具有属性析取萎缩-扩张特征的动态数据智能挖掘

徐凤生[1] 闫立梅[1] 史开泉[2]
(德州学院数学科学学院 德州 253023)[1] (山东大学数学与系统科学学院 济南 250100)[2]

摘 要 S-粗集(singular rough sets)是把动态特征引入到 Z. Pawlak 粗集中对其加以改进而提出的,S-粗集具有动态特征。S-粗集具有3种形式:单向S-粗集(one direction singular rough sets)、单向S-粗集对偶(dual of one direction singular rough sets)与双向S-粗集(two direction singular rough sets);在一定条件下,单向S-粗集、单向S-粗集对偶与双向S-粗集被还原成 Z. Pawlak 粗集。利用单向S-粗集和单向S-粗集对偶给出具有属性析取特征的动态数据智能挖掘与应用;属性析取是数据具有的逻辑特征之一。主要结果是:利用单向S-粗集、单向S-粗集对偶结构,给出属性析取萎缩-扩张特征的动态数据生成与它的属性析取萎缩-扩张关系;给出数据推理与推理模型;利用数据推理给出动态数据智能挖掘定理;利用这些理论结果,给出动态数据智能挖掘-智能认知的应用。

关键词 S-粗集,属性析取,动态数据生成,数据推理,数据智能挖掘,应用

中图法分类号 O144,TP391 **文献标识码** A **DOI** 10.11896/j.issn.1002-137X.2015.5.043

Dynamic Data Intelligent Mining with Attributes Disjunctive Reduction and Expansion Characteristics

XU Feng-sheng[1] YAN Li-mei[1] SHI Kai-quan[2]
(College of Mathematical Sciences, Dezhou University, Dezhou 253023, China)[1]
(School of Mathematics and System Sciences, Shandong University, Jinan 250100, China)[2]

Abstract S-rough sets (singular rough sets) are generated by introducing dynamic characteristics into Z. Pawlak rought sets. S-rough sets have dynamic characteristics. In general, S-rough sets have three formulas including one direction S-rough sets, dual of one direction S-rough sets and two directions S-rough sets. Given certain conditions, one direction S-rough sets and dual of one direction S-rough sets can be reduced to Z. Pawlak rought sets. Dynamic data intelligent mining with attribute disjunctive characteristics was considered. Attribute disjunctive is one of logic characteristics of data. The main results of this paper include generation of dynamic data and the relations between dynamic data and its attribute disjunctive reduction and expansion, data reasoning and reasoning model, theorems of dynamic data intelligent mining, and applications of the obtained results.

Keywords S-rough sets, Attribute disjunctive, Generation of dynamic data, Data reasoning, Data intelligent mining, Application

1 引言

本文给出的研究是以一个简单事实为依据,利用单向S-粗集(one direction singular rough sets)、单向S-粗集对偶(dual of one direction singular rough sets)[1-5]的数学理论与数学模型做支撑得到的。一个简单的事实:A 是一个集团公司,$X=\{x_1,x_2,\cdots,x_k\}\subset U$ 是 A 的产品 x_i 构成的产品集合,$\alpha=\{\alpha_1,\alpha_2,\cdots,\alpha_k\}\subset V$ 是出售 $x_i\in X$ 的合同 α_i 构成的 X 的合同集合,满足:$\forall x_i,x_j\in X,x_i\neq x_j,\alpha_i\neq\alpha_j,\alpha_i,\alpha_j\in\alpha$;$\forall x_i\in X$ 的合同 α_i 满足合同析取范式:$\alpha_i=\alpha_1\vee\alpha_2\vee\cdots\vee\alpha_k=\bigvee_{t=1}^{k}\alpha_t$。I. 若 α 内删除一些合同 α_j,α 变成 $\alpha^{\bar{f}}=\{\alpha_1,\alpha_2,\cdots,\alpha_\lambda\}$,$\alpha^{\bar{f}}\subset\alpha$,则 X 内删除一些 x_j,X 变成 $X^{\bar{f}}=\{x_1,x_2,\cdots,x_\lambda\}$,$X^{\bar{f}}\subset X,\lambda<k$。II. 若 α 内补充一些合同 α_j,α 变成 $\alpha^f=\{\alpha_1,\alpha_2,\cdots,\alpha_k,\alpha_{k+1},\cdots,\alpha_n\}$,$\alpha\subset\alpha^f$,则 X 内补充一些 x_j,X 变成 $X^f=\{x_1,x_2,\cdots,x_k,x_{k+1},\cdots,x_n\}$,$X\subset X^f,k<n$。III. 若在 α 内删除一些合同 α_i,同时补充一些合同 α_j,α 变成 $(\alpha^{\bar{f}},\alpha^f)$,$\alpha^{\bar{f}}\subset\alpha\subset\alpha^f$,则 X 内删除一些 x_i 同时补充一些 x_j,X 变成 $(X^{\bar{f}},X^f)$,$X^{\bar{f}}\subset X\subset X^f$。I—III 是经济系统中的常识,经常被人们遇到。从 I 中得到:$\forall x_i\in X^{\bar{f}}$ 的合同 α_i 满足合同析取范式萎缩 $\alpha_i=\bigvee_{t=1}^{\lambda}\alpha_t$,$\lambda<k$;从 II 中得到:$\forall x_j\in X^f$ 的合同 α_j 满足合同析取范式扩张 $\alpha_j=\bigvee_{t=1}^{n}\alpha_t$,$k<n$。把这个事实中的 X 定义为数据 $[x]$,或者 $[x]=X$,把合同集合 α 定义成 $[x]$ 的属性集合 α,对这个事实给出数学抽象并移植到数据挖掘研究中得到:A. 若在 α 内删除一些属性 α_i,α 变成 $\alpha^{\bar{f}}$,$\alpha^{\bar{f}}\subset\alpha$,则具有属性集合 α 的数据 $[x]$ 变成具有属性集合 $\alpha^{\bar{f}}$ 的数据 $[x]^{\bar{f}}$,$[x]^{\bar{f}}\subset[x]$,$[x]^{\bar{f}}$ 跟随 $\alpha^{\bar{f}}$ 的

到稿日期:2014-06-20 返修日期:2014-12-10 本文受山东省高校科技计划(J12LN92)项目资助。
徐凤生(1965—),男,硕士,教授,主要研究方向为信息系统理论与应用,E-mail:sxxxfs@126.com;**闫立梅**(1969—),女,硕士,教授,主要研究方向为粗系统理论与应用,E-mail:yanlimei9898@163.com;**史开泉**(1945—),男,教授,博士生导师,主要研究方向为粗系统理论与应用,E-mail:shikq@sdu.edu.cn(通信作者)。

变化，$[x]^{\bar{f}}$具有了动态特征，$[x]^{\bar{f}}$在$[x]$内被动态挖掘；$\forall x_i\in[x]^{\bar{f}}$的属性$\alpha_i$是$\forall x_j\in[x]$的属性$\alpha_j$的属性析取范式萎缩。B. 若在$\alpha$内补充一些属性$\alpha_i$，$\alpha$变成$\alpha^f$，$\alpha\subset\alpha^f$，则具有属性集合$\alpha$的数据$[x]$变成具有属性集合$\alpha^f$的数据$[x]^f$，$[x]\subset[x]^f$，$[x]^f$跟随$\alpha^f$的变化，$[x]^f$具有了动态特征，$[x]^f$在$[x]$外被动态挖掘；$\forall x_i\in[x]^f$的属性$\alpha_i$是$\forall x_j\in[x]$的属性$\alpha_j$的属性析取范式扩张。C. 若在$\alpha$内删除一些属性$\alpha_i$同时在$\alpha$内补充一些属性$\alpha_j$，则数据$[x]$变成$([x]^{\bar{f}},[x]^f)$，$[x]^{\bar{f}}\subset[x]\subset[x]^f$；$[x]^{\bar{f}}$，$[x]^f$同时在$[x]$内、外被连续动态挖掘。A与单向S-粗集对偶的动态特征相同，B与单向S-粗集的动态特征相同；单向S-粗集、单向S-粗集对偶在动态信息、动态信息系统中已获得多个应用[1-19,25,26]。S-粗集的函数形式与应用见文献[9-11,13-18,25,26]。

