计算机科学 COMPUTER SCIENCE
http://www.jsjkx.com
DOI:10.11896/jsjkx.191000071

大数据智能检索与大数据区块元智能分离

郝秀梅[1]　史开泉[2]
1 山东财经大学数学与数量经济学院　济南 250014
2 山东大学数学学院　济南 250100
(hxm0912@126.com)

摘　要　文中利用逆 P-集合生成∨型大数据结构,给出∨型大数据的新概念,如大数据区块、区块矩阵、区块元、区块元矩阵与数据元概念;利用这些概念给出区块属性推理结构、区块矩阵推理结构、区块元智能分离定理、区块元智能检索定理、区块与区块元等价类定理;给出区块元智能分离准则、区块智能检索准则;给出区块元智能分离-区块智能检索算法与算法过程;给出大数据智能检索-大数据区块元智能分离-获取的应用。∨型大数据满足"属性析取"的逻辑特征。

关键词:∨型大数据;大数据区块;属性-矩阵推理;智能检索准则;智能检索算法;应用
中图法分类号　O144,TP391

Big Data Intelligent Retrieval and Big Data Block Element Intelligence Separation

HAO Xiu-mei[1] and SHI Kai-quan[2]
1 School of Mathematics and Quantitative Economics,Shandong University of Finance and Economics,Jinan 250014,China
2 School of Mathematics,Shandong University,Jinan 250100,China

Abstract　By using∨-type big data structure generated by inverse P-sets,some new concepts of ∨-type big data are given,such as big data block,block matrix,block element,block element matrix and data element. Based on these concepts,the reasoning structure of block attribute,the reasoning structure of block matrix,the intelligent separation theorems,the intelligent retrieval theorems of block element and the equivalence class theorems of block and block element are given. An intelligent separation criterion of block element and an intelligent retrieval criterion of block are presented. A block element intelligent separation-block intelligent retrieval algorithm and its algorithm process are given. The application of big data intelligent retrieval-big data block element intelligent separation-acquisition is given. ∨-type big data satisfies the logical characteristic of 'attribute disjunction'.

Keywords　∨-type big data,Big data block,Attribute-matrix reasoning,Intelligent retrieval criterion,Intelligent retrieval algorithm,Application

1　引言

文献[1]利用 P-集合模型、逆 P-集合模型(P=Packet)给出两类大数据的结构、基本特征与应用的研究。利用 P-集合、逆 P-集合研究两类大数据的理由与依据是:1)无论来自哪个应用领域中的大数据都具有属性(大数据的特征);或者,大数据与它的属性相伴存在,不同应用领域中的大数据的属性是不同的。2)大数据具有动态特征(大数据中的数据元动态变化);在一定的条件下,一些数据元从大数据中被删除,一些数据元被补充到大数据内。3)在两类大数据中,一类大数据的数据元 x_i 的属性 α_i 满足数理逻辑中的属性"合取范式";另一类大数据的数据元 x_j 的属性 α_j 满足数理逻辑中的属性"析取范式"。4)在两类大数据中,一类大数据的数据元 x_k 从大数据内被删除-补充,另一类大数据的数据元 x_λ 被补充-删除到大数据内,它们各自具有概率特征。5)以时间 $t\in T$ 为参考量,两类大数据跟随 $t\in T$ 的变化而变化;静止不变的大数据只是大数据的特定状态。1)-4)在大数据状态分析、识别、大数据分离、未知大数据检索发现等诸多应用问题研究中都毫无例外地被遇到,1)-4)与 P-集合[2-34]、逆 P-集合[1,35-52]的基本特征相吻合。如果定义有限普通元素集合 X 为大数据的基数据(基础数据),P-集合($X^{\bar{F}},X^F$)、逆 P-集合($\overline{X}^F,\overline{X}^{\bar{F}}$)分别定义成两类大数据,则 P-集合、逆 P-集合可被用作研究两类大数据的基础理论模型与数学方法,从而获得两类大数据的多个应用特征。无论哪一类大数据,它们在应用中都是以族的形式出现。P-集合、逆 P-集合的区别是:在 P-集合中,元素 x_i 的属性满足"属性合取"特征;在逆 P-集合中,元素 x_j

到稿日期:2019-10-12　返修日期:2020-03-27　　本文已加入开放科学计划(OSID),请扫描上方二维码获取补充信息。
基金项目:国家社会科学基金项目(71663010);山东省自然科学基金项目(zr2013aq019)
This work was supported by the National Social Science Foundation of China (71663010) and Natural Science Foundation of Shandong Province,China (zr2013aq019).
通信作者:史开泉(shikq@sdu.edu.cn)

