DISTRIBUTOR SELECTION IN SUPPLY CHAIN MANAGEMENT - A ROUGH SET BASED APPROACH

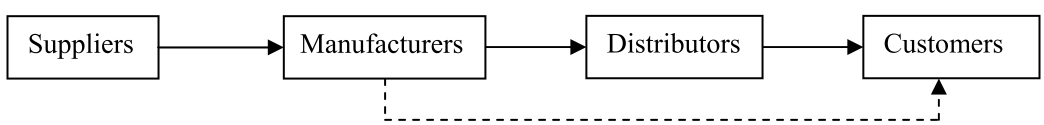
Zhonghai Zou, School of Management, Shanghai University, Shanghai, China
Johnny C. Ho, Turner College of Business, Columbus State University, Columbus, GA 31907
Tzu-Liang (Bill) Tseng, Dept. of Industrial Engr., University of Texas at El Paso, El Paso, TX 79968
Guofang Song, School of Management, Shanghai University, Shanghai, China

ABSTRACT

Distributor's selection is an important issue in Supply Chain Management, particularly in the current competitive environment. The current research works provide only conceptual, descriptive, and simulation results focusing mainly on firms' resources and general marketing factors. The selection and evaluation of distributors generally incorporate qualitative information; however, analyzing qualitative information is difficult by standard statistical techniques. Consequently, a more suitable approach is desired. In this paper, a method based on Rough Set Theory, which has been recognized as a powerful tool in dealing with qualitative data in the literature, is introduced and modified for preferred distributor selection. We derived certain decision rules which are able to facilitate distributor selection and identified several significant features based on an empirical study conducted in China.

1. INTRODUCTION

Industry is now strongly recognizing that total management of the supply chain enhances the competitive edge of all "players" therein. As a result, Supply Chain Management (SCM) has received more attentions from both academicians and practitioners in the past decade. Many articles and books have been published for the methods and opinions about the application of supply chain management. Although there is no generally accepted notion of supply chain, at least it should contain the suppliers' suppliers and the customers' customers. Supply chain in this paper refers to a network of integrated and dependent process through which specifications are transformed to finished deliverables. Figure 1 depicts a conceptual framework for supply chain.



Note: The dash-line indicates those manufacturers which sell directly to their customers.

FIGURE 1: THE CONCEPTUAL FRAMEWORK FOR SUPPLY CHAIN

Supplier selection and evaluation play an important role in the supply chain process and are crucial to the success of manufacturing firms (Sevkli et al., 2008). There are many research done in the supplier selection area, and many methodologies have been applied in practice, including the cost-ratio method, linear or mixed integer programming, and multi-objective linear programming models (Ghodsypour and O'Brien, 1998; Yan et al., 2003; Oliveria and Lourenc¸ 2002). Although these methods have been widely used in the area of supplier selection, there are certain drawbacks associated with the implementation of these methods. More recently, Fuzzy Systems Theory (FST) has been successfully applied to supplier selection problems (Kahraman et al., 2003, 2004; Chan and Kumar, 2007), and Rough Set Theory (RST) has also been applied for preferred suppliers prediction (Tseng et al., 2006).

To date, numerous literatures have explored the issues of supplier selection. Nevertheless, little work has been done in the selection of distributor, particularly via empirical studies. Only conceptual, descriptive and simulation results focused primarily on firms' resources and general marketing/selling factors were discussed (Abratt and Pitt, 1989; Shipley et al., 1989; Cavusgil et al., 1995, Yoeh and Calantone, 1995). It should be noted that distributor selection has not been studied deeply and the theoretical methods developed by academics have not been fully applied in industry. In this paper, we propose a rough set based methodology which is able to perform rule induction effectively. Moreover, the weight of each input feature is incorporated in the proposed approach so as to enhance quality of the derived rules.

The remainder of this paper is organized as follows. The next section provides a literature review on distributor research and introduces the standard rough set–based rule induction problem. Section 3 presents the basic rule identification algorithm to determine the reducts with both equal and unequal weight features. A case study is presented to show how the rule identification approach can be applied to distributor selection in Section 4. Section 5 concludes the paper with discussion of empirical findings.

2. LITERATURE REVIEW AND ROUGH SET-BASED RULE INDUCTION PROBLEM

2.1 Literature Review on Distributor Research

As mentioned above, there are few empirical studies for manufacturers' distributor selection. Ross (1973) studied the selection of the overseas distributor. The author concluded that whether or not the exporter will be able to achieve his goals depends to a great extent on how well he has carried out his analysis of which a firm will do the best possible job for him in a particular market. Lindqvist (1983) reviewed the research trends in distribution in Finland and found that the factors affecting the length of the distribution channel, the variables accounting for dimensions of retail trade in commune level centers, and the influence of the location and size of the automobile dealership on its profitability are at the heart of distribution research. Fram (1992) highlighted the importance of selecting the correct international distributors if a firm wishes to trade effectively in the worldwide market. The author described a study commissioned to explore the steps required to minimize the risk when selecting a distributor, e.g., use of end-user reference and suggestions.

Fonsson and Zineldin (2003) proposed a conceptual model including behavioral dimensions of supplier-dealer relationships and presented hypotheses as to how to achieve satisfactory inter-organizational relationships. Their results showed that good reputation and close relationship are key variables for the achievement of high satisfaction in a "high-trust and commitment relationship". Sharma et al. (2004) proposed a composite Distributor Performance Index (DPI) to evaluate distributors' performance. Based on a case study, Wang and Kess (2006) found that task-related and partner-related dimensions in partner selection of international joint ventures were useful in distributor relationship. A distributor relationship is a product-tied relationship, and product innovation can be used as an approach for performance improvement in distributor relationship. Lin and Chen (2008) derived four key constructs from the marketing, supply chain, and logistics literature to investigate their influences on the distributor selection.

2.2 Rough Set-Based Rule Induction

Rough Set Theory (RST) was originated by Pawlak (1982) and was developed to classify imprecise, uncertain, and incomplete information or knowledge expressed in terms of data acquired from experience; therefore, RST complements fuzzy set theory (Dubois and Prade, 1990). RST is suitable for processing qualitative information that is difficult to analyze by standard statistical techniques (Heckerman et al., 1997). It integrates learning-from-example techniques, extracts rules from a data set of interest, and discovers data regularities (Komorowski and Zytkow, 1997).