Z. Pawlak提出粗集[22-24]，粗集为静态数据挖掘与静态知识发现研究提供了数学理论与数学方法的支持，粗集在静态数据挖掘与静态知识发现研究中获得了应用[20-24]。分析Z. Pawlak粗集结构得到：Z. Pawlak粗集是一个“静态粗集”，理由是：Z. Pawlak粗集是保留了有限普通集合X(Cantor set X)的“静态特性”，用“近似特性”代替有限普通集合X的“精确特性”，改进有限普通集合X得到的，Z. Pawlak粗集不具有“动态特性”。数据不都是“静态”的，大量的动态数据出现在多类信息系统中，利用具有“静态特性”，Z. Pawlak粗集研究具有动态特性的数据的数据挖掘遇到了困难。在这样的背景条件下，S-粗集(singular rough set)被提出，S-粗集是保留了Z. Pawlak粗集的“近似特性”，用动态特性代替Z. Pawlak粗集的静态特性，改进Z. Pawlak粗集得到的。S-粗集具有3种结构：单向S-粗集、单向S-粗集对偶、双向S-粗集；S-粗集为具有动态特征的数据的数据挖掘研究给出数学理论的支持。在一定条件下，单向S-粗集、单向S-粗集对偶、双向S-粗集无一例外地被还原Z. Pawlak粗集。在静态-动态条件下，Z. Pawlak粗集是单向S-粗集、单向S-粗集对偶、双向S-粗集的特例；单向S-粗集、单向S-粗集对偶、双向S-粗集是Z. Pawlak粗集的一般形式。

数据$[x]$与数据$[x]$的属性集合α成对的出现；或者，数据$[x]$对应着$[x]$的属性集合α，利用属性集合α能够找到具有属性集合α的数据$[x]$。数据$[x]$的属性集合α发生变化，引起数据$[x]$的变化，数据$[x]$的动态特性来自数据$[x]$的属性集合的变化。

本文利用一个事实作为证据，利用单向S-粗集与单向S-粗集对偶整合，给出具有属性析取萎缩-扩张特征的动态数据智能挖掘。本文的主要结果：利用单向S-粗集、单向S-粗集对偶结构与动态特性，给出具有属性析取萎缩-扩张特征的动态数据生成与它的属性析取萎缩-扩张关系；给出数据推理结构与推理模型；利用数据推理，给出动态数据的动态智能挖掘定理；给出数据智能挖掘-智能认知的应用；属性析取是数据挖掘中的逻辑关系之一。

为了方便本文的讨论、概念的引用、保持本文内容的完整，把单向S-粗集、单向S-粗集对偶简单地引入到本文的第2节内作为本文讨论的知识准备；单向S-粗集、单向S-粗集对偶的更多特征、概念与应用见文献[25,26]。

2 单向S-粗集，单向S-粗集对偶结构与动态特征

文献[1-5]给出

称X°是$X\subset U$的一个单向S-集合(one direction singular set)，如果

$$X^\circ=X\cup\{u|u\in U,u\notin X,f(u)=x\in X\} \tag{1}$$

称X^f是$X\subset U$的f-扩张，如果

$$X^f=\{u|u\in U,u\notin X,f(u)=x\in X\} \tag{2}$$

给定单向S-集合$X^\circ\subset U$，称$(R,F)_\circ(X^\circ)$，$(R,F)^\circ(X^\circ)$分别是$X^\circ\subset U$的下近似，上近似，如果

$$(R,F)_\circ(X^\circ)=\bigcup[x]=\{x|x\in U,[x]\subseteq X^\circ\} \tag{3}$$

$$(R,F)^\circ(X^\circ)=\bigcup[x]=\{x|x\in U,[x]\cap X^\circ\neq\emptyset\} \tag{4}$$

由$(R,F)_\circ(X^\circ)$，$(R,F)^\circ(X^\circ)$构成的集合对是$X^\circ\subset U$的单向S-粗集(one direction singular rough sets)，如果

$$((R,F)_\circ(X^\circ),(R,F)^\circ(X^\circ)) \tag{5}$$

称$Bnr(X^\circ)$是$X^\circ\subset U$的R-边界，如果

$$Bnr(X^\circ)=(R,F)^\circ(X^\circ)-(R,F)_\circ(X^\circ) \tag{6}$$

称$X'\subset U$是单向S-集合$X^\circ\subset U$的对偶，如果

$$X'=X-\{x|x\in X,\bar{f}(x)=u\notin X\} \tag{7}$$

称$X^{\bar{f}}$是$\subset U$的$\bar{f}$-萎缩，如果

$$X^{\bar{f}}=\{x|x\in X,\bar{f}(x)=u\notin X\} \tag{8}$$

给定单向S-集合X°的对偶$X'\subset U$，称$(R,\bar{F})_\circ(X')$、$(R,\bar{F})^\circ(X')$分别是$X'\subset U$的下近似、上近似，如果

$$(R,\bar{F})_\circ(X')=\bigcup[x]=\{x|x\in U,[x]\subseteq X'\} \tag{9}$$

$$(R,\bar{F})^\circ(X')=\bigcup[x]=\{x|x\in U,[x]\cap X'\neq\emptyset\} \tag{10}$$

由$(R,\bar{F})_\circ(X')$，$(R,\bar{F})^\circ(X')$构成的集合对是$X'\subset U$的单向S-粗集对偶(dual of one direction singular rough sets)，如果

$$((R,\bar{F})_\circ(X'),(R,\bar{F})^\circ(X')) \tag{11}$$

称$Bnr(X')$是$X'\subset U$的R-边界，如果

$$Bnr(X')=(R,\bar{F})^\circ(X')-(R,\bar{F})_\circ(X') \tag{12}$$

应当指出：1. 式(1)，式(2)和式(7)，式(8)中的$X\subset U$是Z. Pawlak粗集$(R_-(X),R^-(X))$中的有限普通集合X。2. 式(3)－式(6)和式(9)－式(12)中的R是等价关系；式(3)，式(4)，式(9)，式(10)中的$[x]$是等价类。3. 式(3)－式(6)中的$F=\{f_1,f_2,\cdots,f_n\}$是元素迁移族，$f\in F$是元素迁移，$f\in F$的特征是：$u\in U,u\notin X$，$f\in F$把u变成$f(u)=x\in X$；式(9)－式(12)中的$\bar{F}=\{\bar{f}_1,\bar{f}_2,\cdots,\bar{f}_n\}$是元素迁移族，$\bar{f}\in\bar{F}$是元素迁移，$\bar{f}\in\bar{F}$的特征是：$x\in X$，$\bar{f}\in\bar{F}$把$x$变成$\bar{f}(x)=u\notin X$。4. 式(7)中$X'\neq\emptyset$。

由单向S-粗集结构式(1)－式(6)，单向S-粗集对偶结构式(7)－式(12)分别给出两个重要结论I与II，I与II在3－5节的讨论中被使用。给定有限普通集合$X=\{x_1,x_2,\cdots,x_k\}$，$\alpha=\{\alpha_1,\alpha_2,\cdots,\alpha_k\}$是$X$的属性集合，$\forall x_i\in X$的属性$\alpha_i$满足$\alpha_i=\alpha_1\vee\alpha_2\vee\cdots\vee\alpha_k$。

