Expert Systems with Applications 38 (2011) 8–15

Contents lists available at ScienceDirect

Expert Systems with Applications

journal homepage: www.elsevier.com/locate/eswa

Combined rough set theory and flow network graph to predict customer churn in credit card accounts

Chiun-Sin Lin [a], Gwo-Hshiung Tzeng [a,b,*], Yang-Chieh Chin [c]

[a] Department of Business and Entrepreneurial Management, Kainan University, No.1 Kainan Road, Luzhu Shiang, Taoyuan 33857, Taiwan
[b] Institute of Management of Technology, National Chiao Tung University, 1001, Ta-Hsueh Road, Hsinchu 30010, Taiwan
[c] Department of International Business, Asia University, 500, Lioufeng Road, Wufeng, Taichung 41354, Taiwan

ARTICLE INFO

Keywords:
Customer churn
Rough set theory
Flow network graph
Credit card
Marketing strategy

ABSTRACT

Customer churn has become a critical issue, especially in the competitive and mature credit card industry. From an economic and risk management perspective, it is important to understand customer characteristics in order to retain customers and differentiate high-quality credit customers from bad ones. However, studies have not yet adequately introduced rules based on customer characteristics and churn forms of original data. This study uses rough set theory, a rule-based decision-making technique, to extract rules related to customer churn; then uses a flow network graph, a path-dependent approach, to infer decision rules and variables; and finally presents the relationships between rules and different kinds of churn. An empirical case of credit card customer churn is also illustrated. In this study, we collect 21,000 customer samples, equally divided into three classes: survival, voluntary churn and involuntary churn. The data from these samples includes demographic, psychographic and transactional variables for analyzing and segmenting customer characteristics. The results show that this combined model can fully predict customer churn and provide useful information for decision-makers in devising marketing strategy.

1. Introduction

Losing a customer is an opportunity for competitors to gain a customer. With so much competition, companies need to focus on keeping existing customers by satisfying their needs because the cost of attracting a new customer is usually considerably more than the cost to retain a current customer (Heskett, Jones, Loveman, Sasser, & Schlesinger, 1994; Reichheld & Sasser, 1990; Van den Poel & Lariviere, 2004). From the risk management perspective, retaining an existing customer lessens the need for a commercial bank to acquire a less credit-worthy customer or one whose ability or willingness to pay is uncertain. Customer characteristics become increasingly important in a competitive and mature credit card industry within which customers can easily switch their accounts and balances from one bank to another. In the past decade, with the help of business intelligence, databases have been growing rapidly. Commercial banks hold enormous amounts of their customers' transaction data in customer relationship management (CRM) databases, including data related to sales, servicing and marketing functions. This data provides abundant information about customers that decision-makers can use to characterize customers for strategic planning and decision-making purposes and to enhance their competitiveness. However, data is only as good as the system that turns it into usable information. Given the proclivity of some customers to switch credit card companies, there is a pressing need for an integrated system that will identify the customer characteristics that lead to churn before the customers are lost so appropriate action can be taken (Chiang, Wang, Lee, & Lin, 2003). This kind of information can also be used to identify and target customers for implementation of customer management strategies to maximize customer value.

The original rough set theory (RST) proposed by Pawlak (1982, 1984) is an effective approach for discovering hidden deterministic rules and associative patterns in all types of data and for handling unknown data distribution and information uncertainty or ambiguity. In other words, it can integrate learning-from-example technology, extract rules from data sets, and identify data regulations (Komorowski & Zytkow, 1997). Recently, RST has been applied in the marketing field because it is of benefit in analyzing and segmenting customer characteristics to formulate efficient and effective marketing strategies (Tseng & Huang, 2007), such as personal investment portfolio analysis (Shyng, Shieh, Tzeng, & Hsieh, 2009), video game customer purchase behavior (Tseng &

* Corresponding author at: Institute of Management of Technology, National Chiao Tung University, 1001, Ta-Hsueh Road, Hsinchu 30010, Taiwan. Tel.: +886 3 3412456.
E-mail addresses: ghtzeng@cc.nctu.edu.tw, jerry110888@gmail.com (G.-H. Tzeng).

0957-4174/$ - see front matter Crown Copyright © 2010 Published by Elsevier Ltd. All rights reserved.
doi:10.1016/j.eswa.2010.05.039

Huang, 2007), insurance market attributes analysis (Shyng, Wang, Tzeng, & Wu, 2007), brand marketing (Beyon, Curry, & Morgan, 2001), supermarket customer loyalty (Lingras, Hogo, Snorek, & West, 2005) and travel demand analysis (Goh & Law, 2003).

Numerous studies have applied customer churn analysis to several areas, like the credit card industry (Kim, Shin, & Park, 2005; Kumar & Ravi, 2008; Lee, Chung, & Kim, 2001; Lee, Chung, & Shin, 2002), the wireless telecom industry (Hwang, Jung, & Suh, 2004; Wei & Chiu, 2002) and the financial services industry (Van den Poel & Lariviere, 2004). Customer churn analysis benefits for managers in decisions about the right marketing strategy to retain their customers. However, RST has not been widely used in predicting customer churn, especially in credit card industry. Therefore, the major objective of this study is to adopt RST to predict the characteristics of credit card customers so decision-makers can understand the rules of customer churn and use multiple indicators to formulate efficient and effective strategies for marketing of reducing churn.

This article is organized as follows. Section 2 provides an overview of previous relevant studies in the domain of customer churn. In Section 3 we use the RST to inducing the rule of credit card customer churn. In Section 4, we design and develop a flow network graph based on the decision rules extracted by RST, and Section 5 concludes.

2. Review of credit card customer churn

Customer churn occurs when customers switch vendors or cancel service altogether and can be categorized as: unavoidable churn, involuntary churn and voluntary churn (Modisette, 1999). Unavoidable churn occurs when a customer dies or moves out the provider's operating area. Involuntary churn occurs when a user fails to pay for service and the provider terminates services as a result. Termination of service resulting from theft, fraudulent service acquisition or fraudulent usage is also classified as involuntary churn. Involuntary churn is detectable, but not yet predictable and absolutely not actionable from a marketing perspective (Burez & Vand den Poel, 2008). Voluntary churn refers to service termination when the customer leaves one operator for another. Decreasing the churn rate is advantageous and should be a goal of marketers because the cost of retaining current customers is much less than the cost of obtaining new one; Karakostas, Kardaras, and Papathanassiou (2005) argued that a 5% increase in customer retention can result in an 18% reduction in operating costs. Therefore, if a customer starts acting in a way that manifests the intent to leave, management should have an anti-churn strategy prepared (Chu, Tsai, & Ho, 2007). Effective customer churn management also plays an important role in enhancing the quality of customer relationships.

