

An Enterprise Competitive Capability Evaluation Based on Rough Sets

Ming-Chang Lee
Department of Information Management
Fooyin University, Kaohsiung, Taiwan, R.O.C.

Email: ming_li@mail2000.com.tw

ABSTRACT

Rough sets are an effective mathematical analysis tool to deal with vagueness and uncertainty in the area of decision analysis and synthesis evaluation. Information Entropy, as a measurement of the average amount of information contained in an information system, is used in the classification of objectives and the analysis of information systems. The weight of synthesis evaluation is determined by expert, lending to subjectivity and without considering the redundancy of attributes exists in traditional synthetic evaluation. Therefore, in this paper we apply the importance measure of attribute based on information entropy to create weight of each attribute. The procedure indicates that the approach is practical and effective.

Keywords: Enterprise competitive capability, Attribute Reduction, Rough set theory, Information entropy

Acknowledgement: I would like to thank the anonymous reviewers for their constructive comments on this paper.

INTRODUCTION

Enterprise competitive capability advantage evaluation is one of business management major direction. Many studies have demonstrated that data mining such as Analytic Hierarchy Process (AHP), Fuzzy synthesis evaluation, neural networks, Data Envelopment Analysis (DEA) and Grey Relation Analysis (GRA) (Ahn, et al. 2000; Hawley et al. 2000; Lin and piessse, 2004; Liu and Shin, 2005; Lin, et al. 2007; Xu and Lin 2010; Zhang and Li 2006; Road et al. 2011; Zhang et al. 2006).

Rough set is a mathematics method that can be used to deal with incomplete and imprecise knowledge. The result is that the connection of data is discovered and useful feature is extracted and concise knowledge expression is gained. The rough set theory has been successfully applied in a variety of fields, including: financial distress classification (Lee, 2007), business failure prediction (Beynon and peel, 2001), travel demand analysis (Goh and Law, 2003), mining stock prices (Wang, 2003), insurance market (Shyng et al. 2007), accident prevention (Wong and chung, 2007), customer relationship management (Liou, 2009) etc. Recently, the rough set theory has become a popular evaluation technique for classification problems because of their strength of handing vague and imprecise data (Jiang and Ruan, 2010). It can extract knowledge from the data itself by mean of indiscernible relations, and generally needs fewer calculations than that of other soft computing techniques. The theory of rough set deals with the approximation of arbitrary subsets of a universe by two definable or observable subsets called lower and upper approximations. By using the concept of lower and upper approximation in rough set theory, the attributes in an information system may be redundant and thus can be eliminated without losing essential classificatory information (Knodo, 2006; Liu and Hu, 2007). In this paper, we define the important measure of attribute and information entropy. Using this definition, we can easily calculate the weight acquisition of attributes.

The purpose of this research is to apply the importance measure of attribute based on information entropy to create weight of each attribute and create enterprise competitive capability evaluation function. The proposed approaches are:

- (1) The establishment of related objects of a knowledge system in enterprise competitive capability.
- (2) Attribute reduction for the knowledge system, and find the attribute core value.
- (3) Calculation the important measure of attributes and calculation of the weight of attributes.
- (4) Create enterprise competitive capability evaluation function.

The rest of this paper is organized as follows. In section 2 we discussed the concept of rough set theory. The algorithm of weight acquisition of attributes and information gain are discussed in section 3. Section 4 gives an illustration. It help in understanding of this procedure, a demonstrative illustration is given to show the key stages involving the use of the introduced concepts. Section 5 is conclusion.

THE CONCEPT OF ROUGH SET THEORY

In this section, we discussed the concept of indiscernible relations. More introductions the main concepts of the theory can be found in Jing and Yanzhi (2007) and through theoretical foundation are illustrated in (Jiang and Ruan, (2010). The basic concepts of rough set theory are as follows:

Concept 1: Knowledge systems

Given an information system (a data set), $S = \{U, A, V, f\}$, where U and A are finite and nonempty sets called the universe, and the sets of attributes, respectively. In information system, there exists a function, such that $f: U \times A \rightarrow V$. A is the union of C and D , the intersection of C and D is empty. C is called as conditional attributes, and D is called as decision attributes. The information system is also called decision system, or knowledge system.