I. 式(1)－式(6)中，X内被补充元素，X变成X°，$X\subset X^\circ$，等价于X的属性集合α内被补充属性，α变成α^f，$\alpha\subset\alpha^f$；或者，$[x]$内被补充元素，$[x]$变成$[x]^f$，$[x]\subset[x]^f$，等价于$[x]$的属性集合α内被补充属性，α变成α^f，$\alpha\subset\alpha^f$。

II. 式(7)－式(12)中，X内被删除元素，X变成X'，$X'\subset X$，等价于X的属性集合α内被删除属性，α变成$\alpha^{\bar{f}}$，$\alpha^{\bar{f}}\subset\alpha$；或者，$[x]$内被删除元素，$[x]$变成$[x]^{\bar{f}}$，$[x]^{\bar{f}}\subset[x]$，等价于$[x]$的属性集合$\alpha$内被删除属性，$\alpha$变成$\alpha^{\bar{f}}$，$\alpha^{\bar{f}}\subset\alpha$；I，II分别潜藏在单向S-粗集的结构式(1)－式(6)，单向S-粗集对偶结构式(7)－式(12)中。

由单向S-粗集、单向S-粗集对偶与Z. Pawlak粗集[22-24]

得到：

定理 1 若 $F=\emptyset$，则单向 S-粗集与 Z. Pawlak 粗集满足

$$((R,F)_{\circ}(X^{\circ}),(R,F)^{\circ}(X^{\circ}))_{F=\emptyset}=(R\_(X),R^{-}(X)) \tag{13}$$

证明：由式(1)得到：若 $F=\emptyset$，则 $X^{f}=\{u|u\in U,u\notin X,f(u)=x\in X\}=\emptyset$，式(1)变成 $X^{\circ}=X\cup\{u|u\in U,u\notin X,f(u)=x\in X\}=X\cup\emptyset=X$，式(3)变成$(R,F)_{\circ}(X^{\circ})=\cup[x]=\{x|x\in U,[x]\subseteq X^{\circ}\}=\{x|x\in U,[x]\subseteq X\}=\cup[x]=R_(X)$；式(4)变成$(R,F)^{\circ}(X^{\circ})=\cup[x]=\{x|x\in U,[x]\cap X^{\circ}\neq\emptyset\}=\{x|x\in U,[x]\cap X\neq\emptyset\}=R^{-}(X)$；或者$((R,F)_{\circ}(X^{\circ}),(R,F)^{\circ}(X^{\circ}))_{F=\emptyset}=(R_(X),R^{-}(X))$，得到式(13)。

定理 2 若 $\bar{F}=\emptyset$，则单向 S-粗集对偶与 Z. Pawlak 粗集满足

$$((R,\bar{F})_{\circ}(X'),(R,\bar{F})^{\circ}(X'))_{\bar{F}=\emptyset}=(R\_(X),R^{-}(X)) \tag{14}$$

证明：由式(7)得到：若 $\bar{F}=\emptyset$，则 $X^{\bar{f}}=\{x|x\in X,\bar{f}(x)=u\notin X\}=\emptyset$，式(7)变成 $X'=X-\{x|x\in X,\bar{f}(x)=u\notin X\}=X-\emptyset=X$，式(9)变成$(R,\bar{F})_{\circ}(X')=\cup[x]=\{x|x\in U,[x]\subseteq X'\}=\{x|x\in U,[x]\subseteq X\}=\cup[x]=R_(X)$；式(10)变成$(R,\bar{F})^{\circ}(X')=\cup[x]=\{x|x\in U,[x]\cap X'\neq\emptyset\}=\{x|x\in U,[x]\cap X\neq\emptyset\}=R^{-}(X)$；或者，$((R,\bar{F})_{\circ}(X'),(R,\bar{F})^{\circ}(X'))_{\bar{F}=\emptyset}=(R_(X),R^{-}(X))$，得到式(14)。这里式(13)，式(14)中$(R_(X),R^{-}(X))$是 Z. Pawlak 粗集。

命题 1 单向 S-粗集是 Z. Pawlak 粗集中$[x]$内被补充元素的生成。

命题 2 单向 S-粗集对偶是 Z. Pawlak 粗集中$[x]$内被删除元素的生成。

命题 3 单向 S-粗集的动态特性与累加器 $T=T+1$ 的动态特性相同。

命题 4 单向 S-粗集对偶的动态特性与累减器 $T=T-1$ 的动态特性相同。

利用第 2 节中的概念，单向 S-粗集、单向 S-粗集对偶的结构在第 3 节中给出。

3 属性析取萎缩-扩张与动态数据生成

约定 在 3—5 节的讨论中，第 2 节中的 X,X°,X'分别记作$[x]$，$[x]^{f}$，$[x]^{\bar{f}}$；或者，$[x]=X$，$[x]^{f}=X^{\circ}$，$[x]^{\bar{f}}=X'$；$[x]$，$[x]^{f}$，$[x]^{\bar{f}}$称作数据，不引起误解。

定义 1 给定数据$[x]=\{x_1,x_2,\cdots,x_k\}$，$\alpha=\{\alpha_1,\alpha_2,\cdots,\alpha_k\}$是$[x]$的属性集合，称$[x]$是具有属性析取特征的数据，如果数据元 $x_i\in[x]$的属性 α_i 满足

$$\alpha_i=\alpha_1\vee\alpha_2\vee\cdots\vee\alpha_k=\vee_{t=1}^{k}\alpha_t \tag{15}$$

定义 2 称$[x]^{\bar{f}}$是$[x]$生成的内-动态数据，$[x]^{\bar{f}}\subseteq[x]$，如果数据元 $x_j\in[x]^{\bar{f}}$的属性α_j 满足属性析取范式萎缩

$$\alpha_j=\bigvee_{t=1}^{n}\alpha_t \tag{16}$$

称$[x]^{f}$ 是$[x]$生成的外-动态数据，$[x]\subseteq[x]^{f}$，如果数据元 $x_j\in[x]^{f}$ 的属性α_j 满足属性析取范式扩张

$$\alpha_j=\bigvee_{t=1}^{m}\alpha_t \tag{17}$$

式(15)—式(17)中，$n<k<m$；$n,k,m\in N^{+}$。

定义 3 由$[x]^{\bar{f}}$，$[x]^{f}$ 构成的数据对，称作$[x]$生成的内-外动态数据

$$([x]^{\bar{f}},[x]^{f}) \tag{18}$$

由定义 1—3 得到：

定理 3 内-动态数据$[x]^{\bar{f}}$存在的充分必要条件是数据元$x_j\in[x]^{\bar{f}}$的属性α_j 是数据元$x_i\in[x]$的属性 α_i 的属性析取范式萎缩，或者

$$\alpha_j=(\bigvee_{t=1}^{k}\alpha_t)-(\bigvee_{t=m+1}^{k}\alpha_t) \tag{19}$$

这里，$\alpha=\{\alpha_1,\alpha_2,\cdots,\alpha_m,\alpha_{m+1},\cdots,\alpha_k\}$是$[x]$的属性集合，$\alpha^{\bar{f}}=\{\alpha_1,\alpha_2,\cdots,\alpha_m\}$是$[x]^{\bar{f}}$的属性集合，$\alpha_i=\bigvee_{t=1}^{k}\alpha_t$ 是数据元 $x_i\in[x]$的属性析取范式。