Many studies have discussed customer churn management in various industries, especially in the mobile telecommunications industry. For instance, Estevez, Held, and Perez (2006) found that analyzing the customer antecedents at the time of application could prevent involuntary churn from subscription fraud in telecommunications, and Wei and Chiu (2002) argued that a mobile services provider needs to be able to predict which of its customers may be at risk of changing services in order to make those users the focus of customer retention efforts. However, there are relatively few studies the analyze customer churn among credit card holders. Customers usually carry one or two bank credit cards, along with other kinds of credit cards, so managing customer churn is a priority for most banks in the credit card sector.

Commercial banks provide credit card services that include account acquisition and activation, funding of receivables, card authorization, private label credit card issuance, statement generation, remittance processing, customer service functions, and marketing services. The tremendous competition among banks for providing these services has resulted in a significant increase in the reliability of and quality of service (Kumar & Ravi, 2008). Therefore, customers have become increasingly conscious of service quality. When a customer feels no particular loyalty to his or her bank, the customer will shift from one bank to another when greater service quality or better rates are perceived as being offered elsewhere.

Understanding what customer characteristics are likely to lead to this kind of churn is important to banks. Lee et al. (2001) described a fuzzy cognitive map approach to integrate explicit knowledge and tacit knowledge for churn analysis of credit card holders in Korea. Lee et al. (2002) compared the neural network approach with logistic analysis and C5.0 for churn analysis of credit card holders in Korea. However, although past research has applied customer churn analysis to discuss how to reduce credit card churn, little has been done to determine what customer characteristics will lead to the churn intention. In this study, we solve the customer credit card churn prediction via rough set theory and outline the most important predictor variables in solving the credit card churn prediction problem. The study determines what customer characteristics will lead to the churn intention, which can contribute to preparing effective anti-churn strategies to retain customers.

3. Rough set theory and flow network graph

The rough set theory was first introduced by Pawlak, 1982. RST has been used by many researchers, and the theory has a long list of achievement (Pawlak & Skowron, 2007). This section reviews the basic concepts of rough sets and flow network graph.

3.1. Information systems

RST may be described as an information table or decision table that can be represented by a set of objects dependent on the multi-valued attributes represented (Pawlak, 1982). Given an information system (IS), $IS = \langle U, Q, V, f\rangle$; $U = \{x_1, x_2, \ldots, x_m\}$, where U is the closed universe of IS, $A = \{a_1, a_2, \ldots, a_n\}$, where A is the set of attributes (features/variables), and for each attribute $a \in A$ (an attribute belonging to the considered set of attributes A), $V = \cup_{a \in A} V_a$, is a set of values of the attributes. According to Pawlak (2002), let $\rho : U \times A \to V$ be a description function, $\rho(x,a) \in V_a$ for each $a \in A$, and $x \in U$. Walczak and Massart (1999) expressed that a decision table T is any IS containing conditional attributes and decision attributes and is denoted $T = (U, A \cup D)$, where $D = \{d_1, d_2, \ldots, d_p\}$, $D \notin A$ is a decision variable or output feature, and the elements of A are condition variables or input features. For any application of RST, the first step is to transfer the original data into decision table.

3.2. Indiscernibility of object and classification

Discretization converts continuous attributes into discrete ones while removing redundant and irrelative attributes. Any set of all indiscernible objects is called an elementary set, and such a set forms a basic granule of knowledge about the universe. Some objects (e.g., a_1 and a_2, where $a_1, a_2 \in U$) in U can hardly be distinguished in an available set of attributes (let it be B in A). It is also said that a_1 and a_2 are indiscernible from a set of attributes B, so the relationship between a_1 and a_2 is determined to be an indiscernible relationship, $IND(B)$, defined as $a \in B$, if $\rho_{x_1}(a) = \rho_{x_2}(a)$ for every $a \in A$. $IND(B)$ divides the given universe U into equivalence classes; then any a_i of U can be determined so that the attributes represented in B are in the same class, called the elementary sets of *IS*. The total process is called classification.

3.3. Attribute dependence and approximation accuracy

A rough set-based rule induction technique can be expressed as a pair of crisp sets, called the lower and the upper approximation. Rough set offers a means to describe vague classes through lower and upper approximations. First, let X be a proper subset of U, as indicated by a specific value of decision attributes d, let $X \in U$, let R be an equivalence relation and let x_i express objects $x_1, x_2, \ldots, x_m$, where $i = 1, 2, \ldots, m$. Then $LX_R = \{x \in U : [x]_R \subseteq X\}$ is the lower approximation for the elements that certainly belong to the set; $UX_R = \{x \in U : [x]_R \cap X \neq \phi\}$ is the upper approximation for the elements that possibly belong to the set; and $BX_R = UX_R - LX_R$, is the boundary region of X in U with respect to R where the objects are inconsistent or vague.

If a nonempty X is a rough set with respect to R, and if LX_R and UX_R are lower and upper approximations of X, respectively, we can obtain a more specific measure of accuracy of approximation by using $\alpha_R(X) = \mathrm{card}(LX_R)/\mathrm{card}(UX_R)$, where the cardinality of a set "card" is the number of members of the set.

3.4. Reduction

An important issue in RST is attribute reduction, which removes the superfluous attributes to make the remaining attributes dependent. First, let A be the attribute set of U and B be the reduced attribute set; therefore, B is included in or equal to A. The equation is

$$\mathrm{RED}(B) \subseteq A, \quad \text{where} \quad \mathrm{RED}(B) \quad \text{is the reduced set composed of a set of attributes} \quad B.$$

However, there may have more than one reduced attribute set. The intersection of all reduced attribute sets is the core attribute set, which is the most important attribute set for decision-making. The equation is

$$\mathrm{COR}(C) = \cap\mathrm{RED}(B), \quad \text{where} \quad \mathrm{COR}(C) \quad \text{is the core composed of a set of attributes} \quad C.$$

After applying the reduction, we can determine the decision rule.

