Knowledge-Based Systems 23 (2010) 586–597

Contents lists available at ScienceDirect

Knowledge-Based Systems

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

An integration method combining Rough Set Theory with formal concept analysis for personal investment portfolios

Jhieh-Yu Shyng [a,b,*], How-Ming Shieh [a], Gwo-Hshiung Tzeng [c,d]

[a] Department of Business Administration, National Central University, No. 300, Chung-da Rd., Chung-Li City 320, Taiwan
[b] Department of Information Management, Lan-Yang Institute of Technology, No. 79, Fu-Shin Rd, To-Chen, I-Lan 621, Taiwan
[c] Department of Business and Entrepreneurial Management, Institute of Project Management, Kainan University, No. 1, Kainan Rd., Luchu, Taoyuan 338, Taiwan
[d] Institute of Management of Technology, National Chiao Tung University, 1001 Ta-Hsueh Road, Hsinchu 300, Taiwan

ARTICLE INFO

Article history:
Received 2 January 2010
Received in revised form 3 April 2010
Accepted 3 April 2010
Available online 13 April 2010

Keywords:
Rough Set Theory (RST)
Formal concept analysis (FCA)
Data mining
Choosing behaviours
Personal investment portfolios

ABSTRACT

The classical Rough Set Theory (RST) always generates too many rules, making it difficult for decision makers to choose a suitable rule. In this study, we use two processes (pre process and post process) to select suitable rules and to explore the relationship among attributes. In pre process, we propose a pruning process to select suitable rules by setting up a threshold on the support object of decision rules, to thereby solve the problem of too many rules. The post process used the formal concept analysis from these suitable rules to explore the attribute relationship and the most important factors affecting decision making for choosing behaviours of personal investment portfolios. In this study, we explored the main concepts (characteristics) for the *conservative portfolio:* the stable job, less than 4 working years, and the gender is male; the *moderate portfolio:* high school education, the monthly salary between NT$30,001 (US$1000) and NT$80,000 (US$2667), the gender is male; and the *aggressive portfolio:* the monthly salary between NT$30,001 (US$1000) and NT$80,000 (US$2667), less than 4 working years, and a stable job. The study result successfully explored the most important factors affecting the personal investment portfolios and the suitable rules that can help decision makers.

1. Introduction

Real-world data may consist of incomplete and inconsistent information. We can process uncertain and/or incomplete information when the information is discovered knowledge. Data pre processing techniques can improve the quality of the data, accuracy, and the efficiency mining process. Since quality decisions must be based on quality data, data pre processing is an important step in the knowledge discovery process. Data mining [35] generates decision rules that can provide business managers with information about the competition in the market.

Research concerning attitudes towards personal wealth has increased in recent years. A well-designed financial plan can help customers achieve good asset allocation and meet their needs. However, few papers have been published on the topic of personal investment portfolios. The most important paper to submit the idea of the choice of portfolio was Markowitz, in 1952 [13]. The personal investment portfolio has been applied to many fields, such as the behavior of financial services consumers [7], management of personal finances [17], retirement plans [6], and the assessment of the impact of customer satisfaction and relationship quality on customer retention [8]. In the paper of Keng and Hwa [10], they propose the residential property as an important component in a household's overall wealth.

The personal investment portfolio belongs to human knowledge which is a natural language. The natural language (or ordinary language) describes as general-purpose communications including speech, writing, or sign language for human in Wikipedia. Machine learning techniques are used to deal with uncertain data in natural language processing. The statistical natural language processing is mainly technology used for machine learning and data mining which both are fields of artificial intelligence.

The fuzzy set and the rough set theories are particularly adequate for the analysis of various data types, especially dealing with inexact, uncertain or vague knowledge. From a computational perspective, this study proposed the Rough Set Theory (RST), which is a rule-based decision-making technique that was developed by Pawlak [14]. Numerous applications of RST are presented in various scientific domains which have more details in the next section.

RST was used to analyze data contents and data features in this study. The results of RST are presented in the form of classification or decision rules derived from a set of data. It is also presented in

* Corresponding author at: Department of Business Administration, National Central University, No. 300, Chung-da Rd., Chung-Li City 320, Taiwan.

E-mail addresses: shyng@mail.fit.edu.tw, shyng@mgt.ncu.edu.tw (J.-Y. Shyng), hmshieh@mgt.ncu.edu.tw (H.-M. Shieh), ghtzeng@cc.nctu.edu.tw, ghtzeng@mail.knu.edu.tw (G.-H. Tzeng).

0950-7051/$ - see front matter © 2010 Elsevier B.V. All rights reserved.
doi:10.1016/j.knosys.2010.04.003

the form of "if..., then..." decision rules which seems to be more understandable for decision support.

The rule selection indices are the support objects of a rule, the compact of a rule, and the accuracy of a rule. It is a useless decision rule with only one support object due to decrease of the decision precision. Too many unqualified rules will decrease the decision precision. RST generates many rules and some of them have the same strength rate and the same number of support objects. These factors make it very difficult for decision makers to choose suitable rules. This study set up a pruning process which is the support object as a user-defined threshold. In this study we set up a decision rule with only one support object as threshold. This threshold can help to select the suitable rules in order to solve the problem of too many decision rules and to improve the decision precision.

Decision rules are the major information source for decision makers to do the data analysis. However, to explore the knowledge among rules is not an easy way. In this study we used formal concept analysis (FCA) to aggregate the suitable decision rules to provide the prior information for decision makers. The lattice diagram was provided by the FCA in order to gather the decision rules, to construct the concept and to explore the relationship among attributes. The FCA provides the mathematical theory, which belongs to algebra and is a branch of lattice theory.

The FCA is a theory of data analysis that constructs the conceptual structures among data sets. It was introduced by Wille [4] and has since grown rapidly. The FCA is a duality notion that can often be observed between two types of items that relate to each other in an application, such as objects, and attributes, or documents and terms. Conceptual relationships are discussed by the data matrices (contexts) formed by attributes and objects. Another, a mathematical model allows us to study mathematically the representation of conceptual knowledge.

In RST, the data for analysis are described by information system (U,A,R), which corresponds to the formal context in FCA and consists of universe U, attributes set A, and the relation R between U and A. RST and FCA are two complementary mathematical tools for data analysis. Knowledge processing and data analysis always uses concepts to elaborate interpretations of given data and information.

In this study, we use two steps to perform the data analysis. The first step is pre process, which focuses on the problem of many decision rules, and sets up a rule threshold using the support object of a rule to reduce the number of rules. The main purpose is to find the suitable rules. The second step is post process, which creates additional values on those suitable rules by the FCA in order to find the relationship among attributes and to construct the conceptual structures among data sets. One of the greatest benefits is that the decision maker can have a first insight before data analysis. The complete process steps are shown in Fig. 1.

For this study, a questionnaire was designed to investigate personal investment portfolios, using real cases of investors in Taiwan as the basis of the empirical study. The questionnaire considered the factors affecting decision making, such as gender, age, the number of family members, monthly income [7,17], and participants' basic data (such as Marriage Status, Education, Number of Working Years, Professional Status), which may serve as a basis for understanding their needs.

The results of the study identify three types of personal investment portfolios: a conservative portfolio, a moderate portfolio, and an aggressive portfolio. The main general concepts (characteristics) of investors who choose conservative portfolios are having a stable job, low working years, and male; investors who have aggressive portfolios have a higher income and more working years; investors who have moderate portfolios are usually high school-educated and male. More details are presented later.

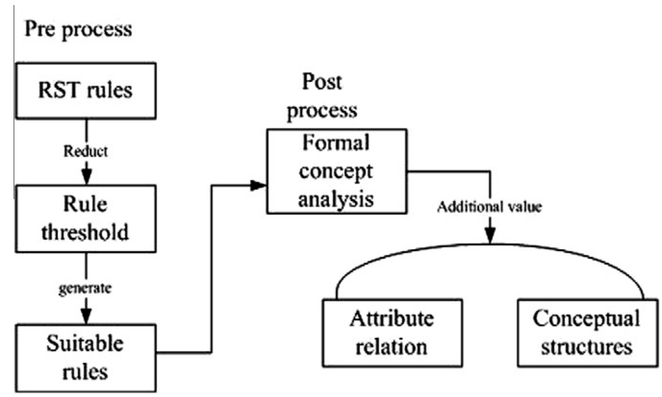


Fig. 1. Study process steps.

In this study, the most important factors affecting the personal investment portfolios were the job type (stable or non-stable), the monthly salary, and education, which carried the greatest affects on the conservative portfolio, moderate portfolio and aggressive portfolio, respectively.

The remainder of this paper is organized as follows. Section 2 describes the concepts to be used in this study. In Section 3, a real case of personal investment portfolio is presented to show the process of this study. In Section 4, we present our conclusions.

2. Concepts about this study

In this section, we briefly introduce RST and FCA, which are used in analyzing the personal investment portfolio. In Section 2.1, the RST is described. In Section 2.2, the FCA is presented.

2.1. Rough set theory and background

RST is a tool for processing uncertain and incomplete information; in this theory, the lower and upper approximations of an arbitrary subset of universe U are the basic operators. In 1982, Pawlak designed RST as a tool to describe the dependencies between attributes, evaluate the significance of the attributes, and deal with inconsistent data. Both fuzzy set theory [28] and RST are used with the indiscernibility relation and perceptible knowledge. The major difference between them is that RST does not need a membership function. A detailed discussion of RST can be found in Walczak and Massart [23]. RST has been applied to the management of a number of issues, including medical diagnosis, engineering reliability, expert systems, empirical study of insurance data [20], machine diagnosis [29], business failure prediction [1], activity-based travel modelling [25], and data mining [2,9,21,34]. Another paper discusses the preference-order of attribute criteria needed to extend the original RST, such as sorting, choice, and ranking problems [5].

RST applies the indiscernibility relation and data pattern comparison based on the concept of an information system with indiscernible data, where the data is uncertain or inconsistent. The data is grouped into classes called elementary sets. More detailed information regarding attributes can be found in the works of Swiniarski and Skowron [22] and Polkowski [18]. The objects in a class may have a relationship with the corresponding features/attributes, and expert knowledge is used to process attribute extraction. Each elementary set is independent of the others. We can extract knowledge from each elementary set used in the real world.

In this section, we will discuss topics, such as the indiscernibility relation, classification, set approximation, reduct and core attribute sets, and decision rules related to RST.

2.1.1. Indiscernibility relation and classification

An information system QM is a 4-tuple $QM = (U,A,V,\rho)$, where U denotes the universal object sets of QM (a finite set of objects, $U = \{x_1,x_2,\ldots,x_n\}$) and A represents the set of attributes (features, variables). Assumed A consists of attributes a_1, a_2, a_3. V_a represents the value of attribute a, then $V = \cup_{a\in A} Va$ is a set of values of the attributes.

Let $\rho: U \times A \rightarrow V$ be a description function such that $\rho(x,a) \in Va$ for each $a \in A$, $x \in U$, where ρx is the description of x in QM [16].