Concept 2: Indiscernible relation

Indiscernible relation is equivalence relation in U , and P is a subset of C ($P \subseteq C$). $IND(P)$, called the indiscernible relation is defined as follows: $IND(P) = \{(x, y) \in U \times U : f(x, a) = f(y, a), \text{ for all } a \in P\}$.

Concept 3: Equivalence classes

Let $U / IND(P)$ be the family of all Equivalence classes of the relation $IND(P)$. For simplicity of notation U/P will be written instead of $U/IND(P)$.

$U/IND(C)$ and $U/IND(D)$ will be called condition and decision classes, respectively.

Concept 4: Attribute reduction and classification

Attribute reduction is one of the central of rough set theory. It is well-known that attribute is not same important in repository, even some attributes are redundant. Some attributes in an information system can be eliminated without losing essential classificatory information. The process of finding a smaller set of attributes with the same or close classificatory power as the original set is called attribute reduction. Through the process of attribute reduction, redundant attributes, called superfluous attributes are removed without losing the classified power of a reduced information system.

Let R be an equivalence relation ($r \in R$). If $IND(R) = IND(R - \{r\})$, then r is thought unnecessary for C , otherwise r is thought necessary for C .

Concept 5: Upper approximation and lower approximation (Jiang and Yanzhi, 2007)

Let U be a non-empty set of finite objects (the universe), R be a subset of A , and X be a subset of U . The lower approximation of set X , denoted by $\underline{R}(X)$ is a union of all elementary sets; objects in this lower approximation unambiguously belong to set X .

$$\underline{R}(X) = \cup \{Y \in U / R \mid Y \subseteq X\} \quad (1)$$

The upper approximation of set X , denoted by $\overline{R}(X)$ is a union of all elementary sets each of which has no-empty intersection with X ; objects in this upper approximation possible belong to set X .

$$\bar{R}(X) = \cup \{Y \in U / R \mid Y \cap X \neq \emptyset\} \quad (2)$$

The R-lower approximation of X is the set of all objects, which can be certainly classified to X using attributes from R. The set $U - \bar{R}(X)$ is the R-outside region of X and consists of those objects, which can be with certainty classified as not belonging to X using attributes from R. The set $BN_R(X) = \bar{R}(X) - \underline{R}(X)$ is the R- boundary region of X.

If $\bar{R}(X) = \underline{R}(X)$, X is exact set, otherwise X is rough set.

Concept 6: Dispensable and Indispensable Features

Given an information system, $S = \{U, A, V, f\}$, $A = C \cup D$. The positive region in D definite as:

$$Pos_C(D) = \cup_{x \in U / D} \underline{R}(X) \quad (3)$$

If $Pos_C(D) \neq Pos_{C-\{a\}}(D)$, the condition attributes “a” is indispensable attribute in C; otherwise the condition attributes “a” is dispensable attribute in C.

A is an independent, if all $c \in C$ are indispensable.

Concept 7: Concepts of attribute reduction and core

Supposing $R \subseteq C$, if R is independent and $IND(R) = IND(C)$, then R is thought as reduction of C. The set that is composed by all necessary relation of C is called core of C and marked CORE (C).

There is relation between CORE and relation as follows.

$$CORE(C) = \cap RED(C) \quad (4)$$

$RED(C)$ means all the reduction of C.

THE ALGORITHM OF WEIGHT ACQUISITION

This section presents the weight acquisition method based on rough set theory and Information gain. The result of the lower approximation can describe the creditable knowledge in information system and the weight of an attribute can be estimated by the variety rating of the lower approximation when the attribute is deleted. One measure to describe the inexactness of approximation classification is called quality of approximation of D by means of the attributes from C.

Definition 1: Quality of approximation and σ - important rating

Given an information system, $S = \{U, A, V, f\}$, the quality of approximation of D by means of the attributes from C is denoted

$$\gamma_c(D) = \frac{\sum_{i=1}^n card(\underline{R}(X_i))}{card(U)} \quad (5)$$

Where $card(U)$ denotes the cardinality of set U

The σ - important rating of attribute a is defined as

$$\sigma_{cd}(a) = \gamma_c(D) - \gamma_{c-\{a\}}(D) \quad (6)$$

Definition 2: Information entropy

Given an information system, $S = \{U, A, V, f\}$, If $P \subseteq A$, $U / \text{IND}(P) = \{X_1, X_2, \dots, X_n\}$ is a n equivalence relation on U . $I(P)$ is called as Information entropy.

