Applied Mathematical Sciences, Vol. 3, 2009, no. 7, 297 - 316

Using Affinity Set on Finding the Key Attributes of Delayed Diagnosis

Yuh-Wen Chen

Institute of Industrial Engineering and Management of Technology, Da-Yeh University, 112 Shan-Jeau Rd., Da-Tsuen, Chang-Hwa 51505, Taiwan
profchen@mail.dyu.edu.tw

Moussa Larbani

Dept. of Busness Administration, Kainan University, No 1, Kainan Road, 338 Taoyuan County, Taiwan
larbani61@hotmail.com
Dept. of Business Administration, Kulliyyah of Economics and Management Sciences IIUM University, Jalan Gombak, 53100 Kuala Lumpur, Malaysia

Chih-Min Shen

Institute of Industrial Engineering and Management of Technology
Da-Yeh University, 112 Shan-Jeau Rd., Da-Tsuen
Chang-Hwa 51505, Taiwan

Chao-Wen Chen

Department of Trauma, Hospital of Kaohsiung Medical University
100 Shi-Chuan 1st Rd., San Ming District, Kaohsiung City, Taiwan

Abstract

According to Institute of Medicine investigation report in the U.S., there are at least

44,000 people die in hospitals each year as a result of medical errors, and these deaths due to medical errors is becoming the 8th-leading cause of death in the United States. These observations point out a serious problem of medical errors, and patients should realize that it is not absolutely safe in the health care system. Thus, medical providers should pay their attentions to reduce the avoidable medical errors and improve patient safety. In this research, medical errors are defined as a delayed diagnosis problem, which means patients' injuries are ignored or missed in Emergency Room (ER), but they are identified by doctors in Intensive Care Unit (ICU). This study is using Affinity Set by topology concept as the data mining tool to classify and focus on the key attributes causing delayed diagnosis. Studying results interestingly indicate that when the patient's triage is resuscitative, and he can breathe normally, but his blood pressure and the pulse are abnormal, it has the high probability of delayed diagnosis. This means two possibilities: (a) there is really no time for doctors in ER to appropriately diagnose the patients; (b) doctors usually ignore the important signs of patients, when they are able to stay consciously and breathe normally. In addition, we also compare the performance between the rough set model (Rosetta) with our affinity set model within this study.

Keywords: Delayed diagnosis, Affinity set, Data mining, Topology

1 Introduction

Curing disease, maintaining health and saving lives are doctor's missions. In the traditional society, people always believe that the doctor's expertise is indisputable, trustworthy, and he never makes mistakes. However, it is not as safe as it should be in a hospital. Based on the results of two studies, "Harvard Medical Practice Study [8]" and "IOM (Institute of Medicine) investigation report-To Err Is Human [9]" published in 1986 and November 1999: the "Harvard Study" shows 3.7% hospitalized patients suffered from medical injury, with 27.6% of these are cause of negligence and 69% are due to human error, 2.6% patients caused permanently disability and 13.6% led to death. On the other hand, "IOM report" also shows that at least 44,000 people and perhaps as many as 99,000 people die in hospitals each year as a result of medical errors which is avoidable [3, 9, 11]. According to these reports, people die in a given year due to the avoidable medical errors are more than the motor vehicle accidents

(43,458), breast cancer (42,297) and AIDS (16,516) [9]. Death due to medical errors is becoming the 8th-leading cause of death in the United States [9].

The aforementioned observations point out reducing the medical errors is critical when most of them are preventable. Medical errors are defined as the delay in diagnosis or the failure of a planned action in operation in this study. The purpose of this paper is finding the key attributes, which may lead to the delayed diagnosis problem by data mining of affinity set. This affinity set [2, 7] developed by Prof. Larbani and Prof. Chen as the data mining tool to classify, analyze and build the relationship between the observed outcomes (consequences) and the possible incomes (causes) of a information system. We actually collect the clinical data from emergency room (ER) in Chung-Ho memorial hospital, Kaohsiung Medical University, Taiwan. After that, we use the affinity data mining model to identify the key attributes of delayed diagnosis. The affinity mining results are also compared with the rough data mining model (Rosetta) for their performances.

This draft paper is organized as follows: in the section 2, we will introduce the affinity set briefly and show how it works in a data mining problem. In the section 3, the actual samples from the memorial hospital of Kaohsiung Medical University, Taiwan are used to validate our affinity data mining idea, from which the key attributes of delayed diagnosis are derived. In addition, we also apply Rosetta of rough set to see if it is able to mining these clinical data. Finally, in the section 4, we give simple conclusions and recommendations based on our achievements of affinity data mining so far.

2 Affinity Set

The original affinity idea comes from the ancient and oriental culture [4, 5, 10]; here, we just briefly give the formal and rigid definitions about the affinity set as follows [2, 7]:

Definition 2.1 Let *e* and A be a subject and an affinity set, respectively. Let I be a subset of the time axis $[0, +\infty)$. The affinity between e and A is represented by a function.

$$\begin{aligned} &M_{\mathrm{A}}^{e}(.) : \mathrm{I} \to [0,1] \\ &t \to M_{\mathrm{A}}^{e}(t) \end{aligned} \tag{1}$$

The value $M_{\mathrm{A}}^{e}(t)$ expresses the degree of affinity between the subject *e* and the affinity set A at time *t*. In fact the affinity function is a fuzzy set defined on the subset *I* of the time axis. When $M_{\mathrm{A}}^{e}(t)$ this means that the affinity of *e* with the affinity set A is complete or at the maximum level at time *t*; it doesn't mean that *e* belongs to A, unless the considered affinity is the belongingness. When $M_{\mathrm{A}}^{e}(t)$=0 this means that e has no affinity with A at time *t*. When $0 < M_{\mathrm{A}}^{e}(t) < 1$, this means that *e* has partial affinity with A at time *t*. Here we emphasize the fact that the notion of affinity is more general than the notion of membership or belongingness. The later is just a particular case of the former.

