

รายงานการวิจัย

เรื่อง

การใช้กราฟสายงานศึกษาข้อมูลคะแนนคณิตศาสตร์  
เพื่อสร้างกฎการจำแนก

Decision Rules Discovery from Mathematics Students' Score by  
Flow Graphs

โดย

ดร. พรรณทิพย์ ภัทรอินทากร

ภาควิชาคณิตศาสตร์และวิทยาการคอมพิวเตอร์

คณะวิทยาศาสตร์

สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง

งานวิจัยนี้ได้รับทุนวิจัยจากทุนวิจัยคณะวิทยาศาสตร์

ประจำปีงบประมาณ พ.ศ. 2551

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า  
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

รายงานการวิจัย

เรื่อง

การใช้กราฟสายงานศึกษาข้อมูลคะแนนคณิตศาสตร์  
เพื่อสร้างกฎการจำแนก

Decision Rules Discovery from Mathematics Students' Score by  
Flow Graphs

เลขหมู่.....  
เลขทะเบียน.....  
วัน, เดือน, ปี.....

120215

โดย

12337651

ดร. พรรณทิพย์ กัทรอินทากร

ภาควิชาคณิตศาสตร์และวิทยาการคอมพิวเตอร์

คณะวิทยาศาสตร์

สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง

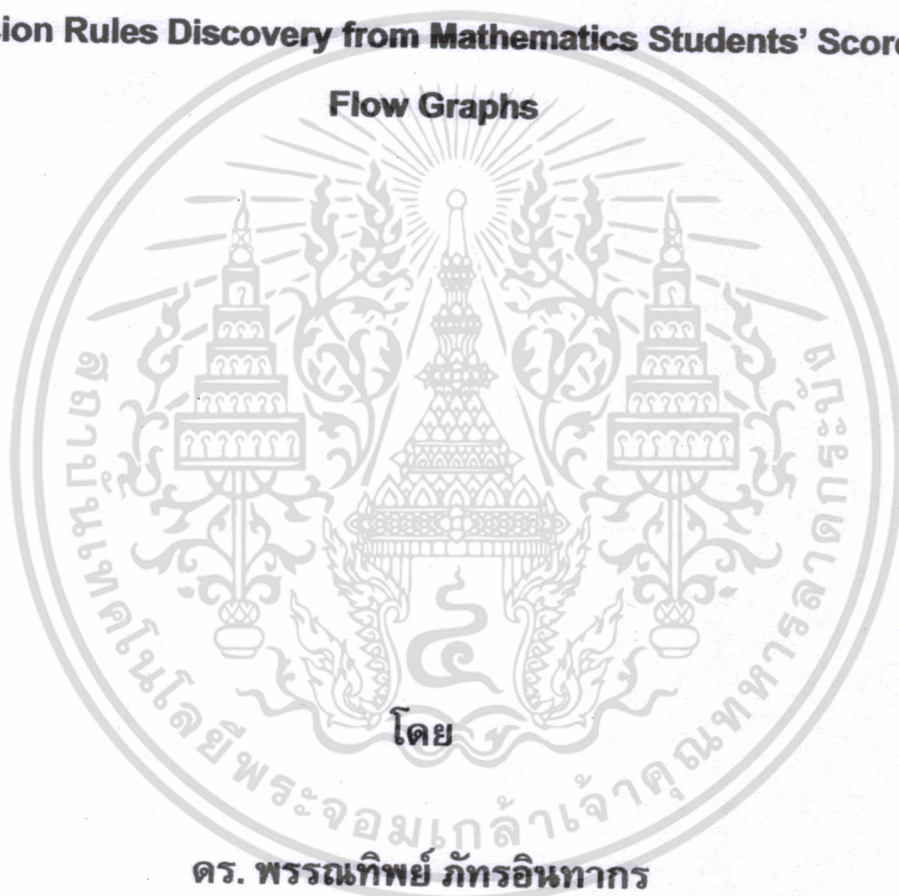
งานวิจัยนี้ได้รับทุนวิจัยจากทุนวิจัยคณะวิทยาศาสตร์

ประจำปีงบประมาณ พ.ศ. 2551

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า  
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

**การใช้กราฟสายงานศึกษาข้อมูลคะแนนคณิตศาสตร์  
เพื่อสร้างกฎการจำแนก**

**Decision Rules Discovery from Mathematics Students' Score by  
Flow Graphs**



**ดร. พรรณทิพย์ ภัทรอินทากร**

**ภาควิชาคณิตศาสตร์และวิทยาการคอมพิวเตอร์**

**คณะวิทยาศาสตร์**

**สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง**



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า...  
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

## กิตติกรรมประกาศ

รายงานการวิจัยเรื่อง “การใช้กราฟสายงานศึกษาข้อมูลคะแนนคณิตศาสตร์เพื่อสร้างกฎการจำแนก” (Decision Rules Discovery from Mathematics Students' Score by Flow Graphs) เป็นโครงการที่จัดทำขึ้นโดยได้รับการสนับสนุนจากทุนวิจัยคณะวิทยาศาสตร์ คณะวิทยาศาสตร์ สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง เพื่อนำข้อมูลไปใช้ประกอบการจัดทำแนวทางการศึกษาวิจัยเพื่อสร้างขีดความสามารถในเรื่องการพัฒนาทรัพยากรมนุษย์และ การวิจัยในด้านการเรียนการสอนด้วยคอมพิวเตอร์ทั้งในส่วนของระบบซอฟต์แวร์และระบบฮาร์ดแวร์ แผนงานวิจัยแห่งชาติด้านวิทยาศาสตร์เทคโนโลยีและอุตสาหกรรม ประเภทการวิจัยประยุกต์ สาขาวิจัยคณิตศาสตร์ประยุกต์

ผู้วิจัยขอขอบคุณคณะกรรมการทุนวิจัยคณะวิทยาศาสตร์ซึ่งมี รศ.ดร.ธีรวัฒน์ มงคลอัครวัฒน์ เป็นประธาน และคณะทำงานอำนวยการ โครงการวิจัยที่ได้ให้ข้อคิดเห็นและข้อเสนอแนะที่เป็นประโยชน์ยิ่งต่อการวิจัย ที่ให้ทุนสนับสนุนในการทำวิจัยและขอขอบคุณ รศ. ไพโรบลย์ พันธรักษ์พงษ์ หัวหน้าภาควิชาคณิตศาสตร์และวิทยาการคอมพิวเตอร์ที่ช่วยติดตามดูแลงานวิจัย โดยเฉพาะอย่างยิ่งขอขอบคุณมูลนิธิส่งเสริมโอลิมปิกวิชาการ และพัฒนามาตรฐานวิทยาศาสตร์ศึกษา ในพระอุปถัมภ์สมเด็จพระเจ้าพี่นางเธอ เจ้าฟ้ากัลยาณิวัฒนา กรมหลวงนราธิวาสราชนครินทร์ (สอวน.) และ รศ.เย็นใจ สมวิเชียร กรรมการและเหรียญกฐลนิธิ สอวน. ที่เปิดโอกาสให้ผู้วิจัยได้ใช้ข้อมูล

ขอขอบคุณทุนวิจัยคณะวิทยาศาสตร์ที่อนุมัติงบประมาณอุดหนุนวิจัย เป็นค่าใช้จ่ายของ โครงการวิจัยนี้และขอขอบคุณฝ่ายเลขานุการคณะทำงานที่ให้ความร่วมมือแก่นักวิจัยอย่างยิ่ง

ดร.พรณทิพย์ ภัทรอินทากร

สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง

2551

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่นิยมนำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

## บทคัดย่อ

การใช้คณิตศาสตร์แขนงเก่า (hard computing) เช่น การใช้ตัวแบบทางคณิตศาสตร์ การวิเคราะห์เชิงตัวเลขในการวิเคราะห์ข้อมูลต่าง ๆ นั้นจะต้องมีการวิเคราะห์ความผิดพลาด (error analysis) ด้วยเสมอว่าในการสร้างตัวแบบนั้นจะให้ความผิดพลาดไม่เกินขอบเขตตามที่ได้พิสูจน์ไว้ แต่อย่างไรก็ดี สำหรับข้อมูลจริงที่มีขนาดใหญ่และมีค่าคลาดเคลื่อนในข้อมูลอยู่แล้วเป็นจำนวนมาก จะให้ค่าความผิดพลาดของตัวแบบดังกล่าวนั้น แบบเส้นโค้งเลขชี้กำลัง ทำให้การใช้คณิตศาสตร์แขนงเก่าดังกล่าวจึงไม่สามารถวิเคราะห์ได้ถูกต้องเพียงพอ

ในปี 2523, Zdzislaw Pawlak ได้คิดค้นคณิตศาสตร์แขนงใหม่ขึ้นจากแนวคิดเรื่องเซต นั่นคือ กราฟสายงาน (flow graph) ซึ่งเป็นคณิตศาสตร์แขนงใหม่ชนิด คณนาแบบเปลี่ยนได้ (soft computing) โดยจะคำนวณค่าคลาดเคลื่อนโดยตรงและไม่ทำให้เกิดความผิดพลาดแบบเส้นโค้งเลขชี้กำลังสำหรับข้อมูลจริงขนาดใหญ่มาก จึงให้ความถูกต้องมากขึ้น อีกทั้งยังสามารถคำนวณด้วยวิธีการประมวลผลแบบขนานได้ จึงทำให้คำนวณได้เร็วขึ้นและใช้ความจุหน่วยเก็บน้อยลง ด้วยประสิทธิภาพดังกล่าว ทำให้วิธีนี้ได้ถูกนำไปใช้ในศาสตร์แขนงต่าง ๆ อย่างกว้างขวาง อาทิเช่น preference analysis, conflict analysis, rule analysis, propositional calculus, data mining, granular computing

จากข้อดีทั้งการใช้ความจุหน่วยเก็บน้อยลง ความเร็วในการคำนวณและความถูกต้องของวิธีนี้ในงานวิจัยเบื้องต้นดังกล่าว ทางผู้วิจัยจึงใช้กราฟสายงานเพื่อวิเคราะห์ข้อมูลคะแนนคณิตศาสตร์โดยคำนวณสัมประสิทธิ์ดังต่อไปนี้ *strength*, *certainty* และ *coverage* เพื่อวิเคราะห์ความสัมพันธ์ของปัจจัยต่าง ๆ รวมถึงสร้างกฎการจำแนกคะแนนของนักเรียนจากความสัมพันธ์ของปัจจัยนั้น

## ABSTRACT

---

---

Mathematical rough set theory has attracted both practical and theoretical researchers. A significant extension of rough set theory is called *flow graphs*, invented by Zdzislaw Pawlak (1980). It is a knowledge representation in the form of information flow. Mathematical flow graphs are in soft computing area which provides less complexity, times and resources. Thus, flow graph is a promising approach to analyze data flow, decision trees, decision rules, probability learning, etc.