The attribute sets in an information system consist of the set of condition attributes (denoted as CA) and the set of decision attributes (denoted as DA), where the information system is called a decision table. A typical decision table is illustrated in Table 1.

If $Y = \{X_1,X_2,\ldots,X_m\}$ is a family of non-empty sets (classification) that $X_i \subseteq U$, $X_i \neq \emptyset$, $X_i \cap X_j = \emptyset$ for $i \neq j$, i, $j = 1,2,\ldots,m$ and $\cup_{i=1}^{m} X_i = U$.

Let B be a subset of A. The indiscernibility relation include by B, $IND(B)$, is defined to be $IND(B) = \{(x_1,x_2) \mid \rho(x_1,a) = \rho\ (x_2,a) \forall a \in B, x_1,x_2 \in U\}$. The set of objects having the same values on B is called an elementary set, and the process is called classification.

2.1.2. Set approximation

If some subsets of objects cannot be distinguished in terms of the available attribute, then they can only be roughly defined. We can classify rough sets into pair approximation sets; one is the lower approximation of set X and the other is the upper approximation of set X.

Let any subset $X \subseteq U$, B be an equivalence relation and x_i be expressed objects $x_1, x_2,\ldots,x_n$ where i was 1 to n. For any element x_i of U, the equivalence class of x_i in relation $IND(B)$ is represented as $[x_i]IND(B)$. $\underline{R}X = \{x \in U: [x_i]IND(B) \subseteq X\}$ represents the lower approximation. The upper approximation expressed as $\overline{R}X = \{x \in U : [x_i]\ IND(B) \cap X \neq \emptyset\}$. $BndB(X) = \overline{R}X - \underline{R}X$ was the boundary region of X that the objects were inconsistent or ambiguous. If a family $Y = \{X_1,X_2,\ldots,X_m\}$ of non-empty sets (classification), then $\underline{R}\ Y = \{\underline{R}\ X_1, \underline{R}\ X_2,\ldots,\underline{R}\ X_m\}$ and $\overline{R}Y = \{\overline{R}X_1, \overline{R}X_2,\ldots,\overline{R}X_m\}$, were also called the R-lower and R-upper approximation of the family Y, respectively.

2.1.3. Reduct and core attribute sets

In an information system, some attributes may be redundant and useless. If those redundant and useless attributes are removed without affecting the classification power of attributes [15], we can call them the superfluous attributes. Assume $Q \subseteq B$ is a reduct of B. If $IND(Q) = IND(B)$ and $\forall a \in Q$, $IND(Q) \neq IND(Q - \{a\})$. Let $RED(B)$ denote the set of all reducts of B which affects the process of decision making that is a minimal set of attributes. The core of B is $COR(B) = \cap RED(B)$ in which they are the most important attributes in the decision-making process. Applying the reduct set to the model, we can induce the decision rules.

Table 1
Decision table (example).

Object #	c_1	c_2	c_3	d
x_1	1	2	1	1
x_2	3	2	1	2
x_3	1	1	1	2
x_4	1	2	3	1
x_5	2	1	2	2
x_6	2	3	2	1
x_7	3	3	1	1
x_8	1	1	1	2
x_9	2	2	1	1
x_{10}	3	1	2	1

2.1.4. Decision rules

Given an attribute space $A = (CA,DA)$, where CA is condition attribute set and DA is the decision attribute set; assume $CA \neq \emptyset$, $DA \neq \emptyset$, then $DA \cap CA = \emptyset$ and $DA \cup CA = A$, which are the elements of the decision table. This assumes an indiscernibility relation IND (DA). Objects that have the same IND (DA) are grouped together and called decision elementary sets (decision classes).

The syntax of the rule is as follows:

If $f(x,a_1)$ *and* $f(x,a_2)$ *and*... *and* $f(x,a_k)$, *then* x belongs to ds_1 or ds_2 or ds_n, where $\{a_1,a_2,\ldots,a_k\} \subseteq CA$ are condition attributes and $\{ds_1,ds_2,\ldots,ds_n\} \subseteq DA$ are decision classes.

According to Pawlak [16], a decision rule in QM is expressed as $\Phi \rightarrow \Psi$, where Φ and Ψ are conditions and decisions of the decision rule, respectively (read as: if Φ then Ψ). There are three measurements for decision rules, the first one is the accuracy of a rule, which means the rule fitting a specific class in which the class should not cover objects of other classes. The second measurement is the support of a rule, which means a good rule fitting a specific class should be supported by most of objects of the same class. The third measurement is the compact of a rule, which means the less the number of attributes is being used is the better of the rule [11,14].

The strength of the decision rule $\Phi \rightarrow \Psi$ in QM is expressed as:

$$\sigma_{QM}(\Phi,\Psi) = \sup pQM(\Phi,\Psi)/card(U).$$

This implies that a stronger rule will cover more objects and that the strength of each decision rule can be calculated in order to decide the appropriate rules.

The *support* of the rule $\Phi \rightarrow \Psi$ in QM can be expressed as:

$\mathrm{supp}_{QM}(\Phi,\Psi) = card(\parallel \Phi \wedge \Psi \parallel_{QM})$, where $card(U)$ is the cardinality of set U. The number of $\mathrm{supp}_{QM}(\Phi,\Psi)$ means the number of object covered in a specific rule.

In this study, the formal concept analysis was used to gather decision rules and to analyze the relationship between rules. The next section contains the detail of the concepts. The relative decision rules for Table 1 are shown in Table 2. A general form expressed the l'th rule for the decision class d as $R^{d,l}$.

2.2. Formal concept analysis and background

The purpose of FCA is to support the user in analyzing and structuring a domain of interest. It is an important mathematical

Table 2
Decision rules.

	Rule #	Decision class	Meaning	Strength (support) object	Strength rate (%)	Accuracy rate (%)
$R^{1,1}$	1	1	If (c_1 = 1) & (c_2 = 2) then d = 1	x_1, x_4	33.33	100
$R^{1,2}$	2	1	If (c_2 = 3) then d = 1	x_6, x_7	33.33	100
$R^{1,3}$	3	1	If (c_1 = 2) & (c_2 = 2) then d = 1	x_9	16.67	100
$R^{1,4}$	4	1	If (c_1 = 3) & (c_2 = 1) then d = 1	x_{10}	16.67	100
$R^{2,1}$	5	2	If (c_1 = 1) & (c_2 = 1) then d = 2	x_3, x_8	50.00	100
$R^{2,2}$	6	2	If (c_1 = 3) & (c_2 = 2) then d = 2	x_2	25.00	100
$R^{2,3}$	7	2	If (c_1 = 2) & (c_2 = 1) then d = 2	x_5	25.00	100

tool for conceptual data analysis and knowledge processing. FCA has been applied to the management of a number of issues, such as linguistics, software engineering, artificial intelligence, environmental databases [32] and information retrieval. The work of Priss [19] contains an overview of FCA in information science. Because different concepts are semantically close, there is a method of measuring the similarity of FCA concepts presented in Formica's work [3]. Some studies related to RST, such as Liu et al. [12], proposed a reduction of the concept lattices based on RST, a kind of attribute and object reduction method for the concept lattices and concept lattices in RST [27].

A rich experience with lattices of concepts has revealed a great variety of applications, mostly supported by graphical representations. Concept lattices can be used for hierarchical classification of objects, representation of the implicational logic of given attributes, construction of concept sequences, identification of objects, recognition of conceptual patterns, aggregation of data and information, and the representation and acquisition of knowledge.

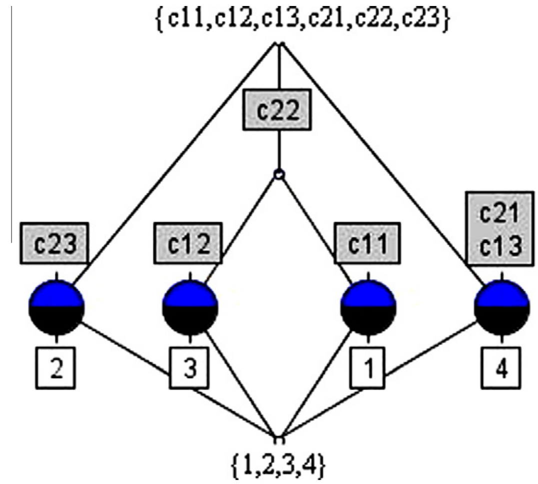


Fig. 2. The lattice diagram of the decision rules for decision class 1 as example.

2.2.1. The concept of FCA

The data for analysis are described by formal context (U,A,R) in FCA which consists of universe U, attributes set A and relation $R \in U \times A$. The formal context can be represented by a cross table called a context table. In RST, the data for analysis are described by information system (U,A,R), which corresponds to the formal context in FCA and consists of universe U, attributes set A, and the relation R between U and A.

In FCA, the formal concept and the concept lattice are two central issues. A formal concept consists of the set of objects and the set of attributes. The set of objects of a formal concept is called its "extension," and the set of attributes is called its "intension." For a given formal context, the extensions and intensions are uniquely defined and fixed for the formal concepts. FCA is based on a set-theoretic model for (formal) contexts, from which concepts and conceptual hierarchies can be formally derived. A basic result is that the formal concepts of a formal context are always form the mathematical structure of a lattice with respect to the subconcept–superconcept relation [24]. The relations can be expressed by a lattice diagram. From the diagram, we can derive concepts, implication sets, and association rules based on the context table.

Statistics and concept analysis complement each other in the field of information science, such as the mathematical lattices that are used in FCA and can be interpreted as classification systems. Formalized classification systems can be analyzed according to the consistency of their relations.

The subconcept–superconcept relation defines an orderly concept relation of all formal concepts in a formal context. All edges in the line diagram of a concept lattice represent subconcept–superconcept relations. The extension of the subconcept is contained in the extension of the superconcept, which is equivalent to the relationship that the intension of the subconcept contains the intension of the superconcept.

Table 3 is the context table converted by the decision rules for decision class 1, the elements on the left side are objects, the elements at the top are attributes, and the relation between them is represented by the cross. It converts the attributes, which are represented in the decision rules, into a binary value form, and the cross mark (×) is expressed as 1 and the blank is expressed as 0. Hereafter, we denoted the rule number as object number. The original decision table and decision rules are shown in the Tables 1 and 2, respectively. Based on the formal context, we can construct formal concepts that form a concept lattice. Fig. 2 is the lattice diagram for the decision rules of decision class 1.

The nodes in Fig. 2 represent formal concepts. If the attribute concepts having c_{22}, c_{12} and c_{11} form two concepts, one concept consists of object 1 and attributes c_{22}, c_{11}, expressed as $(\{1\},\{c_{22},c_{11}\})$; while the other concept consists of object 3 and attributes c_{22}, c_{12}, expressed as $(\{3\},\{c_{22},c_{12}\})$. The superconcept is c_{22} and the subconcepts are c_{12} and c_{11}. The attribute concept c_{22} as intension can follow the lines down to find other subconcepts. In addition, the concepts having objects 1 and 3 as extensions can follow lines up to find the superconcepts which are shared by them. In Fig. 2, lines up can find more general concepts, and lines down can find more specific concepts [26].