Definition 2.2

The universal set, denoted by U, is the affinity set representing the fundamental principle of existence. We have

$$\begin{aligned} &M_{\mathrm{U}}^{e}(.) : [0, +\infty) \to [0,1] \\ &t \to M_{\mathrm{U}}^{e}(t) \end{aligned} \tag{2}$$

and $M_{\mathrm{U}}^{e}(t) = 1$, for all existing objects at time *t*, and for all times *t*.

In other words the affinity set defined by the affinity "existence" has complete affinity with all objects that have existed in the past, that exist in the present and that will exist in the future. In general, in real world situations, some traditional referential set *V*, such that when an object *e* is not in V, $M_{\mathrm{U}}^{e}(t) = 0$ for all *t* in $\mathrm{I} \subset [0, +\infty)$, can be determined. In order to make the notion of affinity set operational and for practical reasons, in the following definitions, instead of dealing with the universal set U, we will deal only with affinity sets defined on a traditional referential set V. Thus, in the rest of the article when we refer to an affinity set, we assume that the sets V and I are given.

Definition 2.3. Let *A* be an affinity set. Then the function defining *A* is

$$\begin{aligned} &F_{\mathrm{A}}(.,.) : \mathrm{V} \times \mathrm{I} \to [0,1] \\ &(e,t) \to F_{\mathrm{A}}(e,t) = M_{\mathrm{A}}^{e}(t) \end{aligned} \tag{3}$$

In real situations it often happens that an element belongs to a set for some times and in some other times not. Using affinity set notion, such behavior can be represented. It has to be noted also that the behavior of an affinity set A over time can be investigated through its function $F_{\mathrm{A}}(.,.)$.

Interpretation 2.1.

i) For a fixed element *e* in V, the function (3) defining the affinity set A reduces to the fuzzy set describing the variation of the degree of affinity of the element *e* over time.

ii) For a fixed time *t*, the function (3) reduces to a fuzzy set defined on V that describes the affinity between the elements of V and the affinity set A at time *t*. Roughly speaking it describes the shape or "content" of the affinity set A at time *t*.

Remark 2.1.

It is easy to see that a traditional set is an affinity set. Indeed, if A is a traditional set then the affinity defining A is the belongingness, then for an element *e* in A we have $M_{\mathrm{A}}^{e}(t)=1$, for all *t*, and for an element *e* not in A, we have $M_{\mathrm{A}}^{e}(t)=0$ for all *t*.

Remark 2.2.

Please notice the affinity set is not merely a dynamic fuzzy set, because we don't need to assume any type of membership function here. The affinity is eventually a distance concept between two objects: this distance could be geometric or abstract: this is why the Topology is very valuable in our theory. Once the decision maker can set up his own distance/affinity degree function, then he is free to develop various models for actual applications.

Definition 2.4.

Let A be an affinity set and $k \in [0,1]$. We say that an element *e* is in the *t*-*k*-Core of the affinity set A at time *t*, denoted by *t-k-Core*(A), if $M_{\mathrm{A}}^{e} \geq k$, that is,

$$t-k-Core(\mathrm{A})=\{e \mid M_{\mathrm{A}}^{e}(t) \geq k\} \tag{4}$$

when *k* =1, *t*-*k*-Core(A) is simply called the core of *A* at time *t*, denoted by *t*-Core(A).

Definition 2.5.

An observation period is defined as the period (continuous or discrete) that one is interested in analyzing the behavior of an element *e* of V with respect to an affinity set A (an illustration is given in Figure 1 below).

Definition 2.6.

Let A be an affinity set and $k \in [0,1]$. A subset T (discrete or continuous) of I is said to be the *k*-life cycle of an element *e* with respect to A if

$$M_{\mathrm{A}}^{e}(t) \geq k \text{, for all } t \in \mathrm{T} \text{ and } M_{\mathrm{A}}^{e}(t)<k \text{, elsewhere in I.}$$

In other words, the period T is the *k*-life cycle of e with respect to A if *e* is in the *t*-*k*-Core(A) for all *t* in T . It is the period of time that the element *e* kept its affinity at least equal to *k* in I. The period of time $T_C=\{t \mid M_{\mathrm{A}}^{e}(t)>0, t \in \mathrm{I}\}$ is called life-cycle of the element *e* with respect to the affinity set A.