3.5. Decision rules

First, assume an attribute space $A = (CA, DA)$, where $CA = \{a_1, a_2, \ldots, a_m\}$, $DA = \{d_1, d_2, \ldots, d_p\}$, $CA \neq \phi$, $DA \neq \phi$, $DA \cap CA \neq \phi$ and $DA \cup CA = A$. Then we assume that there is an indiscernibility relationship, $IND(DA)$, in DA, which is independent of CA. If objects have the same $IND(DA)$, they are grouped together and called decision elementary sets. The attribute sets maintain the main relationship with the decision classes. The expression of the relationship usually uses "if..., then..."; for example, if (x, a_1) and (x, a_2) and ... and (x, a_m) then x belong to d_i, where $i = 1, 2, \ldots, p$. A decision rule in *IS* is expressed as $\Phi \to \Psi$, where Φ and Ψ are conditions and decisions of decision rule, respectively. Then $\sigma_s(\Phi, \Psi) = \mathrm{supp}_s(\Phi, \Psi)/\mathrm{card}(U)$ means the strength of the decision rule $\Phi \to \Psi$ in *IS*, where $\mathrm{supp}_s(\Phi, \Psi)$ is the support rule $\Phi \to \Psi$ in *IS*, and $\mathrm{card}(U)$ is the number of objects in U (Pawlak, 2002). Above implies that a stronger rule will cover more objects, and a decision rule, which can cover more possible rules, has more strength and will decide the more appropriate rule.

3.6. Flow network graph

The flow network graph proposed by Ford and Fulkerson (1962) is a powerful tool that can interpret the path-dependent relationship of each branch of a flow network based on the rough set's decision rules. Based on the flow network graph and Bayes' theorem, the model was used to capture and describe the nature of decision processes within flow network graphs, rather than to describe flow optimization (Pawlak, 2002). Recently, more advanced topics of decision and flow network graphs have been discussed in Pawlak (2002, 2004, 2005) and Ou Yang, Shieh, Tzeng, Yen, and Chan (2008).

A flow network graph is a directed acyclic finite graph $G = (V, \beta, h)$, where V is a set of nodes, $\beta \subseteq V^2$ is a set of directed branches, $h : \beta \to R^+$ is a flow function, and R^+ is the set of non-negative real numbers. A throughflow of a branch $(x, y) \in \beta$ is denoted by $h(x, y)$; then x is an input of y and y is an output of x. The input and output of a graph G are defined as $I(G) = \{x \in V | I(x) \neq \phi\}$ and $O(G) = \{x \in V | O(x) \neq \phi\}$. For every node x in the flow network graph, inflow is defined as $h_+(y) = \sum_{x \in I(y)} h(x, y)$, and outflow is defined as $h_-(y) = \sum_{x \in O(x)} h(x, y)$. Similarly, the inflow and outflow of the whole flow network graph can be defined as $h_+(G) = \sum_{x \in I(G)} h_-(x)$ and $h_-(G) = \sum_{x \in O(G)} h_+(x)$, respectively. We assume that, for any node x in a flow network graph G, $h_+(x) = h_-(x) = h(x)$. In a similar way, a throughflow of the whole flow network graph G is expressed as $h_+(G) = h_-(G) = h(G)$.

To measure the strength of every branch (x, y) in a flow network graph $G = (V, \beta, h)$, we define the strength $\rho(x, y) = h(x, y)/r(G)$, where $0 \leqslant \rho(x, y) \leqslant 1$. The strength of the branch expresses the ratio of total flow through the branch. Every branch (x, y) of a flow network graph G is associated with the certainty and the coverage coefficients. The certainty and the coverage of every branch are defined as $\mathrm{cer}(x, y) = \rho(x, y)/\rho(x)$, and $\mathrm{cov}(x, y) = \rho(x, y)/\rho(y)$, respectively, where $\rho(x, y) = h(x, y)/h(G)$, $\rho(x) = h(x)/h(G)$ and $\rho(y) = h(y)/h(G)$ are normalized throughflow, and $\rho(x) \neq 0$, $\rho(y) \neq 0$, and $0 \leqslant \rho(x, y) \leqslant 1$. The meaning of the certainty coefficient expresses the outflow distribution between outputs of a node, whereas the coverage coefficient exhibits how inflow is distributed between inputs of the node. These coefficients simply explain some properties of flow distribution among the branches in the whole flow network graph, so the flow network graph is a powerful tool for modeling flow information represented by a set of decision rules. Because of the strength of this tool, we decided to adopt the flow network graph to present the decision processes in our study.

4. An empirical credit card case

Based on the global financial tsunami that began in 2008, the number of viable credit card accounts in Taiwan is significantly decreasing, so customer churn is becoming an important issue for commercial banks. Commercial banks that understand the characteristics of customer churn are able to enhance customer loyalty and decrease churn. In this section, we use the ROSE2 (Rough Set Data Explorer) tool to adopt the rough set approach as analytical procedures, include: (1) selecting data and variables; (2) calculating the approximation; (3) finding the reductions and core attributes; (4) creating the decision rules; and (5) importing the decision rules into a flow network graph as the final decision algorithm. The results are used to predict customer churn and to build strategic implementation of churn management in the credit card industry.

4.1. Selection data and variables

In this study, the data were collected from a database of one commercial bank in Taiwan over the period March 2008 – April 2008. Initially, there were 16,866 churn customers consisting of 9,650 voluntary churn customers and 7,216 involuntary churn customers. We collected 7,200 surviving customers in a random sampling. (However, if one class is represented by a large number while the other is represented by only a few, an imbalance problem could distort the predictions (Japkowicz & Stephen, 2002). Therefore, in order to increase the generalization ability and accuracy of this model, we used a down-sizing method in this study.) After removing inaccurate and unreliable data, we had 21,000 customers equally divided into three classes: survival, voluntary churn and involuntary churn. In addition, we limited the card usage and arrears data to at least one year because we believe this period would be sufficient for commercial banks to understand the characteristics of credit card customers.