In this research, we present their connections to pertinent techniques and propose a new extension. We discuss some important properties of flow graphs. Several coefficients: strength, certainty and coverage are analyzed to students' score successfully.

---

## รายงานวิจัยฉบับสมบูรณ์

---

**โครงการ** การใช้กราฟสายงานศึกษาข้อมูลคะแนนคณิตศาสตร์เพื่อสร้างกฎการ  
จำแนก  
Decision Rules Discovery from Mathematics Students' Score by Flow  
Graphs

### ทิศทางการวิจัย (Research Direction):

การศึกษาวิจัยเพื่อสร้างขีดความสามารถในเรื่องการพัฒนาทรัพยากร  
มนุษย์และการวิจัยในด้านการเรียนการสอนด้วยคอมพิวเตอร์ทั้งในส่วนของระบบ  
ซอฟต์แวร์และระบบฮาร์ดแวร์

### แผนวิจัย (Research Plan):

แผนงานวิจัยแห่งชาติด้านวิทยาศาสตร์ เทคโนโลยีและอุตสาหกรรม

### ประเภทการวิจัย:

การวิจัยประยุกต์

### สาขาการวิจัย:

คณิตศาสตร์ประยุกต์

### ระยะเวลาโครงการ: 1 ปี

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า  
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

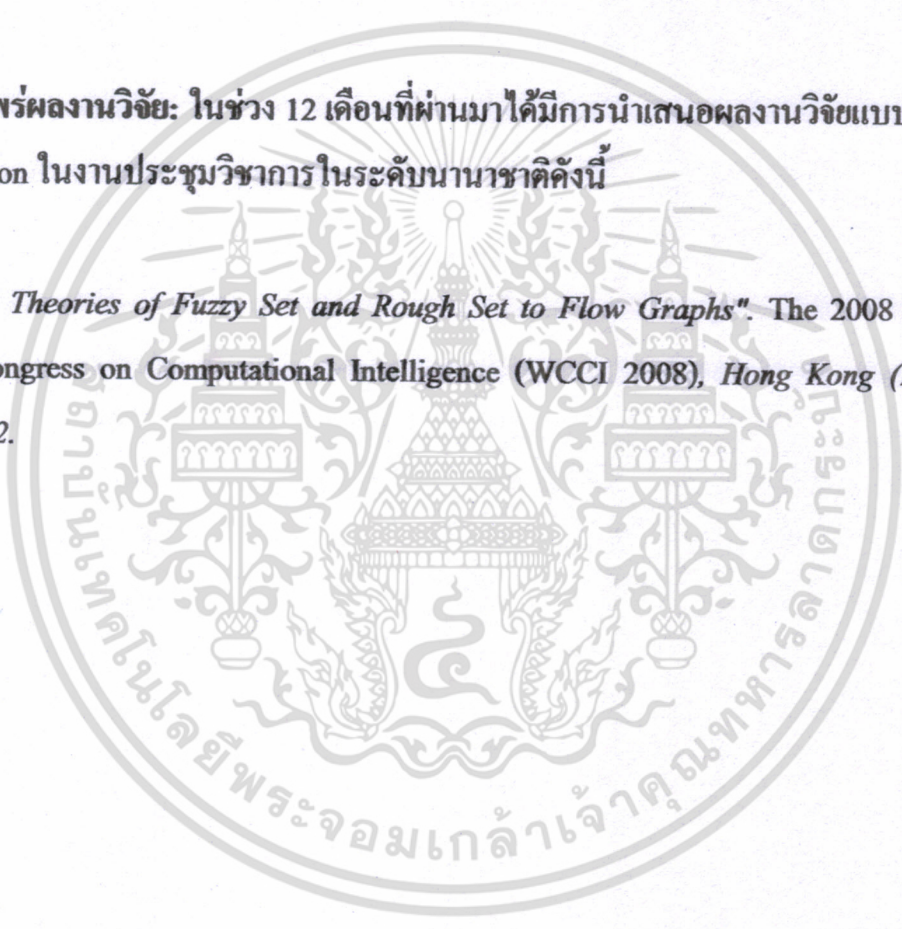
ชื่อหัวหน้าโครงการวิจัยผู้รับทุน: คร. พรรณทิพย์ ภัทรอินทากร

การดำเนินงาน  ได้ดำเนินงานตามแผนที่วางไว้  
 ได้ดำเนินงานล่าช้ากว่าแผนที่วางไว้

คำสำคัญ (Keyword) ของโครงการวิจัย: Flow graphs, Decision rule, Association rule, Data analysis

การเผยแพร่ผลงานวิจัย: ในช่วง 12 เดือนที่ผ่านมาได้มีการนำเสนอผลงานวิจัยแบบ Oral presentation ในงานประชุมวิชาการในระดับนานาชาติดังนี้

"Towards Theories of Fuzzy Set and Rough Set to Flow Graphs". The 2008 IEEE World Congress on Computational Intelligence (WCCI 2008), Hong Kong (2008) 1675-1682.



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า  
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

# สารบัญเรื่อง

เรื่อง	หน้า
I. INTRODUCTION	1
II. ROUGH SET THEORY	3
III. FUZZY SET THEORY	5
IV. FLOW GRAPHS	7
A. NORMALIZED FLOW GRAPHS	7
B. DECISION RULES, DECISION ALGORITHM AND FLOW GRAPH	10
C. NEW CERTAINTY AND COVERAGE COEFFICIENTS FOR DECISION RULES FROM RELAXED MUTUALLY EXCLUSIVE	12
D. FLOW GRAPHS WITH FUZZY ATTRIBUTES	19
V. EXPERIMENTAL RESULTS AND DISCUSSION	22
VI. CONCLUSION AND SUGGESTION	34
REFERENCES	35

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า  
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

# สารบัญตาราง

ตาราง

หน้า

TABLE I. DECISION TABLE WITH FUZZY ATTRIBUTES

6



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า  
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

## สารบัญรูป

รูป

หน้า

FIGURE I. COMPARISON OF *CER*, *COV* OF POSN DATA

33



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า  
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

## I. INTRODUCTION

**R**OUGH set theory was introduced by Zdzislaw Pawlak in the early 1980's [1]. The crux of rough set theory is the novel use of approximations to cope with uncertainty. A significant extension of rough set theory is called *flow graphs*. Rough set theory formed the basis of flow graphs later in the early 2000s.

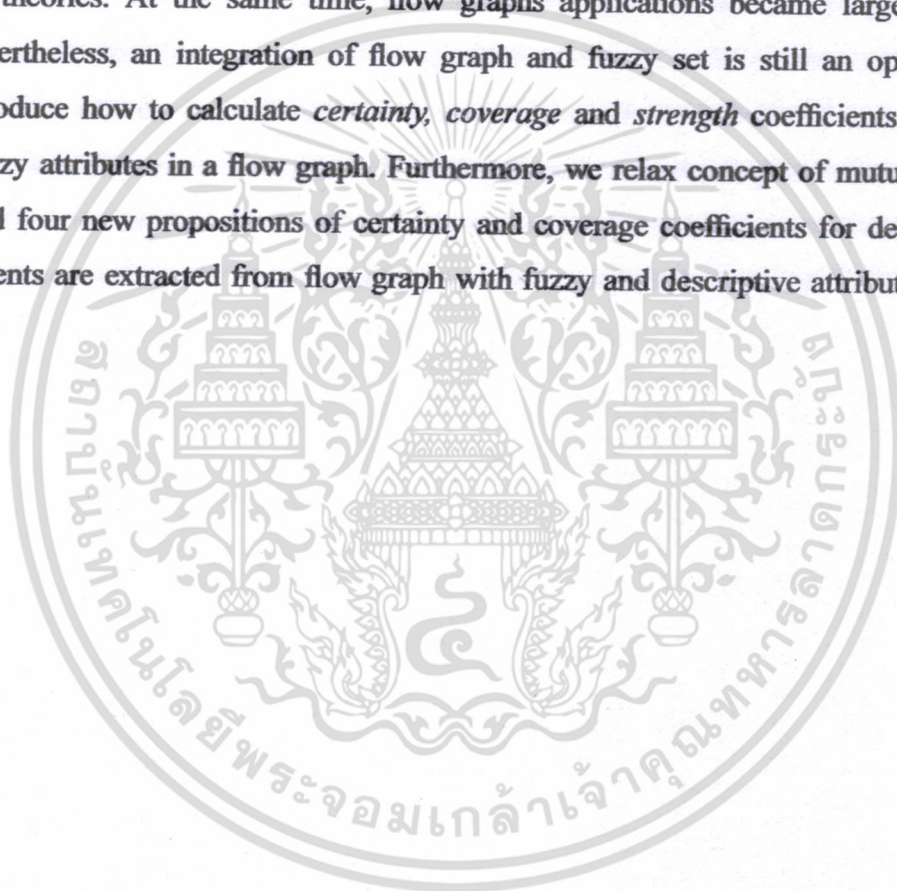
Pawlak's flow graphs are different to *flow networks* introduced by Ford et al. that do not analyze optimal flow [2]. Flow graphs represent the model of information flow from the given data set. The branches of a flow graph can be constructed as decision rules [3]. With every decision rules, there are three associated coefficients: *strength*, *certainty* and *coverage*. These coefficients satisfy Bayes' theorem. We can also discover dependency and correlation within a data set without reference to its probabilistic nature by using flow graphs [4]. Pawlak also proposed a new approach to data mining and knowledge discovery based on information flow distribution. In [5], Pawlak shown that a decision tree can be constructed as a flow graph. The author concluded that such a flow graph provides better structure of data compared to a decision tree. Pawlak revealed the relationship between flow graphs and probability. Pawlak revealed the relationship between flow graphs and probability. Rough sets also closely related with *modus ponens* (MP) and *modus tollens* (MT) inference rules [6], [7]. Preference analysis [8], rule analysis [9], propositional calculus [6], [7], data mining [5], [10] and granular computing [11] were analyzed by flow graphs efficiently. Essentially, Pawlak interpreted the union of all inputs  $x$  of  $y$  as the upper approximation of  $y$  and the union of all inputs  $x$  of  $y$  such that  $cer(x, y) = 1$  as the lower approximation of  $y$  [8].