The subconcept–superconcept relation is transitive, meaning that a concept is the subconcept of any concept that can be reached by travelling upwards from it. As such, attribute c_{22} is inherited by all its subconcepts, such that the attributes c_{11} and c_{12} imply c_{22}. Fig. 2 summarizes the data and explores the concepts from the context table based on the Table 3. In this study we can find the objects 1 and 3 shared by the same attribute c_{22}. The relations for the set of objects {1,3} and the set of attributes $\{c_{22},c_{12},c_{11}\}$ are "closed," because one cannot enlarge the attribute or the object sets. A pair of a set of objects and a set of attributes that is closed in this manner is called a "formal concept." Thus, the set of objects {1,3} and the set of attributes $\{c_{22},c_{12},c_{11}\}$ form a formal concept with the extension {1,3} and the intension $\{c_{22},c_{12},c_{11}\}$. There are more details of a personal investment portfolio empirical case about FCA presented in the Section 3.

Table 3
Context table by the decision rules for decision class 1.

	Objects (rule #)	Attributes					
		c_1			c_2		
		c_{11}	c_{12}	c_{13}	c_{21}	c_{22}	c_{23}
$R^{1,1}$	1	×				×	
$R^{1,2}$	2						×
$R^{1,3}$	3		×			×	
$R^{1,4}$	4			×	×		

3. An empirical case of personal investment portfolios

The questionnaires were distributed to investors in the North and Northeast districts of Taiwan. Data was collected based on nominal and ordinal scales. There were 200 valid questionnaires from a total of 221 received. The percentage of valid questionnaires is 90%. Among the valid respondents, there were 108 females and 92 males.

3.1. Problem description

For the given information system *QM*, expert knowledge is used to process attributes for extraction. There are nine attributes: eight condition attributes, namely Age (c_1), Gender (c_2), Marriage (c_3), Number of Children (c_4), Education (c_5), Number of Working Years (c_6), Professional Status (c_7), Monthly Salary (c_8); and one decision attribute, namely the portfolio type representing the various investment portfolios. The relative details of survey data portfolio types are cited from the FSBT paper [33]. After a reduct process was applied to the condition attributes, we labelled the reduct attribute set as c_1, c_2, c_3, c_5, c_6, c_7, c_8. The attribute c_4 was superfluous and was therefore eliminated [33]. The core attribute set was the same as the reduct attribute set. The original attribute specification is detailed in Table A.1 of Appendix A.

3.2. Process of this study

In this study, the empirical process is divided into two parts. The first part was pre process, which used the support object of a rule as user-defined threshold to prune the rules generated by the ROSE2 [30]. The second part was post process, which used the FCA to aggregate the suitable rules selected from pre process. The post process output the attribute relationship, which helped the decision maker to perform *a priori* predictions.

3.2.1. Result of the pre process

There are three decision classes – decision class 1, decision class 2, and decision class 3 – which expressed different portfolio types: the conservative portfolio, moderate portfolio and aggressive portfolio, respectively. Here, decision class 1 was selected as an example in this section.

In this study, 67 rules were generated by ROSE2. The decision rules are shown in Table B.1 in Appendix B. The decision rules numbered from 52 to 67 are approximate rules, which mean that the rule did not belong to a specific decision class and may overlap more than one decision class. However, too many decision rules impede decision making. In this study, the support object of a rule was set as the user-defined threshold in a pruning process to reduce the rule amount, in order to improve the data analysis. The user-defined thresholds mean to set a rule removing out those unqualified rules. Table 4 shows the details of the number of decision rules for each decision class after pruning process. Based on the survey results, we identified the following types of personal investment portfolio holder:*Conservative*: (1) a stable job, (2) male, (3) married, (4) between 10 and 14 working years experience; *Aggressive*: (1) female, (2) married, (3) a college graduate; *Moderate*: (1) male, (2) married, (3) a service job, and (4) monthly salary between NT$30,000 (US$900) and NT$80,000 (US$2424). More relative details about the three portfolio types of survey data are cited from the FSBT paper [33].

Table 4
The information of decision rules after threshold process.

Decision class #	# of rule			Rule's threshold in support object #	# of unsatisfied threshold rules	# of rules in experiment
	rules	Approximate rules	total			
1	13	12	25	1	13	12
2	14	12	26	1	14	12
3	24	10	34	1	18	16

3.2.2. Results of the post process

Table 5 is the context table, which converted 12 rules of decision class 1 representing the attributes in the rules into a binary form, with the cross mark (×) expressed as 1 and the blank expressed as 0.

The Java-based opensource ConExp [31] program was used in this study. It generated the lattice diagram, and is shown in Fig. 3. The formal concept tool also outputs the association rules and implication sets to aid decision making.

3.3. Discussion and implications

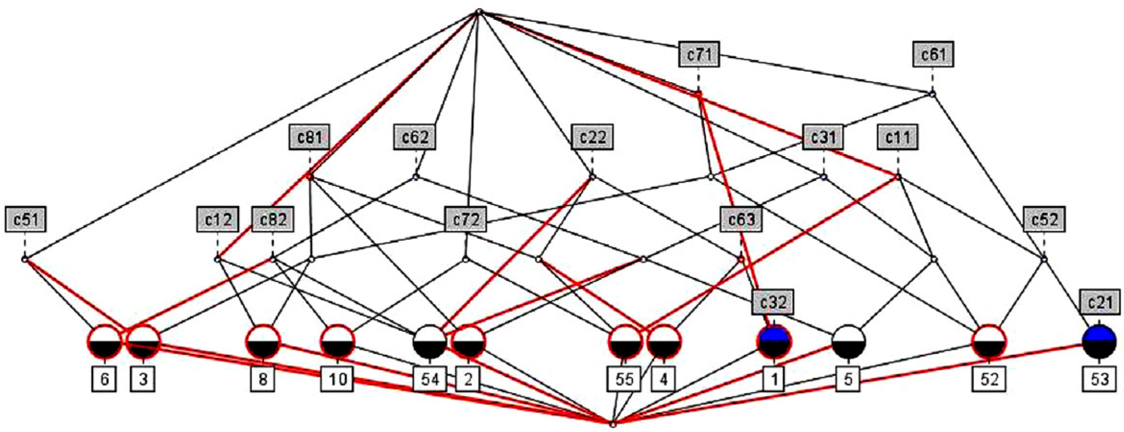
In Fig. 3, the concepts are more general when the lines draw up, and the concepts are more specific when the lines draw down. The details of the decision class 1 (conservative portfolio) results will be presented in Section 3.3.1, followed by the details of the moderate portfolio and aggressive portfolio in Sections 3.3.2 and 3.3.3, respectively.

3.3.1. Conservative portfolio

The FCA can do the data classification. However, data was already classified in this study. The purpose of this study is to use the FCA to aggregate rules and to diagnose the relationship among attributes belonging to the rules in the specific class. From the lattice diagram (Fig. 3), association rules and implication sets (Table C.1) generated by the ConExp program can retrieve general information, such as: (1) the monthly salary between NT$30,000 (US$1000) and NT$80,000 (US$2667) (c_{82}) may have 5–9 working years (c_{62}), i.e. ($c_{82} \Rightarrow c_{62}$); (2) a person with high school education (c_{52}) may be under 29 years old (c_{11}) and have less than 4 working years (c_{61}), i.e. ($c_{52} \Rightarrow c_{11}, c_{61}$); (3) some persons with 10–14 working years (c_{63}) may be male (c_{22}), i.e. ($c_{63} \Rightarrow c_{22}$); and (4) a person

Table 5
Context table for decision rules of decision class 1 as example.

Rule #	Age (c_1)		Gender (c_2)		Marriage (c_3)		Education (c_5)		Working years (c_6)			Prof. (c_7)		Monthly Salary (c_8)	
	c_{11}	c_{12}	c_{21}	c_{22}	c_{31}	c_{32}	c_{51}	c_{52}	c_{61}	c_{62}	c_{63}	c_{71}	c_{72}	c_{81}	c_{82}
1				×		×					×	×			
2					×					×				×	
3							×		×			×		×	
4				×							×			×	
5	×				×					×					
6							×			×					×
8		×							×			×		×	
10										×			×		×
52	×				×			×	×			×			
53	×		×					×	×						
54		×		×	×					×					×
55	×			×									×	×	
Frequency	4	2	1	4	4	1	2	2	4	5	2	4	2	5	3



c_{11}: age under 29; c_{12}: age between 30 and 39; c_{21}: female; c_{22}:male; c_{31}: single; $c_{32;}$:marry; c_{51}: under junior high school; c_{52}: high school; c_{61}: under 4 working years; c_{62}: 5 to 9 working years; c_{63}: 10 to 14 working years; c_{71}: stable job; c_{72}:service job; c_{81}: monthly salary under NT$30,000; c_{82}: monthly salary between NT$30,001 and NT$80,000;

Fig. 3. Lattice diagram for decision rules of decision class 1.

Table 6
Implication relation between attributes for class 1 as example.

Imply (superconcept)	c_{63}	c_{31}	c_{61}	c_{52}	c_{11}	c_{32}	c_{71}	c_{22}	c_{72}	c_{81}	c_{62}	c_{12}	c_{82}
(subconcept)	c_{32} c_{22}, c_{71}	c_{11}, c_{62} c_{11}, c_{71} c_{62}, c_{81} c_{12}, c_{22} c_{12}, c_{62} c_{22}, c_{62}	c_{71}, c_{81}*2 c_{52}*2 c_{21} c_{11}, c_{71} c_{12}, c_{71} c_{31}, c_{71} c_{51}, c_{71} c_{12}, c_{81} c_{51}, c_{81}	c_{11}, c_{61}*2 c_{21} c_{31}, c_{61} c_{11}, c_{71} c_{71}, c_{31} c_{13} c_{21} c_{11}, c_{71}	c_{52}*2 c_{21} c_{31}, c_{61} c_{31}, c_{71} c_{22}, c_{72} c_{72}, c_{81}	c_{22}, c_{71}	c_{61}, c_{81}*2 c_{32} c_{12}, c_{61} c_{31}, c_{61} c_{51}, c_{61} c_{51}, c_{61} c_{12}, c_{81} c_{51}, c_{81}	c_{63}*2 c_{32} c_{11}, c_{72} c_{11}, c_{81} c_{72}, c_{81} c_{62}, c_{31},c_{82} c_{12}, c_{31} c_{12}, c_{62}	c_{11}, c_{22} c_{11}, c_{81}	c_{11}, c_{22} c_{12}, c_{61} c_{51}, c_{61} c_{12}, c_{71} c_{51}, c_{71} c_{11}, c_{72} c_{22}, c_{72}	c_{82}*3 c_{31}, c_{81} c_{12}, c_{22} c_{12},c_{31} c_{22},c_{31}	c_{62}, c_{31}, c_{82} c_{22}, c_{31} c_{22}, c_{62}	c_{12}, c_{22} c_{12}, c_{31} c_{22}, c_{31} c_{12}, c_{62} c_{22}, c_{62} c_{51}, c_{62} c_{62}, c_{72}
Frequency	3	12	18	14	11	2	16	16	4	14	12	7	14

having less than 4 working years (c_{61}) and monthly salary under NT$30,000 (US$1000) (c_{81}) may have a stable job (c_{71}), i.e. (c_{61}, $c_{81} \Rightarrow c_{71}$).