Originally, there were 43 variables in our database, but that number would engender too much complexity for this study. To increase the accuracy of the decision variables, we invited three scholars from a university marketing department and three managers from the banking industry with an average of more than five years' experiences in the credit card business to identify which of the 43 variables were likely to indicate customer churn with credit cards. Our experts chose 16 condition attributes, including demographic, psychographic and transactional variables, and one decision attribution variable, the type of churn form. We employed the logical reasoning analysis of rough set to determine the rules and the credit card customers' characteristics, as shown in Table 1.

4.2. Rule-based prediction of customer churn characteristics

In RST, discretization can convert continuous attributes into discrete attributes while removing redundant and irrelative attributes. In this study, we attempted a k-means algorithm to cluster the continuous variables a_6 and a_8 to a_{11} into three interval sets. Customers who have the values of these attributes in the same intervals have very similar customer churn characteristics. The intervals proposed for discretization are presented in Table 1.

From rough sets analysis of the coded information table, the first results obtained were the approximation of the decision classes and their quality of classification. The accuracy of approximation is used to describe the degree of credit card churn that could be obtained based on the information of those 16 condition attributes, and the accuracy of classification is used to express the percentage of correctly classified cases. As shown in Table 2, we first classified our samples into two classes, survival and churn (voluntary and involuntary), and the accuracy of classification for the two decision classes was 86% and 93%, respectively. In this manner, the results represent that 16 condition attributes play an important role in determining the churn form of credit card customers and are appropriate for predicting whether a customer will survive or churn. The quality of classification was 95%, so most samples of the data were correctly classified.

To improve the classification rate and to use as small a number of attributes as possible, we attempted to determine the superfluous attributes through an exhaustive algorithm. The result of the value of the positive region for this reduction was 1.0, and the reduced attributes set was {a_1, a_2, a_3, a_4, a_5, a_6, a_7, a_8, a_9, a_{10}, a_{11}, a_{12}, a_{13}, a_{14}, a_{15}, a_{16}}, which represents all relevant attributes in the table. This result shows the importance of these 16 variables in predicting credit card customer churn.

Through rough sets analysis, we generated 1,867 rules, of which 497 rules apply to the survival class, 646 rules apply to the voluntary churn class, and 724 rules apply to the in involuntary churn class. However, some rules are low quality in terms of their ability to distinguish the characteristics of customers among the churn classes. In order to find the higher quality rules for decision-making, we set up a threshold for future reduction of rules with minimum condition attributes $\geqslant 6$ and support value $\geqslant 100$ of each rule for each decision class. Thus, we considered only 25 rules in the survival class, 12 rules in the voluntary churn class and 7 rules in the involuntary churn class. These rules are better at discriminating in terms of correctly classifying the credit card churn forms as well as in terms of understanding the characteristics of customers.

Therefore, focusing on the features of customer in determining the credit card churn form, we derived decision rules for customers' characteristics related to survival, voluntary churn and involuntary churn (Tables 3–5, respectively). In Table 3, the top-5 ranking of frequencies and percentages of the variables in surviv-

Table 1
Attribute specification for credit card churn analysis.

Attribute name	Attribute values	Attribute value sets
Condition attributes		
Age (a_1)	<30; 30–39; 40–49; 50–59; >59	{1,2,3,4,5}
Gender (a_2)	Female; male	{1,2}
Marital status (a_3)	Single; married; divorced	{1,2,3}
Education (a_4)	Graduate school; university/college; high school; under junior high school; others	{1,2,3,4,5}
Annual Income (a_5)	Under NT$360,000; NT$360,000–NT$960,000; NT$960, 000–NT$1,440,000; NT$1,440,000–NT$2,400,000; Up NT$ 2,400,000	{1,2,3,4,5}
Holding credit card periods (monthly) (a_6)	<56.37; 56.37–105.00; >105	{1,2,3}
Automatic debit transfers (a_7)	Yes; no	{1,2}
Annual pay-off times (a_8)	<1.84; 1.84–6.70; >6.70	{1,2,3}
Annual no-transactions times (a_9)	<1.63; 1.63–6.39; >6.39	{1,2,3}
Annual average purchase amount (a_{10})	0; NT$1–NT$15,881.44; >NT$15,881.44	{1,2,3}
Annual purchase times (a_{11})	<2.27; 2.27–58.68; >58.68	{1,2,3}
Consumption up times (last six months) (a_{12})	Yes; no	{1,2}
Installment transaction (a_{13})	Yes; no	{1,2}
Call for temporary adjust credit line (a_{14})	Yes; no	{1,2}
Holding credit loan (a_{15})	Yes; no	{1,2}
Holding cash card (a_{16})	Yes; no	{1,2}
Decision attributes		
Churn form (d_1)	Survival; voluntary churn; involuntary churn	{1, 2, 3}

Table 2
Accuracy of classification and quality of classification.

Churn form	Numbers of objects	Lower approximation	Upper approximation	Accuracy of classification	Quality of classification
d_1	–	–	–	0.91	0.95
Survival	7,000	6,331	7,341	0.86	–
Voluntary/involuntary churn	14,000	13,659	14,669	0.93	–

Table 3
Decision rules of credit card survival class.

No.	a_1	a_2	a_3	a_4	a_5	a_6	a_7	a_8	a_9	a_{10}	a_{11}	a_{12}	a_{13}	a_{14}	a_{15}	a_{16}	d_1	S^*	C^*	S^{**} (%)	C^{**} (%)
1	3		2				1	3	1		3						1	349	1	1.66	4.99
2			2		2		1	3		3		1	1				1	327	1	1.56	4.67
3	3	2	2				1	3		3							1	324	1	1.54	4.63
4		2			2	2	1	3		3						2	1	259	1	1.23	3.70
5		2	1		2		1	3		3					2	2	1	218	1	1.04	3.11
6		2		5			1	3		2	3		2	2			1	212	1	1.01	3.03
7		2	2	5	2		1	3	1				1		2		1	186	1	0.89	2.66
8	2		1	5			1	3		3							1	182	1	0.87	2.60
9				5	2		1	3		2	3		2	2			1	177	1	0.84	2.53
10	4	2	2			2	1	3	1			1				2	1	176	1	0.83	2.51
11	3			3	2		1	3	1		3						1	157	1	0.75	2.24
12	2	2		5	2		1	3	1			1	2				1	155	1	0.74	2.21
13	2	2				2	1	3	1		2	1	2	2			1	138	1	0.66	1.97
14	3	2		3		3	1	3	1			1			2		1	128	1	0.61	1.83
15	3	2	2	5	2		1		1		3				2		1	127	1	0.60	1.81
16	1	2			1		1	3							2	2	1	122	1	0.58	1.74
17		2		4	2	2	1					1				2	1	119	1	0.57	1.70
18		2		3	1		1				3				2		1	118	1	0.56	1.69
19			1		1	3	1	3			3						1	117	1	0.56	1.67
20	3	2		3			1	3	1				1		2		1	117	1	0.56	1.67
21	3			5	2		1	3		2			1			2	1	114	1	0.54	1.63
22	1		1	3		2	1	3	1			1					1	108	1	0.51	1.54
23	2					3	1	3		2	3		1		2		1	104	1	0.50	1.49
24				3		2	1	3		3			1				1	103	1	0.49	1.47
25				3		3	1			3		1	2	2	2	2	1	100	1	0.48	1.42
Total times	14	14	10	15	13	10	25	21	10	11	9	8	11	4	9	7	–	–	–	–	–