证明：利用引言中一个简单事实中的 I 得到：1. $\alpha=\{\alpha_1,\alpha_2,\cdots,\alpha_n,\alpha_{n+1},\cdots,\alpha_k\}$，$\alpha^{\bar{f}}=\{\alpha_1,\alpha_2,\cdots,\alpha_n\}$分别是$[x]$，$[x]^{\bar{f}}$的属性集合，$\forall x_i\in[x]$的属性 α_i，$\forall x_j\in[x]^{\bar{f}}$的属性$\alpha_j$ 分别满足：$\alpha_i=\alpha_1\vee\alpha_2\vee\cdots\vee\alpha_n\vee\alpha_{n+1}\vee\cdots\vee\alpha_k=\bigvee_{t=1}^{k}\alpha_t$，$\alpha_j=\alpha_1\vee\alpha_2\vee\cdots\vee\alpha_n=(\alpha_1\vee\alpha_2\vee\cdots\vee\alpha_n\vee\alpha_{n+1}\vee\cdots\vee\alpha_k)-(\alpha_{n+1}\vee\cdots\vee\alpha_k)=\bigvee_{t=1}^{k}\alpha_t-\bigvee_{t=n+1}^{k}\alpha_t$，得到式(19)。2. 若$\forall x_i\in[x]$的属性 α_i 与$\forall x_j\in[x]^{\bar{f}}$的属性$\alpha_j$ 满足式(19)：$\alpha_j=\vee_{t=1}^{k}\alpha_t-\bigvee_{t=n+1}^{k}\alpha_t$，由定义 1、2 得到：$[x]^{\bar{f}}$是$[x]$生成的内-动态数据；由 1 与 2 得到定理 3。

定理 4 外-动态数据$[x]^{f}$ 存在的充分必要条件是数据元$x_j\in[x]^{f}$ 的属性α_j 是数据元$x_i\in[x]$的属性 α_i 的属性析取范式扩张，或者

$$\alpha_j=(\bigvee_{t=1}^{k}\alpha_t)\bigvee_{t=k+1}^{m}\alpha_t \tag{20}$$

这里：$\alpha=\{\alpha_1,\alpha_2,\cdots,\alpha_m,\alpha_{m+1},\cdots,\alpha_k\}$是$[x]$的属性集合，$\alpha^{f}=\{\alpha_1,\alpha_2,\cdots,\alpha_k,\alpha_{k+1},\cdots,\alpha_m\}$是$[x]^{f}$ 的属性集合。

证明：利用引言中一个简单事实中的 II 得到：1. $\alpha=\{\alpha_1,\alpha_2,\cdots,\alpha_m,\alpha_{m+1},\cdots,\alpha_k\}$，$\alpha^{f}=\{\alpha_1,\alpha_2,\cdots,\alpha_k,\alpha_{k+1},\cdots,\alpha_m\}$分别是$[x]$，$[x]^{f}$ 的属性集合，$\forall x_i\in[x]$的属性 α_i，$\forall x_j\in[x]^{f}$ 的属性 α_j 分别满足：$\alpha_i=\alpha_1\vee\alpha_2\vee\cdots\vee\alpha_k=\bigvee_{t=1}^{k}\alpha_t$，$\alpha_j=\alpha_1\vee\alpha_2\vee\cdots\vee\alpha_k\vee\alpha_{k+1}\vee\cdots\vee\alpha_m=(\alpha_1\vee\alpha_2\vee\cdots\vee\alpha_k)\vee(\alpha_{k+1}\vee\alpha_{k+2}\vee\cdots\vee\alpha_m)=(\bigvee_{t=1}^{k}\alpha_t)\bigvee_{t=k+1}^{m}\alpha_t$，得到式(20)。2. 若$\forall x_i\in[x]$的属性 α_i 与$\forall x_j\in[x]^{f}$ 的属性α_j 满足式(20)：$\alpha_j=(\bigvee_{t=1}^{k}\alpha_t)\bigvee_{t=k+1}^{m}\alpha_t$，由定义 1，2 得到：$[x]^{f}$ 是$[x]$生成的外-动态数据；由 1 与 2 得到定理 4。

由定理 3，4 直接得到：

定理 5 内-外动态数据$([x]^{\bar{f}},[x]^{f})$存在的充分必要条件是数据元 $x_i\in[x]^{\bar{f}}$的属性α_i，数据元 $x_j\in[x]^{f}$ 的属性α_j 满足

$$(\alpha_i,\alpha_j)=((\bigvee_{t=1}^{k}\alpha_t)-(\bigvee_{t=n+1}^{k}\alpha_t),(\bigvee_{t=1}^{k}\alpha_t)\bigvee_{t=k+1}^{n}\alpha_t) \tag{21}$$

式(21)表示 $\alpha_i=(\bigvee_{t=1}^{k}\alpha_t)-(\bigvee_{t=n+1}^{k}\alpha_t)$，$\alpha_j=(\bigvee_{t=1}^{k}\alpha_t)\bigvee_{t=k+1}^{n}\alpha_t$。

由定义 1—3，定理 3—5 得到：

推论 1 $[x]^{\bar{f}}$是内-动态数据，当且仅当存在$\nabla\alpha\neq\emptyset$，$[x]^{\bar{f}}$的属性集合$\alpha^{\bar{f}}$，$[x]$的属性集合$\alpha$与$\nabla\alpha$满足

$$\alpha^{\bar{f}}=\alpha-\nabla\alpha \tag{22}$$

推论 2 $[x]^{f}$ 是外-动态数据，当且仅当存在 $\Delta\alpha\neq\emptyset$，$[x]^{f}$ 的属性集合α^{f}，$[x]$的属性集合α与 $\Delta\alpha$ 满足

$$\alpha^{f}=\alpha\cup\Delta\alpha \tag{23}$$

推论 3 $([x]^{\bar{f}},[x]^{f})$是内-外动态数据，当且仅当存在$\nabla\alpha\neq\emptyset,\Delta\alpha\neq\emptyset$，$([x]^{\bar{f}},[x]^{f})$的属性集合$(\alpha^{\bar{f}},\alpha^{f})$与$[x]$的属性集合$\alpha$满足

$$(\alpha^{\bar{f}},\alpha^{f})=(\alpha-\nabla\alpha,\alpha\cup\Delta\alpha) \tag{24}$$

式(24)表示 $\alpha^{\bar{f}}=\alpha-\nabla\alpha$，$\alpha^{f}=\alpha\cup\Delta\alpha$。

推论 4 内-动态数据$[x]^{\bar{f}}$是$[x]$的属性集合 α 内删除属性的生成。

推论 5 外-动态数据$[x]^{f}$是$[x]$的属性集合 α 内补充属性的生成。

推论 6 内-外动态数据$([x]^{\bar{f}},[x]^{f})$是$[x]$的属性集合 α 内删除-补充属性的生成。

命题 5 具有属性析取范式萎缩特征的内-动态数据$[x]_i^{\bar{f}}$构成内-动态数据族$\{[x]_i^{\bar{f}}\mid i\in I,[x]_{i+1}^{\bar{f}}\subseteq[x]_i^{\bar{f}}\}$。

命题 6 具有属性析取范式扩张特征的外-动态数据$[x]_j^{f}$构成外-动态数据族$\{[x]_j^{f}\mid j\in J,[x]_j^{f}\subseteq[x]_{j+1}^{f}\}$。

命题 7 具有属性析取范式萎缩-扩张特征的内-外动态数据$([x]_i^{\bar{f}},[x]_j^{f})$构成内-外动态数据族$\{([x]_i^{\bar{f}},[x]_j^{f})_k\mid k\in K,[x]_{i+1}^{\bar{f}}\subseteq[x]_i^{\bar{f}},[x]_j^{f}\subseteq[x]_{j+1}^{f}\}$。

推论 1－6，命题 5－7 共同指出：

只要在$[x]$的属性集合 α 内连续地删除属性 α_i，内-动态数据$[x]_i^{\bar{f}}$依次从$[x]$内被挖掘，$[x]_i^{\bar{f}}\subset[x]$；只要在$[x]$的属性集合 α 内连续地补充属性 α_j，外-动态数据$[x]_j^{f}$依次从$[x]$外被挖掘，$[x]\subset[x]_j^{f}$；只要在$[x]$的属性集合 α 内连续地删除属性 α_i 同时在$[x]$的属性集合 α 内连续地补充属性 α_j，内-外动态数据$([x]_i^{\bar{f}},[x]_j^{f})$依次从$[x]$内-外被挖掘；"数据挖掘"具有了动态特性。

