the enquiry conversion process for a low involvement, low value item. This study focuses on the process that happen after the potential customer has visited the website and made the enquiry. Several online segmentation approaches utilize the click through behavior of the visitors to a website. This type of analysis helps to understand the effectiveness of a website in enabling a potential customer to achieve his/her objective in visiting the website. Thus, the click stream data analysis mostly concentrates on the process that happens during the visit of the site. The distinction between the focus of the past studies and the present study is indicated in the figure.1. This study attempts to segment and target customers on the basis of their enquiry patterns and certain important demographic characteristics. While this segmentation approach is applicable even for offline behavior, what distinguishes off line enquiry-to-sale behavior and on-line enquiry-to-sale behavior is the presence of search costs. In offline enquiry-to-sale behavior, the customer will have to search for and then visit or call up different travel agents or tour operators to compare rates and make enquiries about the hotel as well as the availability of accommodation. Online search behavior is typically characterized by almost negligible levels of search costs (Moe and Fader, 2004). The relative ease with which a customer can search, visit, enquire and even negotiate with different on-line travel agents makes the enquiry-to-sale conversion phase a challenging task for the on-line travel agent. The typical customer can be expected to make enquiries with several on-line travel agents and then come to a final decision based on the final offer. In this instance, the potential customer can visit even

the service provider's website directly to look at the rates provided and compare it with the rates mentioned in the website of the on-line travel agent. This puts the on-line travel agent in a disadvantageous position. The website of the on-line travel agent could thus turn out to be just a 'yellow page directory' of hotels in a destination. Therefore from an online travel agent's perspective it becomes important to understand the type of customers who eventually transact with them and those who turn away after making the enquires. Based on this knowledge, the online travel intermediary can possibly alter the product profile by including more hotels with a different pricing range or alter their dynamic pricing strategy to increase conversion rates. The segmentation scheme implemented therefore looks at the nature of enquiries and the possibility of targeting the most suitable ones based on a cluster analysis.

### **The Case of Morethailand.com**

Morethailand.com is a website that offers its customers an opportunity to search and book rooms in more than 1000 hotels in Thailand. Based out of Bangkok, Morethailand.com mostly caters to Independent Owned and Operated (IOO) hotels in Thailand that are otherwise not connected to large hotel chains nor part of a Global Distribution System (GDS). Morethailand.com thus occupies a niche by catering to a customer base that desires exotic and value for money hotel rooms in not so well known destinations in Thailand. Morethailand.com was established in 2003 and has since then expanded its customer base as well as the number of hotels with which it is associated. Morethailand.com and its associated websites receive about 5000 hits a day and more than 70 enquiries a week. Since tourism is a seasonal product, the volume and total value of conversions varies across seasons.

The block diagram shown as Fig.2, describes the process through which a customer makes an enquiry and eventually books a room through morethailand.com. For morethailand.com, the volume of the sale depends on two conversion rates. The first is the clicks to enquiry rates. The website attracts more than 5000 clicks every day however hardly 50 of these clicks translate to enquiries. The conversion rate from clicks to

enquiries is quite difficult to predict. The factors that affect this rate may be the variety and the number of hotels presented in the website etc. On the other hand an even more critical conversion is from the enquiry to sale. Here, the conversion rate is roughly about 50%. However it would be ideal to increase it further to more than 75%. This is because, once the room enquired for is available; there are very few factors that should impede the conversion. Even though some of the enquirers could be casual enquirers or those who just want to check the rates in other websites, a majority of those who enquire are expected to buy the product. Possible factors that could impede the enquiry-to-sale conversion rate includes factors like the availability of the same or similar room at a reduced price in other websites, last minute cancellation of travel plans due to the lack of availability of flight tickets etc. It is however important for morethailand.com to increase the conversion rate since, the company incurs substantial investment in such promotional initiatives like search engine marketing (SEM) and search engine optimization (SEO) so as to attract more clicks and enquiries. Further, since the hotels are bound to offer their rooms to morethailand.com only when morethailand.com offer higher sales, lack of sufficient sales will also lead to reduced interest among hotels on morethailand.com. Presently hotels offer on average about 10 to 15% discount on the walk in price to morethailand.com.

After considering these issues, the management decided to analyze this problem in detail. The problem was looked at from different dimensions. It was of course evident at the outset that the enquiry-to-sale conversion rate varied from one destination to another. Morethailand.com offered hotels in various destinations in Thailand. However most of its sales and enquiries were coming from about four main destinations: (i) Bangkok, (ii) Phuket (iii) Krabi and (iv) Koh Samui. The analysis therefore was carried out across each of these four destinations. Of the four destinations, Bangkok and Phuket was already quite well known to the outside world. Offers for these destinations can be considered as 'search good' (Hoontrakul 2004) since most of the outside world is already familiar with these two destinations. Also, these two destinations offered more choice since they had more hotels and these hotels were also well represented in travel websites. Krabi and Samui were destinations which were relatively underdeveloped in terms of tourist

arrivals. Thus, these destinations can be considered as 'experience goods' since the potential tourists would probably be satisfied with the service only after experiencing the product. This fact could actually affect the enquiry-to-sale process considerably. An analysis based on destinations was also advisable from the point of view of implementation. Once actionable strategies are devised based on the analysis it was more practicable to implement them on a destination by destination basis.