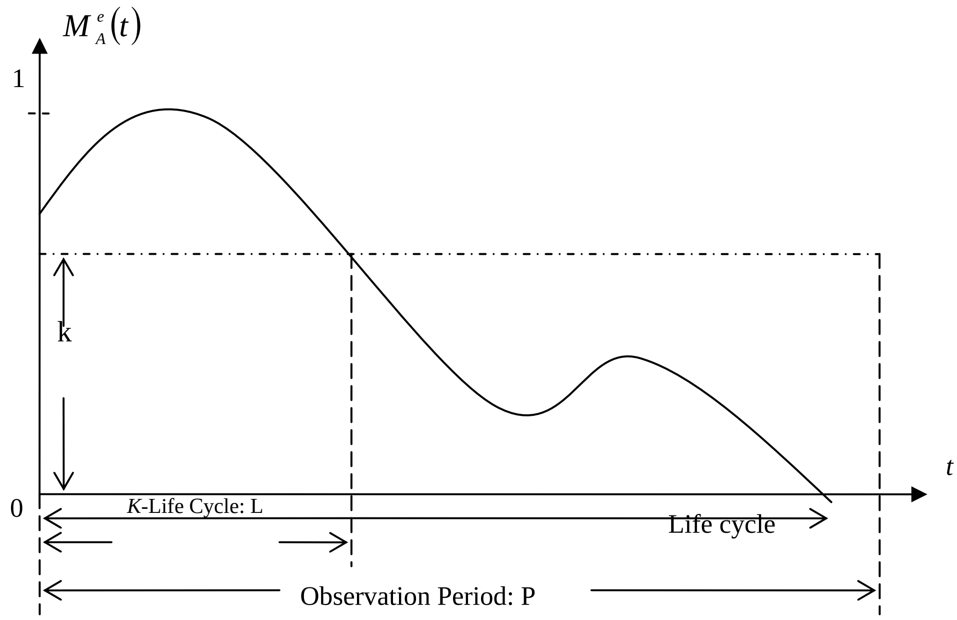

Figure 1 Illustration of the affinity between an element *e* and an affinity set over an observation period P [7].

The basic operations on affinity sets are defined below, these operations will be useful for handling complex systems using the notion of affinity set. In the following definitions 2.7-2.10, we assume that A and B are two given affinity sets defined on I and V.

Definition 2.7.

A and B are considered in an observation period T, then A and B are equal during this period if $M_{\mathrm{A}}^{e}(t) = M_{\mathrm{B}}^{e}(t)$ for all *e* in V and all t in T.

Definition 2.8.

A and B are considered in an observation period T, then A is contained in B during this period, we can write $\mathrm{A} \subset \mathrm{B}$ at tine t if $M_{\mathrm{A}}^{e}(t) \le M_{\mathrm{B}}^{e}(t)$, for all *e* in V.

Definition 2.9.

The union of A and B at time *t*, denoted by $\mathrm{A} \cup \mathrm{B}$, is defined by the function $F_{\mathrm{A}\cup\mathrm{B}}(t,e) = M_{\mathrm{A}\cup\mathrm{B}}^{e}(t) = Max\{M_{\mathrm{A}}^{e}(t), M_{\mathrm{B}}^{e}(t)\}$, for all *e* in V.

Definition 2.10.

The intersection of A and B at time *t*, denoted by $\mathrm{A} \cap \mathrm{B}$, is defined by the function $F_{\mathrm{A}\cap\mathrm{B}}(t,e) = M_{\mathrm{A}\cap\mathrm{B}}^{e}(t) = Min\{M_{\mathrm{A}}^{e}(t), M_{\mathrm{B}}^{e}(t)\}$, for all *e* in V.

It often happens in real-world situations that the affinity of an element *e* with respect to an affinity set A depends implicitly on some other variables than time. In general, these variables express the variability of conditions or constraints that affect the evaluation of affinity. It may be desirable from a practical point of view to study the behavior of an element e with respect to time and also with respect to other variables. It even happens that a decision maker is interested in studying the behavior of an element e at a fixed time with respect to other variables. In the following we extend the definition of affinity set to the case where the desired variables appear explicitly as follows.

Definition 2.11.

Let *e* and A be an element and an affinity set, respectively. Assume that the affinity of e with respect to A depends on some variable *w* that takes its values in a traditional set W. In order to make the variables *w* appear in the definition of the affinity between *e* and A, we introduce the following affinity:

$$\begin{aligned} M_{\mathrm{A}}^{e}(.) &: \mathrm{I} \times \mathrm{W} \to [0,1] \\ (t,w) &\to M_{\mathrm{A}}^{e}(t,w) \end{aligned} \tag{5}$$

The value $M_{\mathrm{A}}^{e}(t,w)$ expresses the degree of affinity between the element *e* and A at time *t* with respect to *w*.

Definition 2.12.

Let A be an affinity set depending on a variable $w \in W$. Then the function defining A is defined by

$$\begin{aligned} R_{\mathrm{A}}(.,.,.) &: \mathrm{V} \times \mathrm{I} \times W \to [0,1] \\ (e,t,w) &\to R_{\mathrm{A}}(e,t,w) = M_{\mathrm{A}}^{e}(t,w) \end{aligned} \tag{6}$$

where *V* is the traditional referential as definition 2.3.

Definition 2.13.