RST is a new mathematical approach to vagueness and uncertainty. The theory has found many real life applications and is considered as a very well suited new mathematical tool to deal with various decision problems. Many articles on rough set theory and decision support have been published recently. RST gives new insight into the decision process and offers new efficient algorithms. The original version of RST has proved to be particularly useful in the analysis of multi-attribute classification problems under inconsistency following from information granulation, i.e., objects having the same description but belonging to different classes.

Greco et al. (2000) extended the original version of RST in a number of directions in order to deal with problems of multi-criteria decision analysis (MCDA). Daubie et al. (2002) compared the rough set and decision tree approaches as techniques for classifying credit applicants. Mickee (2003) applied RST to deal with the problem of apparent indiscemibility between objects in a set. Wei and Zhang (2004) combined the fuzzy set and rough set. Kumar et al. (2005) explored the use of rough-set methods for marketing decision support systems in the retail business. Some other applications are summarized in Table 1.

The main theme of RST is concerned with measuring what may be described as the "ambiguity" inherent in the data. In RST, the essential distinction is made between objects that may definitely be classified into a certain category and those that may possibly be classified. Considering all decision classes yields what is referred to as the "quality of approximation" that measures the proportion of all objects for which definite classification may be achieved.

TABLE 1: ROUGH SET APPLICATION

Applications	Researchers	Description
Human resource management	Chien and Chen (2007)	Exploring and analyzing human resource data for personal selection and human capital enhancement.
Supplier prediction	Tseng et al. (2006)	Presenting a data-mining-based hybrid approach that consists of a novel rough-set algorithm for feature selection and enhanced multi-class support vector machines (SVMs) method for accurate prediction.
Marketing application	Beynon et al. (2001)	Identification of most important attributes and induction of decision rules from market data set.
Medical decision making	Kusiak et al. (2000)	Analysis of large data sets to identify key factors in a medical data set.
Fault diagnosis on diesel engine	Shen et al. (2000)	A new discretization method is developed for discretizing attributes without a priori knowledge.
Risk management	Dimitras et al. (1999)	Rough set based approach to rule extraction to discriminate between healthy and failing firms for risk management.

2.2.1 Information system

According to RST, information can be associated with every object in the universe and thus it can be expressed in a decision table (e.g., see Table 2), in which each row represents an object and each column represents an attribute. The attributes are generally classified into *conditions* and *decisions* (e.g., in Table

2, the four features – *F1*, *F2*, *F3*, and *F4* – define the conditions and *O* describes the decision).

TABLE 2: FIVE-OBJECT DATA SET

Object No.	F1	F2	F3	F4	O
1	1	0	2	0	0
2	1	0	0	1	1
3	0	0	3	1	0
4	1	1	2	0	2
5	0	0	1	0	0

Therefore, knowledge can be described in an information system, containing four components as follows:

$$S = (U, A, V, f) \tag{1}$$

where, called the *universe*, is a nonempty set of all objects and *A* is the finite set of all the *attributes*. *V* is the set of all the attribute values such that

$$V = \bigcup_{a \in A} V_a \tag{2}$$

where V_a is a finite attribute domain of attribute *a*. Finally, *f* denotes an information function such that, for every $a \in A$ and $u_i \in U$,

$$f(u_i, a) \in V_a \tag{3}$$

Table 2 illustrates the information of five objects that are characterized with one decision attribute (*O*) and four condition attributes (*F1*, *F2*, *F3*, *F4*).

3. RULE IDENTIFICATION ALGORITHMS

The proposed conceptual framework to elicit decision rules consists of the following steps: problem definition, data preparation, data partition, reduct generation, and rule validation as shown in Figure 2.

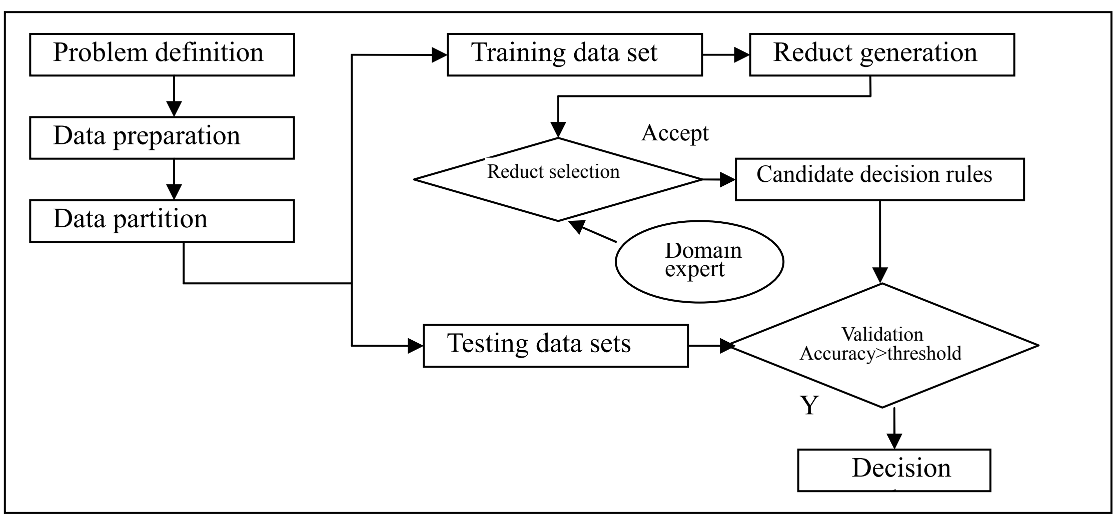


FIGURE 2: THE CONCEPTUAL FRAMEWORK OF ELICITING DECISION RULES

3.1 Problem Definition, Data Preparation and Data Partition

First, the data exploration process starts by identifying the right problems to solve and structuring the corresponding objectives and the associated attributes, that is, to make it clear what we want. Then, the needed data should be collected for the objects and the attributes of the objects. More important, a series of data preprocessing tasks, including consistency checks to detect errors, removing noise or outliers where appropriate, and data completeness checks, should be done to ensure that the data are as accurate as possible. Next, the target dataset is randomly divided into the training data set and the testing data set. Kusiak (2001) suggested the splitting of the data using the bootstrapping method according to the following ratios: 0.632 for training set and 0.368 for testing set. The training data set is used to build the model and derive the rules. The testing data set is used to detect over fitting of the modeling tools.