### **The approach**

After considerable discussions with the management it was felt that a purposive segmentation exercise should be carried out that could identify customer categories which are more prone to cancellations after the enquiry phase and then develop effective strategies oriented towards these segments. This method broadly follows the method advocated by Peltier and Schribrowsky (1997), in which they support a segmentation approach to targeting customer groups based on both the needs underlying a purchase and descriptive buyer data. This made sense because presently the company was unable to put in place meaningful strategies to increase conversion rate due to insufficient information regarding the probability of an enquirer to transact. The main problem was therefore reduced to the identification of segments that could be targeted for further actions. The management wanted to better understand the customer in terms of certain important attributes that were considered to be critical in influencing the enquiry-to-conversion decision. Studies on segmentation of tourists have been carried out extensively in the past [For a general review refer Dolnicar(2004)]. These studies mostly seek to identify segments of tourists that have similar preferences and expectations. Several of these studies use psychographic bases (eg. Kastenholz, Davis and Paul, 1999; Sarigolu and Huang, 2005; Keng and Cheng, 1999) while few other studies deploy demographic characteristics and related attributes (eg. Laesser and Crouch, 2006). It is difficult to find a study that use enquiry based attributes for segmentation.

The next step was to identify variables across which the enquiries could be clustered in order to identify the segments. The company maintained a transaction processing system that captured enquiries across several fields. A large number of variables could have been

considered. However, several of these variables were summarily rejected since they evidently had no association with enquiry-to-conversion process. For instance, variables like the currency used, the date of enquiry, the date of check in etc. were considered to be least influential in determining the enquiry-to-sale conversion behavior. After considerable discussion and analysis, six variables were selected for clustering the enquiries. Further, since the data was captured through different scales for each of these variables it was also decided to use a standard categorization for each of these variables for further analysis. The split of variables into its constituent categories was essential to make the clustering solution actionable. The variables were therefore categorized based on the inputs provided by the reservation staff and also the managers who were concerned with interacting with the hotels for developing the offerings. The variables and the constituent categories are provided in table.1

**||Take in Table.1 Variables included in the analysis||**

In each of the clusters that emerged, it was decided to look at the total volume of enquiry as well as the proportion of these enquires which were converted. This was considered to be important in order to understand the attractiveness of each of the segments as well as the relative strength of morethailand.com in targeting these segments. This was carried out by developing opportunity matrices for each of the destinations. These opportunity matrices guided the management in adopting specific strategies for each of the clusters.

In order to obtain the initial solution, K-means clustering was carried out using Clementine for data from the four destinations separately. K-means methods are very popular for partitioning data especially when the data set is large. Dolincar (2003) in her study on the use of cluster analysis found that in sample of 243 differing clustering studies, about 76% of the partitioning studies used k-means clustering. The K-means method is a clustering method used to group records based on similarity of values for a set of input fields. The basic idea is to try to discover k clusters, such that the records within each cluster are similar to each other and distinct from records in other clusters. The K-means cluster algorithm involves an iterative process wherein each record (here an

enquiry) is assigned to the cluster whose center is the closest. Closeness is measured by the squared Euclidian distance

$$D_{ij} = \| X_i - C_j \|^2 = \sum_{q=1}^Q (x_{qi} - c_{qj})^2$$

Where  $X_i$  is the vector of encoded input fields for record  $i$ ,  $C_j$  is the cluster center vector for cluster  $j$ ,  $Q$  is the number of encoded input fields,  $x_{qi}$  is the value of the  $q^{\text{th}}$  encoded input field for the  $i^{\text{th}}$  record, and  $c_{qj}$  is the value of the  $q^{\text{th}}$  encoded input field for the  $j^{\text{th}}$  record. The process involves assigning records to clusters with the nearest centers and then when all the records are grouped, the cluster centers are updated. This process is repeated till we reach the maximum distance from cluster centers indicated by the proximity distance. The exact number of clusters to be extracted has always been a contentious issue (Dolnicar, 2003). Since the clustering method was used for the purpose of developing customer segments in a practical context, we were guided by the advice given by de Kluyver and Whitlark (p.286), “To be managerially relevant, the number of clusters must be small enough to allow strategy development. At the same time, each cluster or segment must be large enough to warrant such strategic attention and to be reachable and defensible against competitors”. The K-means clustering results for each of the destinations are explained in the following sections.

### **Clustering Process for Bangkok enquiries**

Data that was generated from November 2003 to September 2005 were analyzed. There were a total of 593 enquiries for Bangkok during this time. Initially the entire group was divided into 5 clusters. Based on the analysis of the proximity values the five cluster solution was found to be stable. The details of the cluster solution for Bangkok is provided in table.2

#### **|| Take in Table.2 Cluster output for Bangkok||**

In table.2, clusters 1 and 5 pertain to enquiries which had arrived during the low season and clusters 2 and 4 pertain to enquiries that had arrived during the high season. Cluster 3 is partly high season and partly low season. Across all the clusters, the enquiries are for less than 7 days. Typically Bangkok attracts those who want to spend a few days either

on business trips or for traveling to other destinations in Thailand. Also most of the clusters consist of enquiries that are for less than 50\$ per day rooms. And except for cluster 4 almost all the customers book barely less than a week in advance. In fact it is a very interesting insight that for Bangkok, for the cluster when the customers book in advance, the conversion rate is about 100%. In fact it is fairly clear that almost all the five clusters show similar characteristics in terms of the price range, length of stay and number of advance days. The only characteristic that seems to be useful is the season of booking as well as the status of the customer in terms of their loyalty. Cluster 1. seems to be consisting of new customers during low season while cluster 5 comprises of repeat customers who come during the low season. Cluster 2 and cluster 4 are similar since they consist of predominantly repeat customers during the high season. Cluster 3 comprises of repeat customers spread across both high and low season.