Let A be an affinity set depending an a variable $w \in \mathrm{W}$ and $k \in [0,1]$. We say that an element e in V is the $(t,w^0) - k - Core$ of A at time *t* when $w = w^0$, denoted by $(t,w^0) - k - Core(\mathrm{A})$ if $M_{\mathrm{A}}^{e}(t,w^0) \geq k$, that is,

$$(t,w^0) - k - Core(\mathrm{A}) = \{e \mid M_{\mathrm{A}}^{e}(t,w^0) \geq k\} \tag{7}$$

when *k*=1, $(t,w^0) - k - Core(\mathrm{A})$ is simply called the core of A at time *t* when $w = w^0$ and denoted by $(t,w^0) - Core(\mathrm{A})$.

Remark 2.3.

When the decision maker is interested in studying the behavior of the elements of V with respect to an affinity set A at a fixed time, t can be dropped from the function defining the degree of affinity between any element e and A, thus we can write $M_{\mathrm{A}}^{e}(w)$ instead of $M_{\mathrm{A}}^{e}(t,w)$ the above equation (5), (6) and (7).

The affinity set is far different from the rough set [13, 14] or the fuzzy set [6, 15, 16], because the affinity set is simply constructed on the distance/closeness concept of topology between two objects, and such a distance could be geometric or abstract: that is why we apply the topology concept, which is based on the distance measurement here [12]. We don't need to set up the membership functions (fuzzy set) or the upper/lower bound of a set (rough set). Any interested researcher for affinity set only needs to define his own distance/closeness concept (no mattering it is geometric or abstract), then he can develop/propose many possible applications.

3 Application of Affinity Set to the Delayed Diagnosis Problem

Now, we first use a simple example to show how affinity set can be used for data mining by topology concept.

Example 3.1. (Data Mining Example)

Table 1. Assumed Data of Patients

Sample	x_1 (High Temperature)	x_2 (Sleepy)	y (Death)
P_1	0	1	1
P_2	1	0	1
P_3	1	0	0
P_4	0	1	1
P_5	1	0	0

Assume that a rule base is V ={ r_i, $i=\overline{1,m}$ }, or there are *m* possible rules to explain our observed/collected samples. A rule is a combination of input vector, e.g., *x* vector and output vector, e.g., *y* vector, or *w*=(*x*,*y*) $\in \mathrm{W}$. In addition, these rules $r_i \in \mathrm{V}$, $i=\overline{1,m}$ have been competing for affinity with respect to an affinity set A at a given period: this means we use a cross section data for mining, so that the time dimension (*t*) will be ignored in the following representation. Here this affinity set A is defined as "Appropriateness of explaining observed samples". Consider the following example, some doctors found a new disease, and according to their analysis, there are two possible symptoms within this disease: x_1 and x_2, which could lead to the death of such a disease. Now we collect the patient data from different time periods and classify them into a qualitative (Yes or No) database as in Table 1. Here "1" means we found such an attribute (phenomena) in a patient, "0" means we couldn't find such an attribute (phenomena) in a patient. Thus, according the idea oriented from the topology [12]: this topological way means we assume the behavior of w =(x,y) can be continuously projected/mapped on the plane of rule base V. If a rule or a combination of w =(x_1, x_2, y) of r_i is able to hit the observed samples with higher accuracy (more frequently), then such a combination of w =(x_1, x_2, y) of r_i, or a rule surely has greater affinity with the A, or such a rule r_i is useful to explain the collected/observed samples' behavior. We can easily construct our guesses/rules first for mining these data: there are 8 possible rules we can propose at the beginning for Table 1. For example, we can assume the first rule as: *if* x_1 =1 *and* x_2 =1, *then* y =1; in other words, this implies w^0 =(x_1 =1, x_2 =1, y=1). after that, we check how many actual samples/observations will be hit by this rule in the observation period. In other words, the hit rate is defined as our affinity degree: $M_{\mathrm{A}}^{e}\left(w^0\right)$ here to describing the relationship between the rule $r_i \in \mathrm{V}$ and the observed samples' behavior. We summarize the hit rate by each rule as follows by extending the Definition 2.18 as: (w^0)-*k*-*Core*(A) = $\left\{e \mid M_{\mathrm{A}}^{e}(w^0) \geq k\right\}$, here our observation period is fixed for a given time, e.g., a month, a week, a year, etc, so that the time t is omitted for simplicity.

The M_{A}^{e} (w^0) is defined as the hit frequency of a specified rule divided by the sample size (number of time-interval) n. Thus:

Rule 1 or r_1: if x_1 =1 and x_2 =1, then y =1, hit rate = 0/5 ,

Rule 2 or r_1: if x_1 =1 and x_2 =1, then y =0, hit rate = 0/5

Rule 3 or r_1: if x_1 =1 and x_2 =0, then y =1, hit rate = 1/5

Rule 4 or r_1: if x_1 =1 and x_2 =0, then y =0, hit rate = 2/5

Rule 5 or r_1: if x_1 =0 and x_2 =1, then y =1, hit rate = 2/5

Rule 6 or r_1: if x_1 =0 and x_2 =1, then y =0, hit rate = 0/5

Rule 7 or r_1: if x_1 =0 and x_2 =0, then y =1, hit rate = 0/5

Rule 8 or r_1: if x_1 =0 and x_2 =0, then y =0, hit rate = 0/5

Thus, we can easily organize the (w^0)-0.4-*Core*(A) by two rules: Rule 4 tells us the x_1=1 is not terrible, but Rule 5 warms the doctor that the x_2 =1 has the high death rate in this new disease!!! Of course, as the sample size increases with time, and as the level of classifying the qualitative attribute increases, we can use such a simple thinking to approximate the affinity set of rule: "Appropriateness of explaining observed samples".