3.2 Reduct Generation

The basic construct in RST is called a *reduct*. It is defined as a minimal sufficient subset of features RED $\subseteq$ A such that:

(a) Relation R(RED) = R(A); that is, RED produces the same categorization of objects as the collection A of all features.

(b) For any $g \in$ RED, R(RED – $\{g\}$) $\neq$ R(A); that is, a reduct is a minimal subset of features with respect to the property (a).

The term reduct was initially defined for sets rather than objects with input and output features or for decision tables with decision features (attributes) and outcomes. Reducts of the objects in a decision table have to be computed with consideration given to the value of the output feature. The original definition of reduct considers features only. In this paper, each reduct is viewed from four perspectives – feature, feature value, object, and rule perspective.

The reduct generation algorithm, based on Pawlak 1991, is given as follows:

-Step 0. Initialize object number $i = 1$.

-Step 1. Select object i and find a set of o-reduct with one feature only.
If found, go to Step 3; otherwise go to Step 2.

-Step 2. For object i, find an o-reduct with $m - 1$ features, where m is the number of input features. This step is accomplished by deleting one feature only at a time.

-Step 3. Set $i = i + 1$. If all objects have been considered, stop; otherwise go to step 1.

For example, a subset of all reducts generated for the objects in Table 2 are shown in Table 3. The entry "x" in each reduct implies that the corresponding feature is not considered in determining the feature output of an object.

TABLE 3: REDUCTS OF OBJECTS IN TABLE 2

Object No.	Reduct No.	F1	F2	F3	F4	O
4	1	x	1	x	x	2
2	2	x	x	0	x	1
1	3	x	0	2	0	0
	4	1	0	2	x	0
	5	1	0	x	0	0
3	6	0	x	x	x	0
	7	x	x	3	x	0
5	8	0	x	x	x	0
	9	x	x	1	x	0

3.3 Reduct Selection and Rule Identification Algorithm

The nine reducts in Table 3 could be chose in a number of different ways. As we know that in the real world, the features (attributes) are not the same unique for depicting an object, some are more important, and some are less. In this section, we will choose them first with equal weights for every feature, then the weights were added, and the results were compared.

3.3.1 Equal weight features

When all the features are the same weight, we can choose the reducts based on the following steps:

-Step 1: Select the features used from only a single reduct of the object(s).

-Step 2: Select the features which are used more frequently and may be selected previously in order to get the most similar reducts from other objects.

Based on the steps above, we have selected reducts 1,2,4,6 and 8 from Table 3, as shown in Table 4. It should be noted that in Table 4 only three features *F1*, *F2*, and *F3* out of four features are needed to unambiguously define the objects with output feature *O*. Based on the reducts in Table 4, a few decision rules can be derived. For example, a rule corresponding to object 4 is: IF input feature *F2* = 1, THEN output feature *O* = 2.

TABLE 4: REDUCTS SELECTED BASED ON EQUAL FEATURES

Object No.	Reduct No.	F1	F2	F3	F4	O
4	1	x	1	x	x	2
2	2	x	x	0	x	1
1	4	1	0	2	x	0
3	6	0	x	x	x	0
5	8	0	x	x	x	0

TABLE 5: REDUCTS SELECTED BASED ON UNEQUAL FEATURES

Object No.	Reduct No.	F1	F2	F3	F4	O
4	1	x	1	x	x	2
2	2	x	x	0	x	1
1	3	x	0	2	0	0
3	7	x	x	3	x	0
5	9	x	x	1	x	0
Weight		0.7	0.9	1.0	0.8	

3.3.2 Unequal weight features

The features are frequently unequal in nature. In order to select the features which are decisive for the objects' attributes and the rules identification, we select the reducts based on the following steps:

Step 1: Select the features used from only a single reduct of the object(s), which is the same as that in equal weight features.

Step 2: Select the features whose weight is largest, if the largest weight feature is not used in the reducts, the select the second largest weight feature, or the third, and so on until all objects have the reducts.

Assuming the weight of *F1* is 0.7, the weight of *F2* is 0.9, the weight of *F3* is 1.0, and the weight of *F4* is 0.8, respectively. Based on the aforementioned procedures, the reducts selected are shown in Table 5.

In Table 5, *F2*, *F3* and *F4* can be used to describe the five objects. Comparing with Table 4, Table 5 selects the features with higher weight, and rejects the feature with lower weight (i.e., *F1*). In other words, with the weight incorporated, the higher weight features have priority to be selected.

3.4 The Rule-Validation Procedure

The following steps are applied to examine the objects in the testing data set to estimate the validity of the rules derived from the above algorithm:

Step 1: Compare each decision rule derived from the rule composing algorithm with each new object from the testing data set. Calculate the number of objects that match with the rule.

Step 2: Repeat the comparisons of the decision rules with the objects from the testing data set until no decision rule is left.

Step 3: Calculate the accuracy of each rule by using the total matched objects divided by the summation of the total correctly matched objects and the total incorrectly matched objects. If the accuracy of the rule is greater than the predefined threshold value of confidence, then go to step 4; otherwise, remove the rule. Note that an incorrectly matched object means that the object contains the identical known value of conditional attributes with the rule, yet the outcomes are different from the rule.

Step 4: Stop and output the results of validated rules.

3.5 An Example

An example depicted here is about the distributor's performance indexes, including payment delay, ability of cost control, technical ability, infrastructure and equipment, marketing capability, deliveries/shipment and order quantity. In these seven features, most of the content of the features are continuous. Consequently, the discretization to the continuous feature is required. For example, for the "payment delay", if the delay period of a distributor is less than 2 weeks, then we rank this feature of this distributor as "very low"; if the period is between 2 and 4 weeks, then it is "low" ; between 4 and 6 weeks, it is "middle"; more than 6 weeks, it is "high." Also, all the indexes ranked as very low, low, middle, and high, described by 0, 1, 2 and 3. The output feature (*O*) is the general rank of a distributor, described as bad (0), normal (1), and good (2), see Table 6. Table 7 shows that there are 12 objects with the features.