的属性满足“属性析取”特征。P-集合生成∧型大数据，逆P-集合生成∨型大数据，“∧”是合取算子，“∨”是析取算子。P-集合、逆P-集合是把动态特性引入有限普通元素集合 X 内，改进有限普通元素集合 X 后得到的。

本文是文献[1]中∨型大数据研究的继续，主题是挖掘潜藏在∨型大数据中还未被认识的特征。本文的主要贡献如下：1)引入逆P-集合的结构和特征、逆P-增广矩阵、∨型大数据的结构作为本文知识准备与预备概念；2)给出大数据区块概念与区块矩阵的等价性；3)给出区块属性推理与区块元智能分离、区块元智能检索-过滤；4)给出 $\overline{F}$-区块元智能分离-区块智能检索算法；5)给出大数据智能检索与大数据区块元智能分离-获取的应用，应用例子取自公共安全系统。

2 概念准备

2.1 逆P-集合与它的动态特征

给定有限普通元素集合 $X=\{x_1,x_2,\cdots,x_q\}\subset U$，$\alpha=\{\alpha_1,\alpha_2,\cdots,\alpha_q\}\subset V$ 是 X 的属性集合，称 $\overline{X}^F$ 是 X 生成的内逆P-集合(Internal Inverse Packet Set)，简称 $\overline{X}^F$ 是内逆P-集合。

$$\overline{X}^F=X\cup X^+ \tag{1}$$

X^+ 称作 X 的 F-元素补充集合。

$$X^+=\{u_i\,|\,u_i\in U,u_i\overline{\in}X,f(u_i)=x_i{}'\in X,f\in F\} \tag{2}$$

如果 α^F 是 $\overline{X}^F$ 的属性集合，则：

$$\alpha^F=\alpha\cup\{\beta_i\,|\,f(\beta_i)=\alpha_i{}'\in\alpha,f\in F\} \tag{3}$$

式(3)中，$\beta_i\in V$，$\beta_i\overline{\in}\alpha$，$f\in F$ 把 β_i 生成 $f(\beta_i)=\alpha'$，$\alpha'\in\alpha$；式(1)中，$\overline{X}^F=\{x_1,x_2,\cdots,x_r\}$，$q<r$，$q,r\in\mathbf{N}^+$。

给定有限普通元素集合 $X=\{x_1,x_2,\cdots,x_q\}\subset U$，$\alpha=\{\alpha_1,\alpha_2,\cdots,\alpha_q\}\subset V$ 是 X 的属性集合，称 $\overline{X}^{\overline{F}}$ 是 X 生成的外逆P-集合(Outer Inverse Packet Set)，简称 $\overline{X}^{\overline{F}}$ 是外逆P-集合。

$$\overline{X}^{\overline{F}}=X-X^- \tag{4}$$

X^- 称作 X 的 $\overline{F}$-元素删除集合。

$$X^-=\{x_i\,|\,x_i\in X,\overline{f}(x_i)=u_i\overline{\in}X,\overline{f}\in\overline{F}\} \tag{5}$$

如果 $\alpha^{\overline{F}}$ 是 $\overline{X}^{\overline{F}}$ 的属性集合，则：

$$\alpha^{\overline{F}}=\alpha-\{\alpha_i\,|\,\overline{f}(\alpha_i)=\beta_i\overline{\in}\alpha,\overline{f}\in\overline{F}\} \tag{6}$$

式(6)中，$\alpha_i\in\alpha$，$\overline{f}\in\overline{F}$ 把 α_i 生成 $\overline{f}(\alpha_i)=\beta_i$，$\beta_i\overline{\in}\alpha$；式(4)中，$\overline{X}^{\overline{F}}=\{x_1,x_2,\cdots,x_p\}$，$p<q$，$p,q\in\mathbf{N}^+$；式(4)中，$\overline{X}^{\overline{F}}\neq\emptyset$，式(6)中，$\alpha^{\overline{F}}\neq\emptyset$。

由内逆P-集合 $\overline{X}^F$ 与外逆P-集合 $\overline{X}^{\overline{F}}$ 构成的有限普通元素集合对，称作 X 生成的逆P-集合(Inverse Packet Sets)，简称逆P-集合。

$$(\overline{X}^F,\overline{X}^{\overline{F}}) \tag{7}$$

有限普通元素集合 X 称作逆P-集合 $(\overline{X}^F,\overline{X}^{\overline{F}})$ 的基集合(基础集合)。

在 α 内不断补充属性，由式(3)给出：

$$\alpha_1^F\subseteq\alpha_2^F\subseteq\cdots\subseteq\alpha_{n-1}^F\subseteq\alpha_n^F \tag{8}$$