Since the objective of this research is to find out the key/core attributes which leading to delayed diagnosis, hence we only use the data, of which the diagnosis is exactly delayed by the doctor's judgment. The doctor gives us 100 samples of clinical data, in which there are 95 cases of delayed diagnosis. We use these 100 samples as our training base (derived for rules) and testing base (computing the hit rate). He also suggests us 10 possible influential attributes, which may lead to the delayed diagnosis for affinity data mining. The way of generating rules and calculating the hit rate of each rule are like the topology idea in Example 3.1. According to the doctor's suggestion, we classify the attribute's values from 3 to 5 grades in Table 2.

Table 2 Classification of Attributes

Attribute variable	Values of attribute
x_1: **patient's age**	"1": under 30 year-old "2": 30 to 60 year-old "3": over 60 year-old
x_2: **triage**	"1": resuscitation, injuries require immediate medical care. "2": emergency, injuries require surgery within 10 minutes. "3": urgent, injuries require surgery within 30 minutes.
x_3: **consciousness**	"1": clear "2": to call, the patient has reaction with voice. "3": to pain, the patient has reaction with pain. "4": coma
x_4: **breathe**	"1": 10-24 times per minute, normal. "2": else, abnormal.
x_5: **blood-pressure**	"1": 90-140 mmHg, normal. "2": else, abnormal.
x_6: **pulse**	"1": 60-100 Times per minute, normal. "2": else, abnormal.
x_7: **temperature**	"1": 35.5℃-37.5℃, normal. "2": else, abnormal.
x_8: **doctor's workweek**	"1": less than 24 hrs. "2": 24 to 48 hrs. "3": over 48 hrs.
x_9: **doctor's age**	"1": under 27 year-old. "2": 27 to 44 year-old. "3": over 45 year-old.
x_{10}: **sub-specialist**	"1": doctor has one specialist. "2": doctor has two specialists.

Since there are 10 possible attributes, which may lead to the results of delayed diagnosis, according to the topological method in section 3.4, we define the affinity here as the relationship degree: $M_{\mathrm{A}}^{e}(w^0)$ between a specified rule (w^0 or r_i) and the affinity set A: "Appropriateness of explaining observed samples". In addition, these collected samples have the same outcomes: they are all with delayed diagnosis; thus, the rule of *w*=(*x*,*y*) is simplified as $w = \{x_i, y, i = \overline{1,10}\}$, and this means we use the combination of *w*=(*x*,*y*) to set up each possible guesses (rules) in V. Each rule combination of (*x*,*y*) is coming from 10 causes $\{x_i\}$ and one consequence *y*. We need to find out the key/core rule only base on the change of $\{x_i\}$ and *y* value is ranged in {0,1}. When *y* value is 1 means a delayed diagnosis occurs. If a combination of $\{x_i\}$ has a higher hit rate, then it is more possible/appropriate to approximate the affinity set A. Using the computation spirit of Example 3.1, we summarize some mining results as in Table 3- 5 . Actually, we try all the combinations of $\{x_i\}$, this means the cardinal of $\{x_i\}$ is ranging from 1 to 10 in this study.

Table 3 $\mathrm{M}_{\mathrm{A}}^{e}(w^0)$ **of Cardinal** $\{x_i\}$**=1**

r_j	1	2	3	4	5	6	7	8	9	10
Attribute	x4	x5	x2	x6	x8	x9	x3	x7	x10	x1
The value	1	2	1	2	2	2	1	2	2	2
$\mathrm{M}_{\mathrm{A}}^{e}(w^0)$	0.8947	0.7895	0.7684	0.7474	0.6211	0.6000	0.5579	0.5474	0.5474	0.4947
r_j	11	12	13	14	15	16	17	18	19	20
Attribute	x7	x10	x9	x1	x6	x3	x2	x5	x8	x1
The value	1	1	3	1	1	4	2	1	1	3
$\mathrm{M}_{\mathrm{A}}^{e}(w^0)$	0.4526	0.4526	0.4000	0.3158	0.2526	0.2316	0.2211	0.2105	0.2105	0.1895

Table 4 $M_A^e(w^0)$ **of Cardinal {**x_i**}=2**

r_j	The combination of w^i of r_j	The value		$M_A^e(w^0)$
1	x4,5	1	2	0.7263
2	x5,6	2	2	0.7053
3	x2,5	1	2	0.6947
4	x2,4	1	1	0.6842
5	x4,6	1	2	0.6526
6	x2,6	1	2	0.6105
7	x4,8	1	2	0.5684
8	x5,7	2	2	0.5474
9	x3,4	1	1	0.5474
10	x6,7	2	2	0.5263

Table 5 $M_A^e(w^0)$ **of Cardinal {**x_i**}=3**

r_j	The combination of w^i of r_j	The value			$M_A^e(w^0)$
1	x4,5,6	1	2	2	0.6105
2	x2,4,5	1	1	2	0.6105
3	x2,5,6	1	2	2	0.5895
4	x2,4,6	1	1	2	0.5263
5	x5,6,7	2	2	2	0.5263
6	x4,5,8	1	2	2	0.4632
7	x4,5,7	1	2	2	0.4526
8	x2,5,7	1	2	2	0.4526
9	x2,6,7	1	2	2	0.4421
10	x5,6,8	2	2	2	0.4421