If the segment attractiveness is considered in terms of the percentage of the total volume of enquiries, the second cluster seems to have the highest value though the enquiry-conversion rate is the lowest in the second cluster. The enquiry conversion rate is however creditable in the third, fourth and fifth clusters. Cluster 2 has the most number of enquiries in the 50 to 75\$ hotel category. In fact cluster 2 has the highest percentage of enquiries in above 50\$ hotels, which means that they are probably customer who are affluent and are visiting Thailand on the vacation during the high season and would like to stay in a good hotel even if it is for a few days in transit.

Based on the analysis described above an opportunity matrix was created to help morethailand.com to understand the segments that have to be more focused upon. The opportunity matrix looks at two dimensions, the segment attractiveness in terms of the total volume of enquiries and the firm strength which is based on the conversion rate achieved. The opportunity matrix is presented in Figure 3.

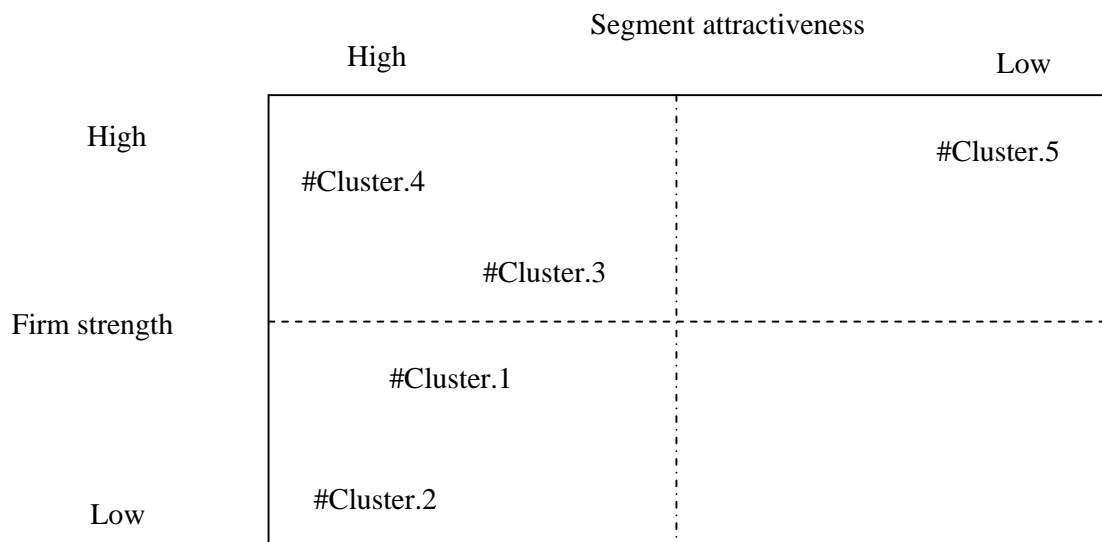


Fig.3 Opportunity matrix for Bangkok

### Cluster output for Phuket Enquiries

In Phuket about 416 enquiries were processed to extract the segmentation model. The analysis of the clusters shows a more clear segmentation. Clusters 1, 4 and 5 are predominantly for short stay and cluster 2 and 3 comprise of enquiries that are typically for long stay. Clusters 2,4 and 5 are dominated by high season enquiries while clusters 1 and 3 are dominated by low season enquiries. Clusters 2 and 3 comprises of enquiries with a lengthier advance booking period while the length of advance booking is high and cluster number 5 is characterized by a large proportion of repeat customers. The price per room does not show much of a variation across the different clusters. When the total volume of enquiries is considered, the conversion rate for cluster2, the most attractive cluster is quite significant, the conversion rate for cluster4, the second most attractive grouping is quite inadequate. The conversion rate for cluster5 is 100% which probably indicates that in Phuket, morethailand.com is able to convert quite a large number of repeat customers. The opportunity matrix for Phuket is given in fig. 4.