Note: S^* means supports number; C^* means certainty; S^{**} means percentage of strength; C^{**} means percentage of coverage.

Table 4
Decision rules of credit card voluntary churn class.

No.	a_1	a_2	a_3	a_4	a_5	a_6	a_7	a_8	a_9	a_{10}	a_{11}	a_{12}	a_{13}	a_{14}	a_{15}	a_{16}	d_1	S^*	C^*	S^{**} (%)	C^{**} (%)
1	2		1	3	1		2	1	3	1					2	2	2	202	0.94	0.96	2.89
2	2		1	2	2		2	1	3	1						2	2	154	0.88	0.73	2.20
3	1		1	2			2	1	3	1			2			2	2	149	0.85	0.71	2.13
4	2		1	2	1		2	1	3	1						2	2	142	0.88	0.68	2.03
5	2		2	3	2	1	2	1	3		1						2	133	1	0.63	1.90
6	2		1	3	2	2	2	1		1					2	2	2	129	0.80	0.61	1.84
7	2		1	3	2	1	2	1		1					2		2	124	0.71	0.59	1.77
8	2		2	2	2	1	2	1		1			2		2	2	2	122	0.62	0.58	1.74
9	1		1	2		1	2	1	3	1			2				2	114	1	0.54	1.63
10	3		2	3		2	2		3		1						2	106	1	0.50	1.51
11	2		1	3	1	2	2		3	1					2	1	2	103	1	0.49	1.47
12	2			3	1	1	2	1	3	1					2	2	2	102	0.95	0.49	1.46
Total times	12	0	11	12	9	8	12	10	9	10	2	0	3	0	6	8	–	–	–	–	–

Note: S^* means supports number; C^* means certainty; S^{**} means percentage of strength; C^{**} means percentage of coverage.

ing customers' decision rules are in automatic debit transfers (a_7) (25 times, 100%, respectively); annual pay-off times (a_8) (21 times, 84%); education (a_4) (15 times, 60%); age (a_1) (14 times, 56%); and gender (a_2) (14 times, 56%). In Table 4, the top-5 ranking of frequencies and percentages of the variables in voluntary churn customers are in age (a_1) (12 times, 100%); gender (a_2) (12 times, 100%); automatic debit transfers (a_7) (12 times, 100%); marital status (a_3) (11 times, 92%); annual pay-off times (a_8) (10 times, 83%); and annual average purchase amount (a_{10}) (10 times, 83%). Finally, in Table 5, the top-5 ranking of frequencies and percentages of the variables in involuntary churn customers are in annual pay-off times (a_8) (7 times, 100%); annual no-transactions times (a_9) (7 times, 100%); marital status (a_3) (5 times, 71%); holding credit card periods (a_6) (5 times, 71%); and consumption up times (a_{12}) (5 times, 71%). Decision maker can use this information to understand which variables may impact each credit card churn form and then to forecast customers' characteristics and develop effective marketing strategies for reducing churn.

4.3. The relationship between characteristics and churn in flow network graph

In order to denote the cause-and-effect relationship between credit card customers' characteristics and churn form, and then

Table 5
Decision rules of credit card involuntary churn class.

No.	a_1	a_2	a_3	a_4	a_5	a_6	a_7	a_8	a_9	a_{10}	a_{11}	a_{12}	a_{13}	a_{14}	a_{15}	a_{16}	d_1	S^*	C^*	S^{**} (%)	C^{**} (%)
1			2		2	2		1	1		1	2				2	3	169	1	0.80	2.41
2				5	2	2	2	1	1			2					3	127	1	0.60	1.81
3	2			2				1	1		2	2	2				3	122	1	0.58	1.74
4	4		2	2				1	1		1						3	111	1	0.53	1.59
5	2		2			2		1	1			2	2			2	3	109	1	0.52	1.56
6	2		1		2	1	2	1	1	1					2	2	3	107	0.93	0.51	1.53
7			1		2	1		1	1		2	2	2		2		3	102	1	0.49	1.46
Total times	4	0	5	3	4	5	2	7	7	1	4	5	3	0	2	3	–	–	–	–	–

Note: S^* means supports number; C^* means certainty; S^{**} means percentage of strength; C^{**} means percentage of coverage.