Table 6 Frequency of Attributes in a Rule

A rule is combined with N attributes	The value	Attribute variable									
		x1	x2	x3	x4	x5	x6	x7	x8	x9	x10
N =2	1	0	6	1	8	0	0	0	0	0	0
	2	0	0	0	0	7	5	4	4	2	3
	3	0	0	0	--	--	--	--	0	0	--
	4	--	--	0	--	--	--	--	--	--	--
N =3	1	0	9	1	12	0	0	0	0	0	0
	2	0	0	0	0	14	9	5	5	1	4
	3	0	0	0	--	--	--	--	0	0	--
	4	--	--	0	--	--	--	--	--	--	--
N =4	1	0	13	2	14	0	0	0	0	0	0
	2	2	0	0	0	17	16	4	4	4	4
	3	0	0	0	--	--	--	--	0	0	--
	4	--	--	0	--	--	--	--	--	--	--
N =5	1	0	14	3	16	0	0	0	0	0	0
	2	4	0	0	0	20	17	11	5	4	5
	3	0	0	0	--	--	--	--	0	1	--
	4	--	--	0	--	--	--	--	--	--	--
N =6	1	0	16	4	17	0	0	0	0	0	0
	2	7	0	0	0	20	19	13	7	4	8
	3	0	0	0	--	--	--	--	0	5	--
	4	--	--	0	--	--	--	--	--	--	--
N =7	1	0	17	10	19	0	0	0	0	0	1
	2	10	0	0	0	20	20	16	9	8	8
	3	0	0	0	--	--	--	--	0	2	--
	4	--	--	0	--	--	--	--	--	--	--
N =8	1	2	18	10	19	0	0	0	0	0	3
	2	12	0	0	0	20	20	17	12	8	10
	3	0	0	0	--	--	--	--	0	6	--
	4	--	--	3	--	--	--	--	--	--	--
N =9	1	9	18	9	19	0	1	1	0	0	3
	2	7	0	0	0	19	18	17	15	7	15
	3	1	0	1	--	--	--	--	0	11	--
	4	--	--	9	--	--	--	--	--	--	--

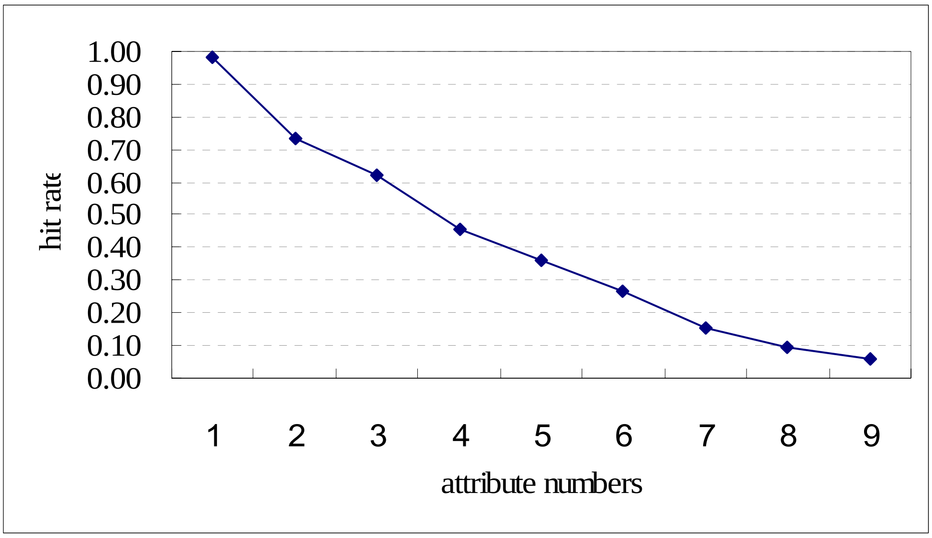

Figure 2 The Hit-rate/Affinity Degree of rule with respect to Card$\{x_i\}$

Fig. 2 has the similar meaning like that in Fig. 1; however, the time dimension t is replaced by the variable dimension $\{x_i\}$. The grey parts in Table 4 and 5 are beyond the doctors' previous expectation: when a patient's vital sign is abnormal, there should be greater chances to delay this patient; however, why the delay possibility is also high when a patient is conscious and when he can breathe normally? What is the hidden message here? Generally speaking, when focusing on the phenomena that goes beyond the doctors' previous expectations by $M_A^e(w^0) \geq 0.4$, we catch the {x2=1, x4=1, x5=2, x6=2} or {triage = resuscitation, breathe=normal, blood-pressure= abnormal, pulse=abnormal} as our key attributes to cause the delayed diagnosis problem. According to such analyzed results, we propose two possibilities: (a) there is really no time for doctors in ER to appropriately diagnose the patients; or (b) doctors usually ignore/miss the important signs of patients, when they are able to stay consciously and breathe normally. No mattering form which possibility of delay, we should set up an efficient mechanism that is able to capture all the vital signs of patient in ER immediately, so that the operation/diagnosis time could be saved in ER, then doctors can have more time to make the right decisions for patients.