$$I(X) = - \sum_{i=1}^n \frac{|X_i|}{|U|} (1 - \log_2 \frac{|X_i|}{|U|}) \tag{7}$$

Where, $|X_i| / |U|$ is card in X_i

Definition 3: Concepts of attribute importance

Given knowledge system $S = (U, A, V, f)$, C is called a set of conditional attribute, D is called a set of decision attribute and $C \cap D = \emptyset$ and $A = C \cup D$. $f: U \times U = V_i$ is an information function. The important measure of attribute ‘‘a’’ is defined:

$$\text{SGF}(a) = I(C) - I(C - \{a\}) \tag{8}$$

Where $a \in C$

When $\text{SGF}(a) > 0$, it is denoted that attribute ‘‘a’’ is need. When $\text{SGF}(a) = 0$, ‘‘a’’ is redundant attribution, that is, a can leave out from the attribution’s set.

If $\text{SGF}(a) > \text{SGF}(b)$, the attribute ‘‘a’’ is more important than attribute ‘‘b’’ in condition C .

Definition 4: The weights based on information entropy

Given information system $S = (U, A, V, f)$, C is called a set of conditional attribute, D is called a set of decision attribute and $C \cap D = \emptyset$ and $A = C \cup D$. $f: U \times U = V_i$ is an information function. $a_i \in A = \{a_1, a_2, \dots, a_n\}$. The attribute weight of a_i denoted:

$$w_i = \text{SGF}(a_i) / \sum_{i=1}^n \text{SGF}(a_i) \tag{9}$$

Definition 5: Create enterprise competitive capability evaluation function

After constructed the weights of each attribute, we create enterprise competitive capability evaluation function.

$$p(y) = \sum_{i=1}^k w_i y_i \tag{10}$$

Where y_i denotes the value of the i^{th} competitive capability evaluation index.

ILLUSTRATION- ENTERPRISE COMPETITIVE CAPABILITY EVALUATION

Step 1: Enterprise Competitive Capability Evaluation Index.

The empirical research which is based on the data Taiwan’s listed company (food stocks). 9 experimental samples are select in 2008. Therefore, $U = \{1, 2, \dots, 9\}$. In this study, we use 12 enterprise competitive capability evaluated indexes are: Return on total assets (x_1), profit ratio of sales (x_2), Profit ratio of total capital (x_3), Working capital to sales ratio (x_4), Inventory turnover ratio (x_5), Account receivable turnover ratio (x_6), Current ratio (x_7), Asset-liability current ratio (x_8), Current Liabilities ratio (x_9), Asset-liability ratio (x_{10}), Equity ratio (x_{11}), pretax profit current debit ratio (x_{12}).

Therefore, $C = \{x_1, x_2, \dots, x_{12}\}$, $D = \{0, 1\}$, where 0 denotes the stock price > 10 , and 1 denotes the stock price ≥ 10 .

Set each competitive capability evaluated index threshold value, which is the average value of 20 enterprises on lately 10 years financial index. The competitive capability evaluated index threshold values are 25.0, 34.0, 330, 0.75, 635.0, 78.0, 130., 24.0, 128.0, 24.8, 256.0, 186.0. For example, the threshold value of attribute “Return on total assets” is 25% (see Table 1). If $(x_1) > 25\%$, we set (x_1) equal to 1; otherwise set (x_1) equal to 0. According to the threshold value, we obtain the enterprise competitive capability evaluated discrete decision table (see Table 2).

Table 1: Enterprise competitive capability evaluated indexes

U	X ₁ (%)	X ₂ (%)	X ₃ (%)	X ₄ (%)	X ₅ (%)	X ₆ (%)	X ₇ (%)
1	22.35	35.12	25.30	3.27	720.00	83.6	160.0
2	18.52	44.12	75.45	6.82	645.01	72.7	132.4
3	14.60	19.20	15.63	2.54	685.54	54.12	112.6
4	22.15	38.34	251.3	6.37	702.02	67.96	148.5
5	40.28	25.30	150.5	2.58	420.05	89.54	88.54
6	20.15	39.65	354.2	7.92	255.08	75.51	137.9
7	48.36	45.80	752.9	21.30	421.79	60.24	156.3
8	35.50	34.85	425.8	10.55	621.02	68.56	135.9
9	42.72	30.12	560.7	8.12	435.12	84.20	120.42
◦	25.0	34.0	330	7.5	635.0	78.0	130.0