Fuzzy set theory was introduced by Lofti Zadeh in 1965 [12]. A fuzzy set is a class of objects with a continuum of grader of membership. The characteristic of *membership function* of a fuzzy set is permitted to have value in the interval  $[0, 1]$ . If  $A$  is a classical set, its membership function can take on only two values  $0$  and  $1$ . It can be understood as a vague boundary set theory. Recently, fuzzy set theory and its applications were developed extensively and fascinated

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า  
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

attention [12]. In real world data analysis, if collected data is in the form of fuzzy attributes, then it can be represented in the decision table with fuzzy attributes. We can analyze this decision table by applying concepts of fuzzy set. Alicja et al. introduced a concept of fuzzy flow graphs in [13]. Their approach allowed computation not only for descriptive attributes but also for fuzzy attributes in flow graphs.

Pawlak's flow graph is a useful tool for knowledge discovery and has been successfully applied in many areas. Within the past decade, there has been much research about theoretical aspects of flow graphs which explored the complementary nature of their properties and other mathematical theories. At the same time, flow graphs applications became larger and more complex. Nevertheless, an integration of flow graph and fuzzy set is still an open problem. Thus, we introduce how to calculate *certainty*, *coverage* and *strength* coefficients of decision rules from fuzzy attributes in a flow graph. Furthermore, we relax concept of mutual exclusion and introduced four new propositions of certainty and coverage coefficients for decision rules. These coefficients are extracted from flow graph with fuzzy and descriptive attributes.



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า  
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

## II. ROUGH SET THEORY

The theory of rough sets, proposed by Pawlak, provides a formal tool to deal with imprecise or incomplete information. The following rough sets preliminary is taken from [1]. This mathematical theory formed the basis of flow graphs later in the early 2000s [3], [5], [7], [8], [14].

In rough set based data analysis, the equivalence relation in an approximation space is commonly interpreted via the notion of information system. Rough sets are based on an *information system (data table)*. More formally, it is a pair  $S = (U, A)$ , where  $U$  is a nonempty finite set of objects called the *universe* and  $A$  is a nonempty finite set of attributes such that  $a: U \rightarrow V_a$  for every  $a \in A$ . The set  $V_a$  is called the *domain* of  $a$ . Any subset  $B$  of  $A$  determines a binary relation  $I(B)$  on  $U$  called an *indiscernibility relation*. It is defined as  $(x, y) \in I(B)$  if and only if  $a(x) = a(y)$  for every  $a \in A$ , where  $a(x)$  denotes the attribute value of element  $x$ . If  $(x, y)$  belongs to  $I(B)$ ,  $x$  and  $y$  are called *B-indiscernible*. Equivalence classes of the relation  $I(B)$  are referred to as *B-elementary sets* (or equivalently *B-granules*, for more granular computing studies please refer to [8], [11]).

If we distinguish an information system to two disjoint classes of attributes, called *condition* and *decision attributes*, then the information system will be called a *decision system*, denoted by  $S = (U, C, D)$ , where  $C \cap D = \emptyset$ .

Suppose we are given an information system  $S = (U, A)$ ,  $X \subseteq U$ ,  $B \subseteq A$  and  $B(X)$  describes  $X$  in the terms of attribute values from  $B$ . We can approximate  $X$  using the information in  $B$  by constructing the *B-lower* and *B-upper approximations* of  $X$ , denoted by  $B_*(X)$  and  $B^*(X)$  respectively. They are defined as

$$B_*(X) = \bigcup_{x \in U} \{B(x) \mid B(x) \in X\},$$

$$B^*(X) = \bigcup_{x \in U} \{B(x) \mid B(x) \cap X \neq \emptyset\},$$

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า  
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

respectively.

In addition, the set

$$BN_B(X) = B^*(X) - B_*(X)$$

will be referred to as *B-boundary region* of  $X$ .

As we can see from above definitions, they are expressed in terms of granules (small pieces) of knowledge. The lower approximation is the union of all granules which are entirely included in the set. The upper approximation is the union of all granules which have non-empty intersection with the set. The boundary region is the difference between upper approximation and lower approximation.

If the boundary region of  $X$  is an empty set

$$BN_B(x) = \emptyset,$$

then  $X$  is *crisp*.

In contrast, if

$$BN_B(X) \neq \emptyset,$$

then  $X$  is *rough*.

Pawlak also discussed two numerical characterizations of the imprecision of a subset  $X$ . More studies involve rough sets are in [1], [3], [5], [6], [7], [8], [10], [14].

### III. FUZZY SET THEORY

Lofti Zadeh proposed a new mathematical approach to handle vagueness called fuzzy set theory in 1965. It provides a better mechanism to describe behavior of data which is too ill-defined to admit precise mathematical analysis by classical approaches. The following fuzzy set theory concept is taken from [12].

Let  $X$  be a space of objects and  $x$  be a generic element of  $X$ . A *fuzzy set*  $A$  in  $X$  is characterized by a *membership function*,  $\mu_A(x) \rightarrow [0,1]$ . The nearer the value of  $\mu_A(x)$  to 1, the higher the grade of membership of  $x$  in  $A$ . Thus, fuzzy sets provide a convenient tool for representing vague concepts by employing the fuzzy membership function. A crucial issue for fuzzy sets applications is to express real-world problem by fuzzy sets attributes in decision table and perform data analysis. In further discussions, we consider decision table with fuzzy attributes in the form introduced in [13], [15].

Let  $(U, C \cup D)$  be a decision table with fuzzy attributes,  $U = \{x_1, x_2, \dots, x_N\}$  be a finite non-empty set of universe with  $N$  elements,  $C = \{c_1, c_2, \dots, c_n\}$  be a set of condition attributes and  $D = \{d_1, d_2, \dots, d_m\}$  be a set of decision attributes.

The domain of the family of linguistic values of condition attribute  $c_i$  denoted by  $V_i = \{V_{i1}, V_{i2}, \dots, V_{in_i}\}$ , where  $n_i$  is the number of the linguistic values of  $i$ -th conditions,  $i = 1, 2, \dots, n$ . The domain of the family of linguistic values of decision attribute  $d_j$  denoted by  $W_j = \{W_{j1}, W_{j2}, \dots, W_{jm_j}\}$ , where  $m_j$  is the number of the linguistic values of  $j$ -th conditions,  $j = 1, 2, \dots, m$ .

In order to describe an element  $x \in U$  with fuzzy attribute, its membership degree should be determined. This process is called fuzzification state. In this state, for any element  $x$ , fuzzy value of condition attribute  $c_i$  and decision attribute  $d_j$  denoted by  $V_i(x)$  and  $W_j(x)$  are defined

as

$$V_i(x) = \left\{ \frac{\mu_{v_{i_1}}(x)}{V_{i_1}}, \frac{\mu_{v_{i_2}}(x)}{V_{i_2}}, \dots, \frac{\mu_{v_{i_{n_i}}}(x)}{V_{i_{n_i}}} \right\} \text{ and}$$

$$W_j(x) = \left\{ \frac{\mu_{w_{j_1}}(x)}{W_{j_1}}, \frac{\mu_{w_{j_2}}(x)}{W_{j_2}}, \dots, \frac{\mu_{w_{j_{m_j}}}(x)}{W_{j_{m_j}}} \right\}.^1$$

Furthermore, for any  $x \in U$  we assume that values of  $V_i(x)$  and  $W_j(x)$  satisfy the following conditions:

$$\text{power}(V_i(x)) = \sum_{k=1}^{n_i} \mu_{V_{i_k}}(x) = 1 \text{ and}$$

$$\text{power}(W_j(x)) = \sum_{k=1}^{m_j} \mu_{W_{j_k}}(x) = 1,$$

where  $i = 1, 2, \dots, n, j = 1, 2, \dots, m$ .

An example of fuzzy values  $V_i(x)$  of  $c_i$  and fuzzy values  $W_j(x)$  of  $d_j$  is demonstrated as shown in Table 1.

TABLE I  
DECISION TABLE WITH FUZZY ATTRIBUTES

	$c_1$	$c_2$	...	$c_n$	$d_1$	$d_2$	...	$d_m$
$x_1$	$V_1(x_1)$	$V_2(x_1)$	...	$V_n(x_1)$	$W_1(x_1)$	$W_2(x_1)$	...	$W_m(x_1)$
$x_2$	$V_1(x_2)$	$V_2(x_2)$	...	$V_n(x_2)$	$W_1(x_2)$	$W_2(x_2)$	...	$W_m(x_2)$
...								...
$x_N$	$V_1(x_N)$	$V_2(x_N)$	...	$V_n(x_N)$	$W_1(x_N)$	$W_2(x_N)$	...	$W_m(x_N)$

<sup>1</sup>If the values of all attributes are singletons or disjoint intervals on the original domain of attributes, then we obtain a classical crisp decision table.

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า  
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

## IV. FLOW GRAPHS

A flow graph was introduced by Pawlak. In this section we present basic definitions and some properties of flow graphs from [3], [5], [7], [8], [10], [14].

### A. Normalized Flow Graphs

A flow graph is an alternative representation of knowledge, more specifically the information flow, in the given data. Several studies on traditional (rough set) flow graphs were in [3], [5], [7], [8], [10], [14] and for fuzzy set theory [13], [15]. In order to demonstrate interesting relationships between flow graphs and other disciplines, we consider normalized version of flow graphs (consider  $\sigma \in [0, 1]$  instead of  $R^+$ ).

A *normalized flow graph* is a *directed, acyclic, finite graph*  $G = (N, B, \sigma)$ , where  $N$  is a set of *nodes*,  $B \subseteq N \times N$  is a set of *directed branches* and  $\varphi: B \rightarrow R^+$  is a flow function,  $\varphi(G)$  is a *throughflow* of flow graph  $G$  and  $\sigma: B \rightarrow [0, 1]$  is a *normalized flow* of  $(x, y)$ . The *strength coefficient* of  $(x, y)$  is given by

$$\sigma(x, y) = \frac{\varphi(x, y)}{\varphi(G)},$$

where  $0 \leq \sigma(x, y) \leq 1$ .

The *strength\_of\_branch*  $\times 100$  can be understood as the percentage of a total flow through the branch.