We also compare the performances between our affinity model and the rough model (Rosetta). When 10 attributes are all input into Rosetta, it generates 122 rules (much more than 100 observed samples) but the best hit rate is only 0.4. When 4 key attributes x2, x4, x5, x6 are input to Rosetta, then we get the best hit rate only with 0.2. These results are summarized in **Appendix**.

4 Conclusions and Recommendations

This is the first attempt of mining data by affinity set. Although the initial achievements of four key attributes: {triage = resuscitation, breathe=normal, blood-pressure= abnormal, pulse=abnormal} are encouraging, many valuable problems are still waiting for resolution from this small beginning. For example, consider other attributes such as number of doctor, the medical resources of hospital, patients' arriving time, trauma in which body regions…, etc, to analyze the relations within medical data. We can also apply other mapping/projection method that is inspired from topology a tool to verify/enhance the accuracy of this research.

This study also validates that an efficient communication between the patients and the doctors are necessary, e.g., building Radio Frequency Identification (RFID) systems capture all the vital signs of patients within seconds will be beneficial to the doctors and patients and reduce the possibility of delayed diagnosis.

References

[1] R. E. Bellman, L. A. Zadeh, Decision making in a fuzzy environment, Management Science 17B (1970), 141-164.

[2] Y. W. Chen, M. Larbani, Developing the Affinity Set and Its Applications, Proceeding of the Distinguished Scholar Workshop by National Science Council, Taiwan, Jul. 14-18, 2006.

[3] Department of Health, Executive Yuan, R.O.C.(Taiwan), 2007.

[4] D. Y. F. Ho, Interpersonal Relationships and Relationship Dominance: An Analysis Based on Methodological Relationism, Asian Journal of Social Psychology, 1(1998), 1-16.

[5] K. Hwang, Face and Favor: The Chinese Power Game, The American Journal of Sociology, 92(1987), 944-974.

[6] L. A. Zadeh, Fuzzy Sets, Information and Control, 8(1965), 338-353.

[7] M. Larbani, Y. W. Chen, Affinity Set and Its Applications, Proceeding of the International Workshop on Multiple Criteria Decision Making, Poland, Apr. 14-18, 2007.

[8] L. L. Leape, T. A. Brennan, N. Laird, A. G. Lawthers, A. R. Localio, B. A. Barnes, L. Hebert, J. P. Newhouse, P. C. Weiler, H. Hiatt, Incidence of Adverse Events and Negligence in Hospitalized Patients, Results of the Harvard Medical Practice Study I, The New England Journal of Medicine, 324(1991), 370-376.

[9] T. L. Kohn, M. J. M. Corrigan, M. S. Donaldson, Editors; To Err is Human: Building a Safer Health System, Washington D. C., 1999.

[10] L. Yadong, Guanxi and Business, World Scientific, 2000.

[11] S. M. Viveros, J. P. Nearhos, M. J. Rothman, Applying Data Mining Techniques to a Health Insurance Information System, 22nd Very Large Data Bases Conference, 1996.

[12] B. Mendelson, Introduction to Topology, Dover Publications, 1990.

[13] Z. Pawlak, Rough Sets. International Journal of Computer and Information Science, 11(1982), 341-356.

[14] Z. Pawlak, Rough Sets. Theoretical Aspects of Reasoning about Data, Kluwer Academic Publishers, 1991.

[15] L. A. Zadeh, Fuzzy Sets, Information and control, 8(1965), 338-353.

[16] L. A. Zadeh, “Making Computer Think Like People,” IEEE Spectrum, 8(1984), 26-32.

Received: July, 2008

Appendix

Computation results of Rosetta

Table A1 The core rules of ROSETTA with 10 attributes

r_j	The combination of x_i	The value				Hit Rate
1	x4,10	1	1	--	--	0.4
2	x3,10	1	1	--	--	0.3053
3	x2,10	1	1	--	--	0.2947
4	x7,10	2	1	--	--	0.2421
5	x1,10	2	1	--	--	0.2316
6	x2,4	2	1	--	--	0.2
7	x2,3	2	1	--	--	0.2
8	x7,9	1	3	--	--	0.1895
9	x3,6	1	1	--	--	0.1789
10	x4,6,8	1	1	2	--	0.1579
11	x1,10	1	1	--	--	0.1474
12	x3,5,8	4	2	2	--	0.1474
13	x8,9	1	2	--	--	0.1368
14	x1,2	2	2	--	--	0.1263
15	x3,8	1	1	--	--	0.1263
16	x3,6,8	4	2	2	--	0.1263
17	x4,5,6	1	1	1	--	0.1263
18	x6,1	1	1	--	--	0.1158
19	x3,7,8	4	2	2	--	0.1158
20	x3,5,9	4	2	2	--	0.1158
21	x1,4,8	3	1	2	--	0.1158
22	x5,10	1	1	--	--	0.1053
23	x1,8,9	1	2	3	--	0.1053
24	x3,4,8	4	1	2	--	0.1053
25	x4,7	2	2	--	--	0.0947
26	x2,7	2	2	--	--	0.0947
27	x3,6,9	4	2	2	--	0.0947
28	x3,7,9	4	2	2	--	0.0947
29	x1,3,7	3	1	1	--	0.0947
30	x1,5,7	2	2	1	--	0.0947
31	x1,4,6	2	1	1	--	0.0947
32	x1,8	2	3	--	--	0.0842
33	x1,8	2	1	--	--	0.0842
34	x2,6	2	1	--	--	0.0842
35	x2,5	2	1	--	--	0.0842
36	x8,10	3	1	--	--	0.0842
37	x6,9	1	3	--	--	0.0842
38	x7,8	1	1	--	--	0.0842
39	x9,10	3	1	--	--	0.0842
40	x2,4,5	1	2	2	--	0.0842