**||Take in Table.3 Cluster output for Phuket ||**

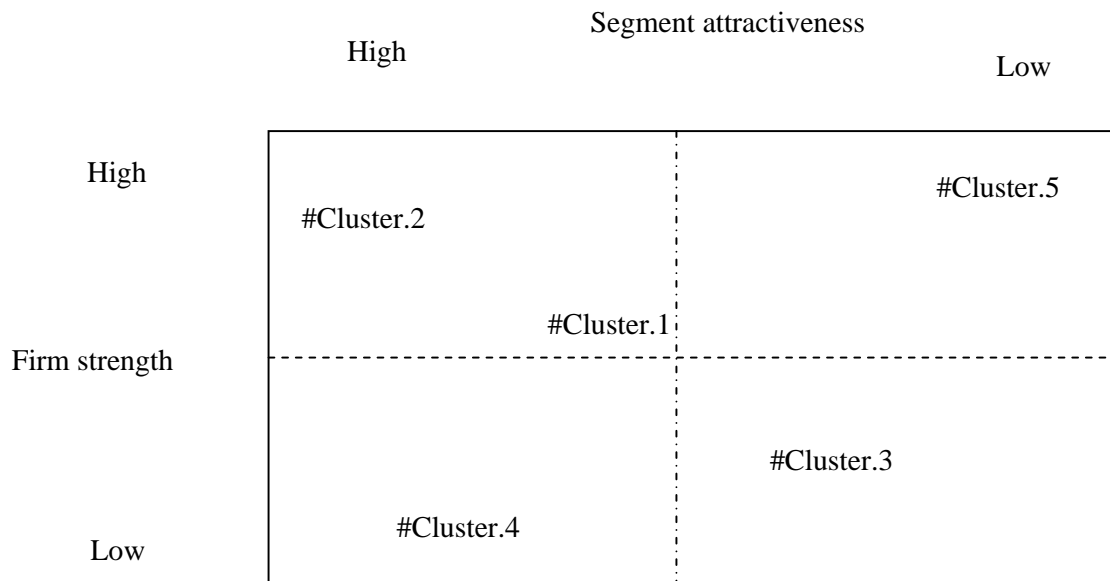


Fig.4 Opportunity matrix for Phuket

### Cluster Output for Koh Samui Enquiries

In Samui a total of 500 enquiries were processed and five clusters were extracted. Based on the proximity values the five cluster solution was found to be stable. Clusters 1, 2 and 3 predominantly comprises of enquiries for short periods of stay while clusters 4 and 5 are for longer periods of stay. Cluster 1, 4 and 5 comprises of enquiries that were attracted in the low season while enquiries in clusters 2 and 3 pertain to the high season. Thus Clusters 4 and 5 typically comprise of enquiries for long stay in the low season when typically the room rates are low. This is explained by the lesser than average room rent per night for Clusters 4 and 5. It is also quite interesting to know that for this segment, the conversion rate is pretty high. Unlike Bangkok and Phuket, the most popular price range for Samui is typically 50 to 75\$ per room night and about 20% of the enquiries pertain to long staying periods. In both the instances, the conversion rate is fairly high. The conversion rate is typically very high except for Cluster2. However, Cluster2 seems to be the most attractive cluster in terms of the total volume of enquiries. Cluster2 mostly comprises of enquiries in the high season, which point towards the firm's

weakness in converting enquiries in the high season. This is a very significant result since, the enquiry conversion rate for low season seems to be quite adequate for Koh Samui for all the other clusters. The opportunity matrix for Koh Samui is given in fig.5

**||Take in Table.4 Cluster output for Koh Samui||**

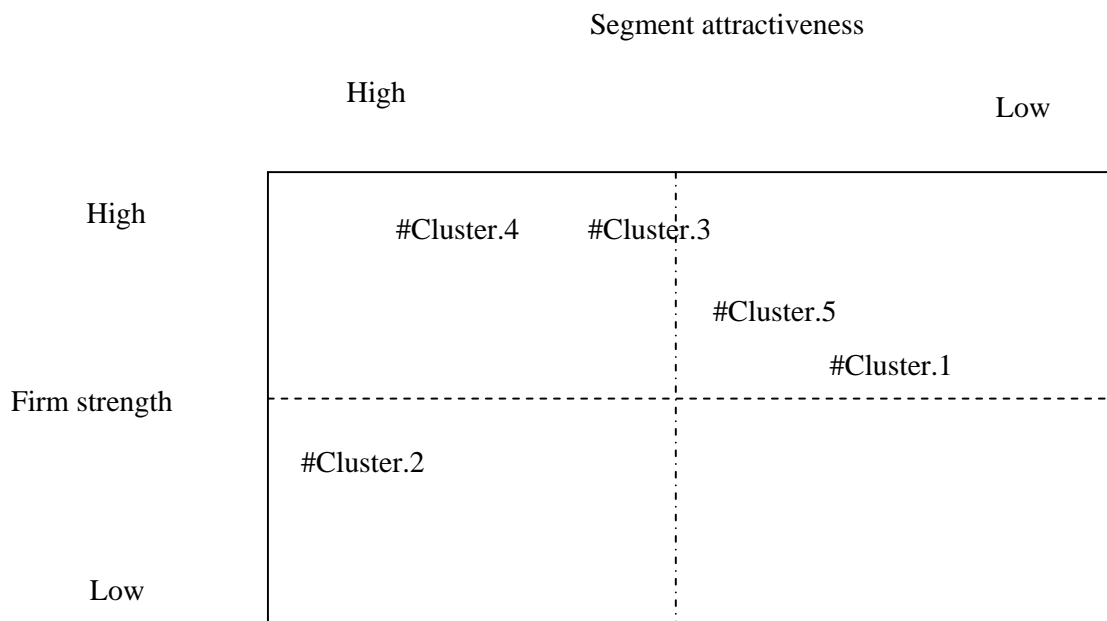


Fig.5 Opportunity matrix for Koh Samui

### Cluster Output for Krabi Enquiries

In Krabi, another destination where morethailand.com has significant presence, about 348 enquiries were processed. Based on the proximity values, a four cluster solution was found to be stable. Clusters 1, 2 and 4 pertain to enquiries for the high season while Cluster 3 comprises of enquiries for the low season. Cluster1 comprises of relatively long stay enquiries with the length of stay being more than one week while the clusters 2,3,and 4 comprises of enquiries with less than one week duration. Apart from Cluster 4, the other three clusters have predominantly had room rates between 25 to 75\$ per night. Also all the clusters consist of enquiries from mainly new customers. In terms of volume of enquiries, Cluster 2 has the highest value and the firm has done well to convert the entire group of enquiries. However Cluster 4 which constitutes the second

highest value in terms of the total volume of enquiry is a matter of concern. The enquiry conversion rate being nil, it is vital that the firm looks into means of converting enquiries in this segment. The main distinction between enquiries in Cluster4 and other clusters is higher incidence of enquiries for rooms with a rent greater than 75\$ per day. It is probably important for the firm to look at the attractiveness of hotels that offer rooms at this rate in Karabi. Further the incidence of new customers is more in Cluster 4 than in other clusters. The opportunity matrix for Krabi is given in fig.6.