With every node  $x$  of a flow graph  $G$ , *normalized inflow* and *outflow* are defined as:

$$\sigma_+(x) = \frac{\varphi_+(x)}{\varphi(G)} = \sum_{y \in I(x)} \sigma(y, x), \text{ and}$$

$$\sigma_-(x) = \frac{\varphi_-(x)}{\varphi(G)} = \sum_{y \in O(x)} \sigma(x, y).$$

Similarly, *normalized inflow* and *outflow* for the flow graph  $G$  are defined as:

$$\sigma_+(G) = \frac{\varphi_+(G)}{\varphi(G)} = \sum_{x \in I(G)} \sigma_+(x), \text{ and}$$

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า  
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

$$\sigma_-(G) = \frac{\varphi_-(G)}{\varphi(G)} = \sum_{x \in O(x)} \sigma_-(x).$$

For any internal node  $x$ ,

$$\sigma_+(x) = \sigma_-(x) = \sigma(x),$$

$\sigma(x)$  is a *normalized throughflow* of  $x$ .

With every branch  $(x, y)$  of a flow graph  $G$ , the *certainty* and *coverage* coefficients of  $(x, y)$  are defined as:

$$cer(x, y) = \frac{\sigma(x, y)}{\sigma(x)} \text{ and}$$

$$cov(x, y) = \frac{\sigma(x, y)}{\sigma(y)},$$

where  $\sigma(x) \neq 0$  and  $\sigma(y) \neq 0$ .

Interestingly, properties:

$$\sum_{y \in O(x)} cer(x, y) = 1 \text{ and}$$

$$\sum_{y \in I(x)} cov(x, y) = 1,$$

are the form of total probability theorem, and

$$\sigma(x) = \sum_{y \in O(x)} cer(x, y)\sigma(x) = \sum_{y \in O(x)} \sigma(x, y) \text{ and}$$

$$\sigma(x) = \sum_{y \in I(x)} cov(x, y)\sigma(y) = \sum_{y \in I(x)} \sigma(x, y)$$

are Bayes' rules [3], [5], [7], [8], [10], [14]. Moreover,  $cer(x, y) = \frac{cov(x, y)\sigma(y)}{\sigma(x)}$ , and  $cov(x, y) = \frac{cer(x, y)\sigma(x)}{\sigma(y)}$  are holded.

Next, it is important to consider path and connection of flow graphs in some situations.

A (directed) *path* from  $x$  to  $y$  ( $x \neq y$ ) in  $G$ , denoted by  $[x \dots y]$ , is a sequence of nodes  $x_1, \dots, x_n$  such that  $x_1 = x$  and  $x_n = y$  and  $(x_i, x_{i+1}) \in B$  for every  $i, 1 \leq i \leq n-1$ .

The *certainty*, *coverage* and *strength of the path*  $[x_1 \dots x_n]$  are defined as:

$$cer[x_1 \dots x_n] = \prod_{i=1}^{n-1} cer(x_i, x_{i+1}),$$

$$cov[x_1 \dots x_n] = \prod_{i=1}^{n-1} cov(x_i, x_{i+1}) \text{ and}$$

$$\sigma[x \dots y] = \sigma(x)cer[x \dots y] = \sigma(y)cov[x \dots y].$$

The set of all paths from  $x$  to  $y$  ( $x \neq y$ ) in  $G$ , denoted by  $\langle x, y \rangle$ , is a *connection* of  $G$  determined by  $x$  and  $y$ .

Its associated *certainty*, *coverage* and *strength of the connection*  $\langle x, y \rangle$  are defined as:

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า  
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

$$cer \langle x, y \rangle = \sum_{[x \dots y] \in \langle x, y \rangle} cer[x \dots y],$$

$$cov \langle x, y \rangle = \sum_{[x \dots y] \in \langle x, y \rangle} cov[x \dots y] \text{ and}$$

$$\sigma \langle x, y \rangle = \sum_{[x \dots y] \in \langle x, y \rangle} \sigma[x \dots y] = \sigma(x)cer \langle x, y \rangle = \sigma(y)cov \langle x, y \rangle.$$



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า  
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้คัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

## B. Decision Rules, Decision Algorithm and Flow Graph

In this section, basic concepts of decision rules and decision algorithms that can be used as classifiers in logical form [4], [8] are provided .

Let  $S = (U, A)$  be an information system, for any subset  $B$  of  $A$ , there is an associated set of formulas  $For(B)$ .  $For(B)$  is built up from attribute value pairs  $(a, v)$  where  $a \in B$  and  $v \in V_a$ . For any  $\Phi \in For(B)$ , let  $\|\Phi\|$  be the set of all objects from universe  $U$  satisfy  $\Phi$  in  $S$ . By means of logical connectives AND ( $\wedge$ ), OR ( $\vee$ ) and NOT ( $\neg$ ), Pawlak [4] defined some inductively properties as follows:

$$\begin{aligned}\|\Phi \vee \Psi\| &= \|\Phi\| \cup \|\Psi\|, \\ \|\Phi \wedge \Psi\| &= \|\Phi\| \cap \|\Psi\| \text{ and} \\ \|\neg\Phi\| &= U - \|\Phi\|.\end{aligned}$$

A decision rule in  $S$  is an expression in the form  $\Phi \rightarrow \Psi$ ,  $\Phi$  and  $\Psi$  are referred to as a *condition* and *decision* of the rule, respectively. For every decision rules, there are three associated coefficients:

$$\begin{aligned}\sigma(\Phi, \Psi) &= \frac{\text{card}(\|\Phi\| \cap \|\Psi\|)}{\text{card}(U)}, \\ \text{cer}(\Phi, \Psi) &= \frac{\text{card}(\|\Phi\| \cap \|\Psi\|)}{\text{card}(\|\Phi\|)} \text{ and} \\ \text{cov}(\Phi, \Psi) &= \frac{\text{card}(\|\Phi\| \cap \|\Psi\|)}{\text{card}(\|\Psi\|)} [3].\end{aligned}$$

A set of all decision rules  $Dec(\Phi, \Psi) = \{\Phi_i \rightarrow \Psi_i\}_{i=1}^n, n \geq 2$ , will be called a *decision algorithm* if all decision rules are satisfied conditions:

- (1) *Mutually exclusive*: for every  $\Phi_i \rightarrow \Psi_i$  and  $\Phi_j \rightarrow \Psi_j$ ,  $\|\Phi_i\| \cap \|\Phi_j\| = \emptyset$  and  $\|\Psi_i\| \cap \|\Psi_j\| = \emptyset$  for  $i \neq j$ ,
- (2) *Admissible*:  $\|\Phi \wedge \Psi\| \neq \emptyset$ ,
- (3) *Covering*:  $\bigcup_{i=1}^n \|\Phi_i\| = U$  and  $\bigcup_{i=1}^n \|\Psi_i\| = U$ .

A flow graph can be interpreted as decision algorithms [3], [10] and will be described as

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า  
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

follows. Let us assume that set of nodes of a flow graph is interpreted as a set of *formulas*, denoted by  $\Phi, \Psi$ <sup>2</sup>. Then every branch  $(\Phi, \Psi)$  can be understood as a *decision rule*  $\Phi \rightarrow \Psi$ .  $\Phi$  will be referred to as a *condition*, whereas  $\Psi$  is *decision*. Such a rule is characterized by three coefficients strength, certainty and coverage,

$$\begin{aligned}\sigma(\Phi, \Psi) &= \frac{\text{card}(\|\Phi\| \cap \|\Psi\|)}{\text{card}(U)}, \\ \text{cer}(\Phi, \Psi) &= \frac{\text{card}(\|\Phi\| \cap \|\Psi\|)}{\text{card}(\|\Phi\|)} \text{ and} \\ \text{cov}(\Phi, \Psi) &= \frac{\text{card}(\|\Phi\| \cap \|\Psi\|)}{\text{card}(\|\Psi\|)} \quad [3].\end{aligned}$$

Next, a *path*,  $[\Phi_1 \dots \Phi_n]$ , is a sequence of decision rules  $\Phi_1 \rightarrow \Phi_2, \Phi_2 \rightarrow \Phi_3, \dots, \Phi_{n-1} \rightarrow \Phi_n$ . As shown in [3], this sequence of decision rules can be interpreted as a single decision rule  $\Phi_1 \Phi_2 \dots \Phi_{n-1} \rightarrow \Phi_n$ , in short  $\Phi^* \rightarrow \Phi_n$ , where  $\Phi^* = \Phi_1 \wedge \Phi_2 \wedge \dots \wedge \Phi_{n-1}$ . It is characterized by three coefficients:

$$\begin{aligned}\text{cer}(\Phi^*, \Phi_n) &= \text{cer}[\Phi_1 \dots \Phi_n], \\ \text{cov}(\Phi^*, \Phi_n) &= \text{cov}[\Phi_1 \dots \Phi_n] \text{ and} \\ \sigma(\Phi^*, \Phi_n) &= \sigma(\Phi_1) \text{cer}[\Phi_1 \dots \Phi_n] = \sigma(\Phi_n) \text{cov}[\Phi_1 \dots \Phi_n],\end{aligned}$$

<sup>2</sup>The formulas can be viewed as *propositional functions* in logic [3].

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า  
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

### C. New Certainty and Coverage Coefficients for Decision Rules from Relaxed Mutually Exclusive

In this section, we apply concept of mutually exclusive in decision algorithm introduced by Pawlak [3], [4] to find new certainty and coverage coefficients.

Let  $\|\Phi_i\|$  and  $\|\Psi_i\|$ ,  $i = 1, 2, \dots, n$  be sets of all objects from universe  $U$  which satisfy these condition and decision attributes. If we want to connect any set of conditions (set of decisions) by logical connectives AND ( $\wedge$ ) or OR ( $\vee$ ) and find certainty and coverage of these decision rules.

Then by Pawlak's definitions of certainty and coverage, the following Remarks 1 and 2 are holded.

