J Intell Inf Syst (2011) 36:347–366
DOI 10.1007/s10844-009-0110-y

Combined rough sets with flow graph and formal concept analysis for business aviation decision-making

Yu-Ping Ou Yang · How-Ming Shieh ·
Gwo-Hshiung Tzeng · Leon Yen ·
Chien-Chung Chan

Received: 21 March 2009 / Revised: 30 July 2009 / Accepted: 26 October 2009 /
Published online: 13 November 2009
© Springer Science + Business Media, LLC 2009

Abstract Although business aviation has been popular in the USA, Europe, and South America, however, top economies in East Asia, including Japan, Korea, and Taiwan, have been more conservative and lag behind in the development of business aviation. In this paper, we hope to discover possible trends and needs of business aviation for supporting the government to make decision in anticipation of eventual deregulation in the near future. We adopt knowledge-discovery tools based on rough set to analyze the potential for business aviation through an empirical study.

Y.-P. Ou Yang · H.-M. Shieh
Department of Business Administration, National Central University,
300 Chung-da Road, Chung-Li City 320, Taiwan

Y.-P. Ou Yang
e-mail: ouyang.ping@msa.hinet.net

G.-H. Tzeng
Department of Business and Entrepreneurial Management, Kainan University,
No. 1, Kainan Road, Luzhu, Taoyuan 338, Taiwan
e-mail: ghtzeng@mail.knu.edu.tw

G.-H. Tzeng
Institute of Management of Technology, National Chiao Tung University,
1001 Ta-Hsueh Road, Hsinchu 300, Taiwan
e-mail: ghtzeng@cc.nctu.edu.tw

L. Yen
Department of International Business, Kainan University, Taoyuan County, Taiwan
e-mail: leonyen@mail.knu.edu.tw

C.-C. Chan (✉)
Department of Computer Science, University of Akron, Akron, OH 44325, USA
e-mail: chan@uakron.edu

C.-C. Chan
Department of Information Communications, Kainan University,
Luzhu, Taoyuan County, Taiwan

Although our empirical study uses data from Taiwan, we are optimistic that our proposed method can be similarly applied in other countries to help governments there make decisions about a deregulated market in the future.

Keywords Rough set theory · Information system · Decision rule · Business aviation · Data mining · Decision making · Formal concept analysis

1 Introduction

General aviation includes all aircrafts not flown by the airlines or the military, and business aviation is a segment of general aviation that consists of companies and individuals using aircrafts as tools in the conduct of their business. The largest market for business aviation is in North America, with Europe second in size. And this is not without reason. The US government actively encourages and promotes the development of aviation. It invests a great deal of resources on the research of general aviation transportation and the establishment of the Advanced General Aviation Transport Experiments (AGATE) in 1994. As a result, there are over 14,000 business aircrafts of various kinds providing services for various types of business in the USA. Likewise, the British government offers special Customs, Immigration, Quarantine, Security (CIQS) to speed up immigration and customs procedures for business aviation passengers. It has also invested and constructed a Jet Center in 2002 that offers expedited passport control for business aviation passengers.

Currently, Asian countries such as China (Hong Kong), Singapore and Philippines are also thinking about liberalizing the skies for business aviation. But unlike most Western countries, business aviation has remained largely undeveloped in East Asian countries for three reasons: (1) lack of landing slots, parking spaces, and facilities at existing airports to support business aviation; (2) excessive regulations regarding the use of airports, pilot qualifications, and maintenance of aircrafts; (3) lack of understanding for the benefits and commercial value of business aviation. In Taiwan, existing air spaces and relevant airports facilities are already congested, so there is no excess capacity for business aviation. But with the completion of the second expressway and the new high-speed rail, it is expected that some airport capacity to become available for the development of business aviation. In addition, as more and more Taiwanese companies become multinational corporations, there should be substantial growth in demand for business aviation.

The purpose of this study is to answer some pertinent questions in anticipation of deregulation in the business aviation industry in Taiwan. Specifically, we would like to address the third reason listed above, which is the lack of understanding for the economic benefits of business aviation, especially by the policymakers and the regulators. Is there a demand for business air travel in Taiwan? If so, which industries want to use business aviation? In addition, what are the most important criteria for business executives regarding business aviation? To find out the answer to these questions, we use the rough set approach to analyze the results from questionnaires given to Taiwanese companies on the subject of business aviation.

The Rough Set Theory (RST), introduced by Pawlak in 1982 (Pawlak 1982, 1991), is a knowledge-discovery tool that can be used to help induce logical patterns hidden in the large amounts of data. This knowledge can then be presented to the decision-maker as convenient decision rules. In fact, RST is a new knowledge-discovery tool in data mining that has many advantages. Its strength lies in its ability to deal with imperfect data and to classify, and its applications have grown in recent years. Some of these advantages as described in Dimitras et al. (1999) and Greco et al. (1998) are as follows: (1) the rough set approach is based on the original data only and does not need any external information, unlike the assumption of distribution in statistics or of the grade of membership in fuzzy set theory (Dubois and Prade 1992; Krusinska et al. 1992; Skowron and Grzymala-Busse 1993). (2) RST is a tool that is suitable for analyzing not only quantitative attributes but also qualitative ones. (3) It can be used to discover important facts hidden in data and express them in the natural language of decision rules that are easy to understand. (4) Furthermore, real world information is often imprecise, incomplete or uncertain. But RST can handle crisp datasets, fuzzy datasets, as well as uncertain, vague, and imprecise datasets. It is a different mathematical approach to vagueness and uncertainty (Pawlak 1982; Walczak and Massart 1999). (5) In fact, RST should be considered as a formal framework for discovering facts from imperfect data.

RST is an effective tool for the analysis of information system representing knowledge gained by experience. The rough set approach has already been applied with success in the management of a large number of the issues, including data mining (Li and Wang 2004; Hu et al. 2003; Chan 1998) machine diagnosis (Zhai et al. 2002), business failure prediction (Beynon and Peel 2001; Dimitras et al. 1999), activity-based travel modeling (Witlox and Tindemans 2004), evaluation of bankruptcy risk (Slowinski and Zopounidis 1995), etc. And the results have all been very positive (Polkowski and Skowron 1998; Shen et al. 2000; Slowinski 1992; Tsumoto et al. 1996; Weiss 1997). However, the RST has not been widely used in analyzing the airline industry, or for that matter, business aviation. Therefore, this study adopts RST combined Flow Graphs and Formal Concept Analysis to analyze above purposes, and the results demonstrate that the combined approaches are well suited for analyzing the market potential for business aviation and the needs of business aviation customers prior to the industry's deregulation. Likewise, the rough set approach can also be used to analyze the underlying market demand in any industry slated for deregulation.

The remainder of this paper is organized as follows: In Section 2, we present the basic concepts of RST, Flow Graphs, and Formal Concept Analysis (FCA). In Section 3, an empirical study is done using the rough set approach, and the results are presented and discussed. Section 4 concludes.

2 Basics of rough set theory

The Rough Set Theory was first introduced by Pawlak (1982). RST has been used by many researchers, and the theory has had a long list of achievements (Pawlak and Skowron 2007). This section reviews the basic concepts of rough sets (Pawlak 1991, 1997, 2002).

