MATEMATIKA, 2009, Volume 25, Number 1, 1–8

Rough Fuzzy Approach in Tourism Demand Analysis

[1]**Ahmad Termimi Abd Ghani,** [2]**Abu Osman Md Tap**
[3]**Mahadzirah Mohamad &** [4]**Mohd Lazim Abdullah**
[1,2,4]Department of Mathematics, Faculty of Science Technology, Universiti Malaysia Terengganu
21030 Kuala Terengganu, Terengganu, Malaysia
[3]Research Management Centre, Universiti Malaysia Terengganu
21030 Kuala Terengganu, Terengganu, Malaysia
e-mail: [1]termimi@umt.edu.my

Abstract The substantial growth of the tourism activities in Malaysia clearly marks tourism as one of the most remarkable economic and social phenomena of the past few years. This paper introduces the rough-fuzzy approach in tourism forecasting. The rough-fuzzy is the extension of rough sets. Its can also be defined when the value of decisions and conditions attribute are uncertain. Within the hybridization process, we can see the strengthens of knowledge by the membership function. The study shows that the membership value of tourist arrivals from Saudi Arabia, Australia and US is 1, while the values for tourist arrivals from Taiwan and UK are 0.6755 and 0.2053 respectively. The degree of tourist arrivals from China and Thailand is also equal to 1, and from Japan's is 0.4167.

Keywords Tourism demand; decision table; membership value; rough-fuzzy sets.

1 Introduction

Malaysia has long been one of the world's best kept tourism secrets. It is an ideal tourism destination in so many different aspects as it offers a vast range of diverse attractions to suit all need and at relatively affordable prices. The future prospects of Malaysia tourism are good. There is widespread recognition of its contribution to the national economy. Malaysia tourism also reaches its enormous potential as a catalyst for future economic and social development across the country.

Witt and Witt [14] forecast tourist arrival by using regression analysis to see the quantitative relation between the tourism demand and determinant factors. On the other hand, Cho [15] applies time-series forecasting technique and Artificial Neural Network to forecast tourist arrivals to Hong Kong. The rough sets approach has been incorporated into tourism and hospitality research by Law & Au [11], Au & Law [8] and Goh & Law [2]. The term 'demand' is subjective in nature since the tourist arrivals are uncertain and depends on numerous factors such as the economic and political condition, weather, population and foreign exchange rate (Tsaur et al. [13]).

In this paper, we will extend Goh and Law [2] using the rough-fuzzy approach for deriving rules from a Decision Table of tourist arrivals to Malaysia. The purpose is to identify the relationship between the determinant factors for tourist arrivals using the rough-fuzzy technique. This paper also determine the membership degree of each arrivals and also to induce patterns in a form of decision rules, which are able to distinguish between the classes of arrival volume based upon differences in the factors that affect tourist arrivals.

2 Tourism Demand Analysis

This study observed the data on tourism demand determining factors for eight main tourist generating countries to Malaysia over a period of 11 years from 1994 to 2004. These eight countries were Saudi Arabia, Australia, China, Japan, Taiwan, Thailand, United Kingdom and United States. Five determining factors were used to forecast whether arrivals increase or decrease in number, these being country of origin, Gross Domestic Product (GDP), Income, population and consumer price index (CPI). Data on arrivals were collected from the Ministry of Tourism Malaysia and other secondary data were collected from the online database. Data for determining factors were recorded in continuous values and then transformed into discrete values to avoid biased estimation and spurious relationship modeling. For the purpose of rough sets data analysis, these values are categorized by using interval-width equivalence method (Chan et al. [1]). In this paper, we utilize the classification and approximation process in rough set theory (Indiscernible Objects, Information Table, Set Approximation, Attribute Reduction, and Decision Rules Induction) to generate the decision rules by considering fuzzy characteristic. Slowinski and Stefanowski [12] suggested handling data in the rough set framework where for each objects (attribute) in Decision Table represented by linguistic form. Salido and Murakami [4] state that for each input will be fuzzified or characterize before approximation process. Information System, S in rough-fuzzy sets can be represented as $S = \langle U,\ Q,\ V,\ f\ \rangle$ where U is universe finite set with N object $\{x_1, x_2, x_3, ..., x_N\}$, Q is finite set with n fuzzy attributes, $\{q_1, q_2, q_3, ..., q_n\}$, $V = \bigcup_{q \in Q} V_q$ where V_q is a domain for attribute q, f: $U \times Q \to V$ is the information function such that $f(x, q) \in V_q$ for every $q \in Q$, $x \in U$.