From the higher frequency of the sub-attribute in Table 5, we can find the main characteristics of each attribute for conservative portfolio, such as a stable job, single marital status, monthly salary under NT$30,000 (US$1000), unspecified education level, between 5 and 9 working years, an age under 29 and gender is male.

Table 7
The most important factor for portfolio.

Portfolio	Most important factor	Following by
Conservative	the stable job (c_{71}), working years under 4 years (c_{61}), male (c_{22})	The salary under NT$30,000 (US$1000) (c_{81}), age under 29 years old (c_{11}), and a single person (c_{31})
Moderate	High school education (c_{52}), salary between NT$30,001 (US$1000) and NT$80,000 (US$2667) (c_{82}), male (c_{22})	Under 4 working years (c_{61}), age between 30 and 39 years old (c_{12}) and a stable job (c_{71})
Aggressive	The salary between NT$30,001 (US$1000) and NT$80,000 (US$2667) (c_{82}), under 4 working years (c_{61}), and a stable job (c_{71})	Age between 30 and 39 years old (c_{12}), and a single person (c_{31})

Table A.1
Attribute specification for the personal investment portfolios analysis.

Attribute name	Attribute values	Attribute value sets
Condition attributes		
Age (c_1)	<29; 30–39; 40–49; 50–59; 60 ~	{1,2,3,4,5}
Gender (c_2)	Female; male	{1,2}
Marriage (c_3)	Single; marry; divorce; others	{1,2,3,4}
Having children (c_4)	Yes; no	{1,2}
Education (c_5)	Under junior high school; high school; college; university; graduate school	{1,2,3,4,5}
Working years (c_6)	Under 4 years; 5–9 years; 10–14 years; 15–19 years; 20–24 years; 25–29 years; 30 years~	{1,2,3,4,5,6,7}
Professional (c_7)	1–6 (categorized according the insurance industry's listed profession)	{1,2,3,4,5,6}
Monthly salary (c_8)	Under NT$30,000; NT$30,001–NT$80,000; NT$80,001–NT$120,000; NT$120,001–NT$200,000; NT$200,001~	{1,2,3,4,5}
Decision attributes		
Portfolio type (d)	Conservative; moderate; aggressive	{1,2,3}

Table B.1
Original rules generated from ROSE2.

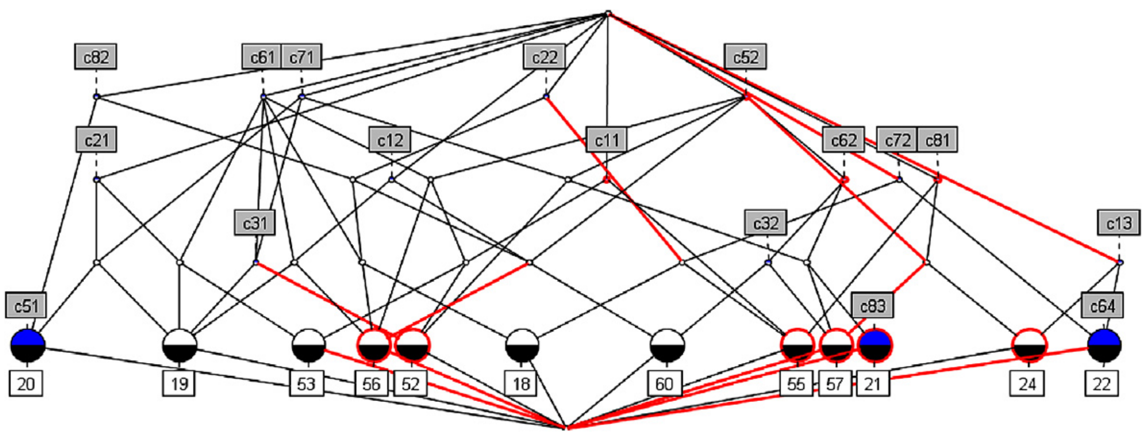
Rule 1. (c2 = 2) & (c3 = 2) & (c6 = 3) & (c7 = 1) ⇒ (d = 1); [5,5,6.49%,100.00%][5,0,0] [{2,41,110,116,145},{},{}]
Rule 2. (c3 = 1) & (c6 = 2) & (c8 = 1) ⇒ (d = 1); [2,2,2.60%,100.00%][2,0,0][{57,109},{},{}]
Rule 3. (c5 = 1) & (c6 = 1) & (c7 = 1) & (c8 = 1) ⇒ (d = 1); [4,4,5.19%,100.00%][4,0,0][{22,40,118,154},{},{}]
Rule 4. (c2 = 2) & (c6 = 3) & (c8 = 1) ⇒ (d = 1); [2,2,2.60%,100.00%][2,0,0][{145,189},{},{}]
Rule 5. (c1 = 1) & (c3 = 1) & (c6 = 2) ⇒ (d = 1); [3,3,3.90%,100.00%][3, 0,0][{109,195,198},{},{}]
Rule 6. (c5 = 1) & (c6 = 2) & (c8 = 2) ⇒ (d = 1); [2,2,2.60%,100.00%][2,0,0][{45,135},{},{}]
Rule 7. (c6 = 2) & (c7 = 2) & (c8 = 1) ⇒ (d = 1); [1,1,1.30%,100.00%][1,0,0][{95},{},{}]
Rule 8. (c1 = 2) & (c6 = 1) & (c7 = 1) & (c8 = 1) ⇒ (d = 1); [3,3,3.90%,100.00%][3,0,0][{27,118,200},{},{}]
Rule 9. (c2 = 1) & (c5 = 2) & (c8 = 5) ⇒ (d = 1); [1,1,1.30%,100.00%][1,0,0][{193},{},{}]
Rule 10. (c6 = 2) & (c7 = 2) & (c8 = 2) ⇒ (d = 1); [3,3,3.90%,100.00%][3,0,0][{45,180,195},{},{}]
Rule 11. (c1 = 5) & (c5 = 1) ⇒ (d = 1); [1,1,1.30%,100.00%][1,0,0][{144},{},{}]
Rule 12. (c1 = 3) & (c2 = 1) & (c5 = 2) & (c6 = 2) ⇒ (d = 1); [1,1,1.30%,100.00%][1,0,0][{114},{},{}]
Rule 13. (c2 = 1) & (c6 = 1) & (c7 = 2) & (c8 = 2) ⇒ (d = 1); [1,1,1.30%,100.00%][1,0,0][{18},{},{}]
Rule 14. (c1 = 2) & (c8 = 4) ⇒ (d = 2); [1,1,1.79%,100.00%][0,1,0][{},{60},{}]
Rule 15. (c1 = 4) & (c2 = 1) & (c8 = 1) ⇒ (d = 2); [1,1,1.79%,100.00%][0,1,0][{},{150},{}]
Rule 16. (c1 = 4) & (c3 = 2) & (c5 = 2) & (c6 = 4) & (c7 = 1) & (c8 = 2) ⇒ (d = 2); [1,1,1.79%,100.00%][0,1,0][{},{158}, {}]
Rule 17. (c1 = 3) & (c3 = 1) & (c6 = 4) ⇒ (d = 2); [1,1,1.79%, 100.00%][0,1,0][{},{108},{}]
Rule 18. (c2 = 2) & (c6 = 1) & (c7 = 2) & (c8 = 2) ⇒ (d = 2); [3,3,5.36%,100.00%][0,3,0][{},{6,96,112},{}]
Rule 19. (c1 = 2) & (c2 = 1) & (c3 = 1) & (c6 = 1) & (c7 = 1) ⇒ (d = 2); [2,2,3.57%,100.00%][0,2,0][{},{43,172},{}]
Rule 20. (c2 = 1) & (c5 = 1) & (c7 = 1) & (c8 = 2) ⇒ (d = 2); [2,2,3.57%,100.00%][0,2,0][{},{42,142},{}]
Rule 21. (c5 = 2) & (c6 = 2) & (c7 = 1) & (c8 = 3) ⇒ (d = 2); [2,2,3.57%,100.00%][0,2,0][{},{48,197},{}]
Rule 22. (c1 = 3) & (c6 = 4) & (c7 = 2) ⇒ (d = 2); [2,2,3.57%,100.00%][0,2,0][{},{62, 190},{}]
Rule 23. (c2 = 2) & (c5 = 3) & (c8 = 5) ⇒ (d = 2); [1,1,1.79%,100.00%][0,1,0][{},{50},{}]
Rule 24. (c1 = 3) & (c5 = 2) & (c8 = 1) ⇒ (d = 2); [3,3,5.36%,100.00%][0,3,0][{},{136,152,190},{}]
Rule 25. (c1 = 2) & (c3 = 2) & (c6 = 4) ⇒ (d = 2); [1,1,1.79%,100.00%][0,1,0][{},{111},{}]
Rule 26. (c3 = 1) & (c5 = 1) & (c7 = 2) ⇒ (d = 2); [1,1,1.79%,100.00%][0,1,0][{},{123},{}]
Rule 27. (c1 = 3) & (c3 = 1) & (c6 = 2) ⇒ (d = 2); [1,1,1.79%,100.00%][0,1,0][{},{74},{}]
Rule 28. (c1 = 2) & (c2 = 1) & (c6 = 2) & (c8 = 2) ⇒ (d = 3); [3,3,4.48%,100.00%][0,0,3][{},{},{3,65,188}]
Rule 29. (c2 = 1) & (c3 = 2) & (c5 = 3) ⇒ (d = 3); [8,8,11.94%,100.00%][0,0,8][{},{},{149,164,167,177,183,185,186,187}]
Rule 30. (c3 = 2) & (c6 = 3) & (c7 = 2) ⇒ (d = 3); [7,7,10.45%,100.00%][0,0,7][{},{},{47,54,63,113,192,194,196}]
Rule 31. (c1 = 4) & (c6 = 3) ⇒ (d = 3); [1,1,1.49%,100.00%][0,0,1][{},{},{168}]
Rule 32. (c1 = 5) & (c2 = 1) ⇒ (d = 3); [2,2,2.99%,100.00%][0,0,2][{},{},{148,167}]
Rule 33. (c6 = 4) & (c8 = 4) ⇒ (d = 3); [2,2,2.99%,100.00%][0,0,2][{},{},{138,165}]
Rule 34. (c1 = 3) & (c3 = 3) ⇒ (d = 3); [1,1,1.49%,100.00%][0,0,1][{},{},{140}]
Rule 35. (c1 = 3) & (c2 = 1) & (c3 = 1) ⇒ (d = 3); [1,1,1.49%, 100.00%][0,0,1][{},{},{76}]
Rule 36. (c3 = 2) & (c5 = 3) & (c8 = 2) ⇒ (d = 3); [5,5,7.46%,100.00%][0,0,5][{},{},{9,149,177,185,187}]
Rule 37. (c1 = 3) & (c6 = 4) & (c8 = 5) ⇒ (d = 3); [1,1,1.49%,100.00%][0,0,1][{},{},{75}]
Rule 38. (c1 = 3) & (c2 = 2) & (c3 = 2) & (c6 = 2) & (c7 = 1) ⇒ (d = 3); [1,1,1.49%,100.00%][0,0,1][{},{},{171}]
Rule 39. (c6 = 3) & (c8 = 3) ⇒ (d = 3); [3,3,4.48%,100.00%][0,0,3][{},{},{54,153,194}]
Rule 40. (c5 = 1) & (c6 = 2) & (c8 = 1) ⇒ (d = 3); [1,1,1.49%,100.00%][0,0,1][{},{},{19}]
Rule 41. (c1 = 2) & (c2 = 1) & (c3 = 2) & (c6 = 3) & (c8 = 2) ⇒ (d = 3); [4,4,5.97%,100.00%][0,0,4][{},{},{44,47,113,192}]
Rule 42. (c1 = 4) & (c3 = 1) ⇒ (d = 3); [1,1,1.49%,100.00%][0,0,1][{},{},{151}]
Rule 43. (c1 = 4) & (c7 = 2) ⇒ (d = 3); [3,3,4.48%,100.00%][0,0,3][{},{},{143,146,147}]
Rule 44. (c1 = 2) & (c2 = 2) & (c3 = 1) & (c6 = 3) & (c8 = 2) ⇒ (d = 3); [1,1,1.49%,100.00%][0,0,1][{},{},{160}]
Rule 45. (c3 = 2) & (c5 = 1) & (c6 = 1) ⇒ (d = 3); [1,1,1.49%,100.00%][0,0,1][{},{},{156}]
Rule 46. (c1 = 1) & (c5 = 3) ⇒ (d = 3); [1,1,1.49%,100.00%][0,0,1][{},{},{155}]
Rule 47. (c7 = 2) & (c8 = 3) ⇒ (d = 3); [3,3,4.48%,100.00%][0,0,3][{},{},{34, 54, 194}]
Rule 48. (c1 = 4) & (c5 = 1) & (c8 = 5) ⇒ (d = 3); [1,1,1.49%,100.00%][0,0,1][{},{},{169}]
Rule 49. (c1 = 2) & (c3 = 1) & (c6 = 4) ⇒ (d = 3); [1,1,1.49%,100.00%][0,0,1][{},{},{71}]
Rule 50. (c1 = 1) & (c3 = 2) & (c6 = 2) ⇒ (d = 3); [1,1,1.49%,100.00%][0,0,1][{},{},{4}]
Rule 51. (c1 = 2) & (c6 = 1) & (c7 = 2) & (c8 = 1) ⇒ (d = 3); [1,1,1.49%,100.00%][0,0,1][{},{},{99}]