**||Take in Table. 5 Cluster Output for Krabi||**

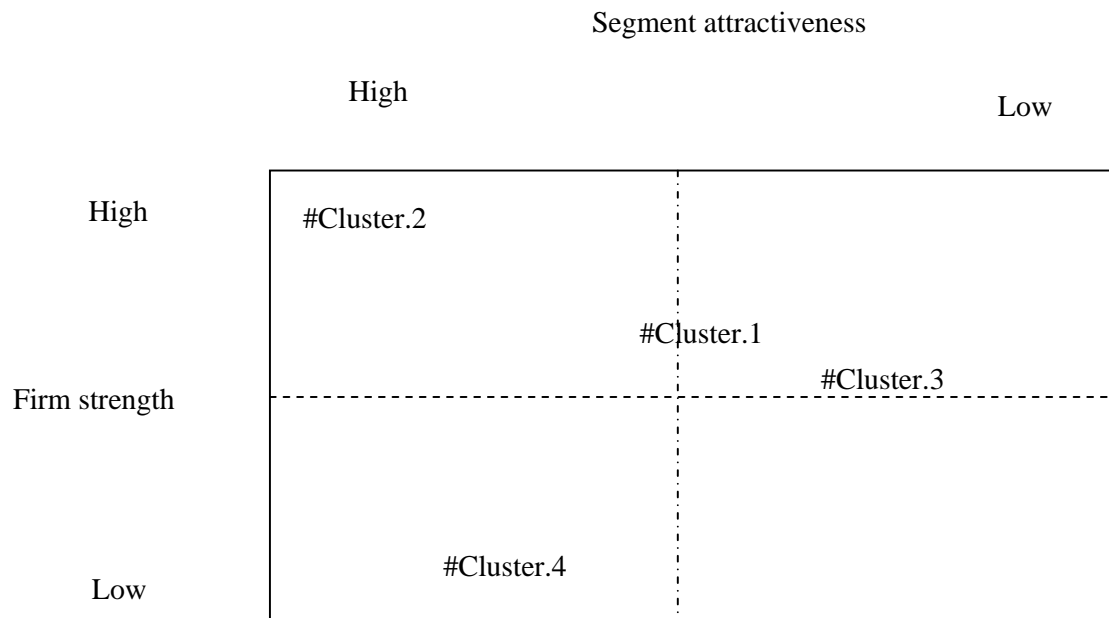


Fig.6 Opportunity matrix for Krabi

**Implications for strategy**

The opportunity matrices for each destination provide meaningful strategic objectives to pursue for morethailand.com. The segment attractiveness of clusters is assumed to be applicable to the entire industry and not just to morethailand.com. This is based on the belief that more or less similar set of enquiries will be received by almost all the on-line travel intermediaries offering similar products to the world outside. Given this fact what varies is actually the firm strength or the conversion rate achieved for each of the

clusters by morethailand.com. Since all the enquiries were replied with 'room available' message, there are very few factors that could have lead to a non-sale. Some of these factors may appear to be random when the customers drop the idea for traveling to Thailand without any apparent reason. However such factors cannot be more than 10 to 20 percent since, the customer had actually taken the trouble to make the enquiry as well as communicate through e-mail in order to prolong the enquiry. The only rational reason is the availability of better price range rooms for the same or a similar hotel in other competitive websites.

As far as strategic responses are concerned, two main possibilities arise: (i) Associating with more hotels within a particular price range so as to offer greater variety as well as gaining price advantage over other online travel intermediaries in that destination or (ii) reworking the dynamic pricing attributes for a particular destination so that customers are more motivated to convert. In the case of Bangkok, Cluster1 and Cluster2 are opportunities since they have not been targeted adequately despite being very high in terms of the segment attractiveness. The segmentation scheme in Phuket has provided mixed results. Although Cluster 4 which has a high segment attractiveness has to be targeted much better. This cluster predominantly consists of high season tourists who would like to visit Phuket for a short stay. What is emerging from the clustering analysis is that morethailand.com is typically not very competitive in terms of high season short stay and low season long stay enquiries. This is probably due to the lack of hotels that can offer competitive rates in these categories. In Koh Samui Cluster 2 needs urgent attention since it is the most lucrative segment in Koh Samui. Though the firm is able to achieve almost 50% conversion, further improvement is always possible. Since these customers are all intending to visit Thailand during the high season, it is quite clear that the firm needs to increase the conversion rate during high season in order to tap this segment. In Krabi, similarly, cluster.4 needs more attention. Cluster.4 comprises of the second most important segment and mostly consists of affluent customers who demand rooms at a price range of 100\$ or more per night. These customers come during the high season. In order to cater to this segment, the company

should look at increasing the number of hotels in its website that could provide more than 100\$ rooms.