满足式(8)的内逆P-集合为：

$$\overline{X}_1^F\subseteq\overline{X}_2^F\subseteq\cdots\subseteq\overline{X}_{n-1}^F\subseteq\overline{X}_n^F \tag{9}$$

在 α 内不断删除属性，由式(6)给出：

$$\alpha_n^{\overline{F}}\subseteq\alpha_{n-1}^{\overline{F}}\subseteq\cdots\subseteq\alpha_2^{\overline{F}}\subseteq\alpha_1^{\overline{F}} \tag{10}$$

满足式(10)的外逆P-集合为：

$$\overline{X}_n^{\overline{F}}\subseteq\overline{X}_{n-1}^{\overline{F}}\subseteq\cdots\subseteq\overline{X}_2^{\overline{F}}\subseteq\overline{X}_1^{\overline{F}} \tag{11}$$

由式(9)、式(11)得到：

$$\{(\overline{X}_i^F,\overline{X}_j^{\overline{F}})\,|\,i\in I,j\in J\} \tag{12}$$

式(12)称作 X 生成的逆P-集合族，式(12)是逆P-集合的一般形式。

逆P-集合 $(\overline{X}^F,\overline{X}^{\overline{F}})$ 是把动态特征引入有限普通元素集合 X 内得到的一个具有动态特性的新集合模型，在什么条件下，逆P-集合 $(\overline{X}^F,\overline{X}^{\overline{F}})$ 被还原成 X？具体结果如命题1、命题2所示。

命题1 在 $F=\overline{F}=\emptyset$ 的条件下，逆P-集合 $(\overline{X}^F,\overline{X}^{\overline{F}})$ 被还原成有限普通元素集合 X，或者：

$$(\overline{X}^F,\overline{X}^{\overline{F}})_{F=\overline{F}=\emptyset}=X \tag{13}$$

命题2 在 $F=\overline{F}=\emptyset$ 的条件下，逆P-集合族 $\{(\overline{X}_i^F,\overline{X}_j^{\overline{F}})\,|\,i\in I,j\in J\}$ 被还原成有限普通元素集合 X，或者：

$$\{(\overline{X}_i^F,\overline{X}_j^{\overline{F}})\,|\,i\in I,j\in J\}_{F=\overline{F}=\emptyset}=X \tag{14}$$

命题1、命题2的证明由式(1)－式(7)、式(12)直接得到，证明略。

逆P-集合的更多特征、应用、逆P-集合存在的事实，以及逆P-集合的逻辑特征请见文献[1，22-41]。

2.2 逆P-增广矩阵结构与生成

给定有限普通元素集合 $X=\{x_1,x_2,\cdots,x_q\}$，$\forall x_i\in X$ 具有 n 个元素值 $y_{1,i},y_{2,i},\cdots,y_{n,i}$；$\boldsymbol{y}_i=(y_{1,i},y_{2,i},\cdots,y_{n,i})^{\mathrm{T}}$ 是 $y_{1,i},y_{2,i},\cdots,y_{n,i}$ 构成的向量，以 $\boldsymbol{y}_i$ 为列得到矩阵 $\boldsymbol{A}$，$\boldsymbol{A}$ 称作 X 生成的元素值矩阵。

$$\boldsymbol{A}=\begin{pmatrix} y_{1,1} & y_{1,2} & \cdots & y_{1,q}\\ y_{2,1} & y_{2,2} & \cdots & y_{2,q}\\ \vdots & \vdots & & \vdots\\ y_{n,1} & y_{n,2} & \cdots & y_{n,q}\end{pmatrix} \tag{15}$$

给定内逆P-集合 $\overline{X}^F=\{x_1,x_2,\cdots,x_r\}$，$\overline{\boldsymbol{A}}^F$ 称作 $\overline{X}^F$ 生成的 $\boldsymbol{A}$ 的内逆P-增广矩阵。

$$\overline{\boldsymbol{A}}^F=\begin{pmatrix} y_{1,1} & y_{1,2} & \cdots & y_{1,r}\\ y_{2,1} & y_{2,2} & \cdots & y_{2,r}\\ \vdots & \vdots & & \vdots\\ y_{n,1} & y_{n,2} & \cdots & y_{n,r}\end{pmatrix} \tag{16}$$

给定外逆P-集合 $\overline{X}^{\overline{F}}=\{x_1,x_2,\cdots,x_p\}$，$\overline{\boldsymbol{A}}^{\overline{F}}$ 称作 $\overline{X}^{\overline{F}}$ 生成的 $\boldsymbol{A}$ 的外逆P-增广矩阵。

$$\overline{\boldsymbol{A}}^{\overline{F}}=\begin{pmatrix} y_{1,1} & y_{1,2} & \cdots & y_{1,p}\\ y_{2,1} & y_{2,2} & \cdots & y_{2,p}\\ \vdots & \vdots & & \vdots\\ y_{n,1} & y_{n,2} & \cdots & y_{n,p}\end{pmatrix} \tag{17}$$