Approximate rules
Rule 52. (c1 = 1) & (c3 = 1) & (c5 = 2) & (c6 = 1) & (c7 = 1) ⇒ (d = 1) OR (d = 2) OR (d = 3); [55,55,76.39%,100.00%][28,17,10][{7,11,14,16,26,29,31,37,46,78,80,81,84, 91,115,119,120,121,124,125,126,127,157,159,161,162,173,184}, {10,20,30,35, 36,38,69,79,82,83,97,117,122,130,133,139,176}, {1,15,25,32,64,89,98,128, 191,199}]
Rule 53. (c1 = 1) & (c2 = 1) & (c5 = 2) & (c6 = 1) ⇒ (d = 1) OR (d = 2) OR (d = 3); [45,45, 62.50%, 100.00%][24, 13, 8][{7,13,16,24,26,29,78,80,81,84,85,87,90,91,94, 119,121,124,125,126,127,157,173,184}, {10,20,23,30,35,38,69,79,82,122,130,131,139}, {12,15,25,32,89,98,191,199}]
Rule 54. (c1 = 2) & (c2 = 2) & (c3 = 1) & (c6 = 2) & (c8 = 2) ⇒ (d = 1) OR (d = 2) OR (d = 3); [8,8,11.11%,100.00%][5,1,2][{56,70,72,132,174}, {68}, {49,58}]
Rule 55. (c1 = 1) & (c2 = 2) & (c7 = 2) & (c8 = 1) ⇒ (d = 1) OR (d = 2); [8,8,42.11%,100.00%][6,2,0][{5,86,88,92,93,129}, {17,28},{}]
Rule 56. (c1 = 2) & (c2 = 2) & (c5 = 2) & (c6 = 1) & (c8 = 2) ⇒ (d = 1) OR (d = 2); [4,4,21.05%,100.00%][1,3,0][{102}, {52,103,182},{}]
Rule 57. (c3 = 2) & (c5 = 2) & (c6 = 2) & (c7 = 1) & (c8 = 1) ⇒ (d = 1) OR (d = 2); [3,3,15.79%,100.00%][1,2,0][{101}, {137,178},{}]
Rule 58. (c1 = 3) & (c2 = 1) & (c3 = 2) & (c5 = 2) & (c6 = 3) & (c7 = 1) & (c8 = 2) ⇒ (d = 1) OR (d = 2); [2,2,10.53%,100.00%][1,1,0][{181}, {53},{}]
Rule 59. (c1 = 3) & (c2 = 2) & (c3 = 2) & (c6 = 4) & (c8 = 2) ⇒ (d = 1) OR (d = 2); [2,2,10.53%,100.00%][1,1,0][{104}, {106},{}]
Rule 60. (c1 = 2) & (c2 = 2) & (c3 = 2) & (c5 = 2) & (c6 = 2) & (c8 = 2) ⇒ (d = 2) OR (d = 3); [5,5,41.67%,100.00%][0,3,2][{}, {55, 77, 141}, {51, 59}]
Rule 61. (c1 = 3) & (c6 = 4) & (c8 = 3) ⇒ (d = 2) OR (d = 3); [3,3,25.00%,100.00%][0,1,2][{},{105}, {100,170}]
Rule 62. (c1 = 4) & (c8 = 3) ⇒ (d = 2) OR (d = 3); [2,2,16.67%,100.00%][0,1,1][{},{166}, {8}]
Rule 63. (c1 = 1) & (c3 = 2) & (c8 = 1) ⇒ (d = 2) OR (d = 3); [2,2,16.67%,100.00%][0,1,1][{},{21}, {39}]
Rule 64. (c1 = 2) & (c2 = 1) & (c3 = 1) & (c6 = 3) ⇒ (d = 1) OR (d = 3); [3,3,30.00%,100.00%][1,0,2][{33},{},{107,134}]
Rule 65. (c3 = 2) & (c5 = 2) & (c6 = 1) & (c8 = 2) ⇒ (d = 1) OR (d = 3); [3,3,30.00%,100.00%][1,0,2][{175},{},{66,179}]
Rule 66. (c1 = 3) & (c2 = 2) & (c3 = 1) & (c6 = 3) ⇒ (d = 1) OR (d = 3); [2,2,20.00%,100.00%][1,0,1][{73},{},{163}]
Rule 67. (c1 = 3) & (c2 = 1) & (c6 = 4) & (c7 = 1) & (c8 = 2) ⇒ (d = 1) OR (d = 3); [2,2,20.00%,100.00%][1,0,1][{61},{},{67}]
**END

Compared the lattice diagram and context table, attributes with the least frequency in the context table will be posited at the bottom of the lattice diagram, and the concepts (attributes) are more specific. This means that those attributes are not important in determining the characteristics of the conservative portfolio, such as the female gender (c_{21}) and married status (c_{32}).

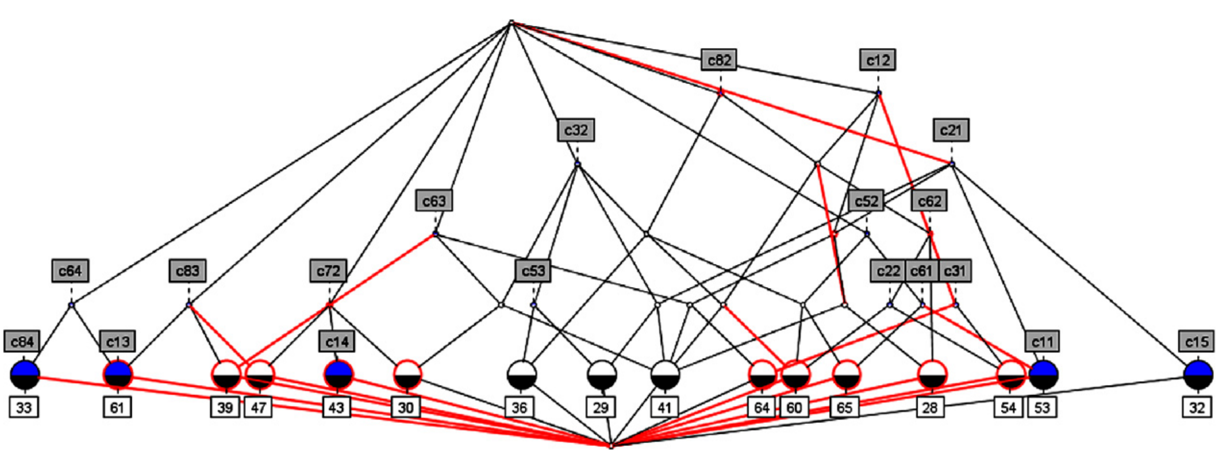
From the lattice diagram, association rules and implication sets generated by the ConExp program can also deduce the attribute relationship. For example, attributes c_{71}, c_{81}, c_{52}, c_{21} imply attribute c_{11}, and attributes c_{61}, c_{81}, c_{32}, c_{12} imply attribute c_{71}. The relative information about implication relation among attributes is shown in Table 6. The relative implication sets are shown in Table C.1 of Appendix C.