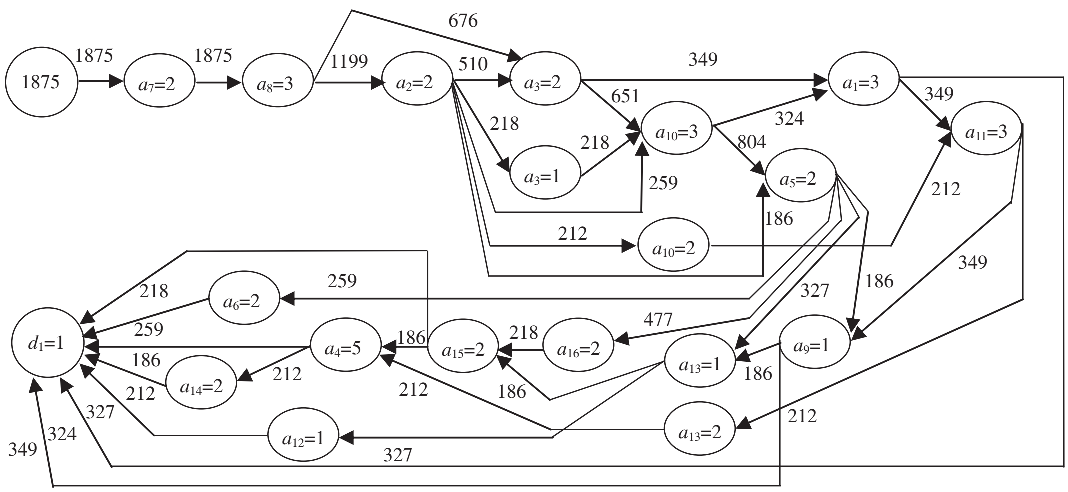

Fig. 1. Decision flow graph for rule-set of survival class. a_7 = 2 means that customer does not attempt automatic debit transfers; a_8 = 3 means that customer annual pay-off times is above 6.7 times; a_2 = 2 means that customer is male; a_3 = 1 means that customer is single; a_3 = 2 means that customer is married; a_{10} = 2 means that annual average purchase amount is from NT$1 to NT$15,881.44; a_{10} = 3 means that annual average purchase amount is above NT$15,881.44; a_5 = 2 means that annual income is from NT$360,000 to NT$960,000; a_1 = 3 means that customer's age is from 40 to 49; a_{11} = 3 means that annual purchase times is above 58.68 times; a_9 = 1 means that annual no-transactions time is below 1.84 times; a_{13} = 1 means that customer has installment transaction; a_{13} = 2 means that customer has no installment transaction; a_{16} = 2 means that customer does not hold cash card; a_{15} = 2 means that customer does not hold credit loan; a_4 = 5 means that customer's education is others; a_{12} = 1 means that consumption up during last six months; a_6 = 2 means that holding credit card periods is from 56.37 to 105.00 monthly; a_{14} = 2 means that customer does not call for temporary adjust credit line; d_1 = 1 means that customer is survival.

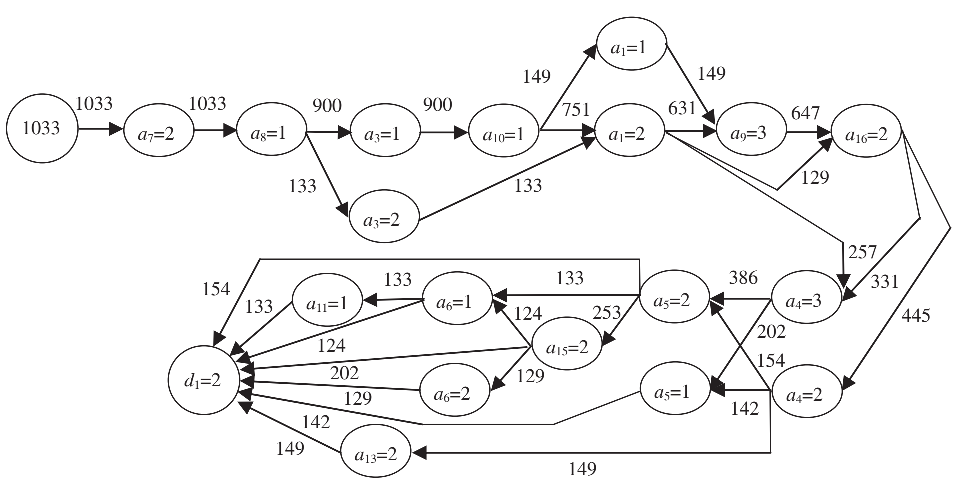

Fig. 2. Decision flow network graph for rule-set of voluntary churn class. a_7 = 2 means that customer does not attempt automatic debit transfers; a_8 = 3 means that customer annual pay-off times is above 6.7 times; a_3 = 1 means that customer is single; a_3 = 2 means that customer is married; a_{10} = 1 means that annual average purchase amount is zero; a_1 = 1 means that customer's age is below 30; a_1 = 2 means that customer's age is from 30 to 39; a_9 = 3 means that annual no-transactions time is above 6.39 times; a_{16} = 2 means that customer does not hold cash card; a_4 = 2 means that customer has university/college degree; a_4 = 3 means that customer has high school degree; a_5 = 1 means that annual income is under NT$360,000; a_5 = 2 means that annual income is from NT$360,000 to NT$960,000; a_{15} = 2 means that customer does not hold credit loan; a_6 = 1 means that holding credit card periods is below 56.37 monthly; a_6 = 2 means that holding credit card periods is from 56.37 to 105.00 monthly; a_{13} = 2 means that customer has no installment transaction; a_{11} = 1 means that annual purchase times is below 2.27 times; d_1 = 2 means that customer is voluntary churn.