Table 1: Tourism Demand Information Table

Objects	Condition Attribute					Decision Attribute
	C_1	C_2	C_3	C_4	C_5	D
	Country	GDP	Income	Population	CPI	Number of Arrivals
x_1	Saudi Arabia	$\{\underline{a}_2\}$	$\{\overline{c}_3\}$	$\{\underline{a}_4\}$	$\{\underline{a}_5\}$	$\{\underline{a}\}$
x_2	Australia	$\{\underline{a}_2\}$	$\{\underline{a}_3\}$	$\{\underline{a}_4\}$	$\{\underline{b}_5\}$	$\{a\}$
x_3	China	$\{\underline{a}_2\}$	$\{\underline{a}_3\}$	$\{\overline{c}_4\}$	n.a	$\{\underline{b}\}$
x_4	Japan	$\{\overline{c}_2\}$	$\{\underline{a}_3\}$	$\{\underline{a}_4\}$	$\{\underline{a}_5\}$	$\{\underline{b}\}$
x_5	Taiwan	$\{\underline{a}_2\}$	$\{\underline{a}_3\}$	$\{\underline{a}_4\}$	$\{\underline{a}_5\}$	$\{\overline{a}\}$
x_6	Thailand	$\{\underline{a}_2\}$	$\{\underline{a}_3\}$	$\{\underline{a}_4\}$	$\{\underline{a}_5\}$	$\{\overline{c}\}$
x_7	U.K	$\{\underline{a}_2\}$	$\{\underline{a}_3\}$	$\{\underline{a}_4\}$	$\{a_5\}$	$\{\overline{a}\}$
x_8	U.S	$\{\underline{a}_2\}$	$\{\underline{a}_3\}$	$\{a_4\}$	$\{\overline{c}_5\}$	$\{a\}$

(VL) = Very Low $\{\underline{a}_i\}$; (SM) = Slightly Medium$\{\ \underline{b}_i\ \}$; (SH) = Slightly High$\{\ \underline{c}_i\}$;
(L) = Low $\{\ a_i\}$; (M) = Medium $\{\ b_i\ \}$; (H) = High $\{\ c_i\ \}$;
(SL) = Slightly Low$\{\ \overline{a}_i\ \}$; (VM) = Very Medium$\{\ \overline{b}_i\ \}$; (VH) = Very High $\{\ \overline{c}_i\ \}$.

Using the concept of rough-fuzzy sets theory, this data set from the Table 1 can be considered as Decision Table, T = (U, $A \cup D$), where universe U, fuzzy attributes A and fuzzy decision feature D respectively. $U = \{x_i|\ i = 1,\ldots,8\}$ and $A = \{C_1, C_2, C_3, C_4, C_5\}$, i.e C_1 = Country of Origin, C_2 = Gross Domestic Product (GDP), C_3 = Income, C_4 = Population, C_5 = Consumer Price Index (CPI) and D = Number of Arrivals. From Table 1, if we are interested in the subset D for demand "Low" or ($X_{\{Low\}} = \{x : D(x) = \underline{a}, a, \overline{a}\}$) and then we wish to use the attributes c_2, c_3, c_4, and c_5 as a possible set of classifiers of objects in X, that is $R= \{c_2, c_3, c_4, c_5\}$, where $R \subseteq A$ and $X \subseteq U$. So, we can identify the upper and lower approximations of X based on the results in Table 1. Elementary sets due to A which are proper subsets of X are $\{x_5\}$, $\{x_7\}$. Hence, the lower approximation can be defined as $\underline{A}(X)= \{x_5, x_7\}$. On the other hand, elementary sets in Table 1 which have at least one element in common with X are $\{x_1\}$, $\{x_2\}$, $\{x_5\}$, $\{x_6\}$, $\{x_7\}$, $\{x_8\}$. So, the upper approximation is $\overline{A}(X)= \{x_1, x_2, x_5, x_6, x_7, x_8\}$.