In Table 6, the highest frequency was c_{61}, meaning that the attribute was the most superconcept implied by other subconcepts. The concept of c_{61} was inherited by all its subconcepts, such as c_{12}, c_{11}, c_{31}, c_{51}, c_{52}, c_{71}, c_{81}. The second tier attributes were c_{71} and c_{22}. Those highest frequency superconcepts expressed the most important information (the common characteristic) for the conservative portfolio, which was less than 4 working years (c_{61}), a stable job (c_{71}) and male gender (c_{22}), which should contain some relationship, such as less monthly salary (c_{81},c_{82}) and lower education (c_{51},c_{52}). The results are reasonable because a person with fewer working years or lower education levels usually lacks the capacity to make aggressive investments; therefore, most of these subjects have a conservative portfolio.



c_{11}: age under 29; c_{12}: age between 30 and 39;c_{13}: age between 40-49; c_{21}: female;c_{22}: male; c_{31}:single; c_{32}:marry; c_{51}: under junior high school; c_{52}: high school; c_{61}: under 4 working years; c_{62}: 5 to 9 working years; c_{64}: 15 to 19 working years; c_{71}: stable job; c_{72}: service job; c_{81}: monthly salary under NT$30,000; c_{82}: monthly salary between NT$30,001 and NT$80,000; c_{83}: monthly salary between NT$80,001~ NT$120,000;

Fig. C.1. Lattice diagram for decision rules of decision class 2.



c_{11}: age under 29; c_{12}: age between 30 and 39; c_{13}: age between 40-49; c_{14}:age between 50~59; c_{15}: age higher than 60; c_{21}: female; c_{22}:male; c_{31}:single; c_{32}: marry; c_{52}:high school; c_{53}: college; c_{61}: under 4 working years; c_{62}: 5 to 9 working years; c_{63}: 10 to 14 working years; c_{64}: 15 to 19 working years; c_{72}: service job; c_{82}: monthly salary between NT$30,001 and NT$80,000; c_{83}: monthly salary between NT$80,001~ NT$120,000; c_{84}: monthly salary between NT$120,001~ NT$200,000;

Fig. C.2. Lattice diagram for decision rules of decision class 3.

Table C.1
The implication sets for decision class 1.

Implication sets	
1 <3> c82 ⇒ c62;	18 <1> c51 c71 ⇒ c61 c81;
2 <2> c52 ⇒ c11 c61;	19 <1> c11 c72 ⇒ c22 c81;
3 <2> c11 c61 ⇒ c52;	20 <1> c22 c72 ⇒ c11 c81;
4 <2> c63 ⇒ c22;	21 <1> c11 c81 ⇒ c22 c72;
5 <2> c61 c81 ⇒ c71;	22 <1> c12 c81 ⇒ c61 c71;
6 <2> c71 c81 ⇒ c61;	23 <1> c31 c81 ⇒ c62;
7 <1> c21 ⇒ c11 c52 c61;	24 <1> c51 c81 ⇒ c61 c71;
8 <1> c11 c22 ⇒ c72 c81;	25 <1> c62 c81 ⇒ c31;
9 <1> c32 ⇒ c22 c63 c71;	26 <1> c72 c81 ⇒ c11 c22;
10 <1> c12 c61 ⇒ c71 c81;	27 <1> c31 c62 c82 ⇒ c12 c22;
11 <1> c31 c61 ⇒ c11 c52 c71;	28 <1> c12 c22 ⇒ c31 c62 c82;
12 <1> c51 c61 ⇒ c71 c81;	29 <1> c12 c31 ⇒ c22 c62 c82;
13 <1> c11 c62 ⇒ c31;	30 <1> c22 c31 ⇒ c12 c62 c82;
14 <1> c11 c71 ⇒ c31 c52 c61;	31 <1> c12 c62 ⇒ c22 c31 c82;
15 <1> c12 c71 ⇒ c61 c81;	32 <1> c22 c62 ⇒ c12 c31 c82;
16 <1> c22 c71 ⇒ c32 c63;	33 <1> c51 c62 ⇒ c82;
17 <1> c31 c71 ⇒ c11 c52 c61;	34 <1> c62 c72 ⇒ c82;

3.3.2. Moderate portfolio

There are 12 rules after the pruning process for the decision class 2. The general information about the attributes relationship for the decision class 2 generated by the ConExp program are (1) a person with 5 and 9 working years (c_{62}) may have a high school education (c_{52}) ($c_{62} \Rightarrow c_{52}$); (2) a person with less than 4 working years (c_{61}) and a stable job (c_{71}) may be single ($c_{61}, c_{71} \Rightarrow c_{31}$); (3) a married person may have high school education (c_{52}) and between 5 and 9 working years ($c_{32} \Rightarrow c_{62}, c_{52}$); and (4) a female ($c_{21}$) with a high school education (c_{52}) may be under 29 years old (c_{11}) and have less than 4 working years (c_{61}) ($c_{21}, c_{52} \Rightarrow c_{11}, c_{61}$). The relative information about implication relations among attributes is shown in Table C.2.1 of Appendix C.

From the higher frequency of the context table, we can find the main characteristics for each attribute of the moderate portfolio, such as: age under 39, single or married unspecific, high school education, less than 4 working years, a stable job, salary between NT\$30,001 (US\$1000) and NT\$80,000 (US\$2667) (c_{82}), and the male gender. The most important information is high school education (c_{52}), and the least important information is attributes c_{51}, c_{64}, and c_{83}. The relative context table and lattice diagram are shown in Table C.2 and Fig. C.1 of Appendix C, respectively.

The most implied attribute is male (c_{22}), followed by high school education (c_{52}), and salary between NT\$30,001 (US\$1000) and NT\$80,000 (US\$2667) (c_{82}). We simply detailed the attribute relationship, such as attribute c_{12}, c_{52} implies attribute c_{22} or c_{82} ($c_{12}, c_{52} \Rightarrow c_{22}, c_{82}$), attribute c_{71}, c_{81} implies attribute c_{52}, c_{32}, or c_{62} ($c_{71}, c_{81} \Rightarrow c_{52}, c_{32}, c_{62}$), and attribute c_{72}, c_{82} could implies either attribute c_{61} ($c_{72}, c_{82} \Rightarrow c_{61}$) or attribute c_{22} ($c_{72}, c_{82} \Rightarrow c_{22}$). The relative information about implication relations among attributes is shown in Tables C.2.1 and C.3 of Appendix C.

Table C.2.1
The implication sets for decision class 2.

Implication sets	
1 <3> c62 ⇒ c52;	21 <1> c52 c62 c81 ⇒ c32 c71;
2 <2> c61 c82 ⇒ c22;	22 <1> c51 ⇒ c21 c71 c82;
3 <2> c61 c71 ⇒ c31;	23 <1> c32 c52 c62c71 ⇒ c81;
4 <2> c52 c82 ⇒ c12 c22;	24 <1> c31 c52c61 c71 ⇒ c11;
5 <2> c32 ⇒ c52 c62;	25 <1> c22 c81 ⇒ c11 c72;
6 <2> c31 ⇒ c61 c71;	26 <1> c21 c82 ⇒ c51 c71;
7 <2> c22 c61 ⇒ c82;	27 <1> c21 c52 ⇒ c11 c61;
8 <2> c22 c52 ⇒ c12 c82;	28 <1> c21 c31c61 c71 ⇒ c12;
9 <2> c12 c82 ⇒ c22 c52;	29 <1> c13 c81 ⇒ c52;
10 <2> c12 c52 ⇒ c22 c82;	30 <1> c13 c72 ⇒ c64;
11 <2> c12 c22 ⇒ c52 c82;	31 <1> c13 c52 ⇒ c81;
12 <2> c11 c61 ⇒ c52;	32 <1> c12 c71 ⇒ c21 c31 c61;
13 <2> c11 c52 ⇒ c61;	33 <1> c12 c22 c52c62 c82 ⇒ c32;
14 <1> c83 ⇒ c52 c62 c71;	34 <1> c12 c21 ⇒ c31 c61 c71;
15 <1> c72 c82 ⇒ c22 c61;	35 <1> c11 c81 ⇒ c22 c72;
16 <1> c72 c81 ⇒ c11 c22;	36 <1> c11 c72 ⇒ c22 c81;
17 <1> c71 c82 ⇒ c21 c51;	37 <1> c11 c71 ⇒ c31 c52 c61;
18 <1> c71 c81 ⇒ c32 c52 c62;	38 <1> c11c22 ⇒ c72 c81;
19 <1> c64 ⇒ c13 c72;	39 <1> c11 c21 ⇒ c52 c61;
20 <1> c61 c72 ⇒ c22 c82;	

3.3.3. Aggressive portfolio

There are 16 rules after the pruning process for decision class 3. The general information about the attribute relationships for the decision class 3 can retrieve from the lattice diagram (Fig. C.2), and implications sets (Table C.3.1) generated by the ConExp program, such as: (1) a person with a stable job (c_{71}) and a monthly salary under UT\$30,000 (US\$1000) (c_{81}) may have less than 4 working years (c_{61}), i.e. ($c_{71}, c_{81} \Rightarrow c_{61}$); (2) a female ($c_{21}$) may have a high school education (c_{52}) ($c_{21} \Rightarrow c_{52}$); (3) a married person (c_{32}) may have between 10 and 14 working years (c_{63}) and a stable job (c_{71}), i.e. ($c_{32} \Rightarrow c_{63}, c_{71}$); and (4) a person under 29 years old (c_{11}) and a high school education (c_{52}) may have less than 4 working years (c_{61}), i.e. ($c_{11}, c_{52} \Rightarrow c_{61}$).

From the higher frequency of the context table, we can find the main characteristics for each attribute of the aggressive portfolio, such as age between 30 and 39, female, married, between 10 and 14 working years, a service job, a salary between NT\$30,001 (US\$1000) and NT\$80,000 (US\$2667), and a high school education. The most important information is female (c_{21}), marriage (c_{32}), the monthly salary between NT\$30,001 (US\$1000) and NT\$80,000 (US\$2667) (c_{82}), and the least important information is attributes

Table C.2
Context Table for decision class 2 as example.

Rule #	Age (c_1)			Gender (c_2)		Marriage (c_3)		Education (c_5)		Working years (c_6)			Prof. (c_7)		Monthly salary (c_8)		
	c_{11}	c_{12}	c_{13}	c_{21}	c_{22}	c_{31}	c_{32}	c_{51}	c_{52}	c_{61}	c_{62}	c_{64}	c_{71}	c_{72}	c_{81}	c_{82}	c_{83}
18					×					×				×		×	
19		×		×		×				×			×				
20				×				×					×			×	
21									×		×		×				×
22			×									×		×			
24			×						×						×		
52	×					×			×	×			×				
53	×			×					×	×							
55	×				×									×	×		
56		×			×				×	×						×	
57							×		×		×		×		×		
60		×			×		×		×		×					×	
Frequency	3	3	2	3	4	2	2	1	7	5	3	1	5	3	3	4	1

c_{11}, c_{13}, c_{14}, c_{15} and c_{84}. The relative context table and lattice diagram are shown in Table C.4 and Fig. C.2 of Appendix C, respectively.

The most implied attribute is a stable job (c_{71}), followed by a monthly salary between NT$30,001 (US$1000) and NT$80,000 (US$2667) (c_{82}) and less 9 working years (c_{61},c_{62}). The attribute relationships are those attribute c_{32} implies attribute c_{71} or c_{63}, attribute c_{12}, c_{22} implies attribute c_{82}, c_{31}, or c_{62}, and attribute c_{31}, c_{52} implies attribute c_{61}, c_{11}, or c_{71}. The relative information about implication relation among attributes is shown in Table C.5 of Appendix C.