2.1 Information system and approximations

RST is founded on the assumption that every object of the universe of discourse is associated with some information (data, knowledge). An *information system/table* can be represented as $S = (U,\ A,\ V,\ f)$, where U is the *universe* (a finite set of objects, $U = \{x_1,\ x_2,\ ...,\ x_n\}$), A is a finite set of attributes (features, variables), $V = \underset{a\in A}{\cup} V_a$, where V_a is the set of values for each attribute a (called the domain of attribute a), and $f : U \times A \to V$ is an information function such that $f(x, a) \in V_a$, for all $x \in U$ and $a \in A$. A single-choice question with value set $V_{a_1} = \{1, 2\}$ for attribute a_1 has a number of 2 values. A multiple-choice question with value set $V_{a_2} = \{1, 2, 1\,\&\,3, 2\,\&\,3\}$ for attribute a_2 has a number of four values, where multiple-choice 1 and 3 is expressed as 1 & 3. The value class of the question with multiple-choice is redefined in order to simplify the value complexity. For example,$V_{a_2} = \{1, 2, 1\,\&\,3, 2\,\&\,3\} = \{1, 2, 1\,\&\,3 \to 3, 2\,\&\,3 \to 4\} = \{1, 2, 3, 4\}$. Let $B \subseteq A$ and $x, y \in U$. We say x and y are indiscernible by the set of attributes B in S iff $f\,(x, b) = f\,(y, b)$ for every $b \in B$. Thus every $B \subseteq A$ generates a binary relation on U, called *B indiscernibility relation*, denoted by I_B. In other words, I_B is an equivalence relation for any B.

In RST, the accuracy and the quality of approximations are very important in extracting decision rules. Let $B \subseteq A$ and $X \subseteq U$. The lower approximation of X in S by B, denoted by $\underline{B}X$, and the upper approximation of X in S by B, denoted by $\overline{B}X$ are defined as:

$$\underline{B}X = \{x \in U|\, I_B\,[x] \subseteq X\}, \tag{1}$$

$$\overline{B}X = \{x \in U\,| I_B[x] \cap X \neq \emptyset\ |\} \tag{2}$$

where the equivalence class of x in relation I_B is represented as $I_B\ [x]$.

The boundary of X in S by B, is defined as

$$BN_B\,(X) = \overline{B}X - \underline{B}X \tag{3}$$

Set $\underline{B}X$ is the set of all elements of U which can be certainly classified as elements of X, employing the set of attributes B, set $\overline{B}X$ is the set of elements of U which can be possibly classified as elements of X, and the set $BN_B(X)$ is the set of elements which cannot be certainly classified to X using the set of attributes B. Therefore, an accuracy measure of the set X in S by $B \subseteq A$ is defined as:

$$\alpha_B\,(X) = card\ \left(\underline{B}X\right)\Big/card\ \left(\overline{B}X\right) \tag{4}$$

where $card(\cdot)$ is the cardinality of a set, which is the number of objects contained in the lower or upper approximation of the set X. Let $F = \{X_1,\ X_2,\ ...,\ X_n\}$ be a classification of U, i.e., $X_i \cap X_j = \emptyset$, $\forall i, j \le n, i \neq j$ and $\overset{n}{\underset{i=1}{\cup}} X_i = U$, X_i are called classes of F. The lower and upper approximations of F by $B \subseteq A$ are defined as: $\underline{B}F = \left\{\underline{B}X_1, \underline{B}X_2, \cdots, \underline{B}X_n\right\}$ and $\overline{B}F = \left\{\overline{B}X_1, \overline{B}X_2, \cdots, \overline{B}X_n\right\}$, respectively. The quality of approximation of classification F by the set B of attributes, or in short, quality of classification F is defined as:

$$\gamma_B\,(F) = \sum_{i=1}^{n} card\ \left(\underline{B}X_i\right)/card\,(U) \tag{5}$$

It expresses the ratio of all B-correctly classified objects to all objects in the system.

2.2 Reductions and core

An important issue in RST is about attribute reduction, which is performed in such a way that the reduced set of attributes B, $B \subseteq A$ provides the same quality of classification $\gamma_B(F)$ (cf. formula 5) as the original set of attributes A. The minimal subset $C \subseteq B \subseteq A$ such that $\gamma_B(F) = \gamma_C(F)$ is called the *F-reduct* of B and is denoted by $RED_F(B)$. A *reduct* is a minimal subset of attributes that has the same classification ability as the whole set of attributes. In other words, attributes that do not belong to a reduct are superfluous in terms of classification of elements of the universe. The *core* is the common part of all reducts. For example, $CORE_F(B)$ is called the F-core of B, if $CORE_F(B) = \cap RED_F(B)$.

2.3 Decision rules

An information table $A = C \cup D$ can be seen as a *decision table*, where C and D are condition and decision attributes, respectively, and $C \cap D = \varnothing$. The set of decision attributes D induces an indiscernibility relation I_D that is independent of the conditional attributes C; objects with the same decision values are grouped together as decision elementary sets (decision classes). The reducts of the condition attribute set will preserve the relevant relationship between condition attributes and decision classes. And this relationship can be expressed by a decision rule.

A decision rule in S is expressed as $\Phi \to \Psi$, read *if* Φ then Ψ (a logical statement). Φ and Ψ are referred to as conditions and decisions of the rule, respectively. In data mining, we usually take into account relevant confirmation measures and apply them within RST to data analysis. They are presented as follows (Pawlak 2002).

The *strength* of the decision rule $\Phi \to \Psi$ in S is expressed as:

$$\sigma_s(\Phi, \Psi) = \mathrm{supp}_s(\Phi, \Psi) \big/ card(U) \tag{6}$$

where $\mathrm{supp}_s(\Phi, \Psi) = card(\|\Phi \wedge \Psi\|_s)$ is called the *support* of the rule $\Phi \to \Psi$ in S and $card(U)$ is the cardinality of U. With every decision rule $\Phi \to \Psi$ we associate a *coverage factor/covering ratio* (CR) defined as:

$$\mathrm{cov}_s(\Phi, \Psi) = \mathrm{supp}_s(\Phi, \Psi) \big/ card(\|\Psi\|_s). \tag{7}$$

CR is interpreted as the frequency of objects having the property Φ in the set of objects having the property Ψ. The strength of the decision rule can simply be expressed as the ratio—the number of facts that can be classified by the decision rule divided by the number of facts in the data table. Both CR and the *strength* of the decision rule are used to estimate the quality of the decision rules. They play an essential role for a decision-maker in considering which decision rule to use.