*Remark 1:* Let  $\Phi$  and  $\Psi_i$ ,  $i = 1, 2, \dots, n$  be the set of condition and decision attributes, then

$$\begin{aligned} 1) \text{cer}(\Phi, \Psi_1 \vee \dots \vee \Psi_n) &= \frac{\text{card}(\|\Phi\| \cap (\|\Psi_1\| \cup \dots \cup \|\Psi_n\|))}{\text{card}(\|\Phi\|)}, \\ 2) \text{cer}(\Phi, \Psi_1 \wedge \dots \wedge \Psi_n) &= \frac{\text{card}(\|\Phi\| \cap (\|\Psi_1\| \cap \dots \cap \|\Psi_n\|))}{\text{card}(\|\Phi\|)}, \\ 3) \text{cov}(\Phi, \Psi_1 \vee \dots \vee \Psi_n) &= \frac{\text{card}(\|\Phi\| \cap (\|\Psi_1\| \cup \dots \cup \|\Psi_n\|))}{\text{card}(\|\Psi_1\| \cup \dots \cup \|\Psi_n\|)}, \\ 4) \text{cov}(\Phi, \Psi_1 \wedge \dots \wedge \Psi_n) &= \frac{\text{card}(\|\Phi\| \cap (\|\Psi_1\| \cap \dots \cap \|\Psi_n\|))}{\text{card}(\|\Psi_1\| \cap \dots \cap \|\Psi_n\|)}. \end{aligned}$$

*Remark 2:* Let  $\Phi_i$  and  $\Psi$ ,  $i = 1, 2, \dots, n$  be the set of condition and decision attributes, then

$$\begin{aligned} 1) \text{cov}(\Phi_1 \vee \dots \vee \Phi_n, \Psi) &= \frac{\text{card}((\|\Phi_1\| \cup \dots \cup \|\Phi_n\|) \cap \|\Psi\|)}{\text{card}(\|\Psi\|)}, \\ 2) \text{cov}(\Phi_1 \wedge \dots \wedge \Phi_n, \Psi) &= \frac{\text{card}((\|\Phi_1\| \cap \dots \cap \|\Phi_n\|) \cap \|\Psi\|)}{\text{card}(\|\Psi\|)}, \\ 3) \text{cer}(\Phi_1 \vee \dots \vee \Phi_n, \Psi) &= \frac{\text{card}((\|\Phi_1\| \cup \dots \cup \|\Phi_n\|) \cap \|\Psi\|)}{\text{card}(\|\Phi_1\| \cup \dots \cup \|\Phi_n\|)}, \\ 4) \text{cer}(\Phi_1 \wedge \dots \wedge \Phi_n, \Psi) &= \frac{\text{card}((\|\Phi_1\| \cap \dots \cap \|\Phi_n\|) \cap \|\Psi\|)}{\text{card}(\|\Phi_1\| \cap \dots \cap \|\Phi_n\|)}. \end{aligned}$$

Remark 3 is the property obtained from set theory.

*Remark 3:* Let  $\Phi_i$ ,  $i = 1, 2, \dots, n$  be a set of attribute, then

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า  
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

$$\begin{aligned}
& \text{card}(\|\Phi_1\| \cup \|\Phi_2\| \cup \dots \cup \|\Phi_n\|) \\
&= \sum_{i=1}^n \text{card}(\|\Phi_i\|) - \sum_{i_1=1}^{n-1} \sum_{i_2=i_1+1}^n \text{card}(\|\Phi_{i_1}\| \cap \|\Phi_{i_2}\|) \\
&+ \sum_{i_1=1}^{n-2} \sum_{i_2=i_1+1}^{n-1} \sum_{i_3=i_2+1}^n \text{card}(\|\Phi_{i_1}\| \cap \|\Phi_{i_2}\| \cap \|\Phi_{i_3}\|) \\
&- \dots + (-1)^{n-1} \text{card}(\|\Phi_{i_1}\| \cap \|\Phi_{i_2}\| \cap \dots \cap \|\Phi_n\|).
\end{aligned}$$

From mutual exclusive of decision algorithm, there are assumptions that for every decision rules  $\Phi_i \rightarrow \Psi_i$  and  $\Phi_j \rightarrow \Psi_j$ ,

$$\|\Phi_i\| \cap \|\Phi_j\| = \emptyset \text{ and } \|\Psi_i\| \cap \|\Psi_j\| = \emptyset$$

for  $i \neq j$ .

In this research, we propose relaxation of this mutual exclusion property as follows.

There exist at least two decision rules  $\Phi_i \rightarrow \Psi$  and  $\Phi \rightarrow \Psi_i$  such that

$$\|\Phi_{i_m}\| \cap \|\Phi_{i_k}\| \neq \emptyset \text{ or } \|\Psi_{i_m}\| \cap \|\Psi_{j_k}\| \neq \emptyset$$

for  $i_m \neq j_k$ .

We can give a brief explanation of the usefulness of this relax mutual exclusion as follows.

For example, if we have decision rules,

IF *score\_of\_Algebra* is high THEN *total\_score* is high and

IF *score\_of\_Number\_Theory* is high THEN *total\_score* is high,

in a decision algorithm, then there may be some elements satisfy both conditions ( $\|\Phi_{i_m}\| \cap \|\Phi_{i_k}\| \neq \emptyset$ ).

Thus, our proposed certainty and coverage coefficients are more flexible with the relaxed mutually exclusive.

If there exist at least two decisions  $\|\Psi_i\|$  and  $\|\Psi_j\|$  such that  $\|\Psi_i\| \cap \|\Psi_j\| \neq \emptyset$ , we can relax mutually exclusive for decisions (as in Propositions 1, 2).

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า  
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

If there exist at least two conditions  $\|\Phi_i\|$  and  $\|\Phi_j\|$  such that  $\|\Phi_i\| \cap \|\Phi_j\| \neq \emptyset$ , we can relax mutually exclusive for conditions (as in Propositions 3, 4).



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า  
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

*Proposition 1:* If we are given a set of decision rules,  $\{\Phi \rightarrow \Psi_i\}_n$ , such that  $\Psi_{i_m} \cap \Psi_{i_k} \neq \emptyset$ , for any  $i_m \neq i_k$ . Then  $cer(\Phi, \Psi_1 \vee \Psi_2 \vee \dots \vee \Psi_n)$

$$\begin{aligned} &= \sum_{i=1}^n cer(\Phi, \Psi_i) - \sum_{i_1=1}^{n-1} \sum_{i_2=i_1+1}^n cer(\Phi, \Psi_{i_1} \wedge \Psi_{i_2}) \\ &+ \sum_{i_1=1}^{n-2} \sum_{i_2=i_1+1}^{n-1} \sum_{i_3=i_2+1}^n cer(\Phi, \Psi_{i_1} \wedge \Psi_{i_2} \wedge \Psi_{i_3}) - \\ &\dots + (-1)^{n-1} cer(\Phi, \Psi_1 \wedge \Psi_2 \wedge \dots \wedge \Psi_n). \end{aligned}$$

*Proof:* If we denote  $cer(\Phi, \Psi_1 \vee \Psi_2 \vee \dots \vee \Psi_n)$  by  $cer(\Phi, \Psi_i)$ , then by Remark 1(1) we have  $cer(\Phi, \Psi_i) = \frac{card(\|\Phi\| \cap (\|\Psi_1\| \cup \|\Psi_2\| \cup \dots \cup \|\Psi_n\|))}{card(\|\Phi\|)}$ . Next, by left-distributive over union, we obtain

$$\begin{aligned} cer(\Phi, \Psi_i) &= \frac{card(\|(\|\Phi\| \cap \|\Psi_1\|) \cup (\|\Phi\| \cap \|\Psi_2\|) \cup \dots \cup (\|\Phi\| \cap \|\Psi_n\|))}{card(\|\Phi\|)} \\ &= \sum_{i=1}^n \frac{card(\|\Phi\| \cap \|\Psi_i\|)}{card(\|\Phi\|)} \\ &- \sum_{i_1=1}^{n-1} \sum_{i_2=i_1+1}^n \frac{card(\|\Phi\| \cap \|\Psi_{i_1}\| \cap \|\Psi_{i_2}\|)}{card(\|\Phi\|)} \\ &+ \sum_{i_1=1}^{n-2} \sum_{i_2=i_1+1}^{n-1} \sum_{i_3=i_2+1}^n \frac{card(\|\Phi\| \cap \|\Psi_{i_1}\| \cap \|\Psi_{i_2}\| \cap \|\Psi_{i_3}\|)}{card(\|\Phi\|)} \\ &- \dots + (-1)^{n-1} \frac{card(\|\Phi\| \cap \|\Psi_1\| \cap \|\Psi_2\| \cap \dots \cap \|\Psi_n\|)}{card(\|\Phi\|)} \text{ (by Remark 3)} \\ &= \sum_{i=1}^n cer(\Phi, \Psi_i) - \sum_{i_1=1}^{n-1} \sum_{i_2=i_1+1}^n cer(\Phi, \Psi_{i_1} \wedge \Psi_{i_2}) \\ &+ \sum_{i_1=1}^{n-2} \sum_{i_2=i_1+1}^{n-1} \sum_{i_3=i_2+1}^n cer(\Phi, \Psi_{i_1} \wedge \Psi_{i_2} \wedge \Psi_{i_3}) \\ &- \dots + (-1)^{n-1} cer(\Phi, \Psi_1 \wedge \Psi_2 \wedge \dots \wedge \Psi_n) \text{ (by Remark 1(2)).} \end{aligned}$$

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า  
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

**Proposition 2:** If we are given a set of decision rules,  $\{\Phi \rightarrow \Psi_i\}_n$ , such that  $\Psi_{i_m} \cap \Psi_{i_k} \neq \emptyset$ , for any  $i_m \neq i_k$ . Then  $\text{cov}(\Phi, \Psi_1 \vee \Psi_2 \vee \dots \vee \Psi_n) < \sum_{i=1}^n \text{cov}(\Phi, \Psi_i)$ .