The FCA can provide further knowledge discovery from suitable rules. RST uses the reduct process to reduce the superfluous attribute in order to produce the reduct attribute set and the most important core attribute set, which may be the most important decision factors for decision making. However, in this study, the reduct attribute set and core attribute set are the same set $\{c_1, c_2, c_3, c_5, c_6, c_7, c_8\}$. RST cannot determine the relationships between attributes, the most important attribute, or the least important attribute. From the table of the implication relation between attributes, we summarized the highest frequency of implied attribute as the most important factor, followed by other the higher frequency of implied attributes as important factors affecting personal investment. Table 7 is the information of the attribute relationship for each personal portfolio.

In this study, we also found if minimum decision rules obtained by rough set theory then the number of rules is minimum and the antecedent of a rule is minimum. There will be no implications between condition attributes. In other words, if there are implications between condition attributes, the antecedents of the decision rules can be further simplified. For example, in the class 1 for the conservative portfolio, the rule 6 (the contents of rule 6 is ($c5 = 1$) & ($c6 = 2$) & ($c8 = 2$) $\Rightarrow$ ($d = 1$)), rule 10 (the contents of rule 10 is ($c6 = 2$) & ($c7 = 2$) & ($c8 = 2$) $\Rightarrow$ ($d = 1$)) and rule 54 (the contents of rule 54 is ($c1 = 2$) and ($c2 = 2$) & ($c3 = 1$) & ($c6 = 2$) & ($c8 = 2$) $\Rightarrow$ ($d = 1$) OR ($d = 2$) OR ($d = 3$)), will exist implication $c82 \Rightarrow c62$ which means that the antecedents of rule 6, 10 and 54 can be further simplified by removing $c6 = 2$ (the working years between 5 and 9 years).

However, the rules generated by ROSE2 cannot really get simplified forms for the facts of too many condition attributes and the number of every attribute value sets affect the class number. Besides, a single object set in condition part that will match one of the decision classes and increase the lower approximation. Both the lower and upper approximations increase, which yields a higher accuracy for the whole classification. The higher accuracy rate indicates that the objects in the class have a higher dependency among the condition attributes of those objects [20]. We call those objects as outlier. A single object set also increases the decision rules which means only one support object of a rule. Too many rules will decrease the precision of decisions. That is why we used the support object of a rule as the rule threshold to reduce the total rules and to increase the precision of decisions. The rule threshold can be user-defined by different method or condition.

Based on the Rough Set Theory, there are some parts belonging vague and indiscernibility relation. Those facts can not really get perfect reduct attribute set and core attribute set under Rough Set Theory. Those are the main reasons which RST always generates many rules. And this is also the main defective that is why we used

Table C.3.1
The implication sets for decision class 3.

Implication sets	
1 <2> c21 ⇒ c52;	24 <1> c31 c81 ⇒ c62;
2 <2> c32 ⇒ c63 c71;	25 <1> c51 c81 ⇒ c61 c71;
3 <2> c11 c52 ⇒ c61;	26 <1> c62 c81 ⇒ c31;
4 <2> c11 c61 ⇒ c52;	27 <1> c63 c81 ⇒ c22;
5 <2> c52 c61 ⇒ c11;	28 <1> c72 c81 ⇒ c11 c22;
6 <2> c63 c71 ⇒ c32;	29 <1> c12 c82 ⇒ c22 c31 c62;
7 <2> c61 c81 ⇒ c71;	30 <1> c22 c82 ⇒ c12 c31 c62;
8 <2> c71 c81 ⇒ c61;	31 <1> c31 c82 ⇒ c12 c22 c62;
9 <1> c11 c22 ⇒ c72 c81;	32 <1> c51 c82 ⇒ c62;
10 <1> c31 c52 ⇒ c11 c61 c71;	33 <1> c52 c82 ⇒ c13 c21 c32 c63 c71;
11 <1> c12 c61 ⇒ c71 c81;	34 <1> c63 c82 ⇒ c13 c21 c32 c52 c71;
12 <1> c31 c61 ⇒ c11 c52 c71;	35 <1> c71 c82 ⇒ c13 c21 c32 c52 c63;
13 <1> c51 c61 ⇒ c71 c81;	36 <1> c72 c82 ⇒ c62;
14 <1> c11 c62 ⇒ c31;	37 <1> c13 ⇒ c21 c32 c52 c63 c71 c82;
15 <1> c11 c71 ⇒ c31 c52 c61;	38 <1> c12 c22 ⇒ c31 c62 c82;
16 <1> c12 c71 ⇒ c61 c81;	39 <1> c12 c31 ⇒ c22 c62 c82;
17 <1> c22 c71 ⇒ c32 c63;	40 <1> c22 c31 ⇒ c12 c62 c82;
18 <1> c31 c71 ⇒ c11 c52 c61;	41 <1> c12 c62 ⇒ c22 c31 c82;
19 <1> c51 c71 ⇒ c61 c81;	42 <1> c22 c62 ⇒ c12 c31 c82;
20 <1> c11 c72 ⇒ c22 c81;	43 <1> c51 c62 ⇒ c82;
21 <1> c22 c72 ⇒ c11 c81;	44 <1> c52 c63 ⇒ c13 c21 c32 c71 c82;
22 <1> c11 c81 ⇒ c22 c72;	45 <1> c21 c52 c71 ⇒ c13 c32 c63 c82;
23 <1> c12 c81 ⇒ c61 c71;	46 <1> c62 c72 ⇒ c82;

Table C.3
implication relation between attributes for class 2.

Imply (superconcept)	c_{52}	c_{61}	c_{31}	c_{71}	c_{72}	c_{81}	c_{82}	c_{62}	c_{21}	c_{22}	c_{11}	c_{12}	c_{13}	c_{64}	c_{32}	c_{51}
(subconcept)	c_{11}, c_{21}	c_{11},c_{21}	c_{12}, c_{21}	c_{12},c_{21}	c_{11}, c_{22}	c_{11},c_{22}	c_{12}, $c_{22}{}^{*}2$	$c_{32}{}^{*}2$	c_{51}	c_{12}, $c_{52}{}^{*}2$	c_{21}, c_{52}	c_{22}, $c_{52}{}^{*}2$	c_{64}	c_{13}, c_{72}	c_{52},c_{62}, c_{81}	c_{21}, c_{82}
	c_{12}, $c_{22}{}^{*}2$	c_{12},c_{21}	c_{11}, c_{71}	$c_{31}{}^{*}2$	c_{64}	c_{13}, c_{52}	c_{51}	c_{71}, c_{81}	c_{12}, c_{71}	c_{11}, c_{72}	c_{31}, c_{52}, c_{61}, c_{71}	c_{21}, c_{31}, c_{61}, c_{71}			c_{71}, c_{81}	c_{71}, c_{82}
	$c_{32}{}^{*}2$	$c_{31}{}^{*}2$	c_{12},c_{71}	c_{51}	c_{11}, c_{81}	c_{32}, c_{52}, c_{62}, c_{71}	c_{12}, $c_{52}{}^{*}2$	c_{83}	c_{71},c_{82}	c_{61}, c_{72}	c_{22},c_{81}	c_{52}, $c_{82}{}^{*}2$			c_{12}, c_{22}, c_{52}, c_{62},c_{82}	
	c_{11}, $c_{61}{}^{*}2$	c_{11},$c_{52}{}^{*}2$	c_{61}, $c_{71}{}^{*}2$	c_{52}, c_{62}, c_{81}	c_{22},c_{81}	c_{11}, c_{72}	c_{22},$c_{52}{}^{*}2$			c_{11}, c_{81}	c_{72}, c_{81}					
	$c_{62}{}^{*}3$	c_{21}, c_{52}		c_{21}, c_{82}			c_{22}, $c_{61}{}^{*}2$			c_{72}, c_{81}						
	c_{11}, c_{71}	c_{11}, c_{71}		c_{83}			c_{61}, c_{72}			c_{12}, $c_{82}{}^{*}2$						
	c_{13}, c_{81}	c_{12}, c_{71}								c_{52}, $c_{82}{}^{*}2$						
	c_{71}, c_{81}	c_{72}, c_{82}								c_{61}, $c_{82}{}^{*}2$						
	c_{12}, $c_{82}{}^{*}2$									c_{72}, c_{82}						
	c_{83}															
Frequency	26	18	10	11	7	10	19	5	5	26	10	12	1	2	10	4

the FCA to diagnose the rules and retrieve the prior knowledge from the group customers' characteristics for the decision makers.

This study found that the formal concept analysis can assist decision makers in further information exploration. It also can help to determine the relationships between attributes. Regardless of the type of soft computing generating the decision rules, we can apply the formal concept analysis to get more information.

4. Conclusions

In this study, RST generates 67 rules. The support object of a rule as the rule threshold can reduce the total rules into 40 suitable rules. These suitable rules can be explored the further information by using the formal concept analysis, such as the most import factors affecting the relationship between personal investment portfolios and its attributes. This attributes relationship can give decision makers *a priori* predictions. The main characteristics of the conservative portfolio are a stable job (c_{71}), less than 4 working years (c_{61}), and male (c_{22}); and the other features are included less monthly salary (c_{81},c_{82}), and lower education (c_{51},c_{52}). The main characteristics of the moderate portfolio are high school education (c_{52}), less than 4 working years (c_{61}), and male (c_{22}); the following features are a stable job (c_{71}), a salary between NT$30,001 (US$1000) and NT$80,000 (US$2667) (c_{82}), and age under 39 years old (c_{11},c_{12}). The main characteristics of the aggressive portfolio are a salary between NT$30,001 (US$1000) and NT$80,000 (US$2667) (c_{82}), under 9 working years (c_{61},c_{62}) and a stable job (c_{71}); the following features are between 30 and 39 years old (c_{12}) and a single person (c_{31}). We believe that, regardless of the type of rule-based soft computing, the formal concept analysis can be used to find further information. We believe that it clearly is a useful method for decision makers.

Appendix A

See Table A.1.

Table C.4
Context Table for decision class 3 as example.

Rule #	Age (c_1)					Gender (c_2)		Marriage (c_3)		Education (c_5)		Working years (c_6)				Prof. (c_7)	Monthly Salary (c_8)		
	c_{11}	c_{12}	c_{13}	c_{14}	c_{15}	c_{21}	c_{22}	c_{31}	c_{32}	c_{52}	c_{53}	c_{61}	c_{62}	c_{63}	c_{64}	c_{72}	c_{82}	c_{83}	c_{84}
28		×				×							×				×		
29						×			×		×								
30									×					×		×			
32					×	×													
33															×				×
36									×		×						×		
39														×				×	
41		×				×			×					×			×		
43				×												×			
47																×		×	
53	×					×				×		×							
54		×					×	×					×				×		
60		×					×		×	×			×				×		
61			×												×			×	
64		×				×		×						×					
65									×	×		×					×		
Frequency	1	5	1	1	1	6	2	2	6	3	2	2	3	4	2	3	6	3	1

Table C.5
implication relation between attributes.