第 3 节中的结果给出且讨论于第 4 节中给出。

4 数据推理与动态数据智能挖掘定理

定义 4 给定内-动态数据$[x]_{k+1}^{\bar{f}}$，$[x]_k^{\bar{f}}$；$\alpha_{k+1}^{\bar{f}}$，$\alpha_k^{\bar{f}}$ 分别是$[x]_{k+1}^{\bar{f}}$，$[x]_k^{\bar{f}}$ 的属性集合，若 $\alpha_{k+1}^{\bar{f}}$ 与 $\alpha_k^{\bar{f}}$，$[x]_{k+1}^{\bar{f}}$ 与$[x]_k^{\bar{f}}$ 满足

$$\text{If } \alpha_{k+1}^{\bar{f}}\Rightarrow\alpha_k^{\bar{f}},\text{ then } [x]_{k+1}^{\bar{f}}\Rightarrow[x]_k^{\bar{f}} \tag{25}$$

式(25)称作内-动态数据生成的内-动态数据推理，简称内-数据推理；$\alpha_{k+1}^{\bar{f}}\Rightarrow\alpha_k^{\bar{f}}$ 称作内-数据推理条件，$[x]_{k+1}^{\bar{f}}\Rightarrow[x]_k^{\bar{f}}$ 称作内-数据推理结论。称$[x]_{k+1}^{\bar{f}}$是满足内-数据推理条件 $\alpha_{k+1}^{\bar{f}}\Rightarrow\alpha_k^{\bar{f}}$，$[x]_k^{\bar{f}}$ 生成的智能挖掘。

式(25)中，$\alpha_{k+1}^{\bar{f}}\Rightarrow\alpha_k^{\bar{f}}$ 与 $\alpha_{k+1}^{\bar{f}}\subseteq\alpha_k^{\bar{f}}$ 等价；$[x]_{k+1}^{\bar{f}}\Rightarrow[x]_k^{\bar{f}}$ 与$[x]_{k+1}^{\bar{f}}\subseteq[x]_k^{\bar{f}}$ 等价。

定义 5 给定外-动态数据$[x]_{k+1}^{f}$，$[x]_k^{f}$；α_{k+1}^{f}，α_k^{f} 分别是$[x]_{k+1}^{f}$，$[x]_k^{f}$ 的属性集合，若 α_{k+1}^{f} 与 α_k^{f}，$[x]_{k+1}^{f}$ 与$[x]_k^{f}$ 满足

$$\text{If } \alpha_k^{f}\Rightarrow\alpha_{k+1}^{f},\text{ then } [x]_k^{f}\Rightarrow[x]_{k+1}^{f} \tag{26}$$

式(26)称作外-动态数据生成的外-动态数据推理，简称外-数据推理；$\alpha_k^{f}\Rightarrow\alpha_{k+1}^{f}$ 称作外-数据推理条件，$[x]_k^{f}\Rightarrow[x]_{k+1}^{f}$ 称作外-数据推理结论。称$[x]_{k+1}^{f}$是满足外-数据推理条件 $\alpha_k^{f}\Rightarrow\alpha_{k+1}^{f}$，$[x]_k^{f}$ 生成的智能挖掘。

定义 6 给定内-外动态数据$([x]_{k+1}^{\bar{f}},[x]_k^{f})$，$([x]_k^{\bar{f}},[x]_{k+1}^{f})$；$(\alpha_{k+1}^{\bar{f}},\alpha_k^{f})$，$(\alpha_k^{\bar{f}},\alpha_{k+1}^{f})$分别是$([x]_{k+1}^{\bar{f}},[x]_k^{f})$，$([x]_k^{\bar{f}},[x]_{k+1}^{f})$的属性集合；若$([x]_{k+1}^{\bar{f}},[x]_k^{f})$与$([x]_k^{\bar{f}},[x]_{k+1}^{f})$，$(\alpha_{k+1}^{\bar{f}},\alpha_k^{f})$与$(\alpha_k^{\bar{f}},\alpha_{k+1}^{f})$满足

$$\text{If } (\alpha_{k+1}^{\bar{f}},\alpha_k^{f})\Rightarrow(\alpha_k^{\bar{f}},\alpha_{k+1}^{f}),\text{ then } ([x]_{k+1}^{\bar{f}},[x]_k^{f})\Rightarrow([x]_k^{\bar{f}},[x]_{k+1}^{f}) \tag{27}$$

式(27)称作内-外动态数据生成的内-外动态数据推理，简称内-外数据推理；$(\alpha_{k+1}^{\bar{f}},\alpha_k^{f})\Rightarrow(\alpha_k^{\bar{f}},\alpha_{k+1}^{f})$称作内-外数据推理条件，$([x]_{k+1}^{\bar{f}},[x]_k^{f})\Rightarrow([x]_k^{\bar{f}},[x]_{k+1}^{f})$称作内-外数据推理结论。称$([x]_k^{\bar{f}},[x]_{k+1}^{f})$是满足内-外数据推理条件$(\alpha_{k+1}^{\bar{f}},\alpha_k^{f})\Rightarrow(\alpha_k^{\bar{f}},\alpha_{k+1}^{f})$，$([x]_{k+1}^{\bar{f}},[x]_k^{f})$生成的智能挖掘。

式(27)中，$(\alpha_{k+1}^{\bar{f}},\alpha_k^{f})\Rightarrow(\alpha_k^{\bar{f}},\alpha_{k+1}^{f})$表示 $\alpha_{k+1}^{\bar{f}}\Rightarrow\alpha_k^{\bar{f}}$，$\alpha_k^{f}\Rightarrow\alpha_{k+1}^{f}$，$([x]_{k+1}^{\bar{f}},[x]_k^{f})\Rightarrow([x]_k^{\bar{f}},[x]_{k+1}^{f})$表示$[x]_{k+1}^{\bar{f}}\Rightarrow[x]_k^{\bar{f}}$，$[x]_k^{f}\Rightarrow[x]_{k+1}^{f}$。

由定义 4－6 得到：

定理 6(内-动态数据智能挖掘定理) $[x]_k^{\bar{f}}$ 在$[x]$内被智能挖掘的充分必要条件是存在$[x]^{*}\neq\varnothing$，$\alpha^{*}\neq\varnothing$，$([x]-[x]^{*})$与$[x]$，$([x]-[x]^{*})$的属性集合$(\alpha-\alpha^{*})$与$[x]$的属性集合 α 满足内-数据推理

$$\text{If } (\alpha-\alpha^{*})\Rightarrow\alpha,\text{ then } ([x]-[x]^{*})\Rightarrow[x] \tag{28}$$

数据元 $x_j\in([x]-[x]^{*})$的属性 α_j 与数据元 $x_i\in[x]$的属性 α_i 满足

$$\alpha_j=(\bigvee_{t=1}^{k}\alpha_t)-\bigvee_{t=n+1}^{k}\alpha_t \tag{29}$$

式(28)中，$([x]-[x]^{*})=[x]_k^{\bar{f}}$，$\alpha-\alpha^{*}=\alpha_k^{\bar{f}}$；式(29)中，$\alpha=\{\alpha_1,\alpha_2,\cdots,\alpha_n,\alpha_{n+1},\cdots,\alpha_k\}$，$\alpha^{\bar{f}}=\{\alpha_1,\alpha_2,\cdots,\alpha_n\}$分别是$[x]$，$([x]-[x]^{*})$的属性集合；$\alpha_i=\alpha_1\vee\alpha_2\vee\cdots\vee\alpha_m\vee\alpha_{m+1}\vee\cdots\vee\alpha_k=\bigvee_{t=1}^{k}\alpha_t$ 是 $x_i\in[x]$的属性。