*Proof:* If we denote  $\text{cov}(\Phi, \Psi_1 \vee \Psi_2 \vee \dots \vee \Psi_n)$  by  $\text{cov}(\Phi, \Psi_i)$ , then by Remark 1(3) we have  $\text{cov}(\Phi, \Psi_i) = \frac{\text{card}(\|\Phi\| \cap (\|\Psi_1\| \cup \|\Psi_2\| \cup \dots \cup \|\Psi_n\|))}{\text{card}(\|\Psi_1\| \cup \|\Psi_2\| \cup \dots \cup \|\Psi_n\|)}$ . Next, by left-distributive over union and inequality property, we obtain

$$\begin{aligned} \text{cov}(\Phi, \Psi_i) &= \frac{\text{card}(\|\Phi\| \cap (\|\Psi_1\| \cup \|\Psi_2\| \cup \dots \cup \|\Psi_n\|))}{\text{card}(\|\Psi_1\| \cup \|\Psi_2\| \cup \dots \cup \|\Psi_n\|)} \\ &< \sum_{i=1}^n \frac{\text{card}(\|\Phi\| \cap \|\Psi_i\|)}{\text{card}(\|\Psi_1\| \cup \|\Psi_2\| \cup \dots \cup \|\Psi_n\|)} \\ &< \sum_{i=1}^n \frac{\text{card}(\|\Phi\| \cap \|\Psi_i\|)}{\text{card}(\|\Psi_i\|)}. \end{aligned}$$

Then  $\text{cov}(\Phi, \Psi_i) < \sum_{i=1}^n \text{cov}(\Phi, \Psi_i)$  (by Remark 1(3)). ■



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า  
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

*Proposition 3:* If we are given a set of decision rules,  $\{\Phi_i \rightarrow \Psi\}_n$ , such that  $\Phi_{i_m} \cap \Phi_{i_k} \neq \emptyset$ , for any  $i_m \neq i_k$ . Then  $cov(\Phi_1 \vee \Phi_2 \vee \dots \vee \Phi_n, \Psi)$

$$= \sum_{i=1}^n cov(\Phi_i, \Psi) - \sum_{i_1=1}^{n-1} \sum_{i_2=i_1+1}^n cov(\Phi_{i_1} \wedge \Phi_{i_2}, \Psi) + \sum_{i_1=1}^{n-2} \sum_{i_2=i_1+1}^{n-1} \sum_{i_3=i_2+1}^n cov(\Phi_{i_1} \wedge \Phi_{i_2} \wedge \Phi_{i_3}, \Psi) - \dots + (-1)^{n-1} cov(\Phi_1 \wedge \Phi_2 \wedge \dots \wedge \Phi_n, \Psi).$$

*Proof:* If we denote  $cov(\Phi_1 \vee \Phi_2 \vee \dots \vee \Phi_n, \Psi)$  by  $cov(\Phi_i, \Psi)$ , then by Remark 2(1) we have  $cov(\Phi_i, \Psi) = \frac{card((\|\Phi_1\| \cup \|\Phi_2\| \cup \dots \cup \|\Phi_n\|) \cap \|\Psi\|)}{card(\|\Psi\|)}$ . Next, by right-distributive over union, we obtain

$$\begin{aligned} cov(\Phi_i, \Psi) &= \frac{card((\|\Phi_1\| \cap \|\Psi\|) \cup (\|\Phi_2\| \cap \|\Psi\|) \cup \dots \cup (\|\Phi_n\| \cap \|\Psi\|))}{card(\|\Psi\|)} \\ &= \sum_{i=1}^n \frac{card(\|\Phi_i\| \cap \|\Psi\|)}{card(\|\Psi\|)} \\ &\quad - \sum_{i_1=1}^{n-1} \sum_{i_2=i_1+1}^n \frac{card(\|\Phi_{i_1}\| \cap \|\Phi_{i_2}\| \cap \|\Psi\|)}{card(\|\Psi\|)} \\ &\quad + \sum_{i_1=1}^{n-2} \sum_{i_2=i_1+1}^{n-1} \sum_{i_3=i_2+1}^n \frac{card(\|\Phi_{i_1}\| \cap \|\Phi_{i_2}\| \cap \|\Phi_{i_3}\| \cap \|\Psi\|)}{card(\|\Psi\|)} \\ &\quad - \dots + (-1)^{n-1} \frac{card(\|\Phi_1\| \cap \|\Phi_2\| \cap \dots \cap \|\Phi_n\| \cap \|\Psi\|)}{card(\|\Psi\|)} \text{ (by Remark 3)} \\ &= \sum_{i=1}^n cov(\Phi_i, \Psi) - \sum_{i_1=1}^{n-1} \sum_{i_2=i_1+1}^n cov(\Phi_{i_1} \wedge \Phi_{i_2}, \Psi) + \sum_{i_1=1}^{n-2} \sum_{i_2=i_1+1}^{n-1} \sum_{i_3=i_2+1}^n cov(\Phi_{i_1} \wedge \Phi_{i_2} \wedge \Phi_{i_3}, \Psi) - \dots \\ &\quad + (-1)^{n-1} cov(\Phi_1 \wedge \Phi_2 \wedge \dots \wedge \Phi_n, \Psi) \text{ (by Remark 2(1)).} \end{aligned}$$

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

*Proposition 4:* If we are given a set of decision rules,  $\{\Phi_i \rightarrow \Psi\}_n$ , such that  $\Phi_{i_m} \cap \Phi_{i_k} \neq \emptyset$ , for any  $i_m \neq i_k$ . Then  $cer(\Phi_1 \vee \Phi_2 \vee \dots \vee \Phi_n, \Psi) < \sum_{i=1}^n cer(\Phi_i, \Psi)$ .

*Proof:* If we denote  $cer(\Phi_1 \vee \Phi_2 \vee \dots \vee \Phi_n, \Psi)$  by  $cer(\Phi_i, \Psi)$ , then by Remark 2(3) we have  $cer(\Phi_i, \Psi) = \frac{card((\Phi_1 \cup \Phi_2 \cup \dots \cup \Phi_n) \cap \Psi)}{card(\Phi_1 \cup \Phi_2 \cup \dots \cup \Phi_n)}$ . Next, by right-distributive over union and inequality property, we obtain

$$\begin{aligned} cer(\Phi_i, \Psi) &= \frac{card((\|\Phi_1\| \cap \|\Psi\|) \cup (\|\Phi_2\| \cap \|\Psi\|) \cup \dots \cup (\|\Phi_n\| \cap \|\Psi\|))}{card(\|\Phi_1\| \cup \|\Phi_2\| \cup \dots \cup \|\Phi_n\|)} \\ &< \sum_{i=1}^n \frac{card(\|\Phi_i\| \cap \|\Psi\|)}{card(\|\Phi_1\| \cup \|\Phi_2\| \cup \dots \cup \|\Phi_n\|)} \\ &< \sum_{i=1}^n \frac{card(\|\Phi_i\| \cap \|\Psi\|)}{card(\|\Phi_i\|)}. \end{aligned}$$

Then  $cer(\Phi_i, \Psi) < \sum_{i=1}^n cer(\Phi_i, \Psi)$  (by Remark 2(1)).



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า  
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

#### D. Flow Graphs with Fuzzy Attributes

Flow graphs with fuzzy attributes have several advantages (e.g., represent data in form of fuzzy attributes) [13]. In further discussions, we apply flow graphs to represent and analyze decision table with fuzzy attributes in the form introduced in [13].

All possible decision rules with fuzzy attributes can be generated by using the *Cartesian product* of the set of attribute values with  $r = \prod_{i=1}^n n_i \prod_{j=1}^m m_j$  as a possible number of rules.

Let  $R_k$  be the  $k$ -th decision rule. It can be expressed by

$R_k$ : IF  $c_1$  is  $V_1^k$ ,  $c_2$  is  $V_2^k$ , ... and  $c_n$  is  $V_n^k$ ,  
THEN  $d_1$  is  $W_1^k$ ,  $d_2$  is  $W_2^k$ , ... and  $d_m$  is  $W_m^k$ ,

where  $k = 1, 2, \dots, r$ , and  $V_i^k \in V_i$ ,  $i = 1, 2, \dots, n$ ,  $W_j^k \in W_j$ ,  $j = 1, 2, \dots, m$ .

Thus, Cartesian products  $C^k = V_1^k \times V_2^k \times \dots \times V_n^k$  and  $D^k = W_1^k \times W_2^k \times \dots \times W_m^k$  are the  $k$ -th decision rules expressed in the form of fuzzy implication,  $C^k \rightarrow D^k$ .

Let  $cda(x, k)$  and  $cdc(x, k)$  be the confirmation degrees of the antecedent and consequent of decision rules and  $cd(x, k)$  be the confirmation degrees of the  $k$ -th decision rules by the element  $x \in U$ .

They are defined respectively as

$$cda(x, k) = T(\mu_{V_1^k}(x), \mu_{V_2^k}(x), \dots, \mu_{V_n^k}(x)) \quad (1)$$

$$cdc(x, k) = T(\mu_{W_1^k}(x), \mu_{W_2^k}(x), \dots, \mu_{W_m^k}(x)) \quad (2)$$

$$cd(x, k) = T(cda(x, k), cdc(x, k)). \quad (3)$$

By using (1), (2) and (3) in fuzzy set domain, the *support (sup)* of decision rule's antecedent, consequent and itself are defined respectively as

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า  
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

$$\begin{aligned} \text{sup}(cda(x, k)) &= \{cda(x_1, k)/x_1, \dots, cda(x_N, k)/x_N\}, \\ \text{sup}(cdc(x, k)) &= \{cdc(x_1, k)/x_1, \dots, cdc(x_N, k)/x_N\} \\ \text{and } \text{sup}(R_k) &= \{cd(x_1, k)/x_1, \dots, cd(x_N, k)/x_N\}. \end{aligned}$$

Moreover, when consider the fuzzy universe, we can define the *throughflow*,  $\varphi(x, y)$  of the branch  $(x, y)$  in  $G$  to the cardinality of the product of fuzzy sets,  $\tilde{X}$  and  $\tilde{Y}$ . The *strength* of  $(x, y)$  is thus given by

$$\sigma(x, y) = \frac{\varphi(x, y)}{\varphi(G)}. \quad (4)$$

With every node  $x$  of a flow graph  $G$ , associated *normalized inflow* and *outflow* are defined respectively as

$$\varphi_-(x) = \text{power}\tilde{x} = \sum_{y \in O(x)} \varphi(x, y) = \sum_{y \in O(x)} \text{power}(\tilde{x} \cap \tilde{y}) \quad (5)$$

$$\varphi_+(x) = \text{power}\tilde{x} = \sum_{y \in I(x)} \varphi(x, y) = \sum_{y \in I(x)} \text{power}(\tilde{x} \cap \tilde{y}) \quad (6)$$

where  $\text{power}\tilde{x}$  denotes fuzzy cardinality. For any internal node  $x$ , it holds that  $\varphi_-(x) = \varphi_+(x) = \varphi(x)$ .

A (directed) *path* from  $x$  to  $y$  ( $x \neq y$ ) in  $G$ , denoted by  $[x \dots y]$ , is a sequence of nodes  $x_1, \dots, x_n$  such that  $x_1 = x$  and  $x_n = y$  and  $(x_i, x_{i+1}) \in B$ ,  $i, 1 \leq i \leq n-1$ .