By definition, the boundary of X in U is the difference between the upper and lower approximations, $BN_A(X) = \overline{A}(X) - \underline{A}(X) = \{x_1, x_2, x_5, x_6, x_7, x_8\} - \{x_5, x_7\} = \{x_1, x_2, x_6, x_8\}$. For the approximation of Medium (M) arrivals attribute D,

$$X_{\{Medium\}} = \{x = D(x) = \underline{b}, b, \overline{b}, \}$$

the collections of elementary set are $X= \{x_3, x_4\}$ with four set attribute $R= \{c_2, c_3, c_4, c_5\}$. So, we can identify the lower$\underline{A}(X)$ and upper $\overline{A}(X)$ approximation as

$$\underline{A}(X) = \{\emptyset\}, \overline{A}(X) = \{x_3, x_4\}.$$

Thus, the boundary of X in U with respect to A is

$$BN_A(X) = \{x_3, x_4\} - \{\emptyset\} = \{x_3, x_4\}.$$

For the "High" tourist demand classification,($X_{\{High\}} = \{x : D(x) = \underline{c}, c, \overline{c}\}$), the upper and lower approximation are respectively $\underline{A}(X) = \{\emptyset\}$, $\underline{A}(X) = \{x_5, x_6, x_7\}$.

3 Decision Rules Induction

Logical rules extracted by rough set approaches are always expressed in the "*if...then*" format. Decision rules can be induced from a reduced set of condition and decision attributes or directly from the lower and upper approximation of sets (Guo [5]). Decision rules derived from the lower approximation will form "certain rules", where as those induced from the upper approximation will form "possible rules". The study shows that all considered attributes for the Low, Medium and High arrivals are important in terms of tourism demand forecasting. These four attributes also can be called core and there are no feature reductions. So, all the attributes must be considered in decision rules making for tourist arrivals forecasting to Malaysia.

The lower approximation for Low arrivals of subset D is $\underline{A}(X) = \{x_5,\ x_7\}$. There are two decision rules that are matching can be developed;

IF(*Country*=Taiwan)$\wedge$(*GDP*$\{\underline{a}_2\}$)$\wedge$(*Income*$\{\underline{a}_3\}$)$\wedge$(*Population*$\{\underline{a}_4\}$)$\wedge$(*CPI*$\{\underline{a}_5\}$),
THEN (*Arrivals* $\{\overline{a}\}$)

IF(*Country*=UK)$\wedge$(*GDP*$\{\underline{a}_2\}$)$\wedge$(*Income*$\{\underline{a}_3\}$)$\wedge$(*Population*$\{\underline{a}_4\}$)$\wedge$(*CPI*$\{a_5\}$),

THEN (*Arrivals* $\{\overline{a}\}$)

Hence, these two rules are said to be a certain rules or in other words, the decision rules are definitely right. Whilst for upper approximation, $\overline{A}(X) = \{x_1, x_2, x_5, x_6x_7, x_8\}$; the rules are

IF$(Country{=}\text{S.Arabia})\wedge(GDP\{\underline{a}_2\})\wedge(Income\{\overline{c}_3\})\wedge(Population\{\underline{a}_4\})\wedge(CPI\{\underline{a}_5\})$,

THEN(*Arrivals* $\{\underline{a}\}$)

IF$(Country{=}\text{Australia})\wedge(GDP\{\underline{a}_2\})\wedge(Income\{\underline{a}_3\})\wedge(Population\{\underline{a}_4\})\wedge(CPI\{\underline{b}_5\})$,

THEN$(Arrivals\{a\})$

IF$(Country{=}\text{Taiwan})\wedge(GDP\{\underline{a}_2\})\wedge(Income\{\underline{a}_3\})\wedge(Population\{\underline{a}_4\})\wedge(CPI\{\underline{a}_5\})$,

THEN$(Arrivals\{\overline{a}\})$

IF$(Country{=}\text{Thailand})\wedge(GDP\{\underline{a}_2\})\wedge(Income\{\underline{a}_3\})\wedge(Population`\{\underline{a}_4\})\wedge(CPI\{\underline{a}_5\})$,

THEN$(Arrivals\{\overline{c}\})$

IF$(Country{=}\text{UK})\wedge(GDP\{\underline{a}_2\})\wedge(Income\{\underline{a}_3\})\wedge(Population\{\underline{a}_4\})\wedge(CPI\{a_5\})$,