TABLE 6: PARTIAL INPUT FEATURE SET OF MEASURE FOR DISTRIBUTOR

Feature	Content	Description	Weight
Fl	Payment delay	Whether there is a long time delay for the payment to the manufacturer.	0.7
F2	Cost control	The percentage of affiliated total cost to the end price.	0.9
F3	Technical ability	Whether the distributor can access to the high level techniques.	0.8
F4	Infrastructure and equipment	Whether the distributor has intranet or internet, communication networks.	0.7
F5	Marketing capability	The brand of the distributor, the scale of coverage of the distributor.	1.0
F6	Deliveries/shipments	How many trucks does the distributor have and how many third parts logistic does it access?	0.6
F7	Order quality	The frequency/quantity of the distributor.	0.7

TABLE 7: DATA SET WITH 12 OBJECTS

Object No.	F1	F2	F3	F4	F5	F6	F7	O
1	0	1	0	2	3	3	1	2
2	0	0	1	3	0	2	0	0
3	1	1	2	2	3	1	2	2
4	1	2	1	0	0	2	1	1
5	1	0	1	1	2	2	1	0
6	0	1	1	1	3	3	2	2
7	2	2	3	2	3	1	0	2
8	2	0	1	1	0	0	2	0
9	1	2	3	2	3	3	1	2
10	1	0	1	2	2	1	0	0
11	1	1	2	1	0	1	1	1
12	0	1	1	1	3	2	2	2

Using the bootstrapping method with ratio 0.632 for training set and 0.368 for testing set, the first eight objects as in training set and the remaining four as in testing objects were selected. Based on the reduct generation algorithms, one feature reducts can be derived as shown in Table 8.

TABLE 8: ONE FEATURE REDUCTS OF DATA SET IN TABLE 7

Object	Reduct	F1F2 F3 F4 F5 F6 F7 O	Object	Reduct	F1F2 F3 F4 F5 F6 F7 O
1	1	x 1 x x x x x 2	6	13	x 1 x x x x x 2
	2	x x 0 x x x x 2		14	x x x x 3 x x 2
	3	x x x 2 x x x 2		15	x x x x x 3 x 2
	4	x x x x x 3 x 2	7	16	x x 3 x x x x 2
	5	x x x x 3 x x 2		17	x x x 2 x x x 2
2	6	x 0 x x x x x 0		28	x x x x 3 x x 2
	7	x x x 3 x x x 0		19	x x x x x 1 x 2
3	8	x 1 x x x x x 2	8	20	x x x x x 0 x 0
	9	x x 2 x x x x 2		21	x 0 x x x x x 0
	10	x x x 2 x x x 2	4	22	x x x 0 x x x 1
	11	x x x x 3 x x 2	5	23	x 0 x x x x x 0
	12	x x x x x 1 x 2			

3.5.1 Equal weight features for rule identification

Using the rule identification algorithm, the favored reducts from Table 8 are chosen and listed in Table 9.

TABLE 9: THE SELECTED REDUTS WITH EQUAL WEIGHT

Reduct No.	F1	F2	F3	F4	F5	F6	F7	O	Matched object
1	x	x	x	0	x	x	x	1	[4]
2	x	0	x	x	x	x	x	0	[2] [5] [8]
3	x	1	x	x	x	x	x	2	[1] [3] [6]
4	x	x	x	3	x	x	x	0	[2]
5	x	x	x	2	x	x	x	2	[1] [3] [7]

After examining the candidate rules through domain experts, three rules can be derived:
(1) IF “cost control ability” is “very low”, THEN the performance is “bad”;
(2) IF “cost control ability” is “low”, THEN the performance is “good”;
(3) IF “Infrastructure and equipment” is “normal”, THEN the performance is “good”.

Using testing data, the accuracy of the rules can be calculated, as shown in Table 10.

TABLE 10: THE TESTING RESULT OF EQUAL WEIGHT FEATURES

Rule No.	Total objects	Matched objects	Matched percentage
Rule 1	1	1	100%
Rule 2	2	1	50%
Rule 3	3	2	66.7%

3.5.2 Unequal weight features for the rule identification algorithm

According to the steps for choosing unequal weight features reducts, the selected reducts are illustrated in Table 11.

TABLE 11: THE SELECTED REDUTS WITH UNEQUAL WEIGHT FEATURES

Reduct No.	F1	F2	F3	F4	F5	F6	F7	O	Matched object
1	x	x	x	0	x	x	x	1	[4]
2	x	0	x	x	x	x	x	0	[2] [5] [8]
3	x	x	x	x	3	x	x	2	[1] [3] [6] [7]
Weight	0.7	0.9	0.8	0.7	1.0	0.6	0.7		

With support of the domain experts, two rules can be derived:
(1) IF “cost control ability” is “very low”, THEN the performance is “bad”;
(2) IF “Marketing capability” is “high”, THEN the performance is “good”.

Using testing data, the accuracy of the rules can be calculated, as shown in Table 12.

TABLE 12: THE TESTING RESULT OF UNEQUAL WEIGHT FEATURES

Rule No.	Total objects	Matched objects	Matched percentage
Rule 1	1	1	100%
Rule 2	2	2	100%

Comparing the rules generated from equal and unequal weight features; the rules reflecting the more decisive features of an object can be concluded in the unequal weight case. More attention should be paid on the features included in the selected rules.

4. CASE STUDIES

The distributor selection is an important issue in supply chain management, especially in the current competitive marketing environment. Although the distributors face increasing challenges in a competitive environment (Kalafatis, 2000; Mudambi and Aggarwal, 2003), the power of distributors’ in marketing channels is getting stronger and stronger, which give much advantages in negotiation with vendors and buyers and makes it more crucial in selecting a good distributor for manufacturers.

Distributor selection involves evaluation and choice (Cavusgil et al., 1995). The evaluation task typically consists of identifying the attributes, criteria or factors relevant to the decision and then measuring or rating eligible distributors on each factor (Patton, 1996). The manufacturer's evaluation reflects an assessment of the value or rewards and risks inherent in the selection. In this study, a manufacturer produces a single product is our focus. It assumes that the manufacturer maintains stable product quality, stable and reliable product supply, and the manufacturer emphasizes much of the direct profit.