## **Conclusion**

This paper presents the process of applying data mining in the context of an on-line service provider. While application of clustering techniques for segmentation is quite common in the case of offline marketing, this paper looks at an instance when a clustering technique is used to understand the patterns for on-line enquiries. The clustering process as applied in this context had two major purposes: (i) understanding the customer profile and (ii) deriving actionable strategies in order to enhance the conversion rate. In terms of the first objective, the clustering technique could unravel the basic pattern of customer enquiries across the attributes considered. For the second approach too, the clustering technique resulted in actionable strategies for all the destinations. This approach endorses some of the claims made by Albers and Manfred (2000) about the utility of portfolio approaches in segmentation. The on-line consumption process varies across different contexts and across different types of customers. The first phase of the on-line consumption process is the conversion from a visit to an enquiry. The second phase, which is very critical in a high-involvement and information intensive product like tourism, is the conversion from enquiry to transaction. In this phase the attributes associated with the enquiry will influence the eventual conversion decision. It is of course debate able whether clustering enquiry patterns without demographic details is useful. However, here we are only focusing on strategies that can enhance the enquiry conversion rate through altering the supply side variables such as providing more attractive room rates, enrolling more hotels etc. Future segmentation approaches could try to combine click stream data and enquiry data to enhance the quality of the clusters.

## References

Albers, S and Manfred, K (2000) "Approaches for the Segmentation of Customers, How suitable are the traditional concepts", *Schmalenbach Business Review*, Vol.52 (3), pp. 312-313.

Chen, H.M and Cooper, M. D (2001), "Using Clustering Techniques to Detect Usage Patterns in a Web-Based Information System", *Journal of the American Society for Information Science and Technology*, Vol.52 (11), pp.888-904.

De Kluyver, C.A and Whitlark, D.B (1986), "Benefit Segmentation for Industrial Products, *Industrial Marketing Management*, pp.286.

Dolnicar, S (2003), "Using Cluster Analysis for Market Segmentation-Typical Misconceptions, Established Methodical Weaknesses and Some Recommendations for Improvement", *Journal of Market Research*, Vol.11 (2), pp.5-12.

Dolnicar(2004) "Beyond Commonsense Segmentation: A Systematics of Segmentation Approaches in Tourism", *Journal of Travel Research*, Vol.42 (February), pp.244-250.

Hoontrakul, Pongsak (2004) "Value Revelation of differentiated Goods in the Travel Industry", working paper, Sasin of Chulalongkorn University, Bangkok, Thailand

Jainszeewski, C (1998),"The Influence of Display Characteristics on Visual Exploratory Search Behavior, *Journal of Customer Research*, Vol.25 (3), pp.290-301.

Kastenholz, E., Davis, D and Paul, G (1999) "Segmenting Tourism in Rural Areas: The Case of North and Central Portugal", *Journal of Travel Research*, Vol.37 (May), pp.353-363.

Keng, K.A. and Cheng, J.L.L. (1999) “Determining Tourist Role Typologies: An exploratory Study of Singapore Vacationers”, *Journal of Travel Research*, Vol.37, pp.382-390

Moe, W and Fader, P. (2004 a), Capturing Evolving Visit Behavior in Click stream Data, *Journal of Interactive Marketing*, Volume 18.(1), pp. 5-19.

Moe, W and Fader, P., S (2004 b), “Dynamic Conversion Behavior at E-Commerce Sites”, *Management Science*, Vol.50 (3), pp.326-335.

Peltier, J., and Schribrowsky, J (1997), “The use of need-based segmentation for developing segment-specific direct marketing strategies”, *Journal of Direct Marketing*, Vol.11 (4). Pp.53-62.

Sarigollu, E and Huang, R., (2005) “Benefit Segmentation of Visitors to Latin America”, *Journal of Travel Research*, Vol.43 (February), pp.277-293.