THEN$(Arrivals\{\overline{a}\})$

IF$(Country{=}\text{US})\wedge(GDP\{\underline{a}_2\})\wedge(Income\{\underline{a}_3\})\wedge(Population\{a_4\})\wedge(CPI\{\overline{c}_5\})$,

THEN$(Arrivals\{a\})$

For the Medium arrivals approximation, the rules can be developing only from upper approximation because there are no objects in the lower approximation (empty set). So, the induced decision rules from upper approximation are

IF$(Country{=}\text{China})\wedge(GDP\{\underline{a}_2\})\wedge(Income\{\underline{a}_3\})\wedge(Population\{\overline{c}_4\})$,

THEN $(Arrivals\{\underline{b}\})$

IF$(Country{=}\text{Japan})\wedge(GDP\{\overline{c}_2\})\wedge(Income\{\underline{a}_3\})\wedge(Population\{\underline{a}_4\})\wedge(CPI\{\underline{a}_5\})$,

THEN$(Arrivals\{\underline{b}\})$

Whilst the induced decision rules from upper approximation for High tourist arrivals are

IF$(Country{=}\text{Taiwan})\wedge(GDP\{\underline{a}_2\})\wedge(Income\{\underline{a}_3\})\wedge(Population\{\underline{a}_4\})\wedge(CPI\{\underline{a}_5\})$,

THEN$(Arrivals\{\overline{a}\})$

IF$(Country{=}\text{Thailand})\wedge(GDP\{\underline{a}_2\})\wedge(Income\{\underline{a}_3\})\wedge(population\{\underline{a}_4\})\wedge(CPI\{\underline{a}_5\})$,

THEN$(Arrivals\{\overline{c}))$

IF$(Country{=}\text{UK})\wedge(GDP\{\underline{a}_2\})\wedge(Income\{\underline{a}_3\})\wedge(Population\{\underline{a}_4\})\wedge(CPI\{a_5\})$,

THEN $(Arrivals\{\overline{a}\})$

4 Membership Function Analysis

Every approximation of objects will be measured by the trapezoidal membership function.

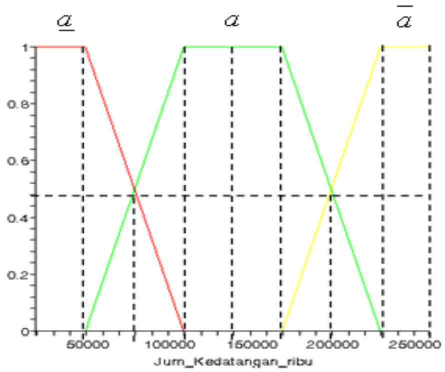

Figure 1(a) Membership Function for Low Arrivals

$$\mu_{\underline{a}}(x) = \begin{cases} 1 & \text{if } x \leq 49000 \\ \dfrac{(109000 - x)}{60000} & \text{if } 49000 < x < 109000 \\ 0 & \text{if } x \geq 109000 \end{cases}$$

$$\mu_{a}(x) = \begin{cases} 0 & \text{if } x \leq 49000 \text{ or } x \geq 229000 \\ \dfrac{(x - 49000)}{60000} & \text{if } 49000 < x < 109000 \\ \dfrac{(229000 - x)}{60000} & \text{if } 169000 < x < 229000 \\ 1 & \text{if } 109000 < x < 169000 \end{cases}$$

$$\mu_{\overline{a}}(x) = \begin{cases} 0 & \text{if } x \leq 169000 \\ \dfrac{(x - 169000)}{60000} & \text{if } 169000 < x < 229000 \\ 1 & \text{if } x \geq 229000 \end{cases}$$