2.4 Flow graphs

The study of flow graphs is not new (Berthold and Hand 1999; Ford and Fulkerson 1962). In this work, we use the approach introduced by Pawlak (2002, 2005). The basic idea is that each branch of a flow graph is interpreted as a decision rule and the entire flow graph describes a decision algorithm. In flow graphs, the Bayesian factors, namely, support, strength, certainty, and coverage factors, associated with each decision rule as defined in previous Section 2.3 are formulated in terms of throughflows. More precisely, a flow graph is a directed acyclic finite graph $G = (V, E, w)$, where V is a set of nodes, $E \subseteq V \times V$, is a set of directed branches, and w: $E \rightarrow R^+$ is a flow function and R^+ is the set of non-negative real numbers. The *throughflow* of a branch (x, y) in E is denoted by $w(x, y)$. For each branch (x, y) in E, x is an input of y and y is an output of x. For x in V, let $I(x)$ denote the set of all inputs of x and $O(x)$ be the set of all outputs of x. The inputs and outputs of a graph G are defined by $I(G) = \{x \text{ in } V |\, I(x) \text{ is empty}\}$ and $O(G) = \{x \text{ in } V |\, O(x) \text{ is empty}\}$.

For every node x in G, *inflow*(x) is the sum of throughflows from all its input nodes, and *outflow*(x) is the sum of throughflows from x to all its output nodes. The inflow and outflow of the whole flow graph can be defined similarly. It is assumed that for any node x in a flow graph, $inflow\,(x) = outflow\,(x) = throughflow\,(x)$. This is also true for the entire flow graph G.

Every branch (x, y) of a flow graph G is associated with the certainty (*cer*) and coverage (*cov*) factors defined as:

$$cer\,(x, y) = \sigma\,(x, y)\big/\sigma\,(x) \text{ and}$$
$$cov\,(x, y) = \sigma\,(\mathrm{x, y})\big/\sigma\,(y)$$

where $\sigma\,(x, y) = w\,(x, y)\big/w\,(\mathrm{G})$, $\sigma\,(x) = w\,(x)\big/w\,(G)$, and $\sigma\,(y) = w\,(y)\big/w\,(G)$ are *normalized throughflows*, which are also called strength of a branch or a node, and we have $\sigma(x) \neq 0$, $\sigma(y) \neq 0$, and $0 \leq \sigma\,(x, y) \leq 1$. Properties of the coefficients were studied in Pawlak (2002, 2005). A directed path from x to y, $x \neq y$ in G is a sequence of nodes $x_1, ..., x_n$ such that $x_1 = x, x_n = y$ and (x_i, x_{i+1}) in E for every i, $1 \leq i \leq \mathrm{n} - 1$. A path from x to y is denoted by $[x \ldots y]$.

The certainty, coverage, and the strength of a path $[x_1 \ldots x_n]$ are defined respectively as:

$$cer\,[x_1 \ldots x_\mathrm{n}] = \Pi\; cer\,(x_\mathrm{i}, x_{\mathrm{i}+1})\,, \text{ for } i = 1, \ldots, n-1,$$
$$cov\,[x_1 \ldots x_\mathrm{n}] = \Pi\; cov\,(x_\mathrm{i}, x_{\mathrm{i}+1})\,, \text{ for } i = 1, \ldots, n-1, \text{ and}$$
$$\sigma\left[x \ldots y\right] = \sigma\,(x)\, cer\left[x \ldots y\right] = \sigma\,(y)\, cov\left[x \ldots y\right].$$

A connection from x to y, denoted by $< x, y >$, is the set of all paths from x to y $(x \neq y)$. Definitions of the certainty, coverage, and strength of a connection $< x, y >$ follow from the above definitions.

2.5 Formal concept analysis

The FCA was first introduced by Wille (1982). It is a theoretical framework for the discovery and design of concept hierarchies from relational information systems. FCA is a theory of data analysis that constructs the conceptual structures among data sets, and it is a mathematical method for analyzing binary relations. Moreover, FCA has been used as a knowledge representation mechanism and as a conceptual clustering method (Carpineto and Romano 1996; Diaz-Agudo and Gonzalez-Calero 2001; Mineau and Godin 1995). Up to now, extensions of FCA include fuzzy concept lattice (Belohlavek 1999; Shao et al. 2007), variable threshold concept lattice (Zhang et al. 2007), and rough formal concept (Zhao et al. 2007). In addition, it has been applied to information retrieval, database management systems, software engineering and other aspects (Carpineto and Romano 1996; Godin and Missaoi 1994; Harms and Deogum 2004; Wille 1989).

In FAC, the data for analysis are represented by formal context (U, A, R), which can be denoted by a cross table (context table), where U is a finite set of objects called the *universe*, A is an attributes set and R is an incidence relation ($R \subseteq U \times A$) or is a binary relation between U and A. Therefore, a formal context may be viewed as a binary relation between the object set and the attribute set with the values 0 and 1 (incidence relation). If an object $x \in U$ and an attribute $a \in A$ are in the relation R, denoted by $(x, a) \in R$ or xRa, it is interpreted as "the object x has the attribute a". Given two sets $X \subseteq U$ and $Y \subseteq A$, we define

$$\mathrm{X}' = \{\mathrm{a} \in \mathrm{A} | xRa \quad \text{for all } x \in X\} \text{ and } \mathrm{Y}' = \{x \in U | xRa \quad \text{for all } a \in Y\}$$

to represent that the set of attributes common to the objects in X and the set of objects which have all attributes in Y, respectively.

A formal concept of the context (U, A, R) is a pair (X, Y) with $X \subseteq U$ and $Y \subseteq A$ satisfying $X' = Y$ and $Y' = X$. Sets X and Y are called the *extent* and the *intent* of the concept, respectively. The set of all the formal concepts of a context (U, A, R) is denoted by $B(U, A, R)$. For any $(X_1, Y_1), (X_2, Y_2) \in B(U, A, R)$, define $(X_1, Y_1) \leq (X_2, Y_2) \Leftrightarrow X_1 \subseteq X_2$ (which is equivalent to $Y_2 \subseteq Y_1$). Concept (X_1, Y_1) is called sub-concept and concept (X_2, Y_2) is called super-concept. The set of all concepts of a context (U, A, R) with relation $\leq$ denoted by $(B(U, A, R), \leq)$ is a complete lattice called concept lattice. See Belohlavek (2004) for more information about its structure. Therefore, a basic result is that the formal concepts of a formal context always form the mathematical structure of a lattice with respect to the subconcept–superconcept relation (Wille 2005). Concept lattice is a form of a hierarchy in which each node (formal concept) represents a subset of objects (extent) with their common attributes (intent), and the intent and the extent of a formal concept uniquely determine each other. Moreover, concept lattices can effectively visualized by labeled line diagrams. The concept lattice is the core of the mathematical theory of FCA.

Rough set theory and formal concept analysis are two complementary mathematical tools for data analysis. In recent years, many efforts have been made to compare and combine the two theories (Yao 2004; Zhang et al. 2007). Combination of FCA and rough set theory provides some new approaches for data analysis and knowledge discovery (Saquer and Deogun 2003; Wang and Liu 2008; Zhao et al. 2007). Therefore, formal concept analysis and rough set theory are two important tools in knowledge representation and knowledge discovery in relational information systems.

3 An empirical case for classifying the business aviation prediction

This study adopts two stages for the rough set approach. Each stage focuses on the problem of classifying data sets into classes, and each stage follows the same analytical procedure, which is as follows: (1) Calculate the approximation; (2) Find the reductions and the core attributes; (3) Create the decision rules; (4) Arrange rules into a decision flow graph as the final decision algorithm induced from data.