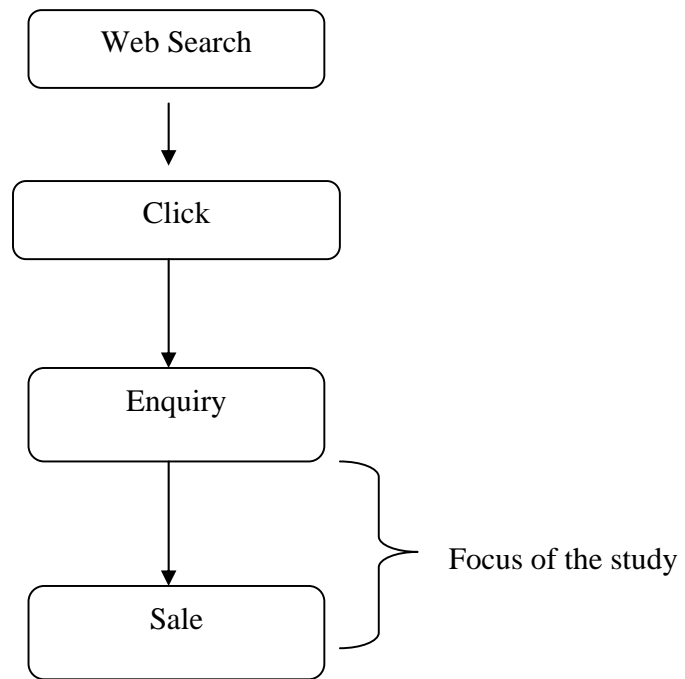


Figure.1 The transaction process in on-line travel sale

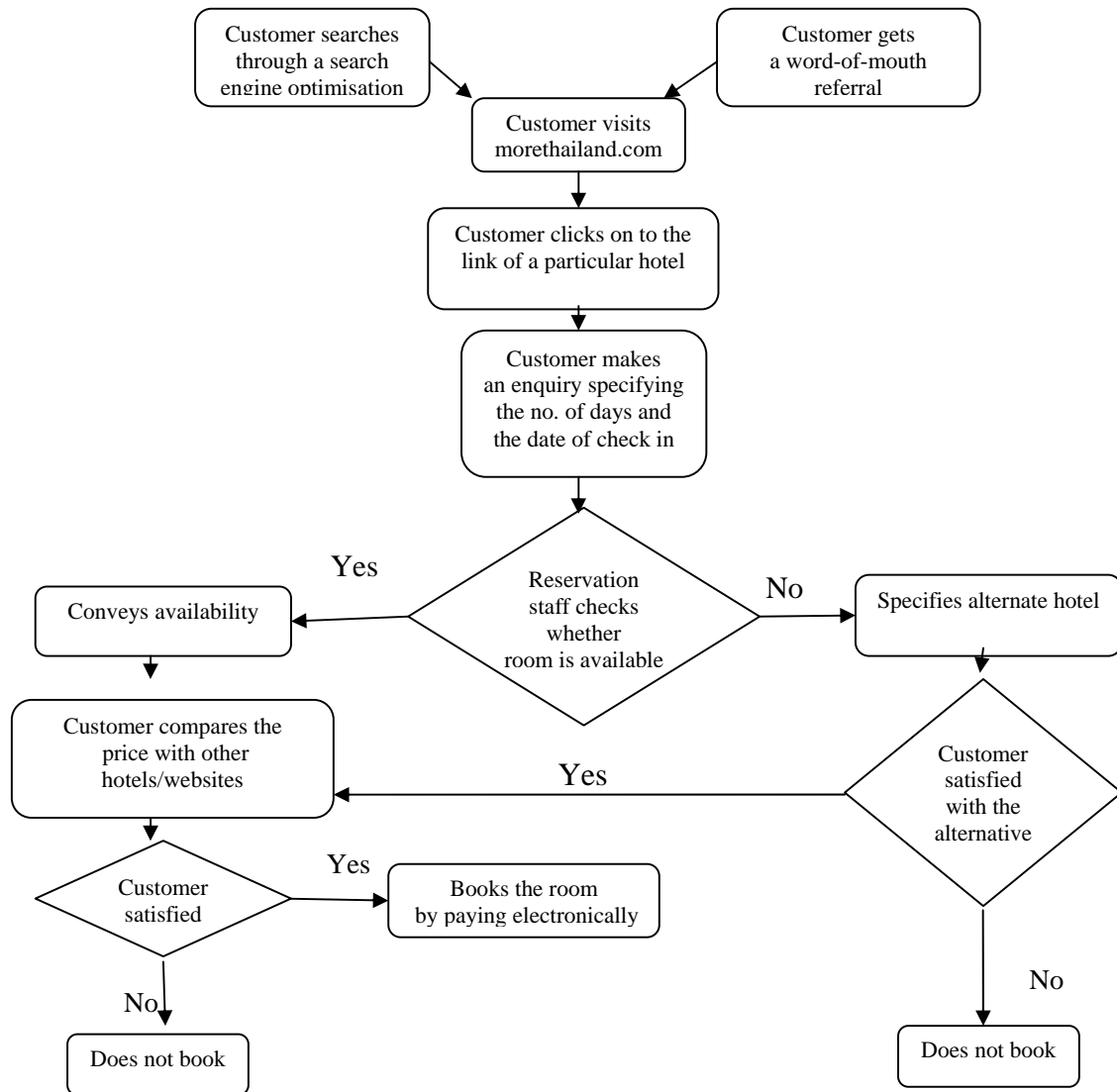


Figure 2: Flow Chart of Transaction Processing in MoreThailand.com

Table.1 Variables considered

Variable name	Categories
The Price of the room desired	Less than 25\$ per day
	25 to 50\$ per day
	50 to 75\$ per day
	70 to 100\$ per day
	Above 100\$ per day
The number of days in advance of the check in date that the enquiry was made	0 to 7 days
	8 to 14 days
	15 to 30 days
	Above 30 days
The length of stay	0 to 7 days
	8 to 14 days
	15 to 30 days
	Above 30 days
The repeat customer status of the enquirer	New customer
	Repeat customer
The nationality group of the enquirer	Northern Europe
	Western Europe
	East Asia
	North America
	Australia and New Zealand
	South Asia
The season when the enquiry was made	High season
	Low season

Table.2 Clustering for Bangkok

Cluster No.		Cluster.1	Cluster.2	Cluster.3	Cluster.4	Cluster.5
Number of members		175	109	93	135	81
Cluster label		New on-transit, economy-seeking	New Holiday makers	Repeat business customers	Economy seeking holiday makers	Repeat on transit economy seeking customers
Check in season	High	0%	100%	53%	100%	0%
	Low	100%	0%	47%	0%	100%
Length of stay	1 to 7 nights	95%	96%	90%	94%	95%
	8 to 15 nights	5%	4%	5%	3%	
	Above 15 nights	0%	-	5%	3%	5%
Nationality status	South-Eastern Asia	23%	15%	16%	27%	27%
	North America	15%	16%	16%	5%	15%
	Western Europe	17%	11%	18%	13%	16%
	United Kingdom	13%	15%	18%	27%	27%
	Australia &NZ	-	17%	18%	28%	6%
	others	32%	26%	14%		9%
No. of advance days	0 to 7 days	35%	30%	43%	26%	32%
	8 to 14 days	13%	14%	19%	14%	16%
	15 to 30 days	33%	18%	12%	23%	20%
	Above 30 days	9%	38%	26%	37%	32%
Repeat customer status	New customers	100%	67%	0	65%	0%
	Repeat customers	0%	33%	100%	35%	100%
Price range	Less than 25\$	30%	11%	8%	18%	30%
	25 to 50\$	44%	47%	48%	62%	40%
	50 to 75\$	14%	24%	14%	6%	18%
	75 to 100\$	6%	7%	12%	7%	8%
	Above 100\$	6%	11%	18%	7%	4%
Percentage of total number of enquiries		29%	18%	15%	22%	13.6%
% of the total volume of enquiries		23%	24%	19.5%	22%	9%