由内逆P-增广矩阵 $\overline{\boldsymbol{A}}^F$ 与外逆P-增广矩阵 $\overline{\boldsymbol{A}}^{\overline{F}}$ 构成的矩阵对，称作逆P-集合 $(\overline{X}^F,\overline{X}^{\overline{F}})$ 生成的 $\boldsymbol{A}$ 的逆P-增广矩阵。

$$(\overline{\boldsymbol{A}}^F,\overline{\boldsymbol{A}}^{\overline{F}}) \tag{18}$$

称

$$\{(\overline{\boldsymbol{A}}_i^F,\overline{\boldsymbol{A}}_j^{\overline{F}})\,|\,i\in I,j\in J\} \tag{19}$$

是逆P-集合族 $\{(\overline{X}_i^F,\overline{X}_j^{\overline{F}})\,|\,i\in I,j\in J\}$ 生成的 $\boldsymbol{A}$ 的逆P-增广矩阵族，式(19)是逆P-增广矩阵的一般形式。

式(15)－式(17)中，$p<q<r$，$p,q,r\in\mathbf{N}^+$。

特别说明：1)在普通数学中，$\boldsymbol{A}$ 是 $m\times n$ 阶矩阵，如果在 n 列中增加 t 列，$\boldsymbol{A}_{m\times n}$ 生成 $\boldsymbol{A}_{m\times(n+t)}$，$\boldsymbol{A}_{m\times(n+t)}$ 称作 $\boldsymbol{A}_{m\times n}$ 的增广矩阵，$\boldsymbol{A}_{m\times(n+t)}$ 记作 $\boldsymbol{A}^*$，$\boldsymbol{A}^*$ 是 $\boldsymbol{A}$ 的增广矩阵。如果在 n 列中删去 λ 列，$\boldsymbol{A}_{m\times n}$ 生成 $\boldsymbol{A}_{m\times(n-\lambda)}$，$\lambda<n$。$\boldsymbol{A}_{m\times(n-\lambda)}$ 是否是 $\boldsymbol{A}$ 的增广矩阵？如果 $\boldsymbol{A}_{m\times(n-\lambda)}$ 与 $\boldsymbol{A}_{m\times(n+t)}$ 构成一个矩阵对 $(\boldsymbol{A}_{m\times(n-\lambda)},\boldsymbol{A}_{m\times(n+t)})$，

$(\boldsymbol{A}_{m\times(n-\lambda)},\boldsymbol{A}_{m\times(n+t)})$ 是否是 $\boldsymbol{A}$ 的增广矩阵? $\boldsymbol{A}_{m\times(n-\lambda)}$,$(\boldsymbol{A}_{m\times(n-\lambda)},\mathrm{A}_{m\times(n+t)})$ 在普通数学中找不到它们的定义;$\boldsymbol{A}_{m\times(n-\lambda)}$,$(\boldsymbol{A}_{m\times(n-\lambda)},\boldsymbol{A}_{m\times(n+t)})$ 在大数据智能分析-识别、信息智能变换中经常出现。根据这样的应用背景,文献[30]利用P-集合的结构与动态特征提出P-增广矩阵;文献[35]利用逆P-集合的结构与动态特征提出逆P-增广矩阵。P-增广矩阵、逆P-增广矩阵为两类大数据的智能分析提供新的数学工具。2) P-增广矩阵由内P-增广矩阵 $\boldsymbol{A}^{\bar{F}}_{m\times(n-\lambda)}$ 与外P-增广矩阵 $\boldsymbol{A}^{F}_{m\times(n+t)}$ 构成,或者 $(\boldsymbol{A}^{\bar{F}}_{m\times(n-\lambda)},\boldsymbol{A}^{F}_{m\times(n+t)})$ 是P-增广矩阵;逆P-增广矩阵由内逆P-增广矩阵 $\overline{\boldsymbol{A}}^{F}_{m\times(n+t)}$ 与外逆P-增广矩阵 $\overline{\boldsymbol{A}}^{\bar{F}}_{m\times(n-\lambda)}$ 构成,或者 $(\overline{\boldsymbol{A}}^{F}_{m\times(n+t)},\overline{\boldsymbol{A}}^{\bar{F}}_{m\times(n-\lambda)})$ 是逆P-增广矩阵。3)外P-增广矩阵 $\boldsymbol{A}^{F}_{m\times(n+t)}$、内逆P-增广矩阵 $\overline{\boldsymbol{A}}^{F}_{m\times(n+t)}$ 与普通增广矩阵 $\boldsymbol{A}^*$ 是同一个概念。