We use an example in business aviation to illustrate the strength of the rough set approach. The results are used to identify and predict the willingness of companies to use business aviation; they can also be used to propose improved attributes levels through plans to satisfy customers' needs by the government agencies or airline companies. The background of the empirical study, the statement of problem, and the experimental setup are discussed below.

3.1 Background and the statement of problem

In this age of global competition, regularly scheduled commercial flights can no longer satisfy companies' demands for fast, reliable, safe, comfortable, and private business aviation. As a result, business aviation is a rapidly growing sector of general aviation. Currently, the largest market for business aviation is in the USA, then Europe, and South America respectively (NBAA Fact Book 2004). However, the top economies in East Asia, including Japan, Korea, and Taiwan, have been more conservative and lag behind in the development of business aviation. Here we examine the case of Taiwan. Taiwan's business aviation sector has not yet been deregulated. But with Taiwan's multinational businesses having production facilities strewn all over Asia, there is a clear demand for business aviation. Our research aims to find out the underlying demand for business aviation if the industry were to be deregulated in the future. We will use the rough set approach to analyze the demand for business aviation and make suggestions for government agencies to start making plans for a deregulated future today.

3.2 Experimental setup

In order to investigate the possibility of utilizing business aviation in the future, we randomly select 200 companies from the top 500 corporations in Taiwan (according to the annual ranking by a notable business magazine in Taiwan) as our survey subjects and mail out a questionnaire to these selected companies. The content of the questionnaire, the criteria or attribute about business aviation that we are interested in, is based on a research report done by Taiwan's Civil Aeronautics Administration (Research Program of CAA in 2004). Of the 88 questionnaires we got back, 76 of them are qualified replies (the other 12 were deemed unqualified replies because decision attribute D1 was not answered). Among the 76 replies, two (2.6%) companies said that they would definitely use business aviation when business aviation becomes available, 40 (52.6%) companies said that they are considering the possibility of using business aviation, and 34 (44.7%) companies said they would not use business aviation. In order to operate multiple-choice question value and to improve the classification rate, the value class has been redefined by the Research Program of CAA in 2004.

Our analysis includes two stages based on the rough set approach. The first stage is to find out the market potential for business aviation. The companies' basic attributes and their present experiences of flying are stated as conditional attributes (Appendix 1, Table 9, A1∼A7) and the classification is defined as a decision attribute (D1). The decision classes include the set {Definitely yes, Considering the possibility, No}. The second stage is to understand the factors elaborated by companies in considering the use of business aviation. These factors are the six condition attributes.

They include: (B1) the purpose of utilizing business aviation; (B2) the estimated flying trips in a year; (B3) what special services are needed; (B4) which Taiwan airports are preferred; (B5) what is the maximum price the company is willing to pay for a business chartered plane; (B6) whether the business is willing to buy an aircraft or rent instead. Four factors that influence the companies in considering whether to use business aviation are: (D2-1) direct flight between Taiwan and China; (D2-2) whether the cost is too expensive or not; (D2-3) the quality of ground services and maintenance; (D2–4) the scheduled flights cannot meet your needs. These four factors serve as decision attributes (Appendix 1, Table 10). The details of our analysis of this survey data are as follows.

3.3 First stage results

In the first stage, we use question D1 to define a classification on our sample observations, and we calculate the accuracy of approximations of each class. Here class 1 corresponds to the set of companies who answered "Definitely yes" to using business aviation on question D1, class 2 corresponds to the set of companies who answered "Considering the possibility", and class 3 corresponds to the set of companies who answered "No" to using business aviation. And together, the three subsets define our classification where we account for all our sample observations, with no overlaps in each class. The result is presented in Table 1. Both the accuracy and the quality of the whole classification are equal to 1 (by Eqs. 4 and 5), which means that the results of the classification are very satisfactory.

In the next step, the indiscernibility relation is used to deal with the reduction of attributes. And here we used the ROSE2 (Rough Set Data Explorer) tool (Predki et al. 1998; Predki and Wilk 1999) to find all the potential reducts and the core attributes of these reducts. Two reduct sets are generated below:

$$\begin{aligned}\text{Set1} &= \{\text{A1, A2, A3, A4, A5}\};\\ \text{Set2} &= \{\text{A2, A3, A4, A5, A7}\}\end{aligned}$$

The intersection of all the reducts is the core. So in this case, the core is the set of attributes {A2, A3, A4, A5}, which means that these attributes are the most significant attributes for classification purposes. This study assumes that the decision maker selects the best set that fits his or her preferences. Dimitras et al. (1999) proposed the viewpoints of selected criteria: (1) the reduct should contain as small a number of attributes as possible, (2) the reduct should not miss the attributes judged by the decision maker as the most significant for evaluation of the objects. Based on the above statements, we assume that the decision maker wants to know the reduct Set1: A1, A2, A3, A4 and A5.

Table 1 Approximation of decision classes (the first stage)

Class number	Number of objects	Lower approximation	Upper approximation	Accuracy	Quality
D1				1	1
1	2	2	2	1	
2	40	40	40	1	
3	34	34	34	1	

Table 2 Stage 1 rules for class 2, companies considering utilize business aviation with minimum support value ≥2

A1	A2	A3	A4	A5	D1	Support	Certainty	Strength	Coverage
4					2	5	1	0.0658	0.125
	5	3		2	2	3	1	0.0395	0.075
	2	4			2	2	1	0.0263	0.05
	4			1	2	3	1	0.0395	0.075
1	3	1			2	4	1	0.0526	0.1
1	5	1			2	2	1	0.0263	0.05
1	3	3			2	2	1	0.0263	0.05

Based on reduct Set1, we used the BLEM2 tool (Chan and Santhosh 2003) to generate decision rules with Bayesian factors described in "Section 2.3", and these results are shown in Table 11 of Appendix 2. There are 51 rules generated by BLEM2 for classes 1, 2, and 3. There are two rules with total support of two for class 1 (D1 = "1"), which denotes the companies definitely will utilize business aviation, there are 26 rules with total support of 40 for class 2 (D1 = "2"), which denotes companies that are considering the possibility of utilizing it, and there are 23 rules with total support of 34 for class 3 (D1 = "3"), denoting companies have no willingness at all to use business aviation.

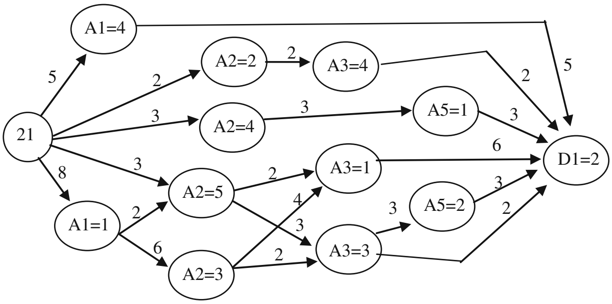
We have applied a minimum support threshold for further reduction of rules. For class 2 rules, with minimum support value ≥2, the number of rules is reduced from 26 to seven as shown in Table 2. In addition, condition attribute A4 is no longer used among these rules as shown in the table that entries in A4 are all blanks.