Table.3 Clustering for Phuket

Cluster No.		Cluster.1	Cluster.2	Cluster.3	Cluster.4	Cluster.5
Number of members		101	85	59	99	72
Cluster label		New economy seeking weekenders	High season, rich long stayers	Low season long stayers	High season short stay holiday makers	High season loyal holiday makers
Check in season	High	0%	100	34%	100%	84%
	Low	100	0%	66%	0%	16%
Length of stay	1 to 7 nights	88%	37%	37%	93%	90%
	8 to 15 nights	12%	51%	56%	4%	5%
	Above 15 nights	0%	12%	7%	3%	5%
Nationality status	South-Eastern Asia	16%	6%	0%	17%	15%
	North America	10%	10%	25%	14%	15%
	Western Europe	13%	10%	24%	7%	20%
	United Kingdom	27%	32%	27%	18%	20%
	Australia & NZ	8%	4%	24%	8%	12%
	others	26%	38%		36%	18%
No. of advance days	0 to 7 days	30%	4%	17%	28%	20%
	8 to 14 days	23%	6%	12%	17%	25%
	15 to 30 days	27%	15%	15%	27%	29%
	Above 30 days	20%	75%	56%	28%	26%
Repeat customer status	New customers	67%	88%	40%	78%	10%
	Repeat customers	33%	12%	60%	22%	90%
Price range	Less than 25\$	22%	12%	14%	16%	23%
	25 to 50\$	45%	22%	39%	23%	34%
	50 to 75\$	16%	28%	17%	24%	11%
	75 to 100\$	11%	18%	9%	12%	10%
	Above 100\$	6%	18%	23%	25%	22%
Percentage of total enquiries		24%	20%	12%	24%	17%
% of the total volume of enquiries		19%	26%	17%	22%	14%

Table.4 Clustering output for Koh Samui

Cluster No.		Cluster.1	Cluster.2	Cluster.3	Cluster.4	Cluster.5
Number of members		140	142	85	85	48
Cluster label		New low season short stayers	High seasoners	Low season loyal customers	Low season economy seeking long stayers	Low season affluent long stayers
Check in season	High	0	100%	34%	13%	27%
	Low	100%		65%	87%	72%
Length of stay	1 to 7 nights	99%	78%	100%	15%	
	8 to 15 nights	1%	16%	0	77%	96%
	Above 15 nights		6%	0	8%	4%
Nationality status	South-Eastern Asia	13%	8%	16%	-	18%
	North America	10%	19%	10%	3%	17%
	Western Europe	15%		22%	23%	21%
	United Kingdom	20%	25%	24%	32%	31%
	Australia &NZ	9%	12%	8%	6%	
	others	33%	36%	20%	36%	13%
No. of advance days	0 to 7 days	18%	17%	21%	12%	5%
	8 to 14 days	15%	12%	16%	6%	18%
	15 to 30 days	35%	15%	29%	13%	21%
	Above 30 days	32%	56%	33%	69%	56%
Repeat customer status	New customers	73%	65%	10%	82%	
	Repeat customers	27%	35%	90%	18%	
Price range	Less than 25\$	4%	8%	5%	3%	10%
	25 to 50\$	40%	24%	10%	26%	9%
	50 to 75\$	34%	29%	44%	54%	50%
	75 to 100\$	6%	17%	13%	9%	10%
	Above 100\$	15%	22%	28%	8%	21%
Percentage of the total number of enquiries		28%	28%	17%	17%	10%
% of the total volume of enquiries		12%	31%	18%	24%	13%

Table.5 Clustering output for Krabi

Cluster No.		Cluster.1	Cluster.2	Cluster.3	Cluster.4
Number of members		49	146	78	85
Cluster label		New high season long stayers	High season economy seeking short holidaymakers	Low seasoners	High season affluent short holiday makers
Check in season	High	85%	100%		100%
	Low	15%	0%	100%	
Length of stay	1 to 7 nights	-	100%	100%	100%
	8 to 15 nights	100%			
	Above 15 nights	-	-	-	-
Nationality status	South-Eastern Asia	-	10%	30%	10%
	North America	8%	10%	14%	13%
	Western Europe	4%	8%	4%	12%
	United Kingdom	35%	26%	14%	22%
	Australia & NZ	7%	6%	4%	10%
	others	46%	40%	34%	33%
No. of advance days	0 to 7 days	4%	6%	21%	13%
	8 to 14 days	10%	10%	17%	13%
	15 to 30 days	4%	20%	33%	22%
	Above 30 days	82%	64%	29%	52%
Repeat customer status	New customers	65%	60%	60%	76%
	Repeat customers	35%	40%	40%	24%
Price range	Less than 25\$	8%	9%	21%	4%
	25 to 50\$	24%	32%	47%	20%
	50 to 75\$	37%	28%	18%	20%
	75 to 100\$	20%	12%	14%	26%
	Above 100\$	11%	19%	-	30%
Percentage of total enquiries		14%	40%	22%	24%
Percentage of volume of enquiries in each cluster		18.5%	39%	14%	28%