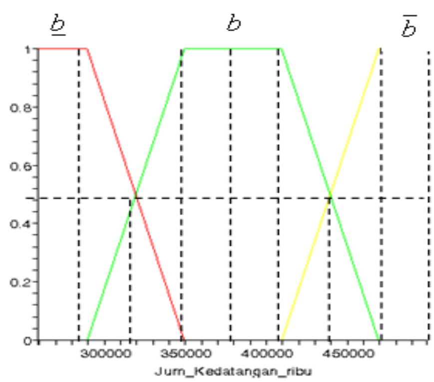

Figure 1(b) Membership Function for Medium Arrivals

$$\mu_{\underline{b}}(x) = \begin{cases} 1 & \text{if } x \leq 289000 \\ \dfrac{(349000 - x)}{60000} & \text{if } 289000 < x < 349000 \\ 0 & \text{if } x \geq 349000 \end{cases}$$

$$\mu_{b}(x) = \begin{cases} 0 & \text{if } x \leq 289000 \text{ or } x \geq 469000 \\ \dfrac{(x - 289000)}{60000} & \text{if } 289000 < x < 349000 \\ \dfrac{(469000 - x)}{60000} & \text{if } 409000 < x < 469000 \\ 1 & \text{if } 349000 < x < 409000 \end{cases}$$

$$\mu_{\overline{b}}(x) = \begin{cases} 0 & \text{if } x \leq 409000 \\ \dfrac{(x - 409000)}{60000} & \text{if } 409000 < x < 469000 \\ 1 & \text{if } x \geq 469000 \end{cases}$$

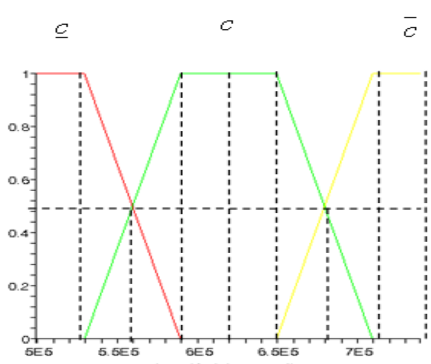

Figure 1(c) Membership Function for High Arrivals

$$\mu_{\underline{c}}(x) = \begin{cases} 1 & \text{if } x \leq 529000 \\ \dfrac{(589000 - x)}{60000} & \text{if } 529000 < x < 589000 \\ 0 & \text{if } x \geq 589000 \end{cases}$$

$$\mu_{c}(x) = \begin{cases} 0 & \text{if } x \leq 529000 \text{ or } x \geq 709000 \\ \dfrac{(x - 529000)}{60000} & \text{if } 529000 < x < 589000 \\ \dfrac{(709000 - x)}{60000} & \text{if } 649000 < x < 709000 \\ 1 & \text{if } 589000 < x < 649000 \end{cases}$$

$$\mu_{\overline{c}}(x) = \begin{cases} 0 & \text{if } x \leq 649000 \\ \dfrac{(x - 649000)}{60000} & \text{if } 649000 < x < 709000 \\ 1 & \text{if } x \geq 709000 \end{cases}$$

Both fuzzy and rough set theory represented two different approaches to vagueness. Fuzzy set theory address *gradualness* of knowledge, expressed by the fuzzy membership, whereas rough set theory addresses *granularity* of knowledge, expressed by the indiscernibility relation. Hence, within the hybridization process, we can see the strengthens of knowledge by the indiscernible of objects.

Table 2: Membership Degree for Arrival Low, Medium and High

Objects	Membership Value	Objects	Membership Value	Objects	Membership Value
$\mu_{\underline{a}}(x_1)$	1	$\mu_{\underline{b}}(x_3)$	1	$\mu_{\overline{c}}(x_6)$	1
$\mu_a(x_2)$	1	$\mu_{\underline{b}}(x_4)$	0.4167	-	-
$\mu_{\overline{a}}(x_5)$	0.6755	-	-	-	-
$\mu_{\overline{a}}(x_7)$	0.2053	-	-	-	-
$\mu_a(x_8)$	1	-	-	-	-

Refer to Table 2, the membership value of tourist arrivals from Saudi Arabia, Australia and US is 1. While the value of Taiwan and UK is 0.6755 and 0.2053 respectively. From the Table also shows that the degree of tourist arrivals from China and Thailand is equal to 1, and the membership degree for Japan is 0.4167.

5 Results and Conclusion

Table 3 (a), Table 3 (b) and Table 3 (c) showed the differences between forecasting result and the actual value of tourist arrivals volume to Malaysia for year 2004. We can see that the forecasted value for the number of arrivals "Low" are accurate that is equal to the actual value. If there were exists the increasing or decreasing in tourist arrivals, but these changing still in the range. So, it means the forecasting from the lower approximations is definitely accurate as apparent in rough sets theory.