证明：设$[x]_k^{\bar{f}}=([x]-[x]^{*})$，$\alpha_k^{\bar{f}}=\alpha-\alpha^{*}$；$\alpha$，$\alpha_k^{\bar{f}}$ 分别是$[x]$，$[x]_k^{\bar{f}}$ 的属性集合。1. 由定义 4 得到$([x]-[x]^{*})$是$[x]$生成的智能挖掘，则$([x]-[x]^{*})$与$[x]$，$([x]-[x]^{*})$的属性集合$(\alpha-\alpha^{*})$与$[x]$的属性集合 α 满足内-数据推理：If $(\alpha-\alpha^{*})\Rightarrow\alpha$，then $([x]-[x]^{*})\Rightarrow[x]$得到式(28)。2. 若$([x]-[x]^{*})$与$[x]$，$([x]-[x]^{*})$的属性集合$(\alpha-\alpha^{*})$与$[x]$的属性集合 α 满足式(28)：If $(\alpha-\alpha^{*})\Rightarrow\alpha$，then $([x]-[x]^{*})\Rightarrow[x]$，由定义 4 得到：$([x]-[x]^{*})$是$[x]$生成的智能挖掘，定理中式(28)成立。式(29)的证明见定理 3。

定理 7(外-动态数据智能挖掘定理) $[x]_k^{f}$ 在$[x]$外被智能挖掘的充分必要条件是存在$[x]^{\circ}\neq\varnothing$，$\alpha^{\circ}\neq\varnothing$，$([x]-[x]^{\circ})$与$[x]$，$([x]\cup[x]^{\circ})$的属性集合$(\alpha\cup\alpha^{\circ})$与$[x]$的属性集合 α 满足外-数据推理

$$\text{If } \alpha\Rightarrow(\alpha\cup\alpha^{\circ}),\text{ then } [x]\Rightarrow([x]\cup[x]^{\circ}) \tag{30}$$

数据元 $x_j\in([x]\cup[x]^{\circ})$的属性 α_j 与数据元 $x_i\in[x]$的属性 α_i 满足

$$\alpha_j=(\bigvee_{t=1}^{k}\alpha_t)\bigvee_{t=k+1}^{m}\alpha_t \tag{31}$$

式(30)中，$([x]\cup[x]^{\circ})=[x]_k^{f}$，$\alpha\cup\alpha^{\circ}=\alpha_k^{f}$；式(31)中，$\alpha=\{\alpha_1,\alpha_2,\cdots,\alpha_k\}$，$\alpha^{f}=\{\alpha_1,\alpha_2,\cdots,\alpha_k,\alpha_{k+1},\cdots,\alpha_m\}$分别是$[x]$，$([x]\cup[x]^{\circ})$的属性集合；$\alpha_i=\alpha_1\vee\alpha_2\vee\cdots\vee\alpha_k=\bigvee_{t=1}^{k}\alpha_t$ 是 $x_i\in[x]$的属性。

证明：设$[x]_k^{f}=([x]\cup[x]^{\circ})$，$\alpha_k^{f}=\alpha\cup\alpha^{\circ}$；$\alpha$，$\alpha_k^{f}$ 分别是$[x]$，$[x]_k^{f}$ 的属性集合。1. 由定义 5 得到$([x]\cup[x]^{\circ})$是$[x]$生成的智能挖掘，则$([x]\cup[x]^{\circ})$与$[x]$，$([x]\cup[x]^{\circ})$的属性集合$(\alpha\cup\alpha^{\circ})$与$[x]$的属性集合 α 满足外-数据推理：If $\alpha\Rightarrow(\alpha\cup\alpha^{\circ})$，then $[x]\Rightarrow([x]\cup[x]^{\circ})$，得到式(30)。2. 若$([x]\cup[x]^{\circ})$与$[x]$，$([x]\cup[x]^{\circ})$的属性集合$(\alpha\cup\alpha^{\circ})$与$[x]$的属性集合 α 满足式(30)：If $\alpha\Rightarrow(\alpha\cup\alpha^{\circ})$，then $[x]\Rightarrow([x]\cup[x]^{\circ})$，由定义 5 得到：$([x]\cup[x]^{\circ})$是$[x]$生成的智能挖掘，定理中式(30)成立。式(31)的证明见定理 4。

由定理 6，7 直接得到：

定理 8(内-外动态数据智能挖掘定理) $([x]_k^{\bar{f}},[x]_k^{f})$在$[x]$内-外被智能挖掘的充分必要条件是存在$([x]^{*},[x]^{\circ})\neq\varnothing$，$(\alpha^{*},\alpha^{\circ})\neq\varnothing$，$([x]-[x]^{*},[x]\cup[x]^{\circ})$与$[x]$，$([x]-$

$[x]^*, [x]\cup[x]^\circ$)的属性集合($\alpha-\alpha^*, \alpha\cup\alpha^\circ$)与$[x]$的属性集合 α 满足内-外数据推理

If $(\alpha-\alpha^*, \alpha)\Rightarrow(\alpha, \alpha\cup\alpha^\circ)$, then $([x]-[x]^*, [x])\Rightarrow([x], [x]\cup[x]^\circ)$ (32)

数据元 $x_i\in([x]-[x]^*)$的属性 α_i 与数据元 $x_\lambda\in[x]$的属性 α_λ,数据元 $x_j\in([x]\cup[x]^\circ)$的属性 α_j 满足

$$(\alpha_i, \alpha_j)=((\bigvee_{t=1}^{k}\alpha_t)-\bigvee_{t=n+1}^{k}\alpha_t, (\bigvee_{t=1}^{k}\alpha_t)\bigvee_{t=k+1}^{m}\alpha_t) \quad (33)$$

推论 7 具有属性析取范式萎缩特征的内-动态数据 $[x]_i^{\bar f}$ 构成的内-动态数据族$\{[x]_i^{\bar f}\mid i\in I, [x]_{i+1}^{\bar f}\subseteq[x]_i^{\bar f}\}$是内-数据推理族$\{(\text{If } \alpha_{k+1}^{\bar f}\Rightarrow\alpha_k^{\bar f}, \text{then } [x]_{k+1}^{\bar f}\Rightarrow[x]_k^{\bar f})_i\mid i\in I\}$的智能挖掘生成;$[x]_k^{\bar f}-[x]_{k+1}^{\bar f}\neq\varnothing$。

推论 8 具有属性析取范式扩张特征的外-动态数据 $[x]_j^f$ 构成的外-动态数据族$\{[x]_j^f\mid j\in J, [x]_j^f\subseteq[x]_{j+1}^f\}$是外-数据推理族$\{(\text{If } \alpha_k^f\Rightarrow\alpha_{k+1}^f, \text{then } [x]_k^f\Rightarrow[x]_{k+1}^f)_j\mid j\in J\}$的智能挖掘生成;$[x]_{k+1}^f-[x]_k^f\neq\varnothing$。

推论 9 具有属性析取范式萎缩-扩张特征的内-外动态数据($[x]_i^{\bar f}, [x]_j^f$)构成的内-外动态数据族$\{([x]_i^{\bar f}, [x]_j^f)_k\mid k\in K, [x]_{i+1}^{\bar f}\subseteq[x]_i^{\bar f}, [x]_j^f\subseteq[x]_{j+1}^f\}$是内-外数据推理族$\{(\text{If } (\alpha_{k+1}^{\bar f}, \alpha_k^f)\Rightarrow(\alpha_k^{\bar f}, \alpha_{k+1}^f), \text{then } ([x]_{k+1}^{\bar f}, [x]_k^f)\Rightarrow([x]_k^{\bar f}, [x]_{k+1}^f))_t\mid t\in T\}$的智能挖掘生成;$[x]_k^{\bar f}-[x]_{k+1}^{\bar f}\neq\varnothing, [x]_{k+1}^f-[x]_k^f\neq\varnothing$。