TABLE 13: THE WEIGHTED FEATURES FOR DISTRIBUTORS' SELECTION

Signal	Attribute	Description	Weight
F1	Financial strength	Distributors in good financial positions are likely to be well established and capable of selling many products for their manufacturing clients.	0.80
F2	Physical facilities	Adequate physical facilities, including modern technology and equipment may indicate a firm's capacity to carry out channel/supply chain task.	0.70
F3	Logistic capabilities	Capabilities in logistics provide an opportunity to achieve substantial cost savings while enhancing operational flexibility and creating value for customers.	0.65
F4	Sunk cost	Some cost that will never be gotten back such as the fee paid for the exclusive contract, the extra discount for the distributor and so on.	0.63
F5	Product line	Manufacturers typically prefer distributors who handle compatible and complementary products, rather than substitute products, especially avoiding distributors carrying directly competitive products.	0.54
F6	Market coverage	Adequate market coverage has been found necessary to gain an optimum volume of sales in each market, secure a reasonable market share and attain satisfactory market penetration, and therefore is important for manufacturers' distributor/channel member selection.	0.78
F7	Marketing experience	The market experience of a firm influences its competitive position, with experience helping the firm obtain better information, decrease uncertainty, and better handle managerial resources.	0.92
F8	Relationship intensity	Relationship intensity is defined as the degree of perceived reciprocity, closeness and friendliness in the relationship between the manufacturer and prospective distributor.	1.0
F9	Management ability	Management ability relates to management quality and operational competency. Many manufacturers feel that a supply chain member should only be considered if its management capabilities are good.	0.85

4.1 Attributes Identification

To determine the attributes for distributors' selection, we collected 15 attributes from the literature and practitioners in manufacturing companies. Then, we posted the 15 attributes to nine experts – five coming from different universities and four from several well known companies. These experts chose nine attributes which they believe the attributes are critical. After receiving the responses from the experts, the top nine attributes for distributors' selection were determined. Later, the top nine attributes were posted and the nine experts were asked for providing a weight of each attribute, the average weight from all nine experts' responses was taken as the final weight of the attribute. The final result is given in Table 13.

4.2 Data Preparation

In order to derive the rules for distributors' selection, top ten manufacturers have been selected and the sales or marketing department of each company has been requested to provide the scores of the 9 features of their distributors. 345 distributors' score associated with 9 features have been received. After data cleaning operation, 285 objects are left for further investigation. According to the ratio 0.632 for training set and 0.368 for testing set, 180 objects were selected at randomly for training and 105 objects for testing. Each attribute (feature) has been classified into three levels: "Very low", "Middle", and "High", and represented by "0", "1", and "2". For instance, if *F8* = 0, which means the relationship intensity is low; then *F3* = 2, which means logistic capabilities is high. For the output feature *O*, we also classified it into good (2), normal (1), and bad (0). Table 14 is the typical data set for 15 distributors.

4.3 Computational Results

4.3.1 The result of equal weight features

For the equal weight features study, the RSES software v.2.2 (Warsaw University Rough Set Exploration System (RSES) version 2.2, Logic Group, Inst. Mathematics, Warsaw Univ., http://logic.mimuw.edu.pl/~rses/) was applied for data analysis. Figures 3 and 4 demonstrate partial results through the computation procedure. Using "object related discernibility" and the "exhaustive algorithm," 104 reducts have been derived. After the evaluation conducted by the domain experts, seven candidate rules are selected (see Table 15). Then the candidate rules' accuracy has been examined using testing data set. The results are shown in Table 16.

TABLE 14: THE TYPICAL DATA SET

Object	F1	F2	F3	F4	F5	F6	F7	F8	F9	O
1	2	2	2	0	0	2	2	2	2	2
2	1	1	1	0	1	2	2	2	1	2
3	1	1	0	0	0	0	1	2	1	1
4	2	2	2	2	1	2	2	0	2	1
5	1	1	1	0	1	2	2	2	1	2
6	1	0	0	1	0	0	1	1	1	1
7	0	0	0	1	0	0	0	1	0	0
8	1	2	2	2	2	1	1	1	1	1
9	2	1	2	2	0	2	2	2	1	2
10	1	2	1	1	1	1	1	2	1	2
11	0	0	0	0	1	0	1	2	1	1
12	2	2	2	2	2	2	2	0	2	0
13	1	1	1	0	0	1	1	2	2	2
14	2	1	1	2	2	2	2	0	1	0
15	1	1	0	2	1	1	1	0	1	0

TABLE 15: CANDIDATE RULES WITH EQUAL WEIGHT FEATURES

Reduct	F1	F2	F3	F4	F5	F6	F7	F8	F9	O	Supporting objects
1	0	x	x	x	x	x	x	x	x	0	21
2	x	x	2	x	x	x	x	x	x	2	14
3	x	x	x	x	x	1	x	x	x	1	38
4	x	x	x	x	x	x	1	x	x	1	40
5	x	x	x	x	x	x	2	x	2	2	77
6	x	x	x	x	x	x	x	2	x	2	80
7	x	x	x	x	x	x	x	0	x	0	43

TABLE 16: RULE VALIDATION TESTING RESULT WITH EQUAL WEIGHT FEATURES

Rules	Matched	Total
Rule 1 IF “Financial strength” is “Very low” THEN “Bad” distributor.	13 accuracy	46 28.3%
Rule 2 IF “Logistic capabilities” is “High” THEN “Good” distributor.	8 accuracy	39 20.5%
Rule 3 IF “Market coverage” is “Normal” THEN “Normal” distributor.	13 accuracy	56 23.2%
Rule 4 IF “Marketing experience” is “Normal” THEN “Normal” distributor.	46 accuracy	48 95.8%
Rule 5 IF “Marketing experience” is “High” and “Management ability” is “High” THEN “Good” distributor.	32 accuracy	36 88.9%
Rule 6 IF “Relationship intensity” is “High” THEN “Good” distributor.	50 accuracy	55 90.9%
Rule 7 IF “Relationship intensity” is “Very low” THEN “Bad” distributor.	24 accuracy	36 66.7%