2.1节和2.2节给出了下列结论:

(1)从2.1节中的式(1)－式(3)得到:若定义 $(\bar{x})^F_k=\overline{X}^F_k$ 是满足属性析取扩展大数据的数据元,α^F_k 是 $(\bar{x})^F_k$ 的属性集合,则在 α^F_k 内补充属性的条件下,$(\bar{x})^F_{k+1}$ 在 $(\bar{x})^F_k$ 之外被检索获取,$(\bar{x})^F_k\subseteq(\bar{x})^F_{k+1}$。

(2)从2.1节中的式(4)－式(6)得到:若定义 $(\bar{x})^{\bar{F}}_k=\overline{X}^{\bar{F}}_k$ 是满足属性析取收缩大数据的数据元,$\alpha^{\bar{F}}_k$ 是 $(\bar{x})^{\bar{F}}_k$ 的属性集合,则在 $\alpha^{\bar{F}}_k$ 内删除属性的条件下,$(\bar{x})^{\bar{F}}_{k+1}$ 在 $(\bar{x})^{\bar{F}}_k$ 内被检索获取,$(\bar{x})^{\bar{F}}_{k+1}\subseteq(\bar{x})^{\bar{F}}_k$。

(3)从2.2节中的式(15)、式(16)得到:内逆P-增广矩阵 $\overline{\boldsymbol{A}}^F$ 与矩阵 $\boldsymbol{A}$ 满足 $\boldsymbol{A}\subseteq\overline{\boldsymbol{A}}^F$,或者 $\boldsymbol{A}\Rightarrow\overline{\boldsymbol{A}}^F$;由式(15)、式(17)得到:外逆P-增广矩阵 $\overline{\boldsymbol{A}}^{\bar{F}}$ 与矩阵 $\boldsymbol{A}$ 满足 $\overline{\boldsymbol{A}}^{\bar{F}}\subseteq\boldsymbol{A}$,或者 $\overline{\boldsymbol{A}}^{\bar{F}}\Rightarrow\boldsymbol{A}$。

(4)从2.2节中的式(18)得到:逆P-增广矩阵 $(\overline{\boldsymbol{A}}^F_k,\overline{\boldsymbol{A}}^{\bar{F}}_{k+1})$ 与 $(\overline{\boldsymbol{A}}^F_{k+1},\overline{\boldsymbol{A}}^{\bar{F}}_k)$ 满足 $(\overline{\boldsymbol{A}}^F_k,\overline{\boldsymbol{A}}^{\bar{F}}_{k+1})\subseteq(\overline{\boldsymbol{A}}^F_{k+1},\overline{\boldsymbol{A}}^{\bar{F}}_k)$ 或者 $(\overline{\boldsymbol{A}}^F_k,\overline{\boldsymbol{A}}^{\bar{F}}_{k+1})\Rightarrow(\overline{\boldsymbol{A}}^F_{k+1},\overline{\boldsymbol{A}}^{\bar{F}}_k)$。这里,$(\overline{\boldsymbol{A}}^F_k,\overline{\boldsymbol{A}}^{\bar{F}}_{k+1})\subseteq(\overline{\boldsymbol{A}}^F_{k+1},\overline{\boldsymbol{A}}^{\bar{F}}_k)$ 表示 $\overline{\boldsymbol{A}}^F_k\subseteq\overline{\boldsymbol{A}}^F_{k+1}$,$\overline{\boldsymbol{A}}^{\bar{F}}_{k+1}\subseteq\overline{\boldsymbol{A}}^{\bar{F}}_k$;$(\overline{\boldsymbol{A}}^F_k,\overline{\boldsymbol{A}}^{\bar{F}}_{k+1})\Rightarrow(\overline{\boldsymbol{A}}^F_{k+1},\overline{\boldsymbol{A}}^{\bar{F}}_k)$ 表示 $\overline{\boldsymbol{A}}^F_k\Rightarrow\overline{\boldsymbol{A}}^F_{k+1}$,$\overline{\boldsymbol{A}}^{\bar{F}}_{k+1}\Rightarrow\overline{\boldsymbol{A}}^{\bar{F}}_k$。