Table 1: Enterprise competitive capability evaluated indexes (continue)

U	X ₈ (%)	X ₉ (%)	X ₁₀ (%)	X ₁₁ (%)	X ₁₂ (%) ₂	D
1	21.5	105.9	21.9	278.2	189.5	0
2	12.38	153.4	19.8	198.5	335.2	1
3	16.72	110.2	18.54	214.6	19.58	0
4	14.32	190.5	11.24	225.1	210.5	1
5	22.17	88.5	20.19	345.8	124.7	0
6	32.54	154.2	41.80	238.7	188.55	0
7	32.45	161.6	28.58	262.1	190.84	1
8	25.87	115.1	34.20	214.5	250.85	1
9	45.12	120.3	25.95	259.8	155.85	0
◦	24.0	128.0	24.80	256.0	186.0	

◦ : Threshold value

Table 2: The Enterprise Competitive Capability Evaluated Discrete Decision Table

U	X ₁ (%)	X ₂ (%)	X ₃ (%)	X ₄ (%)	X ₅ (%)	X ₆ (%)	X ₇ (%)	X ₈ (%)	X ₉ (%)	X ₁₀ (%)	X ₁₁ (%)	X ₁₂ (%)	D
1	0	1	0	0	1	1	1	0	0	0	1	1	0
2	0	1	0	0	1	0	1	0	1	0	0	1	1
3	0	0	0	0	1	0	0	0	0	0	0	0	0
4	0	1	0	0	1	0	1	0	0	0	0	1	1
5	1	0	0	0	0	1	0	0	0	0	1	0	0
6	0	1	1	1	0	0	1	1	1	1	0	1	0
7	1	1	1	1	0	0	1	1	1	1	0	1	1
8	1	1	1	1	1	0	1	1	0	1	0	1	1
9	1	0	1	1	0	1	0	1	0	1	1	0	0

Step 2: Reduction of Attributes

We find out that the attributes X₃, X₄, X₈, X₁₀ have the same value in table 2, and therefore we remove attributes X₄, X₈, X₁₀; the attributes X₂, X₇, X₁₂ have the same value in table 2, and therefore we remove attributes X₇, X₁₂; the attributes X₆, X₁₁ have the same value in table 2, and therefore we remove attributes X₆. Based algorithm of reduction, the reduce attributes are {X₁, X₂, X₃, X₅, X₉, X₁₁}.

We calculate the minimum attributes set by using concept 2 and 3.

$$\begin{aligned}
 \text{IND}(C) &= \{1, 2, 3, 4, 5, 6, 7, 8, 9\} \\
 \text{IND}(D) &= \{(1, 3, 5, 6, 9), (2, 4, 7, 8)\} \\
 \text{IND}(C-(x_1)) &= \{1, 2, 3, 4, 5, (6, 7), 8, 9\} \neq \text{IND}(C) \\
 \text{IND}(C-(x_2)) &= \{1, 2, (4, 5), 3, 6, 7, 8, 9\} \neq \text{IND}(C) \\
 \text{IND}(C-(x_3)) &= \{1, 2, 3, 4, (5, 9), 6, 7, 8\} \neq \text{IND}(C) \\
 \text{IND}(C-(x_5)) &= \{1, 2, 3, 4, 5, 6, 7, 8, 9\} = \text{IND}(C) \\
 \text{IND}(C-(x_9)) &= \{1, 3, 5, 6, (2, 4), 7, 8, 9\} \neq \text{IND}(C) \\
 \text{IND}(C-(x_{11})) &= \{(1, 4), 2, 3, 5, 6, 7, 8, 9\} \neq \text{IND}(C)
 \end{aligned}$$

Under concept 5, we calculate the positive region in D.