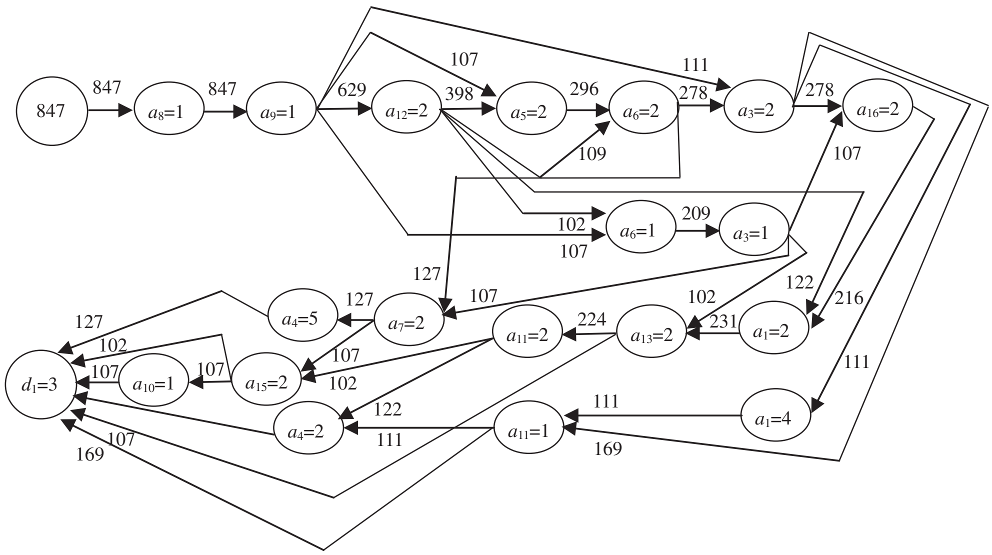

Fig. 3. Decision flow graph for rule-set of involuntary churn class. $a_8 = 1$ means that customer annual pay-off times is below 6.7 times; $a_9 = 1$ means that annual no-transaction times is below 1.63 times; $a_{12} = 2$ means that consumption does not up during last six months; $a_5 = 2$ means that annual income is from NT$360,000 to NT$960,000; $a_6 = 1$ means that holding credit card monthly is below 56.37 monthly; $a_6 = 2$ means that holding credit card monthly is from 56.37 to 105.00 monthly; $a_3 = 1$ means that customer is single; $a_3 = 2$ means that customer is married; $a_{16} = 2$ means that customer does not hold cash card; $a_1 = 2$ means that customer's age is from 30 to 39; $a_1 = 4$ means that customer's age is from 50 to 59; $a_{13} = 2$ means that customer has no installment transaction; $a_{11} = 1$ means that annual purchase times is below 2.27 times; $a_{11} = 2$ means that annual purchase times is from 2.27 to 58.68 times; $a_7 = 2$ means that customer does not attempt automatic debit transfers; $a_4 = 2$ means that customer has university/college degree; $a_4 = 5$ means that customer's education is others; $a_{15} = 2$ means that customer does not hold credit loan; $a_{10} = 1$ means that annual average purchase amount is zero; $d_1 = 3$ means that customer is involuntary churn.

to provide information for further model refinement, we use a flow network graph to represent this relationship (Figs. 1–3). By using the flow network graph, we can view the entire flow network graph as a decision algorithm, where each branch describes a decision rule. However, in practice, it is too complex to show the relationship among the characteristics of credit card customers if all rules from Tables 3 to 5 are considered. To reduce the complexity of the flow network graph, we selected the top 7 rules of each class in order to provide clearer decision-making information. The coefficients of support, certainty, strength and coverage associated with each branch of each churn form are shown in Tables 3–5, and these rules can be translated into one decision algorithm represented by the decision flow network, shown in Figs. 1–3. Strength represents the ratio of total flow through the branch, and coverage exhibits how inflow is distributed between inputs of the node. To simplify the flow network graph, only supports are shown in the figure, and certainty, strength and coverage are omitted.

In the survival class, the total inflow of the graph is 1,875, which is the sum of the supports corresponding to the rules in Table 3. To introduce the importance of condition attributes, we present the flow network graph by the number of supports. Fig. 1 illustrates the relationship between customer characteristics and churn form; the top-5 characteristics of surviving credit card customer, according to the number of supports, are: (1) does not attempt automatic debit transfers (1,875 supports); (2) number of annual pay-off time is more than 6.7 times (1,875 supports); (3) customer is male (1,199 supports); (4) marital status is married (1,186 supports); and (5) annual average purchase amount is more than NT$15,881.44 (1,128 supports).

Based on the same procedure, we can obtain the flow network graph of the voluntary churn class, as shown in Fig. 2. In the voluntary churn class, the total inflow of the graph is 1,033, which is the sum of the supports corresponding to the rules in Table 4. The top-5 characteristics of voluntary churn credit card customers, according to the number of supports, are: (1) does not attempt automatic debit transfers (1,033 supports); (2) number of annual pay-off times is under 1.84 times (1,033 supports); (3) marital status is single (900 supports); (4) annual average purchase amount is zero (900 supports); and (5) customers' age is between 30 and 39 (884 supports).

Finally, in the involuntary churn class, the total inflow of the graph is 847 supports corresponding to the rules in Table 5, as shown in Fig. 3. The top-5 characteristics of involuntary churn credit card customer, based on the number of supports, are: (1) number of annual pay-off times is under 1.84 times (847 supports); (2) number of annual no-transactions times is under 1.63 times (847 supports); (3) no consumption up times during the last 6 months (629 supports); (4) annual income is between NT$360,000 and NT$960,000 (505 supports); and (5) holding credit card periods (monthly) is between 56.37 and 105 (405 supports).

Table 6
The hit rates of new data.

Original number of objects	Decision rules		Hit decision				Hit rate (%)
	Class	Number of new test objects	Decision number	Survival	Voluntary churn	Involuntary churn	
7,000	Survival	1,000	Survival	962	21	17	95.2
7,000	Voluntary churn	1,000	Voluntary churn	13	887	100	88.7
7,000	Involuntary churn	1,000	Involuntary churn	18	59	923	92.3
21,000	–	3,000	–	–	–	–	–

4.4. Discussion and managerial implications

With new data, sets will generate new rules by themselves; the limitation of RST is whether the original rules can fit with the new data. Thus, 3,000 validation sample data sets were added to the retest to check the feasibility of the decision rules in this study. The results in Table 6 show that the hit rates are 95.2% for the survival class, 88.7% for the voluntary churn class, and 92.3% for the involuntary churn class. It is clear, then, that new objects will probably fit into the existing decision classes.

We found several interesting patterns from our data: (1) males usually fall into the survival class; (2) married customers usually fall into the survival class, while single customers usually fall into the voluntary churn class; (3) customers aged 30–39 are more likely to be in the voluntary churn class; (4) the number of annual pay-offs is higher for survival customers than for churn customers; (5) the greater the annual average purchase amount, the greater the chance that the customer will survive; (6) involuntary churn customers usually do not have consumption up time during the last 6 months; and (7) survival and voluntary churn customers do not attempt automatic debit transfers.

The results of this study have implications for decision-makers. First, commercial banks may use demographic variables, such as age, gender and marital status, to determine a customer's likely churn class. For instance, a customer who is male, married and older is likely to be more loyal. Commercial banks should also pay attention to customers' transaction information, such as number of annual pay-off times, annual average purchase amount, and whether there is any consumption up time during the last six months, and combine this data with demographic variables to predict customer characteristics in terms of churn forms. Commercial banks can use RST and a flow network graph of their CRM database to establish an integrated system for predicting customer churn in order to develop appropriate actions to retain good customers and let problem customers go.