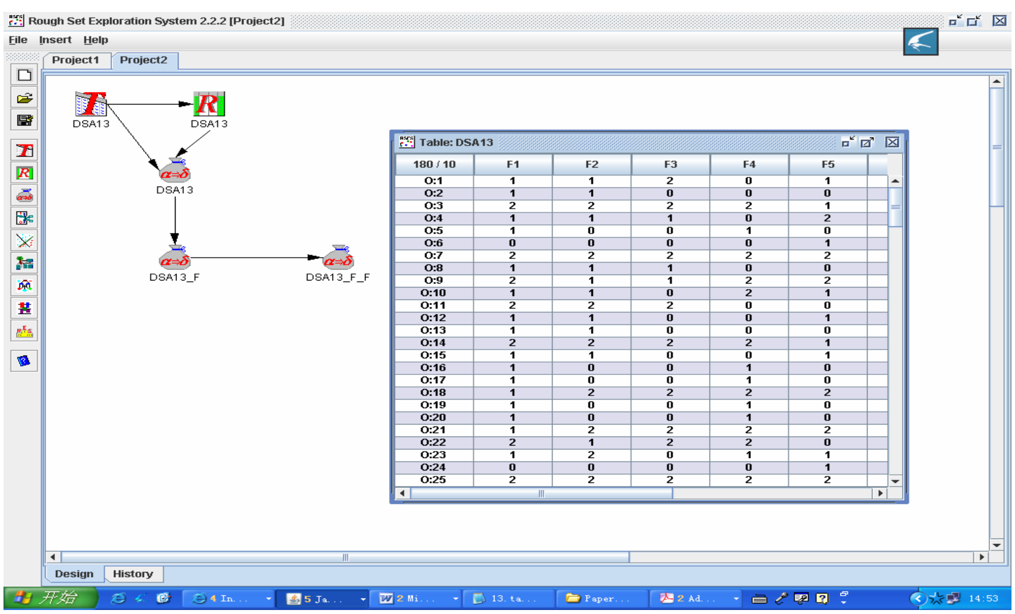
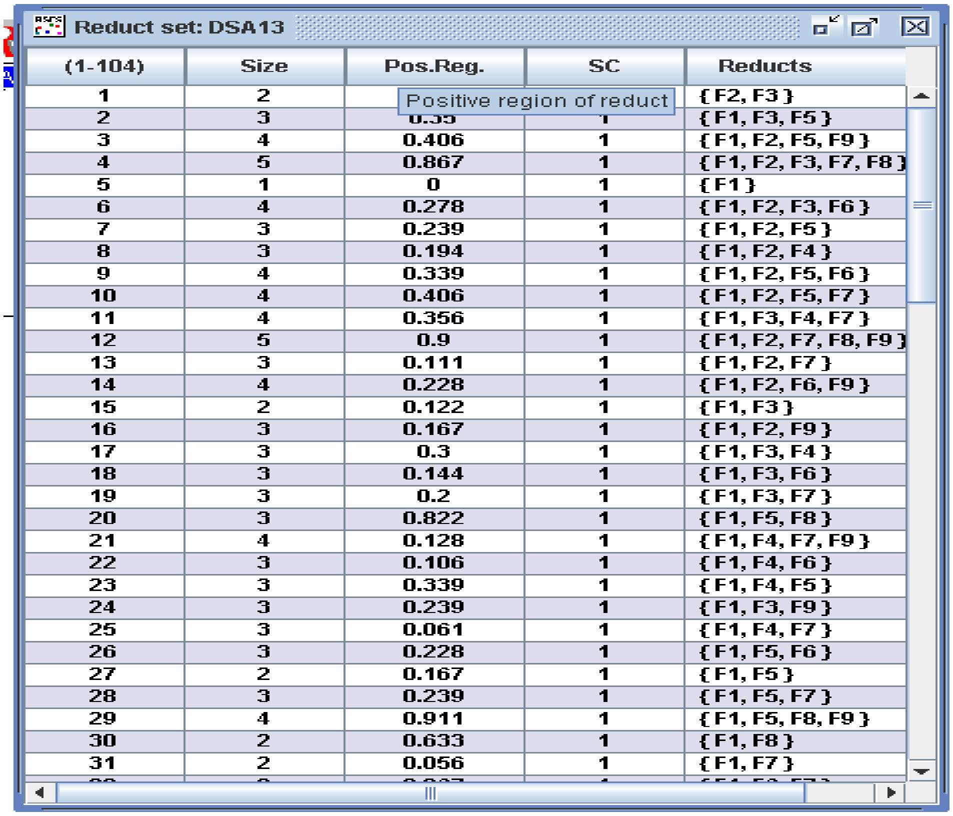


FIGURE 3: THE OBJECTS FOR ANALYSIS



Reduct set: DSA13

(1-104)	Size	Pos.Reg.	SC	Reducts
1	2	Positive region of reduct		{ F2, F3 }
2	3	0.35	1	{ F1, F3, F5 }
3	4	0.406	1	{ F1, F2, F5, F9 }
4	5	0.867	1	{ F1, F2, F3, F7, F8 }
5	1	0	1	{ F1 }
6	4	0.278	1	{ F1, F2, F3, F6 }
7	3	0.239	1	{ F1, F2, F5 }
8	3	0.194	1	{ F1, F2, F4 }
9	4	0.339	1	{ F1, F2, F5, F6 }
10	4	0.406	1	{ F1, F2, F5, F7 }
11	4	0.356	1	{ F1, F3, F4, F7 }
12	5	0.9	1	{ F1, F2, F7, F8, F9 }
13	3	0.111	1	{ F1, F2, F7 }
14	4	0.228	1	{ F1, F2, F6, F9 }
15	2	0.122	1	{ F1, F3 }
16	3	0.167	1	{ F1, F2, F9 }
17	3	0.3	1	{ F1, F3, F4 }
18	3	0.144	1	{ F1, F3, F6 }
19	3	0.2	1	{ F1, F3, F7 }
20	3	0.822	1	{ F1, F5, F8 }
21	4	0.128	1	{ F1, F4, F7, F9 }
22	3	0.106	1	{ F1, F4, F6 }
23	3	0.339	1	{ F1, F4, F5 }
24	3	0.239	1	{ F1, F3, F9 }
25	3	0.061	1	{ F1, F4, F7 }
26	3	0.228	1	{ F1, F5, F6 }
27	2	0.167	1	{ F1, F5 }
28	3	0.239	1	{ F1, F5, F7 }
29	4	0.911	1	{ F1, F5, F8, F9 }
30	2	0.633	1	{ F1, F8 }
31	2	0.056	1	{ F1, F7 }

FIGURE 4: THE REDUCTS OF THE DSA MODEL

4.3.2 The result of unequal weight features

For unequal weight features analysis, the "rough set-based decision support system" software (see Figure 5) was applied. With the help of the domain experts, five candidate rules are set as shown in Table 17.

TABLE 17: CANDIDATE RULES OF UNEQUAL WEIGHT FEATURES

Reduct	F1	F2	F3	F4	F5	F6	F7	F8	F9	O	Supporting objects
1	x	x	x	x	x	x	x	2	x	2	80
2	x	x	x	x	x	x	2	x	2	2	77
3	x	x	x	x	x	x	x	0	x	0	43
4	x	x	x	x	x	0	x	0	x	0	59
5	x	x	x	x	x	x	1	x	x	1	40
Weight	0.80	0.70	0.65	0.63	0.54	0.78	0.92	1.0	0.85		

Then the candidate rules' accuracy was examined using testing data set. The results are shown in Table 18.