$$\begin{aligned}
 \text{POS}_C(D) &= \{1, 2, 3, 4, 5, 6, 7, 8, 9\} \\
 \text{POS}_{C-(x_1)}(D) &= \{1, 2, 3, 4, 5, 8, 9\} \neq \text{POS}_C(D) \\
 \text{POS}_{C-(x_2)}(D) &= \{1, 2, 3, 6, 7, 8, 9\} \neq \text{POS}_C(D) \\
 \text{POS}_{C-(x_3)}(D) &= \{1, 2, 3, 4, 6, 7, 8, 9\} \neq \text{POS}_C(D) \\
 \text{POS}_{C-(x_5)}(D) &= \{1, 2, 3, 4, 5, 6, 7, 8, 9\} = \text{POS}_C(D) \\
 \text{POS}_{C-(x_9)}(D) &= \{1, 3, 5, 6, 7, 8, 9\} \neq \text{POS}_C(D) \\
 \text{POS}_{C-(x_{11})}(D) &= \{2, 3, 5, 6, 7, 8, 9\} \neq \text{POS}_C(D)
 \end{aligned}$$

The attributes X₅ is dispensable in C by using concept 6 and the minimum attributes set is {X₁, X₂, X₃, X₉, X₁₁}. The enterprise competitive capability evaluated reduction decision table is showed as Table 3.

Table 3: The Enterprise Competitive Capability Evaluated Reduction Decision Table

U	X ₁	X ₂	X ₃	X ₉	X ₁₁	D
1	0	1	0	0	1	0
2	0	1	0	1	0	1
3	0	0	0	0	0	0
4	0	1	0	0	0	1
5	1	0	0	0	1	0
6	0	1	1	1	0	0
7	1	1	1	1	0	1
8	1	1	1	0	0	1
9	1	0	1	0	1	0

Step 3: Calculation the Importance Measure of Attributes and Calculation of the Weight of Attributes

We calculate the important measure of attribute by using definition 2 and 3

$$SGF(x_1) = I(C) - I(C - \{x_1\}) = 4.1699 - 3.3176 = 0.8523$$

$$SGF(x_2) = SGF(x_3) = SGF(x_9) = SGF(x_{11}) = 0.8523$$

We calculate the weight of attributes x_i by using definition 4:

$$w_i = SGF(a_i) / \sum_{i=1}^n SGF(a_i) = 0.2 \quad (i = 1, 2, \dots, 5)$$

Table 4: The Enterprise Competitive Capability Evaluation Function

Company	X ₁	X ₂	X ₃	X ₉	X ₁₁	Evaluation function of Weight w_i	Rank
1	21.5	35.12	25.30	105.9	278.2	81.88	8
2	12.38	44.12	75.45	153.4	198.5	85.47	7
3	16.72	19.20	15.63	110.2	214.6	68.086	9
4	14.32	38.34	251.3	190.5	225.1	133.38	5
5	22.17	25.30	150.5	88.5	345.8	116.96	6
6	32.54	39.65	354.2	154.2	238.7	149.42	4
7	32.45	45.80	752.9	161.6	262.1	235.32	1
8	25.87	34.85	425.8	115.1	214.5	151.08	3
9	45.12	30.12	560.7	120.3	259.8	188.16	2

Step 4: Create Enterprise Competitive Capability Evaluation Function

We calculate the enterprise competitive capability evaluation function by using eq. 10. Where w_i is called important measure of attribute, in this example $w_i = 0.2$ ($i=1,2, \dots, 5$).

Therefore, we obtain the order sequence of company (see Table 4). In addition, the process of the weight acquisition is based on the data from the information system and don't add any expert's preference.

CONCLUSIONS

The weight of synthesis evaluation is determined by expert, lending to subjectivity and without considering the redundancy of attributes exists in traditional synthetic evaluation. The advantage of this method is that it eliminates the personal subjectivity as and deals with the redundancy of attributes properly. This study presents an approach based on rough set theories, which can find out evaluation to enterprise competitive capability. The experimental above can prove that the approach is practical and effective. We think that rough set theory will do more achievement in this domain. Our future research will work more effective method for weight generation such as fuzzy rough set theory.