5. Conclusion

This study uses a rough set approach and a flow network graph to predict credit card customer churn in a commercial bank in Taiwan. RST is used to discover hidden information in data and to exploring the rules and characteristics of customer churn. The decision rules can be transferred into a flow network graph to represent the connections of pathways and the degrees of their interdependency. Our empirical results, which illustrate customer characteristics as a predictor of churn forms, show that combining RST and a flow network graph is an effective tool for supporting CRM decision-making. The advantage of this hybrid model is that it expresses data in the natural language of decision rules, and these rules can be easily deduced and easily understood for decision-makers to make accurate prediction and act accordingly. This model should be welcomed for its ability to represent the importance of variables on customer churn and to provide abundant information for decision-makers to distinguish which customer characteristics are the key factors that influence customer churn for subsequent strategic planning.

Since limited studies exist on credit card customer churn, numerous possible research avenues remain. First, the present study considers only 1-month period data, so future studies could use longer period data for more accurate results. Second, this study was conducted in Taiwan using Taiwanese objects, so the results may not be applicable to customers in other countries and future studies could expand the application described here to other countries. Finally, future research could compare this combined model with other approaches or modify the model to be even more effective in predicting customer churn. Exploring these mediators is likely to provide a fruitful extension to this work.

References

Beyon, M. J., Curry, B., & Morgan, P. (2001). Knowledge discovery in marketing – An approach through rough set theory. *European Journal of Marketing, 35*(1), 915–933.

Burez, J., & Vand den Poel, D. (2008). Separating financial from commercial customer churn: A modeling step towards resolving the conflict between the sales and credit department. *Expert Systems with Applications, 35*(1), 497–514.

Chiang, D. A., Wang, Y. F., Lee, S. L., & Lin, C. J. (2003). Goal-oriented sequential pattern for network banking churn analysis. *Expert Systems with Applications, 25*(3), 293–302.

Chu, B. H., Tsai, M. S., & Ho, C. S. (2007). Toward a hybrid data mining model for customer retention. *Knowledge-Based Systems, 20*(1), 703–718.

Estevez, P. A., Held, C. M., & Perez, C. A. (2006). Subscription fraud prevention in telecommunications using fuzzy rules and neural networks. *Expert Systems with Applications, 31*(1), 337–344.

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

Goh, C., & Law, R. (2003). Incorporating the rough sets theory into travel demand analysis. *Tourism Management, 24*(5), 511–517.

Heskett, J. L., Jones, T. O., Loveman, G. W., Sasser, W. E., Jr., & Schlesinger, L. A. (1994). Putting the service–profit chain to work. *Harvard Business Review, 72*(2), 164–175.

Hwang, H., Jung, T., & Suh, E. (2004). An LTV model and customer segmentation based on customer value: A case study on the wireless telecommunication industry. *Expert Systems with Applications, 26*(1), 181–188.

Japkowicz, N., & Stephen, S. (2002). The class imbalance problem: A systematic study. *Intelligent Data Analysis, 6*(1), 429–449.

Karakostas, B., Kardaras, D., & Papathanassiou, E. (2005). The state of CRM adoption by the financial services in the UK: An empirical investigation. *Information & Management, 42*(1), 853–863.

Kim, S., Shin, K. S., & Park, K. (2005). An application of support vector machines for customer churn analysis: Credit card case. *Lecture Notes in Computer Science, 3611*(1), 636–647.

Komorowski, J., & Zytkow, J. (1997). *Principles of data mining and knowledge discovery*. New York: Springer-Verlag.

Kumar, D. A., & Ravi, V. (2008). Predicting credit card customer churn in banks using data mining. *International Journal of Data Analysis Techniques and Strategies, 1*(1), 4–28.

Lee, K. C., Chung, N. H., & Kim, J. K. (2001). A fuzzy cognitive map approach to integrating explicit knowledge and tacit knowledge: Emphasis on the churn analysis of credit card holders. *Information Systems Review, 11*(4), 113–133.

Lee, K. C., Chung, N. H., & Shin, K. S. (2002). An artificial intelligence-based data mining approach to extracting strategies for reducing the churning rate in credit card industry. *Journal of Intelligent Information Systems, 8*(2), 15–35.

Lingras, P., Hogo, M., Snorek, M., & West, C. (2005). Temporal analysis of clusters of supermarket customers: Conventional versus interval set approach. *Information Sciences, 172*(1), 215–240.

Modisette, L. (1999). Milking wireless churn for profit, telecommunications. *Dedham, 33*(2), 73–74.

Ou Yang, Y. P., Shieh, H. M., Tzeng, G. H., Yen, L., & Chan, C. C. (2008). Business aviation decision-making using rough sets. In C.-C. Chan et al. (Eds.), *RSCTC 2008, LNAI 5306* (pp. 329–338).

Pawlak, Z. (1982). Rough set. *International Journal of Computer and Information Sciences, 11*(1), 341–356.

Pawlak, Z. (1984). Rough classification. *International Journal of Man–Machine Studies, 20*(5), 469–483.

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

Pawlak, Z. (2004). Decision rules and flow networks. *European Journal of Operation research, 154*(1), 184–190.

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

Pawlak, Z., & Skowron, A. (2007). Rudiments of rough sets. *Information Science, 117*(1), 3–27.

Reichheld, F. F., & Sasser, W. E. Jr., (1990). Zero defections: Quality comes to service. *Harvard Business Review, 68*(5), 105–111.

Shyng, J. J., Shieh, H. M., Tzeng, G. H., & Hsieh, S. H. (2009). Using FSBT technique with rough set theory for personal investment portfolio analysis. *European Journal of Operational Research, 201*(2), 601–607.

Shyng, J. J., Wang, F. K., Tzeng, G. W., & Wu, K. S. (2007). Rough set theory in analyzing the attributes of combination values for the insurance market. *Expert Systems with Applications, 32*(1), 56–64.

Tseng, T. L., & Huang, C. C. (2007). Rough set-based approach to feature selection in customer relationship management. *The International Journal of Management Science, 35*(1), 365–383.

Van den Poel, D., & Lariviere, B. (2004). Customer attrition analysis for financial services using proportional hazard models. *European Journal of Operational Research, 157*(1), 196–217.

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

Wei, C. P., & Chiu, I. T. (2002). Turning telecommunications call details to churn prediction: A data mining approach. *Expert Systems with Applications, 23*(1), 103–112.