The rules in Table 2 can be translated into one decision algorithm represented by the decision flow graph shown in Fig. 1. For simplicity, the certainty, strength, and coverage factors associated with each branch are omitted; only throughflows (supports) are shown in the figure. The total inflow of the graph is 21, which is the sum of the supports corresponding to rules appeared in Table 2.

In addition, we used FCA to further analyze the decision rules for class 2 with minimum support value ≥2. Firstly, we selected the rules with support value ≥2 from Table 11 (given in Appendix 2). There were 15 rules selected. Then, we used Java-based open source tool ConExp (http://sourceforge.net/projects/conexp/) to generate the lattice diagram. The result is shown in Fig. 2.

In Fig. 2, the nodes represent formal concepts. In FCA, a formal concept is a pair (O, A), where O is a set of objects and A is a set of attributes. In a lattice diagram, a

Fig. 1 Decision flow graph for rule-set shown in Table 2



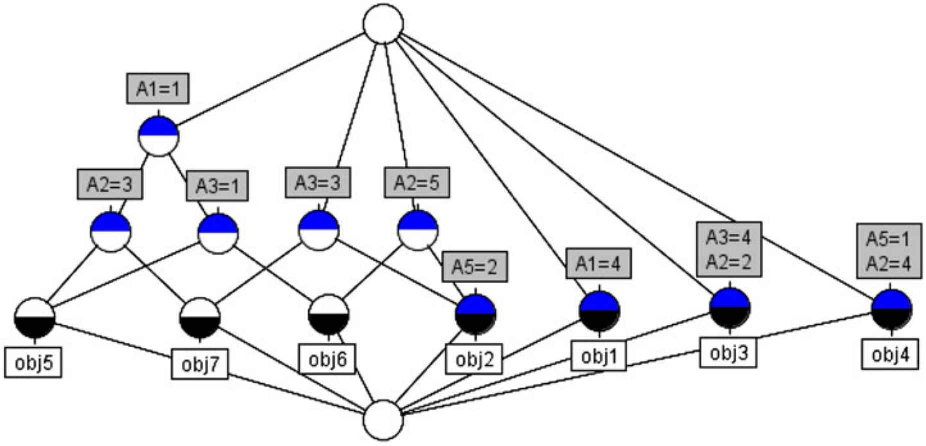


Fig. 2 Lattice diagram for decision rules on class 2 (with minimum support value ≥2)

concept node drawn as blue filled upper semicircle indicates that there is an attribute attached to this concept. Attributes are drawn as rectangles with labels such as A1 = 4, A5 = 1, A2 = 4, etc. If a concept node is drawn as black filled lower semicircle, it means that there is an object attached to this concept. Furthermore, a lattice diagram shows each object's related attributes. For example, the related attributes of object 5 (obj5) are A2 = 3 and A3 = 1. Conversely, we can also see the related objects of each attribute. For instance, the related objects of attribute A3 = 1 are obj5 and obj6, and the related objects of attribute A2 = 3 are obj5 and obj7. In addition, the attribute of node A1 = 1 implies attributes A2 = 3 and A1 = 1. Hence, the node (formal concept) A2 = 3 can be expressed as ({obj5, obj7}, {A2 = 3, A1 = 1}). Similarly, the node A3 = 1 is expressed as ({obj5, obj6}, {A3 = 1, A1 = 1}). The node A1 = 1 can be expressed as ({obj5, obj6, obj7}, {A1 = 1}). Therefore, node A1 = 1 is called super-concept and node A3 = 1 and A2 = 3 are called sub-concepts.

In Figs. 1 and 2, we find the generated results are similar. In Fig. 2, each classification rule is shown as an individual object. In Fig. 1, each path in the flow graph is shown and the degree of support is also listed. In short, the FCA method is similar to the flow graph method in classification. Furthermore, the two methods provide different information, which may be applied according to manager's needs.

3.4 Second stage results

In the second stage, we isolate the 40 companies who said they would consider using business aviation and analyze them separately. Four decision tables are generated, with each decision table made up of the same six conditional attributes (B1∼B6) and one decision attribute D2-X (where X is assigned 1, 2, 3, 4, respectively). Similarly, the approximations of decision classes are calculated the same way as in the first stage, and the results are presented in Table 3. The classification accuracy of D2-1 and D2-4 is 0.6667 and their classification quality is 0.8; both of the classification accuracy and quality of D2-2 are 1; the classification accuracy of D2-3 is 0.8182 and

Table 3 Approximation of decision classes (the second stage)

Class number	Number of objects	Lower approximation	Upper approximation	Accuracy	Quality
D2-1				0.6667	0.8
0	18	15	23	0.6522	
1	22	17	25	0.6800	
D2-2				1	1
0	3	3	3	1	
1	37	37	37	1	
D2-3				0.8182	0.9
0	18	16	20	0.8	
1	22	20	24	0.8333	
D2-4				0.6667	0.8
0	19	16	24	0.6667	
1	21	16	24	0.6667	

Table 4 Reduct sets and core sets of the second stage

Class number	Reduct sets	Core sets
D2-1	{B1, B2, B3, B4, B5}	{B1, B2, B3, B4, B5}
D2-2	{B2, B3, B4}	{B2, B3, B4}
D2-3	{B1, B2, B3, B4, B5}	{B1, B2, B3, B4, B5}
D2-4	{B1, B2, B3, B4, B5}	{B1, B2, B3, B4, B5}

Table 5 The support value more than 2 from stage 2 (D2-1 = 1)

B1	B2	B3	B4	B5	D2-1	Support	Certainty	Strength	Coverage
	3				1	3	1	0.075	0.1364
2	1	H	A		1	2	1	0.05	0.0909
1	1	H		1	1	3	0.75	0.075	0.1364

Table 6 The support value more than 2 from stage 2 (D2-2 = 1)

B2	B3	B4	D2-2	Support	Certainty	Strength	Coverage
	J		1	9	1	0.225	0.2432
	K		1	4	1	0.1	0.1081
	N		1	3	1	0.075	0.0811
N	H	A	1	4	1	0.1	0.1081
1	H		1	11	1	0.275	0.2973
2	H	B	1	2	1	0.05	0.0541

Table 7 The support value more than 2 from stage 2 (D2-3 = 1)

B1	B2	B3	B4	B5	D2-3	Support	Certainty	Strength	Coverage
		N			1	3	1	0.075	0.1364
2	N				1	2	1	0.05	0.0909
2	1	H	A		1	2	1	0.05	0.0909
1	1	H		1	1	4	1	0.1	0.1818

Table 8 The support value more than 2 from stage 2 (D2-4 = 1)

B1	B2	B3	B4	B5	D2-4	Support	Certainty	Strength	Coverage
	2	N			1	2	1	0.05	0.0952
2	N				1	2	1	0.05	0.0952
	2			N	1	2	1	0.05	0.0952
2	1	H	A		1	2	1	0.05	0.0952
1	1	H		1	1	3	0.75	0.075	0.1429

its classification quality is 0.9. These results of the classification are acceptable, since all the accuracies for D2-1 through D2-4 are better than 0.6, and their qualities are close to 1.

Next, we find the reductions and the core attributes. These results are shown in Table 4.