命题 8 内-数据推理族$\{(\text{If } \alpha_{k+1}^{\bar f}\Rightarrow\alpha_k^{\bar f}, \text{then } [x]_{k+1}^{\bar f}\Rightarrow[x]_k^{\bar f})_i\mid i\in I\}$生成内-动态数据智能挖掘序列$[x]_n^{\bar f}\Rightarrow[x]_{n-1}^{\bar f}\subseteq\cdots\subseteq[x]_2^{\bar f}\subseteq[x]_1^{\bar f}$。

命题 9 外-数据推理族$\{(\text{If } \alpha_k^f\Rightarrow\alpha_{k+1}^f, \text{then } [x]_k^f\Rightarrow[x]_{k+1}^f)_j\mid j\in J\}$生成外-动态数据智能挖掘序列$[x]_1^f\subseteq[x]_2^f\subseteq\cdots\subseteq[x]_{n-1}^f\subseteq[x]_n^f$。

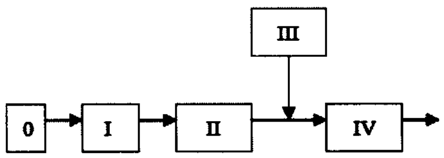
命题 10 内-外数据推理族$\{(\text{If } (\alpha_{k+1}^{\bar f}, \alpha_k^f)\Rightarrow(\alpha_k^{\bar f}, \alpha_{k+1}^f), \text{then } ([x]_{k+1}^{\bar f}, [x]_k^f)\Rightarrow([x]_k^{\bar f}, [x]_{k+1}^f))_t\mid t\in T\}$生成内-外动态数据智能挖掘序列$([x]_n^{\bar f}, [x]_1^f)\subseteq([x]_{n-1}^{\bar f}, [x]_2^f)\subseteq\cdots\subseteq([x]_2^{\bar f}, [x]_{n-1}^f)\subseteq([x]_1^{\bar f}, [x]_n^f)$。

利用 3,4 节的结果,5 节中给出。

5 动态数据智能挖掘-智能认知与应用

这一节中只给出内-动态数据的动态智能挖掘与应用,应用例子取自计算机数据采集-分检系统,系统的简单工作过程是:系统的输入端具有多个传感器,接收来自系统前端信号,传感器对这些信号依据给定的规则,不加选择地对这些信号进行转换,生成数据元 x_i,x_i 构成数据$[x]$;数据$[x]$内有用的数据元 x_i,无用的数据元 x_j 相互混杂。在数据阈值 λ_i 的限定下,通过把$[x]$内的冗余数据元 x_i 删除-过滤得到有用的数据$[x]_i^{\bar f}$,或者,$[x]_i^{\bar f}$ 从$[x]$内被挖掘,对$[x]_i^{\bar f}$ 给出认知。

图 1 给出了计算机数据自动采集-分检系统的简化框图。



0 是传感器;I 是数据元 x_i 生成单元;II 构成数据$[x]$单元;III 数据分检设定入库;IV 数据元删除-过滤与数据$[x]_i^{\bar f}$ 生成-$[x]_i^{\bar f}$ 认知单元

图 1 数据自动采集-分检系统简化框图

为了讨论的简化,又不引起混乱,也不失应用的一般性;数据元 x_i 的属性 α_i 用 x_i 的特征值(传感器的输出值)$y_i\in R^+$ 代替;或者,$\alpha_i=y_i$;II 给出数据$[x]$,与$[x]$的属性集合 α

$$[x]=\{x_1, x_2, x_3, x_4, x_5, x_6, x_7, x_8\} \quad (34)$$

$$\alpha=\{\alpha_1, \alpha_2, \alpha_3, \alpha_4, \alpha_5, \alpha_6, \alpha_7, \alpha_8\}=\{1.63, 1.78, 1.21, 1.69, 1.54, 1.63, 1.10, 1.47\} \quad (35)$$

显然,$\forall x_i\in[x]$的属性 α_i 满足属性析取范式:

$$\alpha_i=\alpha_1\vee\alpha_2\vee\alpha_3\vee\alpha_4\vee\alpha_5\vee\alpha_6\vee\alpha_7\vee\alpha_8=\bigvee_{t=1}^{8}\alpha_t$$

取 $\lambda_1\geqslant1.60$,$\lambda_2\geqslant1.40$,由式(34),式(35)分别得到$[x]_1^{\bar f}$,$[x]_2^{\bar f}$以及$[x]_1^{\bar f}$,$[x]_2^{\bar f}$ 的属性集合 $\alpha_1^{\bar f}$,$\alpha_2^{\bar f}$,而且

$$[x]_1^{\bar f}=\{x_1, x_2, x_4, x_6\}_{\lambda_1} \quad (36)$$

$$\alpha_1^{\bar f}=\{\alpha_1, \alpha_2, \alpha_4, \alpha_6\}=\{1.63, 1.78, 1.69, 1.63\} \quad (37)$$

$$[x]_2^{\bar f}=\{x_1, x_2, x_4, x_5, x_6, x_8\}_{\lambda_2} \quad (38)$$

$$\alpha_2^{\bar f}=\{\alpha_1, \alpha_2, \alpha_4, \alpha_5, \alpha_6, \alpha_8\}_{\lambda_2}=\{1.63, 1.78, 1.69, 1.54, 1.63, 1.47\} \quad (39)$$

在 λ_1,λ_2,$\lambda_2<\lambda_1$ 的条件下,数据$[x]$被分离成$[x]_1^{\bar f}$,$[x]_2^{\bar f}$;$\forall x_i\in[x]_1^{\bar f}$ 的属性 α_i,$\forall x_j\in[x]_2^{\bar f}$ 的属性 α_j 分别是 $\forall x_k\in[x]$的属性 α_k 的属性析取范式萎缩;$[x]_1^{\bar f}$ 与$[x]$,$[x]_1^{\bar f}$ 的属性集合 $\alpha_1^{\bar f}$ 与$[x]$的属性集合 α;$[x]_2^{\bar f}$ 与$[x]$,$[x]_2^{\bar f}$ 的属性集合 $\alpha_2^{\bar f}$ 与$[x]$的属性集合 α 分别满足内-数据推理:

$$\begin{aligned}&\text{If } \alpha_1^{\bar f}\Rightarrow\alpha, \text{then } [x]_1^{\bar f}\Rightarrow[x]\\&\text{If } \alpha_2^{\bar f}\Rightarrow\alpha, \text{then } [x]_2^{\bar f}\Rightarrow[x]\end{aligned} \quad (40)$$

满足式(40)的$[x]_1^{\bar f}$,$[x]_2^{\bar f}$ 分别连续调用内-数据推理,在$[x]$内被动态智能挖掘;$[x]_1^{\bar f}$,$[x]_2^{\bar f}$ 依据$[x]$被认知;或者,$IDE([x]_1^{\bar f}, [x])$,$IDE([x]_2^{\bar f}, [x])$;IDE=identification。

6 讨论

Z. Pawlak 粗集是研究数据挖掘与知识发现的一个重要的理论工具与数学方法;Z. Pawlak 粗集解决了在静态条件下,静态数据挖掘的多个应用问题。数据$[x]$与数据$[x]$的属性集合 α 构成数据的整体概念;或者,数据$[x]$与数据$[x]$的属性集合 α 成对的出现,这个现象在众多的实际应用问题中经常被遇到,人们自然要问:从数据$[x]$中挖掘(寻找)未知的数据$[x]^*$,数据元 $x_i\in[x]^*$ 的属性 α_i 与数据元 $x_j\in[x]$的属性 α_j 之间满足什么样的属性逻辑关系?本文把数理逻辑中的"析取"概念引入到数据挖掘研究中,给出数据挖掘中的"属性析取"结构,给出数据挖掘与属性析取的联系。数据挖掘中的属性析取结构没有引起人们的太多注意。