结论(1)－(4)在本文第3－5节的讨论中被直接使用,不再做特别的声明。

2.3 ∨型大数据结构

称式(20)是 (x) 生成的具有属性析取特征的大数据,简称∨型大数据;(x) 称作∨型大数据的基数据,(x) 是具有静态特性的大数据。$(\bar{x})^F_i$,$(\bar{x})^{\bar{F}}_j$ 分别是∨型大数据的析取前项、析取后项;∨是 α 上的"析取"运算;α 是基数据 (x) 的属性集合;$\bar{\rho}^F_i$ 是 (x) 内被补充数据元 x_i 的可能性($(\bar{x})^F_i$ 存在的概率值),$\bar{\rho}^{\bar{F}}_j$ 是 (x) 内删除数据元 x_j 的可能性($(\bar{x})^{\bar{F}}_j$ 存在的概率值),$0<\bar{\rho}^F_i\leqslant 1$,$0<\bar{\rho}^{\bar{F}}_j\leqslant 1$。

$$\{(\bar{x})^F_i,(\bar{x})^{\bar{F}}_j\mid i\in I,j\in J:\alpha,\vee,(\bar{\rho}^F,\bar{\rho}^{\bar{F}})\}\tag{20}$$

3 大数据区块与区块矩阵等价性

称式(21)是大数据式(20)的前项生成的 F-区块;$(\bar{x})^F_i$ 是 F-区块元,$(\bar{x})^F_i\in\overline{G}^F$。

$$\overline{G}^F=\{(\bar{x})^F_i\mid i\in I:(\bar{x})^F_i\subseteq(\bar{x})^F_{i+1},\alpha^F_i\subseteq\alpha^F_{i+1}\}\tag{21}$$

称式(22)是大数据式(20)的后项生成的 $\bar{F}$-区块;$(\bar{x})^{\bar{F}}_j$ 是 $\bar{F}$-区块元,$(\bar{x})^{\bar{F}}_j\in\overline{G}^{\bar{F}}$。

$$\overline{G}^{\bar{F}}=\{(\bar{x})^{\bar{F}}_j\mid j\in J:(\bar{x})^{\bar{F}}_{j+1}\subseteq(\bar{x})^{\bar{F}}_j,\alpha^{\bar{F}}_{j+1}\subseteq\alpha^{\bar{F}}_j\}\tag{22}$$

由 $\overline{G}^F$ 与 $\overline{G}^{\bar{F}}$ 构成的区块对,称作大数据式(20)生成的区块,即:

$$(\overline{G}^F,\overline{G}^{\bar{F}})\tag{23}$$

F-区块元 $(\bar{x})^F_i$ 与 $\bar{F}$-区块元 $(\bar{x})^{\bar{F}}_j$ 构成的区块元对称作 $(\overline{G}^F,\overline{G}^{\bar{F}})$ 的区块元,即:

$$((\bar{x})^F_i,(\bar{x})^{\bar{F}}_j)\tag{24}$$

其中,$(\bar{x})^F_i\in\overline{G}^F$,$(\bar{x})^{\bar{F}}_j\in\overline{G}^{\bar{F}}$。

称 $\overline{\boldsymbol{B}}^F_k$ 是 F-区块元矩阵,如果 $\overline{\boldsymbol{B}}^F_k$ 是 $(\bar{x})^F_k\in\overline{G}^F$ 的数据元 $x_i\in(\bar{x})^F_k$ 的生成。

$$\overline{\boldsymbol{B}}^F_k=\begin{pmatrix}y_{1,1}&y_{1,2}&\cdots&y_{1,r}\\y_{2,1}&y_{2,2}&\cdots&y_{2,r}\\\vdots&\vdots&&\vdots\\y_{n,1}&y_{n,2}&\cdots&y_{n,r}\end{pmatrix}\tag{25}$$

其中,F-区块元 $(\bar{x})^F_k=\{x_1,x_2,\cdots,x_r\}$,数据元 $x_i\in(\bar{x})^F_k$ 具有 n 个数据值 $y_{1,i},y_{2,i},\cdots,y_{n,i}$,$i=1,2,\cdots,r$;$\boldsymbol{y}_i=(y_{1,i},y_{2,i},\cdots,y_{n,i})^{\mathrm{T}}$ 是 $y_{1,i},y_{2,i},\cdots,y_{n,i}$ 构成的向量,$\boldsymbol{y}_i$ 是 $\overline{\boldsymbol{B}}^F_k$ 的第 i 列。

称式(26)是 F-区块 $\overline{G}^F=\{(\bar{x})^F_1,(\bar{x})^F_2,\cdots,(\bar{x})^F_\lambda\}$ 生成的 F-区块矩阵。

$$\overline{\boldsymbol{B}}^F=\bigcup_{i=1}^{\lambda}\overline{\boldsymbol{B}}^F_i\tag{26}$$