As a result, the reduct and the core for D2-2 are obtained as {B2, B3, B4}; the reduct and the core for the others are {B1, B2, B3, B4, B5}. The next step of data analysis is to generate decision rules by using BLEM2 based on the selected reducts. The resulting rule sets with minimum support value ≥2 are shown in Tables 5, 6, 7 and 8.

4 Discussions

Our empirical study on the market potential for business aviation using the rough set approach has uncovered some important facts. In the first stage, by correlating their decision to use business aviation with their basic attributes, including type of industry and the regions they are doing business with, our results match up closely to the present economic realities in Taiwan: (1) the majority of businesses who replied to our survey said they would consider using business aviation if it became available; (2) the times that the companies need to dispatch employees to go abroad for business every year are between 61 and 80; (3) the regions that these companies are doing business with are mainly China region, North America, and Northeast Asia; (4) the companies' type of industry is mainly traditional industries, manufacturing industries, and information technology industries. The RST analysis clearly shows that there is clearly a market potential for business aviation in Taiwan. Therefore, if the Taiwan government wants to develop its business aviation industry, it should start thinking about how to satisfy companies' flying needs before business aviation is deregulated in the future. Our study has also addressed this issue through a second stage RST analysis of just the companies that said they would consider using business aviation in the future, and we found five common patterns which could be of interest to government policymakers as they plan relative resources distribution, provide special services, select available airports, and set reasonable price, etc. Based on Tables 5, 6, 7 and 8 with higher strength, we find several proposals could be made:

(1) Government agency should consider the regions that the companies are doing business with when they plan air traffic rights in the future.
(2) Government agency should plan a way for these passengers to pass through the immigration & customs quickly. In addition, airports need to construct a busi-

ness aviation center that can deal with related services such as transportation and hotel accommodation.

(3) The most preferred airport for business aviation is Taipei Songshan Airport. Previously, Songshan Airport was an international airport. Nowadays, it is a domestic airport. Although the land this airport sits on isn't that big, it is big enough for smaller aircrafts. And it is situated right in the heart of Taipei, the capital of Taiwan, where most Taiwanese companies are headquartered. Also, Taipei's subway system is being extended to Songshan Airport right now. This makes Taipei Songshan Airport an ideal airport for business aviation.

(4) The maximum price that businesses are willing to pay for business chartered flights should be limited to no more than double the price of first-class tickets in future planning.

The above proposals are made for the consideration of Taiwan government agencies. Moreover, government agencies should consider the resource distribution between existing scheduled commercial flights and business aviation, so that they can both thrive in the market. Currently, business aviation has not yet been deregulated in many countries. The rough set approach should be able to help these countries predict the underlying market demand and trends for business aviation and help their government agencies plan for a future where business aviation is allowed to thrive.

In addition, we adopt FCA to further analyze the decision rules for class 2 with minimum support value ≥ 2. The results indicate the combined rough sets with formal concept analysis method are as suitable as the combined rough sets with flow graph. We could adopt them according to manager's needs.

5 Conclusions

This paper uses the combined rough sets with flow graph and FCA approach to analyze the underlying market demand and needs for business aviation in Taiwan. We discover several interesting patterns from our survey data: (1) More than 55% of the respondents say they would consider using business aviation; (2) Most Taiwanese companies do business with China region, North America, and Northeast Asia; (3) The traditional industries, manufacturing industries, and information technology industries are the main industries that would consider using business aviation; (4) Companies plan to use business aviation for meetings and conducting business and treat their clients; (5) Most companies estimate that the number of business chartered flights they would use is less than 40 in a year; (6) Business travelers prefer to pass through the immigration and customs speedily, in addition, government must take account of transport and accommodation services; (7) The most preferred airport for business aviation is Songshan Airport; (8) The maximum price companies are willing to pay for business chartered flights is less than double the price of first-class tickets. In sum, we have shown that the combined rough sets with flow graph and FCA approach is a promising method for discovering important facts hidden in data and minimal sets of relevant data (data reduction) for the business aviation. Rough set analysis expresses data in the natural language of decision rules, and these rules can be easily deduced and easily understood. Our results provide several useful proposals for government agencies. For example, government agency should consider the regions that the companies are doing business with in the future when they plan air traffic rights, plan a way for these passengers to pass through the immigration and customs quickly, and construct business aviation centers. Also, the most preferred airport is Taipei Songshan Airport and the maximum price that businesses are willing to pay for business chartered flights should be limited to no more than double the price of first-class tickets. In addition, although our empirical study uses data from Taiwan, we are optimistic that rough set theory can be similarly applied in other countries to help governments there make decisions today about a deregulated market in the future based on an RST analysis of the underlying market potential and needs.

Appendix 1

Table 9 Attribute specifications in the information table (the first stage)

	Code	Code description	Response number of questionnaire data
Condition attributes			
A1			
How many times does your company need to dispatch employees to go abroad for business every year?	1	Between 1 and 20, [1,20]	34 (44.74%)
	2	Between 21 and 40, [21,40]	13 (17.11%)
	3	Between 41 and 60, [41,60]	11 (14.47%)
	4	Between 61 and 80, [61,80]	5 (6.58%)
	5	More than 81, [81,∞]	13 (17.11%)
A2			
How many people need to go abroad every time?	1	One person	4 (5.26%)
	2	Two persons	25 (32.89%)
	3	Three persons	28 (36.84%)
	4	Four persons	11 (14.47%)
	5	Five persons	8 (10.53%)
A3			
The region your company is doing business with are mainly	1	China region and North America	26 (34.21%)
	2	China region and Northeast Asia	18 (23.68%)
	3	North America and Northeast Asia	19 (25.00%)
	4	[China regain and Southeast Asia] or [Central Asia and China region] or [Europe and China region]	5 (6.58%)
	5	[North America and Europe] or [North America and Southeast Asia]	6 (7.89%)
	6	[Northeast Asia and Europe] or [Southeast Asia and Northeast Asia]	2 (2.63%)
A4			
How many employees does your company have?	1	Between 1 and 50, [1,50]	8 (10.53%)
	2	Between 51 and 200, [51,200]	6 (7.89%)
	3	More than 200, [200,∞]	61 (81.58%)

Table 9 (continued)

	Code	Code description	Response number of questionnaire data
A5			
What type of industry is your company?	1	Information technology industries	19 (25.00%)
	2	Traditional and manufacturing industries	20 (26.32%)
	3	International trade and logistics	13 (17.11%)
	4	Business and services industries	19 (25.00%)
	5	Real estate and related industries	0 (0.00%)
	6	Consumer goods industry	0 (0.00%)
	7	Medicine and farming industries	0 (0.00%)
	8	Others	5 (6.58%)
A6			
Is your company one of top 100 firms in Taiwan?	1	Yes	20 (26.32%)
	0	No	56 (73.68%)
A7			
Your company's annual revenue is	1	Between NT$ 100 and 10,000 million	37 (48.68%)
	2	Between NT$ 10,100 and 30,000 million	18 (23.68%)
	3	Between NT$ 30,100 and 50,000 million	7 (9.21%)
	4	Between NT$ 50,100 and 70,000 million	4 (5.26%)
	5	More than NT$ 70,100 million	10 (13.16%)
Decision attribute			
D1			
Will your company consider the possibility of utilizing business aviation (a business charter or a private aircraft) after business aviation is deregulated?	1	Definitely yes	2 (2.63%)
	2	Considering the possibility	40 (52.63%)
	3	No	34 (44.74%)