Imply (superconcept)	c_{52}	c_{63}	c_{71}	c_{61}	c_{11}	c_{32}	c_{72}	c_{81}	c_{31}	c_{22}	c_{62}	c_{12}	c_{21}	c_{13}	c_{82}
(subconcept)	$c_{21}{}^{*}2$	$c_{32}{}^{*}2$	$c_{32}{}^{*}2$	c_{11} $c_{52}{}^{*}2$	c_{52} $c_{61}{}^{*}2$	c_{63} $c_{71}{}^{*}2$	c_{11} c_{22}	c_{11} c_{22}	c_{11} c_{62}	c_{11} c_{72}	c_{31} c_{81}	c_{22} c_{82}	c_{52} c_{82}	c_{52} c_{82}	c_{13}
	c_{11}, $c_{61}{}^{*}2$	c_{22},c_{71}	c_{61}, $c_{81}{}^{*}2$	c_{71},$c_{81}{}^{*}2$	c_{31}, c_{52}	c_{22},c_{71}	c_{11}, c_{81}	c_{12},c_{61}	c_{11}, c_{71}	c_{11},c_{81}	c_{12}, c_{82}	c_{31},c_{82}	c_{63}, c_{82}	c_{63},c_{82}	c_{12}, c_{22}
	c_{31}, c_{61}	c_{52}, c_{82}	c_{31}, c_{52}	c_{31}, c_{52}	c_{31}, c_{61}	c_{52}, c_{82}		c_{51}, c_{61}	c_{62}, c_{81}	c_{63}, c_{81}	c_{22}, c_{82}	c_{22}, c_{31}	c_{71}, c_{82}	c_{71}, c_{82}	c_{12}, c_{31}
	c_{11}, c_{71}	c_{71}, c_{82}	c_{12}, c_{61}	c_{11}, c_{71}	c_{31}, c_{71}	c_{63}, c_{82}		c_{12}, c_{71}	c_{12}, c_{82}	c_{72}, c_{81}	c_{31}, c_{82}	c_{22}, c_{62}	c_{13}	c_{52}, c_{63}	c_{22}, c_{31}
	c_{31}, c_{71}	c_{13}	c_{31}, c_{61}	c_{12}, c_{71}	c_{22}, c_{72}	c_{71}, c_{82}		c_{31}, c_{71}	c_{22}, c_{82}	c_{12}, c_{82}	c_{51}, c_{82}		c_{52}, c_{63}	c_{21}, c_{52}, c_{71}	c_{12}, c_{62}
	c_{63}, c_{82}	c_{21},c_{52}, c_{71}	c_{51}, c_{61}	c_{31}, c_{71}	c_{72}, c_{81}	c_{13}		c_{11}, c_{72}	c_{12},c_{22}	c_{31}, c_{82}	c_{72},c_{82}				c_{22}, c_{62}
	c_{71}, c_{82}		c_{12},c_{81}	c_{31}, c_{71}		c_{52},c_{63}		c_{22}, c_{72}	c_{12},c_{62}	c_{12}, c_{31}	c_{12},c_{22}				c_{51}, c_{62}
	c_{13}		c_{51}, c_{81}	c_{12}, c_{81}		c_{21}, c_{52}, c_{71}			c_{22}, c_{62}	c_{12}, c_{62}	c_{12}, c_{31}				c_{52}, c_{63}
			c_{52}, c_{82}	c_{51},c_{81}							c_{22}, c_{31}				c_{21}, c_{52}, c_{71}
			c_{63}, c_{82}												c_{62}, c_{72}
			c_{13}												
			c_{52}, c_{63}												
Frequency	17	12	25	22	14	18	4	14	16	16	18	8	9	11	20

Appendix B

The original rules generated from ROSE2 are described in Table B.1. The rule syntax represents as follow:

(attribute$_\wedge$relation$_\wedge$value)$_\wedge$&...$_\wedge$(attribute$_\wedge$relation$_\wedge$value) ⇒ (decision = value)$_\wedge$ OR...OR$_\wedge$ (decision = value);[support, strength, relative strength,level_of_descrimination] [support_class1,support_class2,...,support_classN] [{class1_objects},{class2_objects},...,{classN_objects}]

The line begins with sequence of "rule no". Attribute is the name of conditional attribute, value is its value and relation is one of the following: "=", ">", "<", ">=", "<=", "in". Next the conditional part of the rule comes assignment to the decision class (es). Others details about the rule syntax can check the manual of the software ROSE2.

Appendix C

See Tables C.1, C.2, C.2.1, C.3, C.3.1, C.4 and C.5 and Figs. C.1 and C.2.

References

[1] M.J. Beynon, M.J. Peel, Variable precision rough set theory and data discretisation: an application to corporate failure prediction, Omega 29 (6) (2001) 561–576.
[2] C.C. Chan, A rough set approach to attribute generalization in data mining, Journal of Information Sciences 107 (1–4) (1998) 169–176.
[3] A. Formica, Concept similarity an formal concept analysis: an information content approach, Knowledge-Based Systems 21 (1) (2008) 80–87.
[4] R. Wille, Concept lattices and conceptual knowledge systems, Computers and Mathematics with Applications 23 (6-9) (1992) 493–515.
[5] S. Greco, B. Matarazzo, R. Slowinski, Rough sets theory for multicriteria decision analysis, European Journal of Operational Research 129 (1) (2001) 1–47.
[6] K.A. Hanisch, Reasons people retire and their relations to attitudinal and behavioral correlates in retirement, Journal of Vocational Behavior 45 (1) (1994) 1–16.
[7] T. Harrison, Understanding the behaviour of financial services consumers: a research agenda, Journal of Financial Service Marketing 8 (1) (2003) 6–9.
[8] T. Hennig-Thurau, A. Klee, The impact of customer satisfaction and relationship quality on customer retention: a critical reassessment and model development, Psychology & Marketing 14 (8) (1997) 737–764.
[9] Y.C. Hu, R.S. Chen, G.H. Tzeng, Finding fuzzy classification rules using data mining techniques, Pattern Recognition Letters 24 (1–3) (2003) 509–519.
[10] T. Keng, T. Hwa, The role of residential property in personal investment portfolios: the case of MALAYSIA, Pacific Rim Property Research Journal 10 (4) (2004) 466–486.
[11] H.L. Li, M.H. Chen, Induction of multiple criteria optimal classification rules for biological and medical data, Computers in Biology and Medicine 38 (1) (2008) 42–52.
[12] M. Liu, M. Shao, W. Zhang, C. Wu, Reduction method for concept lattices based on rough set theory and its application, Computer and Mathematics with Applications 53 (9) (2007) 1390–1410.
[13] H. Markowitz, Portfolio selection, Journal of Finance Paper 7 (1) (1952) 77–91.
[14] Z. Pawlak, Rough sets, International Journal of Computer and Information Science 11 (5) (1982) 341–356.
[15] Z. Pawlak, Rough classification, International Journal of Man–Machine Studies 20 (5) (1984) 469–483.
[16] Z. Pawlak, Rough sets, decision algorithms and Bayes' theorem, European Journal of Operational Research 136 (1) (2002) 181–189.
[17] D.A. Plath, T.H. Stevenson, Financial services consumption behavior across Hispanic American consumers, Journal of Business Research 58 (8) (2005) 1089–1099.
[18] L.T. Polkowski, Toward rough set foundations, mereological approach, in: Shusaku Tsumoto, Roman Slowinski, Jan Komoroski, Jerzy W. Grzymala-Busse (Eds.), Rough Sets and Current Trends in Computing, Lecture Notes in Artificial Intelligence (LNAI), vol. 3066(1), Springer Verlag, Berlin, 2004, pp. 8–25.
[19] U. Priss, Formal concept analysis in information science, in: Cronin, Blaise (Eds.), Annual Review of Information Science and Technology, vol. 40, Information Today, Inc., Medford, NJ, 2006, pp. 521–543.
[20] J.Y. Shyng, F.K. Wang, G.H. Tzeng, K.S. Wu, Rough set theory in analyzing the attributes of combination values for the Insurance Market, Expert System with Application 32 (1) (2007) 56–64.
[21] J.Y. Shyng, G.H. Tzeng, S.H. Hsieh, F.K. Wang, Data mining for multi-domain decision-making based on rough set theory, in: 2006 IEEE International Conference on Systems, Man, and Cybernetics, Taipei, Taiwan.
[22] R.W. Swiniarski, A. Skowron, Rough set methods in feature selection and recognition, Pattern Recognition Letters 24 (6) (2003) 833–849.
[23] B. Walczak, D.L. Massart, Tutorial rough sets theory, Chemometrics and Intelligent Laboratory Systems 47 (1) (1999) 1–16.
[24] R. Wille, Formal concept analysis as methodical theory of concepts and concept hierarchies, in: B. Ganter et al. (Eds.), Formal Concept Analysis, LNAI 3626, 2005, pp. 1–3.
[25] F. Witlox, H. Tindemans, The application of rough sets analysis in activity-based modeling, opportunities and constraints, Expert Systems with Application 27 (2) (2004) 171–180.
[26] B. Wormuth, P. Becker, Introduction to formal concept analysis, in: 2nd International Conference of Formal Concept Analysis February 23–February 27, 2004, Sydney, Australia.
[27] Y.Y. Yao, Concept lattices in rough set theory, in: Proceedings of 23rd International Meeting of the North American Fuzzy Information Processing Society, 2004.
[28] L.A. Zadeh, Fuzzy sets, Information and Control 8 (3) (1965) 338–353.
[29] L.Y. Zhai, L.P. Khoo, S.C. Fok, Feature extraction using rough set theory and genetic algorithms an application for the simplification of product quality evaluation, Computers & Industrial Engineering 43 (4) (2002) 661–676.
[30] ROSE 2 – rough set data explorer, aboratory of Intelligent Decision Support Systems, Poznan University of Technology, Poznan, 2005, <http://www-idss.cs.put.poznan.pl/software>.
[31] ConExp, http://sourceforge.net/projects/conexp>.
[32] R. Brüggemann, K. Voigt, C.E.W. Steinberg, Application of formal concept analysis to evaluate environmental databases, Pergamon 35 (3) (1997) 479–486.
[33] J.Y. Shyng, H.M. Shieh, G.H. Tzeng, S.H. Hsieh, Using FSBT technique with rough set theory for personal investment portfolio analysis, European Journal of Operational Research 201 (2010) 601–607.
[34] T. Herawan, M.M. Deris, J.H. Abawajy, A rough set approach for selecting clustering attribute, Knowledge-Based Systems 23 (2010) 220–231.
[35] H.L. Lim, C.S. Lee, Processing online analytics with classification and association rule mining, Knowledge-Based Systems 23 (2010) 248–255.