4.3.3 Comparison between equal and unequal weight features

Compared the results in Tables 16 and 18, we find that the weight incorporated features rules identification can give out more accurate rules than that of equal ones. In addition, we discussed the derived rules with domain experts and found that these 5 rules in Table 18 are corresponding to the distributors' selection in China, especially rule 1, which demonstrates the importance of "relationship". It also is a crucial part of the Chinese tradition. The results also demonstrate the important attributes of a distributor are its software, such as relationship intensity, marketing experience and management ability, rather than its hardware, such as financial strength, physical facilities or logistic capabilities.

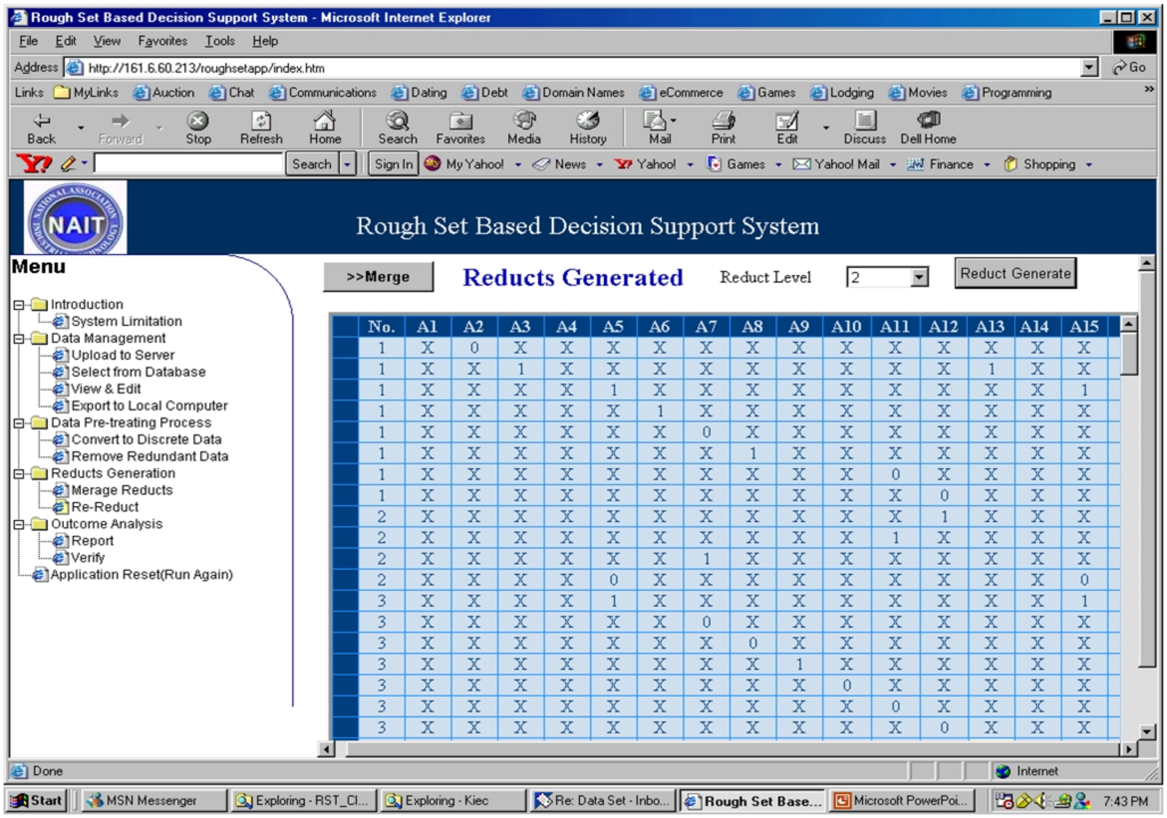


FIGURE 5: ROUGH SET-BASED DECISION SUPPORT SYSTEM

TABLE 18: RULE VALIDATION TESTING RESULT OF EQUAL WEIGHT FEATURES

Rules	Matched	Total
Rule 1 IF "Relationship intensity" is "High" THEN "Good" distributor	50 accuracy	55 90.9%
Rule 2 IF "Marketing experience" is "High" and "Management ability" is "High" THEN "Good" distributor	32 accuracy	36 88.9%
Rule 3 IF "Relationship intensity" is "Very low" THEN "Bad" distributor	24 accuracy	36 66.7%
Rule 4 IF "Marketing coverage" and "Relationship intensity" are both "Very low" THEN "Bad" distributor	24 accuracy	25 96%
Rule 5 IF "Marketing experience" is "Normal" THEN "Normal" distributor	46 accuracy	48 95.8%

5. CONCLUSIONS

In this paper, distributors' selection is analyzed based on the Rough Set Theory approach with both equal and unequal weight features. Through this method, several rules are generated for distributors' evaluation and selection. Our result not only shows the effectiveness of unequal weight incorporated rules identification, but also it shows the importance of the relationship intensity, marketing experience, and the management ability in selecting the distributors. These rules have been shown to be useful and convenient to conduct a selection process for the manufacturers. Moreover, the derived rules provide an important implication – all constituencies in the supply chain should maintain an intensity relationship with each other.

REFERENCES

[1] Abratt, R. and Pitt, L.F. "Selection and Motivation of Industrial Distributors: A Comparative Analysis." *European Journal of Marketing*, 1989, 23, 144-53.

[2] Beynon, M., Curry, B., and Morgan, P. "Knowledge Discovery in Marketing: An Approach through Rough Set Theory." *European Journal of Marketing*, 2001, 35, 915-935.

[3] Cavusgil, S.T., Yeoh, P., and Mitri, M. "Selecting Foreign Distributors: An Expert Systems Approach." *Industrial Marketing Management*, 1995, 24, 297-304.

[4] Chan, F.T.S. and Kumar, N. "Global Supplier Development Considering Risk Factors using Fuzzy Extended AHP-based Approach." *Omega*, 2007, 35, 417-431.

[5] Chien, C.-F. and Chen, L.-F. "Using Rough Set Theory to Recruit and Retain High-Potential Talents for Semiconductor Manufacturing." *IEEE Transactions on Semiconductor Manufacturing*, 2007, 20, 528-541.