称 $\overline{\boldsymbol{B}}^{\bar{F}}_k$ 是 $\bar{F}$-区块元矩阵,如果 $\overline{\boldsymbol{B}}^{\bar{F}}_k$ 是 $(\bar{x})^{\bar{F}}_k\in\overline{G}^{\bar{F}}$ 的数据元 $x_j\in(\bar{x})^{\bar{F}}_k$ 的生成。

$$\overline{\boldsymbol{B}}^{\bar{F}}_k=\begin{pmatrix}y_{1,1}&y_{1,2}&\cdots&y_{1,p}\\y_{2,1}&y_{2,2}&\cdots&y_{2,p}\\\vdots&\vdots&&\vdots\\y_{n,1}&y_{n,2}&\cdots&y_{n,p}\end{pmatrix}\tag{27}$$

其中,$\bar{F}$-区块元 $(\bar{x})^{\bar{F}}_k=\{x_1,x_2,\cdots,x_p\}$,数据元 $x_i\in(\bar{x})^{\bar{F}}_k$ 具有 n 个数据值 $y_{1,i},y_{2,i},\cdots,y_{n,i}$,$i=1,2,\cdots,p$;$\boldsymbol{y}_i=(y_{1,i},y_{2,i},\cdots,y_{n,i})^{\mathrm{T}}$ 是 $y_{1,i},y_{2,i},\cdots,y_{n,i}$ 生成的向量,$\boldsymbol{y}_i$ 是 $\overline{\boldsymbol{B}}^{\bar{F}}_k$ 的第 i 列。式(25)、式(27)中,$p<r$,$p,r\in\mathrm{N}^+$。

称式(28)是 $\bar{F}$-区块 $\overline{G}^{\bar{F}}=\{(\bar{x})^{\bar{F}}_1,(\bar{x})^{\bar{F}}_2,\cdots,(\bar{x})^{\bar{F}}_t\}$ 生成的 $\bar{F}$-区块矩阵。

$$\overline{\boldsymbol{B}}^{\bar{F}}=\bigcap_{i=1}^{t}\overline{\boldsymbol{B}}^{\bar{F}}_i\tag{28}$$

由 $\overline{\boldsymbol{B}}^F$ 与 $\overline{\boldsymbol{B}}^{\bar{F}}$ 构成的矩阵对,称作大数据区块 $(\overline{G}^F,\overline{G}^{\bar{F}})$ 生成的区块矩阵。

$$(\overline{\boldsymbol{B}}^F,\overline{\boldsymbol{B}}^{\bar{F}})\tag{29}$$

由式(21)－式(29)得到:

命题3 F-区块元矩阵 $\overline{\boldsymbol{B}}^F_k$ 是矩阵 $\boldsymbol{A}$ 的内逆P-增广矩阵 $\overline{\boldsymbol{A}}^F$,$\overline{\boldsymbol{B}}^F_k=\overline{\boldsymbol{A}}^F$。

命题4 $\bar{F}$-区块元矩阵 $\overline{\boldsymbol{B}}^{\bar{F}}_k$ 是矩阵 $\boldsymbol{A}$ 的外逆P-增广矩阵 $\overline{\boldsymbol{A}}^{\bar{F}}$,$\overline{\boldsymbol{B}}^{\bar{F}}_k=\overline{\boldsymbol{A}}^{\bar{F}}$。

命题5 区块矩阵 $(\overline{\boldsymbol{B}}^F_k,\overline{\boldsymbol{B}}^{\bar{F}}_k)$ 是矩阵 $\boldsymbol{A}$ 的逆P-增广矩阵 $(\overline{\boldsymbol{A}}^F,\overline{\boldsymbol{A}}^{\bar{F}})$,$(\overline{\boldsymbol{B}}^F_k,\overline{\boldsymbol{B}}^{\bar{F}}_k)=(\overline{\boldsymbol{A}}^F,\overline{\boldsymbol{A}}^{\bar{F}})$。

这里,命题3－命题5中的 $\boldsymbol{A}$,$\overline{\boldsymbol{A}}^F$,$\overline{\boldsymbol{A}}^{\bar{F}}$ 与 $(\overline{\boldsymbol{A}}^F,\overline{\boldsymbol{A}}^{\bar{F}})$ 分别是第2节中的式(15)－式(18)。

定理1(F-区块矩阵属性定理) 若 α^F_k 是 F-区块元矩阵 $\overline{\boldsymbol{B}}^F_k$ 的属性集合,则 F-区块矩阵 $\overline{\boldsymbol{B}}^F$ 的属性集合 α^F 满足:

$$\alpha^F=\bigcup_{k=1}^{\lambda}\alpha^F_k\tag{30}$$

证明:由式(21)得到,给定 F-区块 $\overline{G}^F=\{(\bar{x})^F_1,(\bar{x})^F_2,\cdots,(\bar{x})^F_\lambda\}$,$\alpha^F_i$ 是 $(\bar{x})^F_i$ 的属性集合,或者 α^F_i 是 $\overline{\boldsymbol{B}}^F_i$ 的属性集合,$i=1,2,\cdots,\lambda$。由第2节中的式(1)－式(3)得到,$(\bar{x})^F_1$,$(\bar{x})^F_2$,$\cdots$,