41	x2,4,6	1	2	2	--	0.0842
42	x3,4,6	4	2	2	--	0.0842
43	x3,4,5	4	2	2	--	0.0842
44	x3,8,9	4	2	3	--	0.0842
45	x1,3,5	2	4	2	--	0.0842
46	x1,9	3	3	--	--	0.0737
47	x8,9	3	3	--	--	0.0737
48	x1,6	1	1	--	--	0.0737
49	x1,5	1	1	--	--	0.0737
50	x1,2,8	3	1	2	--	0.0737
51	x5,6,8	2	1	2	--	0.0737
52	x1,3,6	2	4	2	--	0.0737
53	x3,10	4	1	--	--	0.0632
54	x2,10	2	2	--	--	0.0632
55	x8,10	1	1	--	--	0.0632
56	x3,9	2	3	--	--	0.0632
57	x1,3,8	3	1	2	--	0.0632
58	x1,8,1	3	2	2	--	0.0632
59	x1,3,7	2	4	2	--	0.0632
60	x3,4,9	4	1	2	--	0.0632
61	x1,7,8	3	2	2	--	0.0632
62	x1,3	2	2	--	--	0.0526
63	x1,8	1	3	--	--	0.0526
64	x2,8	2	1	--	--	0.0526
65	x4,6,1	2	2	2	--	0.0526
66	x3,7,8	1	1	3	--	0.0526
67	x1,3,8	1	4	2	--	0.0526
68	x1,3,9	1	4	2	--	0.0526
69	x1,3,4	2	4	1	--	0.0526
70	x4,5,1	2	2	2	--	0.0526
71	x1,6,8	3	1	2	--	0.0526
72	x2,3,7,8	1	2	1	2	0.0526
73	x3,4,7,8	2	1	1	2	0.0526
74	x4,8	2	1	--	--	0.0421
75	x1,8	3	1	--	--	0.0421
76	x1,2	1	2	--	--	0.0421
77	x1,4	1	2	--	--	0.0421
78	x5,9	1	3	--	--	0.0421
79	x6,8	1	1	--	--	0.0421
80	x1,3,9	2	4	3	--	0.0421
81	x1,3,9	1	1	3	--	0.0421
82	x1,4,6	2	2	2	--	0.0421
83	x1,4,5	2	2	2	--	0.0421
84	x3,5,7	4	2	1	--	0.0421
85	x3,4,7	4	1	1	--	0.0421
86	x3,7,8,10	2	1	2	2	0.0421

87	x1,4,6,7	3	1	2	1	0.0421
88	x1,3	3	4	--	--	0.0316
89	x4,9	2	3	--	--	0.0316
90	x2,8	2	3	--	--	0.0316
91	x5,8	1	3	--	--	0.0316
92	x5,8	1	1	--	--	0.0316
93	x2,9	2	3	--	--	0.0316
94	x3,4,6	4	1	1	--	0.0316
95	x3,6,8	2	1	2	--	0.0316
96	x1,3,7	1	2	1	--	0.0316
97	x1,5,6	2	2	1	--	0.0316
98	x6,7,8	2	1	3	--	0.0316
99	x3,5,6	4	2	1	--	0.0316
100	x1,6,7,10	3	2	1	2	0.0316
101	x3,4,6,7	2	1	2	1	0.0316
102	x2,3,6,7	1	2	2	1	0.0316
103	x3,8	4	3	--	--	0.0211
104	x3,5	2	1	--	--	0.0211
105	x6,7	1	2	--	--	0.0211
106	x3,6,8	2	2	3	--	0.0211
107	x1,3,7	1	1	2	--	0.0211
108	x3,6,7	4	2	1	--	0.0211
109	x3,6,7,10	2	2	1	2	0.0211
110	x1,2,6,7	3	1	2	1	0.0211
111	x4,8	2	3	--	--	0.0105
112	x2,3	2	4	--	--	0.0105
113	x3,8	2	1	--	--	0.0105
114	x1,5	3	1	--	--	0.0105
115	x3,4	1	2	--	--	0.0105
116	x1,2,4	3	1	2	--	0.0105
117	x3,7,8	2	2	3	--	0.0105
118	x4,5,6	2	2	1	--	0.0105
119	x1,3,7	1	4	1	--	0.0105
120	x2	3	--	--	--	0.0105
121	x9	1	--	--	--	0.0105
122	x3	3	--	--	--	0.1053

Table A2 The core rules of ROSETTA with 4 attributes

r_j	The combination of x_i	The value			Hit Rate
1	x2	3	--	--	0.0105
2	x2,4	2	1	--	0.2000
3	x2,5	2	1	--	0.0842
4	x2,6	2	1	--	0.0842
5	x2,4,5	1	2	2	0.0842
6	x2,4,6	1	2	2	0.0842
7	x4,5,6	1	1	1	0.1263
8	x4,5,6	2	2	1	0.0105