REFERENCES

- Ahn, B. S., Cho, S. & Smkim, C. Y. (2000). The integrated methodology of rough sets and artificial neural network for business failure prediction. *Expert system with application*, 18(4), 65-74.
- Beynon, M. J. & peel, M. J. (2001). Variable precision rough set theory and data discretization: an application to corporate failure prediction. *Omega*, 29(6), 561-576.
- Goh, C. & Law, R. (2003). Incorporating the rough sets theory into travel demand analysis. *Tourism Management*, 24(5), 511-517.
- Hawley, D. D., Johnson, J. D., & Rsina, D., (2000). Artificial Neural Network Systems: A Network tool for financial Decision Making. *Financial Analysis Journal*, Nov/Dec, 63-72,
- Jiang, H. & Ruan, H. (2010). Applying rough set theory to evaluate network marketing performance of china's agricultural products. *Journal of Computers*, 5(8), 1264-1272.
- Jiang, H. & Yanzhi, D. (2007). Application of association rules based on rough set in human resource management. *The sixth Wuhan International Conference on E-business – Innovation Management Track*, 2407-2413.
- Knodo, M. (2006). On the structure of generalized rough sets. *Information Science*, 176, 589-600.
- Kusiak, A. (2001). Rough set Theory: A Data Mining Tool for Semiconductor Manufacturing. *IEEE Transactions on Electronic Packaging Manufacturing*, 24(1), January. 44-50.
- Lee, M. C. (2007). Developing financial distress evaluation based on fuzzy-rough

- approach. *Proceeding of Second International Conference on Innovative Computing, Information and Control (ICICIC 2007)*, 00, 296-299.
- Lin, L. & Piesse, J., (2004). Identification of corporate distress in UK industrials: A conditional probability analysis approach. *Applied Financial Economics*, 14, 73-82.
- Lin, S. J., Lu, I. J. & Lewis, C. (2007). Grey relation performance correlations among economics, energy use and carbon dioxide emission in Taiwan, *Energy Policy*, 33, 1948-1955.
- Liou, James J. H. (2009). A novel decision rules approach for customer relationship management of the airline market. *Expert System with Applications*, 36(3), 4374-4371.
- Liu, D. & Hu, P., (2007). Hierarchical cluster analysis methodology based on rough sets Theory and its application. *Static and Decision*, 236, 26-28.
- Liu, D. R. & Shin, Y. Y., (2005). Integrating AHP and Data mining for product recommendation based on customer lifetime value. *Information & management*, 42, 387-400.
- Pal, S. K., (2004). Soft data mining computational theory of perceptions and rough fuzzy approach. *Information Science*, 163, 5-12.
- Road, A., Naderi, B. & Soltani, M. (2011). Clustering and ranking university majors using data mining and AHP algorithm: A case study in Iran. *Expert System with Application*, 38(1), 755-765.
- Shyng, J. Y., Wang, F. K., Tzeng, G. H., & Wu, K. (2007). Rough set theory in analyzing the attributes of combination values for insurance market. *Expert System with Applications*, 32(1), 56-64.
- Wang, Y. F. (2003). Mining stock price using fuzzy rough set system. *Expert System with application*, 24(1), 13-23.
- Wong, J. C. & chung, Y. S. (2007). Rough set approach for accident chains exploration. *Accident Analysis and Prevention*, 39(3), 629-637.
- Xu, X. & Lin, J. (2010). Strategic supplier network for supplier selection. *Journal of Computers*, 5(6), 979-986.
- Zhang, L. J. & Li, Z. J. (2006). Gene selection for classifying microarray data using Grey relation Analysis. *Lecture Notes in Computer Science*, 4265, 378-382.
- Zhang, L. j., Li, Z. j., Chen, H. W., & Wen, J. (2006). Minimum redundancy gene selection based on grey relational analysis. *2006 IEEE International Conference on Data mining workshops*, 120-124.

Ming-Chang Lee is Assistant Professor of Department of Information Management at Fooyin University and National Kaohsiung University of Applied Sciences. His qualifications include a Master degree in applied Mathematics from National Tsing Hua University and a PhD degree in Industrial Management from Nation Cheng Kung University. His research interests include knowledge management, parallel computing, and data analysis. His publications include articles in the journal of Computer & Mathematics with Applications, International Journal of Operation Research, Computers & Engineering, American Journal of Applied Science and Computers, Industrial Engineering, International Journal innovation and Learning, Int. J. Services and Standards, Lecture Notes in computer Science (LNCS), International Journal of Computer Science and Network Security, Journal of Convergence Information Technology and International Journal of Advancements in computing Technology.