Table 3(a): Forecasting Accuracy for Low Arrivals

Country	Approximation	
	Actual Value (2004)	Forecast Value
Australia	$\{\overline{a}\}$	$\{a\}$
Taiwan	$\{\overline{a}\}$	$\{\overline{a}\}$
UK	$\{\overline{a}\}$	$\{\overline{a}\}$
Saudi Arabia	$\{\underline{a}\}$	$\{\underline{a}\}$
US	$\{a\}$	$\{a\}$

Whilst for "Medium" and "High" arrivals forecasting, the result showed that there were occur some mistake (*bias*), which is forecasting for Thailand is High (H), even though the actual value is Very High (VH). Likewise for forecasting of China that give the Medium (M) value, while the actual value is High (H). This happens because of the forecasting are from the upper approximation. Hence, we can summarize this approach is very useful and

Table 3(b): Forecasting Accuracy for Medium Arrivals

Country	Approximation	
	Actual Value (2004)	Forecast Value
China	$\{\underline{c}\}$	$\{\underline{b}\}$
Japan	$\{\underline{b}\}$	$\{\underline{b}\}$

Table 3(c): Forecasting Accuracy for High Arrivals

Country	Approximation	
	Actual Value (2004)	Forecast Value
Thailand	$\{\overline{c}\}^*$	$\{\overline{c}\}$

successful tools when dealing with vagueness especially in decision making process. This study has provides a new method in rules induction in forecasting for tourism demand.

Acknowledgements

This research was supported by the Kolej Universiti Sains dan Teknologi Malaysia, under the Fundamental Research Grant (No. 57012).

References

[1] C. Chan, C. Batur & A. Srinivasasn, *Determination of quantization intervals in rule based model for dynamic systems*, Proceedings of the IEEE Conference on Systems, Man, and Cybernetics, Va., 1719-1723, 1991.

[2] C. Goh & R. Law, *Incorporating the Rough Sets Theory into Travel Demand Analysis,* Journal of Tourism Management, 24(2003) 511-517.

[3] D. Dubois & H. Prade, *Rough Fuzzy Sets and Fuzzy Rough Sets,* International Journal of General Systems, 17(2-3)(1990), pp. 191-209.

[4] F.J.M. Salido & S. Murakami, *Rough Set analysis of a general type of fuzzy data using transitive aggregation of fuzzy similarity relations,* Fuzzy Sets and System, 139(2003), 635-660.

[5] J-Y. Guo, *Rough Set-Based Approach To Data Mining.* Ph.D Dissertation, Case Western Reserve University, Cleveland, Ohio, USA, 2003.

[6] L.A. Zadeh, *Fuzzy sets,* Inform. Control, 8(1965), 338-353.

[7] M. Sarkar, *Rough-Fuzzy Functions in Classification,* Journal of Fuzzy Sets and Systems, 132(3)(2002), 353-369.

[8] N. Au & R. Law, *The application of rough sets to sightseeing expenditures,* Journal of Travel Research, 39(1)(2000), 70-77.

[9] P. Ligras, *Fuzzy-Rough and Rough-Fuzzy serial Combinations in Neurocomputing,* Journal of Neurocomputing, 36(2001), 29-44.

[10] Q. Shen & A. Chouchoulas, *A rough-fuzzy approach for generating classification rules,* Journal of Pattern Recognition, 35(2002), 2425-2438.

[11] R. Law & N. Au, *A rough set approach to hotel expenditure decision rules induction,* Journal of Hospitality and Tourism Research, 22(4), (1998) 359-375.

[12] R. Slowinski & J. Stefanowski, *Rough Set reasoning about uncertain data,* Fund. Inform., 27(1996), 229-243.

[13] S-H. Tsaur, G-H. Tzeng & K-C. Wang, *Evaluating Tourist Risk from Fuzzy Perspectives,* Annals of Tourism Research, Vol. 24, No. 4, (1997), 796-812.

[14] S.F. Witt & C.A. Witt, *Forecasting Tourism Demand: A Review of Empirical Research,* International Journals of Forecating, 11(1995), 447-475.

[15] V. Cho, *A Comparison of Three Different Approaches to Tourist Arrival Forecasting,* Journal of Tourism Management, 24(2003), 323-330.

[16] Z. Pawlak, *Rough Sets,* International Journal of Computer and Information Science, 11(1982), 341-356.