Table 10 Attribute specifications in the information table (the second stage)

Attributes	Code	Code description	Response number of questionnaire data
B1			
What is the purpose of utilizing business aviation?	1	Meeting and business	28 (70%)
	2	Meeting and business and treat our clients	8 (20%)
	3	Meeting and business and recreation and leisure	3 (7.5%)
	4	Meeting and business and treat our clients and recreation and leisure	1 (2.5%)
B2			
If utilizing business aviation, what could be the estimated flying trips in a year?	1	Between 1 and 20, [1, 20]	19 (47.5%)
	2	Between 21 and 40, [21, 40]	7 (17.5%)
	3	Between 41 and 60, [41, 60]	3 (7.5%)

Table 10 (continued)

Attributes	Code	Code description	Response number of questionnaire data
	4	Between 61 and 80, [61, 80]	3 (7.5%)
	5	More than 81, [81, ∞)	1 (2.5%)
	N	No response	7 (17.5%)
B3			
While utilizing business aviation, what kind of special services do you need the most?	H	Pass the immigration and customs fast	23 (57.5%)
	I	Given VIP treatment	1 (2.5%)
	J	Transport and accommodation services	9 (22.5%)
	K	Quick access to the aircraft	4 (10%)
	N	No response	3 (7.5%)
B4			
Which airport do you prefer to fly in Taiwan?	A	Taipei Songshan Airport	21 (52.5%)
	B	CKS International Airport	7 (17.5%)
	C	Taichung domestic airport	4 (10%)
	D	Kaohsiung International Airport	6 (15%)
	E	Hualien domestic airport	0 (0%)
	F	Kinmen domestic airport	0 (0%)
	G	Magong domestic airport	0 (0%)
	N	No response domestic airport	2 (5%)
B5			
What is the maximum price that you can afford for the business charter?	1	Less than 2 times the first-class price	28 (70%)
	2	Less than 5 times the first-class price	6 (15%)
	3	Less than 10 times the first-class price	1 (2.5%)
	4	Less than 15 times the first-class price	0 (0%)
	5	Less than 20 times the first-class price	0 (0%)
	N	No response	5 (12.5%)
B6			
Will your company prefer to purchase an aircraft or rent a business charter in the future?	1	Purchase an aircraft	4 (10%)
	2	Rent a business charter	36 (90%)
Decision attributes			
On question D1, choice 2, the possibility of utilizing business aviation is			
D2-1			
Direct flight between Taiwan and China	0	No	18 (45%)
	1	Yes	22 (55%)
D2-2			
Whether the cost is too expensive or not	0	No	3 (7.5%)
	1	Yes	37 (92.5%)
D2-3			
The quality of ground services and maintenance	0	No	18 (45%)
	1	Yes	22 (55%)
D2-4			
The scheduled flights cannot meet your needs	0	No	19 (47.5%)
	1	Yes	21 (52.5%)

Appendix 2

Table 11 Rules for decision attribute D1, the possibility of utilizing business aviation (for the reduct Set1)

A1	A2	A3	A4	A5	D1	Support	Certainty	Strength	Coverage
	5	4			1	1	1	0.0132	0.5
2				8	1	1	1	0.0132	0.5
4					2	5	1	0.0658	0.125
	2			8	2	1	1	0.0132	0.025
	5	3		2	2	3	1	0.0395	0.075
	2	4			2	2	1	0.0263	0.05
2			2		2	1	1	0.0132	0.025
	4			1	2	3	1	0.0395	0.075
		6		3	2	1	1	0.0132	0.025
			2	2	2	1	1	0.0132	0.025
5		3		2	2	1	1	0.0132	0.025
1	3	1			2	4	1	0.0526	0.1
3				3	2	1	1	0.0132	0.025
		4		2	2	1	1	0.0132	0.025
	2	2	2		2	1	1	0.0132	0.025
3	3			2	2	1	1	0.0132	0.025
2		2	3	2	2	1	1	0.0132	0.025
1	5	1			2	2	1	0.0263	0.05
	1			3	2	1	1	0.0132	0.025
5		2	3		2	1	1	0.0132	0.025
	3	1	3	2	2	1	1	0.0132	0.025
1		5		8	2	1	1	0.0132	0.025
3	2			1	2	1	1	0.0132	0.025
3	2			2	2	1	1	0.0132	0.025
1	3	3			2	2	1	0.0263	0.05
	1			4	3	3	1	0.0395	0.0882
2			1		3	2	1	0.0263	0.0588
5				3	3	3	1	0.0395	0.0882
	3	5			3	2	1	0.0263	0.0588
		6	2		3	1	1	0.0132	0.0294
		5		4	3	1	1	0.0132	0.0294
2	2	3			3	3	1	0.0395	0.0882
3	4			8	3	1	1	0.0132	0.0294
2	4	4			3	1	1	0.0132	0.0294
3			1		3	1	1	0.0132	0.0294
5		2	2		3	1	1	0.0132	0.0294
			2	4	3	1	1	0.0132	0.0294
	2	1	3	4	3	2	1	0.0263	0.0588
	4	3		2	3	2	1	0.0263	0.0588
3	5	1			3	1	1	0.0132	0.0294
5			1		3	1	1	0.0132	0.0294
	2	1		3	3	2	1	0.0263	0.0588
	4	2		2	3	1	1	0.0132	0.0294
5	5				3	1	1	0.0132	0.0294
2		2	3	1	3	1	1	0.0132	0.0294
5	3	1		1	3	1	1	0.0132	0.0294
	3	2		3	3	1	1	0.0132	0.0294
1	2	3			3	1	1	0.0132	0.0294

References

Belohlavek, R. (1999). Fuzzy Galois connections. *Mathematical Logic Quarterly, 45*, 497–504.

Belohlavek, R. (2004). Concept lattices and order in fuzzy logic. *Annals of Pure and Applied Logic, 128*(1–3), 277–298.

Berthold, M., & Hand, D. J. (1999). *Intelligent data analysis—An introduction.* Berlin: Springer.

Beynon, M. J., & Peel, M. J. (2001). Variable precision rough set theory and data discretisation: An application to corporate failure prediction. *Omega, International Journal of Management Science, 29*(6), 561–576.

Carpineto, C., & Romano, G. (1996). A lattice conceptual clustering system and its application to browsing retrieval. *Machine Learning, 24*, 95–122.

Chan, C. C. (1998). A rough set approach to attribute generalization in data mining. *Journal of Information Sciences, 107*(1–4), 169–176.

Chan, C. C., & Santhosh, S. (2003). BLEM2: Learning Bayes' rules from examples using rough sets. In *Proc. NAFIPS 2003, 22nd int. conf. of the North American fuzzy information processing society*, 24–26 July 2003, Chicago, Illinois, pp. 187–190.

Diaz-Agudo, B., & Gonzalez-Calero, P. A. (2001). Formal concept analysis as a support technique for CBR. *Knowledge-Based Systems, 14*(3–4), 163–171.