[6] Daubie, M., Levecq, P., and Meskens, N. "A Comparison of the Rough Sets and Recursive Partitioning Induction Approaches: An Application to Commercial Loans." *International Transactions in Operational Research*, 2002, 9, 681-694.

[7] Dimitras, A.I., Slowinski, R., Susmaga, R., and Zopounidis, C. "Business Failure Prediction using Rough Sets." *European Journal of Operational Research*, 1999, 114, 263-280.

[8] Dubois, D. and Prade, H. "Rough Fuzzy Sets and Fuzzy Rough Sets." *International Journal of General Systems*, 1990, 17, 191-209.

[9] Fonsson, P. and Zineldin, M. "Achieving High Satisfaction in Supplier-Dealer Working Relationships." *Supply Chain Management: An International Journal*, 2003, 8, 224-240.

[10] Fram, E.H. "We can do a Better Job of Selecting International Distributors." *Journal of Business & Industrial Marketing*, 1992, 7, 61-70.

[11] Ghodsypour, S.H. and O'Brien, C. "A Decision Support System for Supplier Selection using an Integrated Analytical Hierarchy Pprocess and Linear Programming." *International Journal of Production Economics*, 1998, 56/57, 199-212.

[12] Greco, S., Matarazzo, B., and Slowinski, R. "Extension of the Rough Set Theory Approach to Multicriteria Decision Support." *INFOR*, 2000, 38, 161-196.

[13] Heckerman, D., Mannila, H., Pregibon, D., and Uthurusamy, R. *Proceedings of 3rd International Conference on Knowledge Discovery and Data Mining*. Menlo Park, CA: AAAI Press, 1997.

[14] Kahraman, C., Cebeci, U., and Ruan, D. "Multi-attribute Comparison of Catering Service Companies using Fuzzy AHP: The Case of Turkey." *International Journal of Production Economics*, 2004, 87, 171-184.

[15] Kahraman, C., Cebeci, U., and Ulukan, Z. "Multi Criteria Supplier Selection using Fuzzy AHP." *Logistics Information Management*, 2003, 16, 382-394.

[16] Kalafatis, S.P. "Buyer-Seller Relationships along Channels of Distribution." *Industrial Marketing Management*, 2000, 31, 215-228.

[17] Komorowski, J. and Zytkow, J. *Principles of Data Mining and Knowledge Discovery*. New York: Springer-Verlag, 1997.

[18] Kumar, A, Agrawal, D.P., Joshi, S.D. "Advertising Data Analysis using Rough Sets Model." *International Journal of Information Technology & Decision Making*, 2005, 4, 263-276.

[19] Kusiak, A. "Rough Set Theory: A Data Mining Tool for Semiconductor Manufacturing." *IEEE Transactions on Electronics Packaging Manufacturing*, 2001, 24, 44-50.

[20] Kusiak, A., Kern, J.A., Kernstine, K.H., and Tseng, T.-L. "Autonomous Decision-Making: A Data Mining Approach." *IEEE Transactions on Information Technology in Biomedicine*, 2000, 4, 274-284.

[21] Lin, F.-S. and Chen, C.-R. "Determinants of Manufacturers' Selection of Distributor." *Supply chain management: An International Journal*, 2008, 13, 356-365.

[22] Lindqvist, L.J. "Current Trends in Distribution Research in Finland." *International Journal of Physical & Logistics Management*, 1983, 13, 105-116.

[23] Mickee, T.E. "Rough Sets Bankruptcy Prediction Models versus Auditor Signaling Rates." *Journal of Forecasting*, 2003, 22, 569-586.

[24] Mudambi, S. and Aggarwal, R. "Industrial Distributors: Can they Survive in the New Economy?" *Industrial Marketing Management*, 2003, 32, 317-325.

[25] Oliveria, R.C. and Lourenco, J.C. "A Multicriteria Model for Assigning New Orders to Service Suppliers." *European Journal of Operational Research*, 2002, 139, 390-399.

[26] Patton, W.E. "Use of Human Judgment Models in Industrial Buyers' Vendor Selection Decisions." *Industrial Marketing Management*, 1996, 25, 135-149.

[27] Pawlak, Z. "Rough Sets." *International Journal of Computer and Information Sciences*, 1982, 11, 341-356.

[28] Pawlak, Z. *Rough Sets*, Kluwer Academic Publishers, Dordrecht, the Netherlands, 1991.

[29] Ross, R.E. "Selection of the Overseas Distributor: An Empirical Framework." *International Journal of Physical & Logistics Management*, 1973, 3, 83-90.

[30] Sevkli, M., Lenny Koh, S.C., Zaim, S., Demirbag, M., and Tatoglu, E. "Hybrid Analytical Hierarchy Process Model for Supplier Selection." *Industrial Management & Data System*, 2008, 108, 122-142.

[31] Sharma, D., Sahay, B.S., and Sachan, A. "Modeling Distributor Performance Index Using System Dynamics Approach." *Asia Pacific Journal of Marketing and Logistics*, 2004, 16, 37-67.

[32] Shen, L., Tay, F.E.H., Qu, L., and Shen, Y. "Fault Diagnosis using Rough Set Theory." *Computers in Industry*, 2000, 43, 61-72.

[33] Shipley, D.D., Cook, D., and Barnett, E. "Recruitment, Motivation, Training and Evaluation of Overseas Distributors." *European Journal of Marketing*, 1989, 23, 79-93.

[34] Tseng, T.-L., Huang, C.-C., Jiang, F., and Ho, J.C. "Applying a Hybrid Data-mining Approach to Prediction Problems: A Case of Preferred Suppliers Prediction." *International Journal of Production Research*, 2006, 44, 2935-2954.

[35] Wang, L. and Kess, P. "Partnership Motives and Partner Selection – Case studies of Finnish Distributor Relationships in China." *International Journal of Physical & Logistics Management*, 2006, 36, 466-478.

[36] Wei, L.-L. and Zhang, W.-X. "Probabilistic Rough Sets Characterized by Fuzzy Sets." *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, 2004, 12, 47-60.

[37] Yan, H., Yu, Z., and Cheng, T.C.E. "A Strategic Model for Supply Chain Design with Logical Constraints: Formulation and Solution." *Computers & Operations Research*, 2003, 30, 2135-2155.

[38] Yeoh, P. and Calantone, R.J. "An Application of the Analytical Hierarchy Process to International Marketing: Selection of a Foreign Distributor." *Journal of Global Marketing*, 1995, 8, 39-65.