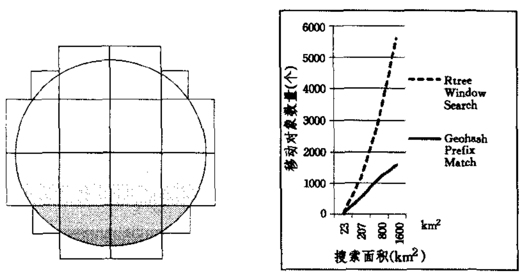
事实上,静态数据并不是数据的全部,动态数据却是数据的真实面貌,静态数据只是数据的一种特殊形式。利用 Z. Pawlak粗集研究具有动态特征的数据与它的动态挖掘遇到困难,理由是:Z. Pawlak 粗集不具有动态特征。本文利用具有动态特征的单向 S-粗集、单向 S-粗集对偶与它们的整合,给出动态数据与它的动态挖掘,使得挖掘模型(单向 S-粗集、单向 S-粗集对偶)与动态数据的动态特征相匹配。单向 S-粗集、单向 S-粗集对偶是改进 Z. Pawlak 粗集被提出的。动态数据智能挖掘依赖于推理,推理表达智能,智能来自推理;利用单向 S-粗集、单向 S-粗集对偶提出数据推理,给出推理结构。把数据推理与动态数据的动态挖掘交叉、共享,给出动态数据的动态智能挖掘与应用,使得数据挖掘智能化。本文给出的讨论可以移植到其他问题的应用研究中。

参 考 文 献

[1] Shi Kai-quan. S-rough sets and its applications in diagnosis-rec-

ognition for disease [J]. IEEE Proceedings of the First International Conference on Machine Learning and Cybernetics, 2002, 1(4): 50-54
[2] 史开泉,崔玉泉. S-粗集与它的结构 [J]. 山东大学学报:理学版,2002,37(6):471-474
[3] Shi Kai-quan, Cui Yu-quan. F-decomposition and $\bar{F}$-reduction of S-rough sets [J]. An International Journal of Advances in Systems Science and Applications, 2004, 4(4): 487-477
[4] Shi Kai-quan, Chang Ting-cheng. One direction S-rough sets [J]. International Journal of Fuzzy Mathematics, 2005, 13(2): 319-334
[5] Shi Kai-quan. Two directions S-rough sets [J]. International Journal of Fuzzy Mathematics, 2005, 13(2): 335-347
[6] 史开泉,崔玉泉. S-粗集与它的分解-还原 [J]. 系统工程与电子技术,2005,27(4):644-651
[7] Hu Hai-qing, Wang Hong-yu, Shi Kai-quan. Two directions S-rough recognition of knowledge and recognition model [J]. An International Journal of Advances in Systems Science and Applications, 2005, 5(3): 368-374
[8] 史开泉. S-粗集与新材料发现-识别 [J]. 系统工程与电子技术,2006,28(3):383-388
[9] Shi Kai-quan. Function S-rough sets and function transfer [J]. An International Journal of Advances in Systems Science and Applications, 2005, 5(1): 1-8
[10] 史开泉. 函数 S-粗集 [J]. 山东大学学报:理学版,2005,40(1):1-10
[11] 张萍,史开泉,卢昌荆. 函数 S-粗集与粗规律-分离 [J]. 系统工程与电子技术,2005,27(11):1899-1902
[12] Zhang Ling, Shi Kai-quan. Security transmission and recognition of F-knowledge [J]. Journal of Systems Engineering and Electronics, 2009, 20(4): 877-882
[13] Qiu Jin-ming, Shi Kai-quan. F-rough law and the discovery of rough law [J]. Journal of Systems Engineering and Electronics, 2009, 20(1): 81-89
[14] 史开泉,姚炳学. 函数 S-粗集与规律辨识 [J]. 中国科学 E:信息科学,2008,38(4):553-564
[15] 史开泉,赵建立. 函数 S-粗集与隐藏规律安全-认证 [J]. 中国科学 E:信息科学,2008,38(8):1234-1243
[16] Shi Kai-quan, Yao Bing-xue. Function S-rough sets and law identification [J]. Science in China F: Information Sciences, 2008, 51(5): 499-510
[17] Shi Kai-quan, Zhao Jian-li. Function S-rough sets and security-authentication of hiding law [J]. Science in China F: Information Sciences, 2008, 51(7): 924-935
[18] 史开泉. 函数 S-粗集,函数粗集与信息系统规律拆分-合成 [J]. 计算机科学,2010,37(10):1-10
[19] 史开泉. S-粗集与数据挖掘单位圆特征 [J]. 计算机科学,2010,37(5):1-8
[20] Yoon Y, Lee J, Park S, et al. Direct integration of microarrays for selecting informative gene and phenotype classification [J]. Information Science, 2008, 178(12): 88-105
[21] Vogiatzis D, Tsapatsoulis N. Active learning for microarray data [J]. International Journal of Approximate Reasoning, 2008, 47(1): 85-96
[22] Pawlak Z. Rough set approach to knowledge-based decision support [J]. European Journal of Operational Research, 1997, 87(99): 48-57
[23] Pawlak Z. Rough sets and fuzzy sets [J]. Fuzzy Sets and Systems, 1985, 79(17): 99-102
[24] Pawlak Z. Rough sets [J]. International Journal of Computer and Information Sciences, 1982, 66(11): 341-356
[25] 史开泉,崔玉泉. S-粗集与粗决策 [M]. 北京:科学出版社,2006,155-165
[26] 史开泉,姚炳学. 函数 S一粗集与系统规律挖掘[M]. 北京:科学出版社,2007:147一198

(上接第 214 页)



(a)分割方法　　(b)范围查询误差

图 6

结束语　本文在 FNR 索引结构的基础上提出了 RNR 索引结构,通过增加辅助索引结构,使得索引结构在更新操作、范围查询、轨迹查询等操作上的效率有大大提升,并且融合 Geohash 算法,使得索引结构能够提供高效的最近邻查询和窗口查询等功能。性能评估实验中使用能够提供真实经纬度数据的 San Francisco 地图数据,验证了近邻查询效果,也对比验证了 Geohash 算法的效率。但是研究中对道路之间的关系考虑不多,下一步将着重研究道路之间的关系,以进一步完善索引结构。

参 考 文 献

[1] 廖巍,熊伟,景宁. 移动对象索引技术研究进展[J]. 计算机科学,2006,33(8):166-169
[2] 肖晖,李清泉. 移动对象数据库索引研究综述[J]. 计算机应用,2010,30(4):1064-1067,1071
[3] Yu X, Chen Y. A moving object database model based on road network [J]. Journal of Software, 2003, 1498(9): 1600-1607
[4] Chen J, Meng X. Update-efficient indexing of moving objects in road networks[J]. GeoInformatica, 2009, 13(4): 397-424
[5] 曾倩,金敏. 基于道路分布的移动对象动态组合索引方法[J]. 计算机应用,2008,28(1):3251- 3253
[6] 李峰,罗磊. 基于道路网络的时空索引方法 IMon-tree[J]. 计算机应用,2012,32(8):2205-2208
[7] Kyoung-S Kim. Fast indexing and updating method for moving objects on road networks[C]//Proc. 4th Int'l Conf. Web Information Systems Engineering. Los Alamitos, CA: IEEE Computer Society Press, 2003: 34-42
[8] 丁治明. 一种适合于频繁位置更新的网络受限移动对象轨迹索引[J]. 计算机学报,2012,35(7):1448-1461
[9] 宋广军,郝忠孝,王丽杰. 一种基于受限网络的移动对象索引[J]. 计算机科学,2009,36(12):138-141
[10] Zhu Y, Zheng V, Yang Q. Computing with Spatial Trajectories [M]. New York: Springer New York, 2011
[11] Brinkhoff T. Generating network-based moving objects [C]// Proc of the 12th Int'l Conf. on Scientific and Statistical Database Management(SSDBM00). 2000: 253-255