Dimitras, A. I., Slowinski, R., Susmaga, R., & Zopounidis, C. (1999). Business failure prediction using rough sets. *European Journal of Operational Research, 114*(2), 263–280.

Dubois, D., & Prade, H. (1992). Putting rough sets and fuzzy sets together. In R. Slowinski (Ed.), *Intelligent decision support: Handbook of applications and advances of the rough sets theory* (pp. 203–232). Dordrecht: Kluwer.

Ford, L. R., & Fulkerson, D. R. (1962). *Flows in networks*. Princeton: Princeton University Press.

Godin, R., & Missaoi, R. (1994). An incremental concept formation approach for learning from databases, in: Formal methods in databases and software engineering. *Theoretical Computer Science, 133*(special issue), 387–419.

Greco, S., Matarazzo, B., & Slowinski, R. (1998). A new rough set approach to evaluation of bankruptcy risk. In C. Zopounidis (Ed.), *Operational tools in the management of financial risks* (pp. 121–136). Dordrecht: Kluwer.

Harms, S. K., & Deogum, J. S. (2004). Sequential association rule mining with time lags. *Journal of Intelligent Information Systems, 22*(1), 7–22.

Hu, Y. C., Chen, R. S., & Tzeng, G. H. (2003). Finding fuzzy classification rules using data mining techniques. *Pattern Recognition Letters, 24*(1–3), 509–519.

Krusinska, E., Slowinski, R., & Stefanowski, J. (1992). Discriminant versus rough set approach to vague data analysis. *Applied Stochastic Models and Data Analysis, 8*(1), 43–56.

Li, R., & Wang, Z. O. (2004). Mining classification rules using rough sets and neural networks. *European Journal of Operational Research, 157*(2), 439–448.

Mineau, G., & Godin, R. (1995). Automatic structuring of knowledge bases by conceptual clustering. *IEEE Transactions on Knowledge and Data Engineering, 7*(5), 824–829.

NBAA Fact Book (2004). National Business Aviation Association, INC. http://www.gaservingamerica.org/library_pdfs/BUSINE_1.PDF.

Pawlak, Z. (1982). Rough sets. *International Journal of Computer and Information Sciences, 11*(5), 341–356.

Pawlak, Z. (1991). *Rough sets: Theoretical aspects of reasoning about data*. Dordrecht: Kluwer.

Pawlak, Z. (1997). Rough sets. In T. Y. Lin, & N. Cercone (Eds.), *Rough sets and data mining* (pp. 3–8). Dordrecht: Kluwer.

Pawlak, Z. (2002). Rough sets, decision algorithms and Bayes' theorem. *European Journal of Operational Research, 136*(1), 181–189.

Pawlak, Z. (2005). Flow graphs and intelligent data analysis. *Fundamenta Informaticae, 64*(1/4), 369–377.

Pawlak, Z., & Skowron, A. (2007). Rudiments of rough sets. *Information Sciences, 177*(1), 3–27.

Polkowski, L., & Skowron, A. (Eds.) (1998). *Rough sets in knowledge discovery 1: Approach and applications; rough sets in knowledge discovery 2: Applications, case studies, and software systems*. Wurzburg: Physica.

Predki, B., & Wilk, S. Z. (1999). Rough set based data exploration using ROSE system. In Z. W. Ras, & A. Skowron (Eds.), *Foundations of intelligent systems, lecture notes in artificial intelligence* (Vol. 1609, pp. 172–180). Berlin: Springer.

Predki, B., Slowinski, R., Stefanowski, J., Susmaga, R., & Wilk, S. Z. (1998). ROSE—Software implementation of the rough set theory. In L. Polkowski, & A. Skowron (Eds.), *Rough sets and* *current trends in computing, lecture notes in artificial intelligence* (Vol. 1424, pp. 605–608). Berlin: Springer.

Saquer, J., & Deogun, J. (2003). Approximating monotone concepts. *Design and Application of Hybrid Intelligent System, 105*, 605–613.

Shao, M., Liu, M., & Zhang, W. (2007). Set approximations in fuzzy formal concept analysis. *Fuzzy Sets and Systems, 158*(23), 2627–2640.

Shen, L., Tay, E. H., Qu, L., & Shen, Y. (2000). Fault diagnosis using rough sets theory. *Computers in Industry, 43*(1), 61–72.

Skowron, A., & Grzymala-Busse, J. W. (1993). From the rough set theory to the evidence theory. In M. Fedrizzi, J. Kacprzyk, & R. R. Yager (Eds.), *Advances in the Dempster–Shafer theory of evidence* (pp. 295–305). New York: Wiley.

Slowinski, R. (Ed.) (1992). *Intelligent decision support—Handbook of applications and advances of the rough sets theory*. Dordrecht: Kluwer.

Slowinski, R., & Zopounidis, C. (1995). Application of the rough set approach to evaluation of bankruptcy risk. *International Journal of Intelligent Systems in accounting, Finance and Management, 4*(1), 27–41.

Tsumoto, S., Kobayashi, S., Tanaka, H., & Nakamura, A. (Eds.) (1996). *Proceedings of the fourth international workshop on rough sets, fuzzy sets and machine discovery, RSDF'96*. Tokyo.

Walczak, B., & Massart, D. L. (1999). Rough sets theory. *Chemometrics and Intelligent Laboratory Systems, 47*(1), 1–16.

Wang, L., & Liu, X. (2008). A new model of evaluating concept similarity. *Knowledge-Based Systems, 21*, 842–486. doi:10.1016/j.knosys.2008.03.042.

Weiss, E. (1997). Rough sets, rough neurons, induction and data mining #2. *Journal of Computational Intelligence in Finance, 5*(3), 10–11.

Wille, R. (1982). Restructuring lattice theory: An approach based on hierarchies of concepts. In I. Rival (Ed.), *Ordered sets* (pp. 445–470). Dordrecht: Reidel.

Wille, R. (1989). *Knowledge acquisition by methods of formal concept analysis. Knowledge acquisition by methods of formal concept analysis* (pp. 365–380). New York: Nova Science.

Wille, R. (2005). Formal concept analysis as methodical theory of concepts and concept hierarchies. In B. Granter et al. (Eds.), *Formal concept analysis, LNAI 3626* (pp. 1–3). Heidelberg: Springer.

Witlox, F., & Tindemans, H. (2004). The application of rough sets analysis in activity-based modeling, opportunities and constraints. *Expert Systems with Application, 27*(2), 171–180.

Yao, Y. (2004). Concept lattices in rough set theory. In *Proceedings of 2004 annual meeting of the north american fuzzy information processing society 2* (pp. 796–801).

Zhai, L. Y., Khoo, L. P., & Fok, S. C. (2002). Feature extraction using rough set theory and genetic algorithms an application for the simplification of product quality evaluation. *Computers & Industrial Engineering, 43*(4), 661–676.

Zhang, W., Ma, J., & Fan, S. (2007). Variable threshold concept lattices. *Information Sciences, 177*(22), 4883–4892.

Zhao, Y., Halang, W. A., & Wang, X. (2007). Rough ontology mapping in E-business integration. *Studies in Computational Intelligence, 37*, 75–93.