We obtain the *certainty*, *coverage* and *strength* coefficients of the path  $[x_1 \dots x_n]$  [13], [15] as:

$$\begin{aligned} \text{cer}[x_1 \dots x_n] &= \prod_{i=1}^{n-1} \text{cer}(x_1 \dots x_i, x_{i+1}), \text{ where } \text{cer}(x_1 \dots x_i, x_{i+1}) = \frac{\text{power}(\tilde{x}_1 \cap \tilde{x}_2 \cap \dots \cap \tilde{x}_{i+1})}{\text{power}(\tilde{x}_1 \cap \tilde{x}_2 \cap \dots \cap \tilde{x}_i)}, \\ \text{cov}[x_1 \dots x_n] &= \prod_{i=1}^{n-1} \text{cov}(x_1 \dots x_i, x_{i+1}), \text{ where } \text{cov}(x_1 \dots x_i, x_{i+1}) = \frac{\text{power}(\tilde{x}_1 \cap \tilde{x}_2 \cap \dots \cap \tilde{x}_{i+1})}{\text{power}(x_{i+1})}, \text{ and} \\ \sigma[x_1 \dots x_n] &= \sigma(x_1) \text{cer}[x_1 \dots x_n] = \sigma(x_n) \text{cov}[x_1 \dots x_n]. \end{aligned}$$

The decision rule  $R_k, \Phi^* \rightarrow \Phi_n$ , where  $\Phi^*$  is condition and  $\Phi_n$  is decision can be characterized by three coefficients:

$$\text{cer}(\Phi^*, \Phi_n) = \text{cer}(R_k) = \frac{\text{power}(\text{sup}(R_k))}{\text{power}(\text{sup}(cda(x, k)))} \quad (7)$$

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่นอนุญาตให้นำไปใช้ประโยชน์ด้านการค้า  
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

$$\text{cov}(\Phi^*, \Phi_n) = \text{cov}(R_k) = \frac{\text{power}(\text{sup}(R_k))}{\text{power}(\text{sup}(\text{cdc}(x, k)))} \quad (8)$$

$$\sigma(\Phi^*, \Phi_n) = \text{strength}(R_k) = \frac{\text{power}(\text{sup}(R_k))}{\text{card}(U)}. \quad (9)$$



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า  
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

## V. EXPERIMENTAL RESULTS AND DISCUSSION

In this section, we describe experimental results from the scores of Olympiad mathematics' students' analysis. Data contains 75 records collected from POSN.

In this research, we study decision rules between condition attributes and decision attribute as the following.

Condition	Symbol	Condition	Symbol
score of Geometry	<i>score_Geo</i>	preference of Algebra	<i>Pref_Alg</i>
score of Number theory	<i>score_Num</i>	preference of Combinatorics	<i>Pref_Com</i>
score of Algebra	<i>score_Alg</i>	preference of Inequality	<i>Pref_I.E.</i>
score of Combinatorics	<i>score_Com</i>	practiced of Geometry	<i>Prac_Geo</i>
score of Inequality	<i>score_I.E.</i>	practiced of Number theory	<i>Prac_Num</i>
preference of Geometry	<i>Pref_Geo</i>	practiced of Algebra	<i>Prac_Alg</i>
preference of Number theory	<i>Pref_Num</i>	practiced of Combinatorics	<i>Prac_Com</i>
practiced of Inequality	<i>Prac_I.E.</i>		
Decision	Symbol		
student total score	<i>total_score</i>		

Next, in Cases 1–2, we consider attributes in POSN data set as descriptive attributes. Possible values of first five condition attributes and decision attribute are low score and high score.

Several case studies are then presented according to the combination of these attributes. The threshold for decision rule extraction is  $cer > 0.6$ .

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า  
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

**Case 1.** Decision rules between score of any subject and total score. The selected decision rules are:

IF *Score\_Num* is low THEN *total\_score* is low,  $cer = 0.867$ ,  $cov = 0.812$ ,<sup>3</sup>

IF *Score\_Geo* is low THEN *total\_score* is low,  $cer = 0.792$ ,  $cov = 0.775$ ,

⋮

IF *Score\_I.E.* is high THEN *total\_score* is high,  $cer = 0.956$ ,  $cov = 0.814$ ,

with average  $cer = 0.791$  and average  $cov = 0.731$ .



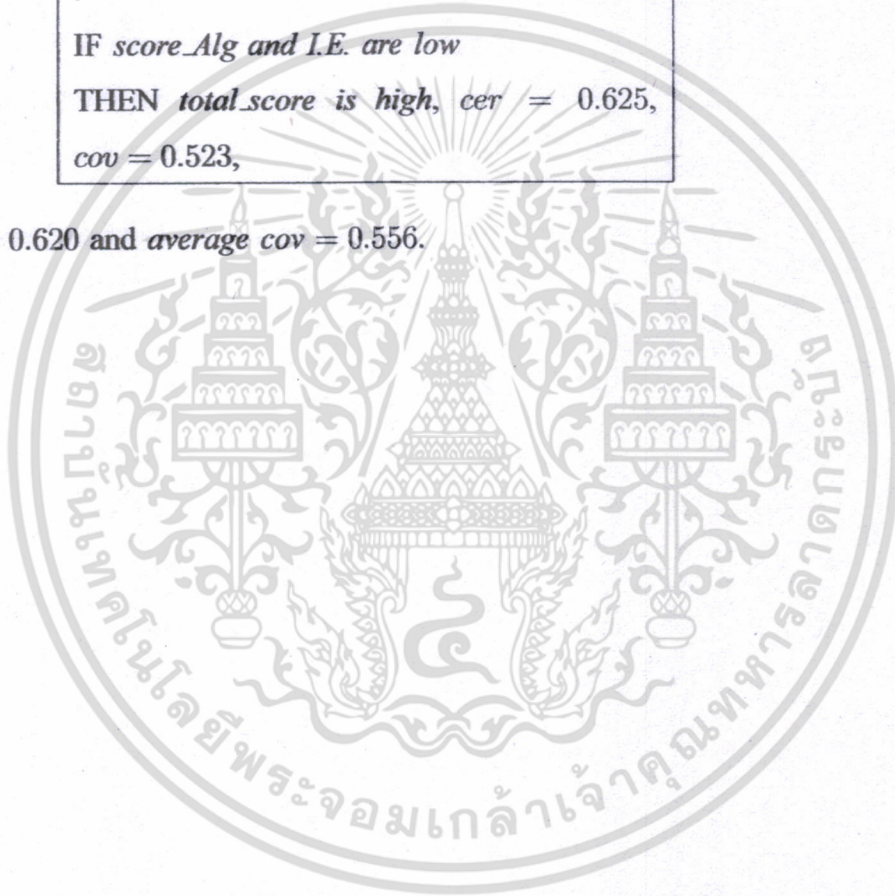
<sup>3</sup>Due to round off errors in computations, the equation may be not satisfied exactly in our examples.

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า  
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

*Case 2.* Decision rules between score of any two subjects and total score. Then the selected decision rules are:

IF *score\_Geo* and *I.E.* are low  
 THEN *total\_score* is low, *cer* = 0.633, *cov* = 0.581,  
 IF *score\_Geo* and *Alg* are low  
 THEN *total\_score* is low, *cer* = 0.622, *cov* = 0.571,  
 ∴  
 IF *score\_Alg* and *I.E.* are low  
 THEN *total\_score* is high, *cer* = 0.625, *cov* = 0.523,

with *average cer* = 0.620 and *average cov* = 0.556.



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า  
 ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

From Cases 1–2, we can summarize that,

*if score of one or two subjects are low, then total score tends to be low too and*

*if score of one or two subjects are high, then total score tends to be high too.*

We see that decision rules obtain from Cases 1–2 depend on only two possible values of attributes (high and low). Thus, we next consider cases that scores are high and low with their degrees which are fuzzy attributes.

Next, in Cases 3–8, we transform POSN data set to be in fuzzy attributes form by triangular membership function. Several case studies are then presented according to the combination of these attributes. The threshold for decision rule extraction is  $cer > 0.6$ .



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า  
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

**Case 3.** Decision rules between score of any subject and total score. Then the selected decision rule is:

IF *Score\_Num* is *high* THEN *total\_score* is *high*,  
 $cer = 0.627$ ,  $cov = 0.504$

with  $cer = 0.627$  and  $cov = 0.504$ .

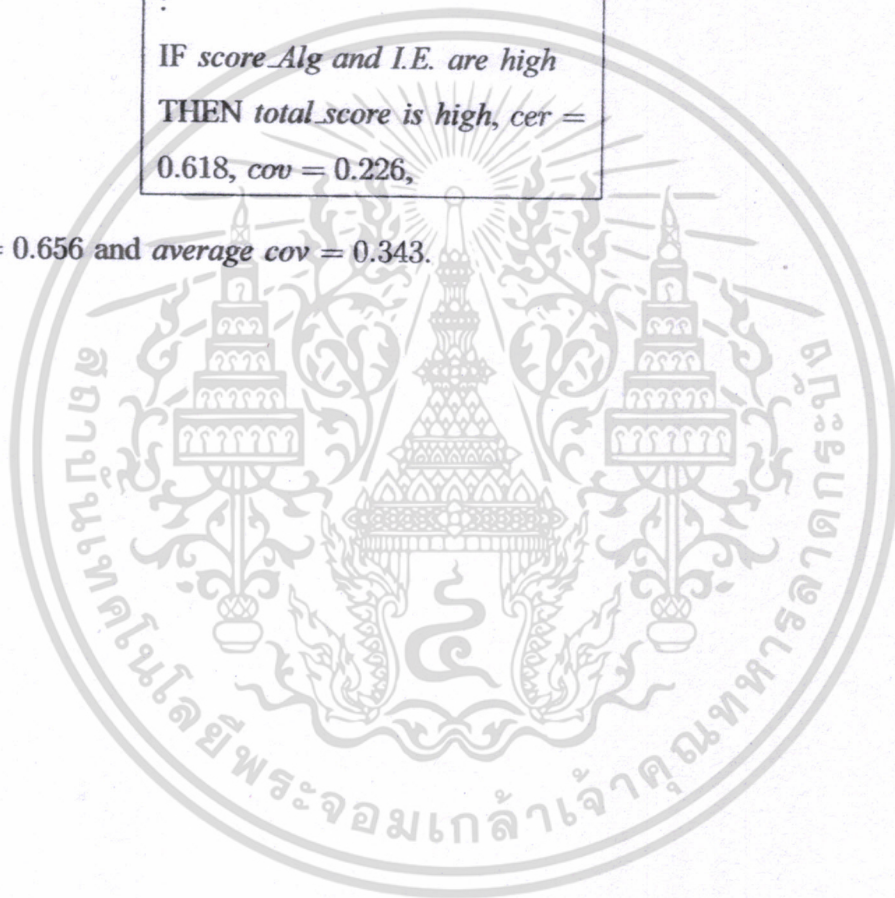


เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า  
 ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

**Case 4.** Decision rules between score of any two subjects and total score. Then the selected decision rules are:

IF *score\_Geo* and *I.E.* are low  
 THEN *total\_score* is low, *cer* =  
 0.651, *cov* = 0.409,  
 IF *score\_Geo* and *I.E.* are high  
 THEN *total\_score* is high, *cer* =  
 0.68, *cov* = 0.301,  
 ∴  
 IF *score\_Alg* and *I.E.* are high  
 THEN *total\_score* is high, *cer* =  
 0.618, *cov* = 0.226,

with *average cer* = 0.656 and *average cov* = 0.343.



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า  
 ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

*Case 5.* Decision rules between score of any five subjects and total score. Then the selected decision rules are:

IF *score\_Geo, Num, Alg, Com*  
and *score\_I.E* are low  
THEN *total\_score* is low, *cer* =  
0.782, *cov* = 0.139,  
IF *score\_Geo, Num, Alg, and*  
*Com* are low, *I.E* is high  
THEN *total\_score* is low, *cer* =  
0.645, *cov* = 0.09,  
:  
IF *score\_Geo, Num, Alg, Comb*  
and *I.E.* are high  
THEN *total\_score* is high, *cer* =  
0.81, *cov* = 0.024,

with *average cer* = 0.691 and *average cov* = 0.047.

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า  
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

**Case 6.** Decision rules between practiced and preferred subjects, score of practiced and preferred subject and total score. The selected decision rules are:

IF *Pref* and *Prac\_Geo* and  
*score\_Geo* is low

THEN *total\_score* is low, *cer* =  
0.638, *cov* = 0.719,

IF *Pref* and *Prac\_Num* and  
*score\_Num* is high

THEN *total\_score* is high, *cer* =  
0.834, *cov* = 0.6,

with *average cer* = 0.736 and *average cov* = 0.660.



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า  
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

**Case 7.** Decision rules between practiced subjects, score of practiced subjects and total score.

The selected decision rules are:

IF *Prac\_Geo and score\_Geo is low*

THEN *total\_score is low, cer = 0.602, cov = 0.731,*

IF *Prac I.E. and score I.E. is low*

THEN *total\_score is low, cer = 0.665, cov = 0.623,*

with *average cer = 0.634 and average cov = 0.677.*



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า  
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

**Case 8.** Decision rules between preferred subjects, score of preferred subjects and total score.

The selected decision rules are:

IF *Pref\_Geo* and *score\_Geo* is  
*high*

*total\_score* is *high*, *cer* = 0.671,

*cov* = 0.541,

IF *Pref\_Num* and *score\_Num* is  
*high*

THEN *total\_score* is *high*, *cer* =

0.633, *cov* = 0.602,

IF *Pref\_I.E.* and *score\_I.E.* is *low*

THEN *total\_score* is *low*, *cer* =

0.829, *cov* = 0.72,

IF *Pref\_I.E.* and *score\_I.E.* is *high*

THEN *total\_score* is *low*, *cer* =

0.831, *cov* = 0.28,

with *average cer* = 0.741 and *average cov* = 0.536.

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า  
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

From these selected decision rules, we can distil useful knowledge from this flow graph as follows.

Number theory affects total score substantially. Because in Case 3, 62.7% of students with high score of Number theory also have high total score. Moreover, 50.4% of students which have high total score are from the group in which Number theory score is high. Obviously, students who get high score from any two (or more) subjects will successively get high total score as shown in Cases 4 and 5.



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า  
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

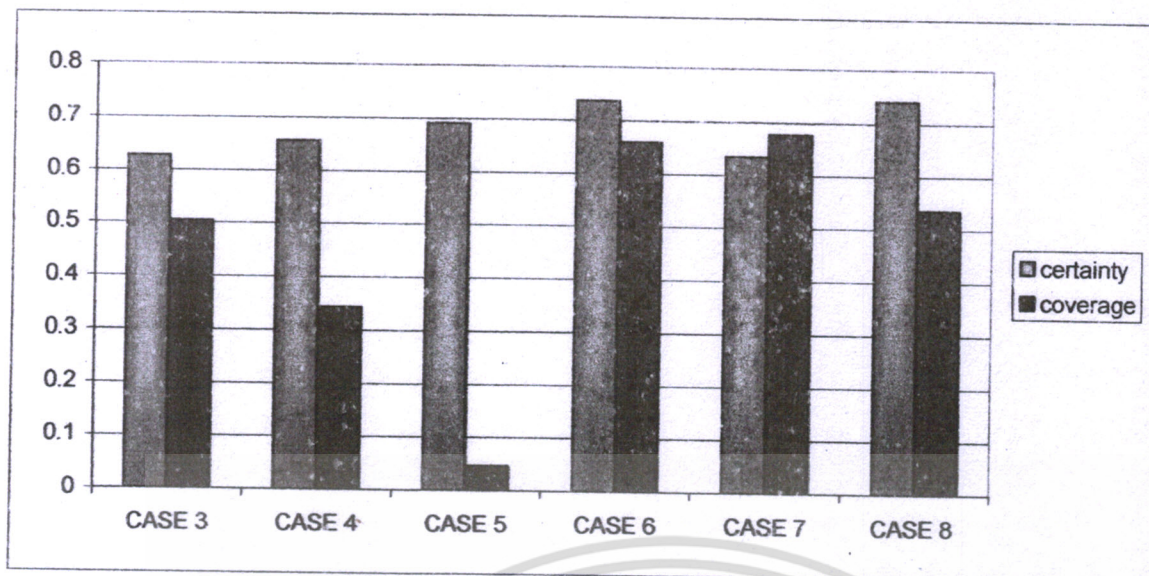


Fig. 1. Comparison of *cer*, *cov* of POSN data.

Interestingly, the combinations of either preferred subject or practiced subject and high score for such subject resulted in high total scores from Cases 6, 7 and 8. In Case 8, we obtain maximum certainty of selected decision rules as one can see in Fig. 1. The results are desirable values of *cer* and *cov*.

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า  
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

## VI. CONCLUSION AND SUGGESTION

Pawlak's flow graph is a useful tool for knowledge discovery and has been successfully applied in many areas. Within the past decade, there has been much research about theoretical aspects of flow graphs which explored the complementary nature of their properties and other mathematical theories. At the same time, flow graphs applications became larger and more complex. Nevertheless, an integration of flow graph and fuzzy set is still an open problem. Thus, in this research, an extension of mathematical fuzzy set theory and rough set theory, namely flow graphs, is considered. We introduce how to calculate *certainty*, *coverage* and *strength* coefficients of decision rules from fuzzy attributes in a flow graph. Furthermore, we relax concept of mutual exclusion and introduced four new propositions of certainty and coverage coefficients for decision rules. These coefficients are extracted from flow graph with fuzzy and descriptive attributes. New certainty and coverage coefficients are introduced. They are successively applied to POSN data set efficiency.

Our future work is to combine rough sets and fuzzy sets flow graphs for better suit specific type of data set and invent new method to allow extensively computation for large data set.

## REFERENCES

- [1] Z.Pawlak, *Rough Sets. Theoretical Aspects of Reasoning about Data*. (Dordrecht: Kluwer Academic Publishers, 1991)
- [2] L.R. Ford, D.R. Fulkerson, *Flows in Networks*. Princeton University Press, Princeton. New Jersey, 1962
- [3] Z.Pawlak, Decision Algorithms and Flow Graphs: A Rough Set Approach. *J. of Telecom. and Informa Tech.* Vol. 3, 2003, 98–101.
- [4] Z. Pawlak, “Rough Sets, Decision Algorithms and Bayes’ Theorem.” *Eur. J.of Oper. Res.* Vol. 136, 2002, pp 181–189.
- [5] Z.Pawlak, Decision Trees and Flow Graphs, in Greco, S. (Eds.) *Rough Sets and Current Trends in Computing, LNAI 4259* (Berlin: Springer-Verlag, 2006) 1–11.
- [6] S. Greco, Z.Pawlak, R. Slowinski, Generalized Decision Algorithms, Rough Inference Rules, and Flow Graphs, in Alpigini, J.J. et al. (Eds) *Rough Sets and Current Trends in Computing, LNAI, 2475* (Berlin: Springer-Verlag, 2002) 93–104.
- [7] Z.Pawlak, Probability, Truth and Flow Graphs, in *Proc of Rough Sets in Knowledge Discovery and Soft Computing*, Warsaw, Poland, 2003, 1–9.
- [8] Z. Pawlak, Rough Set and Flow Graphs, In Slezak, D. et al.(Eds.) *Rough Sets, Fuzzy Sets, Data Mining and Granular Computing, LNAI 3641* (Berlin: Springer-Verlag, 2005) 1–11.
- [9] P. Pattaraintakorn, N. Cercone, K. Naruedomkul, Rule Learning: Ordinal Prediction based on Rough Set and Soft-computing. *Appl. Math. Lett.* 19. (12), 2006, 1300–1307.
- [10] Z.Pawlak, Flow Graphs and Data Mining, in Peters, J.F., Skowron, A. (Eds.) *Transactions on Rough Sets III, LNCS 3400* (Berlin: Springer-Verlag, 2005) 1–36.
- [11] J. Sun, H. Liu, C. Qi, H. Zhang, An Interpretation of Flow Graphs by Granular Computing in Greco, S. (Eds.) *Rough Sets and Current Trends in Computing, LNCS 4259* (Berlin: Springer-Verlag 2006) 448–457.
- [12] L.A. Zadeh, Fuzzy set. *Internation and Control* 8, 1965, 338–353
- [13] M. Alicja, M. Leszek, Flow Graph and Decision Tables with Fuzzy Attributes, *International Conference Artificial Intelligence and Soft Computing , LNAI 4029* (Berlin: Springer-Verlag, 2006) 268–277.
- [14] Z.Pawlak, Flow Graphs and Decision Algorithms, in Wang, G. et al. (Eds.) *Rough Sets, Fuzzy Sets, Data Mining and Granular Computing, LNAI, 2639* (Berlin: Springer-Verlag 2003) 1–10.
- [15] M. Alicja, M. Leszek, Fuzzy Implication Operators in Variable Precision Fuzzy Rough Sets Model, in L. Rutkowski et al. (eds.) *International Conference Artificial Intelligence and Soft Computing , LNAI 3037* (Berlin: Springer-Verlag, 2006) 498–503.
- [16] C.J. Butz, W. Yan, B. Yang, The Computational Complexity of Inference using Rough Set Flow Graphs, in Slezak, D. et al. (Eds.) *Rough Sets, Fuzzy Sets, Data Mining and Granular Computing, LNAI, 364* (Berlin: Springer-Verlag, 2005), 335–344.

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า  
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้