$(\bar{x})_\lambda^F$ 满足 $(\bar{x})_1^F \subseteq (\bar{x})_2^F \subseteq \cdots \subseteq (\bar{x})_\lambda^F$，或者 $\alpha_1^F \subseteq \alpha_2^F \subseteq \cdots \subseteq \alpha_\lambda^F$。因为 $\bar{\boldsymbol{B}}_1^F, \bar{\boldsymbol{B}}_2^F, \cdots, \bar{\boldsymbol{B}}_\lambda^F$ 分别是 $(\bar{x})_1^F, (\bar{x})_2^F, \cdots, (\bar{x})_\lambda^F$ 的生成，则有 $\bar{\boldsymbol{B}}_1^F \subseteq \bar{\boldsymbol{B}}_2^F \subseteq \cdots \subseteq \bar{\boldsymbol{B}}_\lambda^F$。由式(26)得到，$\bar{\boldsymbol{B}}^F$ 的属性集合 $\alpha^F = \alpha_1^F \cup \alpha_2^F \cup \cdots \cup \alpha_\lambda^F = \bigcup_{k=1}^{\lambda} \alpha_k^F$，显然 $\bar{\boldsymbol{B}}^F = \bigcup_{k=1}^{\lambda} \bar{\boldsymbol{B}}_k^F$ 具有属性集合 $\alpha^F = \bigcup_{k=1}^{\lambda} \alpha_k^F$，得到定理 1。

定理 2($\bar{F}$-区块矩阵属性定理)　若 $\alpha_k^{\bar{F}}$ 是 $\bar{F}$-区块元矩阵 $\bar{\boldsymbol{B}}_k^{\bar{F}}$ 的属性集合，则 $\bar{F}$-区块矩阵 $\bar{\boldsymbol{B}}^{\bar{F}}$ 的属性集合 $\alpha^{\bar{F}}$ 满足：

$$\alpha^{\bar{F}} = \bigcap_{k=1}^{t} \alpha_k^{\bar{F}} \tag{31}$$

定理 2 的证明与定理 1 类似，证明略。

定理 3(F-区块元属性析取扩展定理)　若 α_k^F 是 F-区块元 $(\bar{x})_k^F \in \bar{G}^F$ 的属性集合，则数据元 $x_i \in (\bar{x})_k^F$ 的属性 α_i 满足：

$$\alpha_i = (\bigvee_{t=1}^{p} \alpha_t) \bigvee_{t=p+1}^{q} \alpha_t \tag{32}$$

其中，$(x) = \{x_1, x_2, \cdots, x_p\}$，$\alpha = \{\alpha_1, \alpha_2, \cdots, \alpha_p\}$ 是 (x) 的属性集合；$(x) \subseteq (\bar{x})_k^F = \{x_1, x_2, \cdots, x_p, x_{p+1}, \cdots, x_q\}$，$\alpha_k^F$ 是 $(\bar{x})_k^F$ 的属性集合，$\alpha_k^F = \{\alpha_1, \alpha_2, \cdots, \alpha_p, \alpha_{p+1}, \cdots, \alpha_q\}$。

由第 2 节中的逆 P-集合的逻辑特征，直接得到式(32)，证明略。

定理 4($\bar{F}$-区块元属性析取收缩定理)　若 $\alpha_k^{\bar{F}}$ 是 $\bar{F}$-区块元 $(\bar{x})_k^{\bar{F}} \in \bar{G}^{\bar{F}}$ 的属性集合，则数据元 $x_j \in (\bar{x})_k^{\bar{F}}$ 的属性 α_j 满足：

$$\alpha_j = (\bigvee_{t=1}^{q} \alpha_t) - \bigvee_{t=p+1}^{q} \alpha_t \tag{33}$$

其中，$(x) = \{x_1, x_2, \cdots, x_p, x_{p+1}, \cdots, x_q\}$，$\alpha = \{\alpha_1, \alpha_2, \cdots, \alpha_p, \alpha_{p+1}, \cdots, \alpha_q\}$ 是 (x) 的属性集合；$(\bar{x})_k^{\bar{F}} = \{x_1, x_2, \cdots, x_p\} \subseteq (x)$，$\alpha_k^{\bar{F}}$ 是 $(\bar{x})_k^{\bar{F}}$ 的属性集合，$\alpha_k^{\bar{F}} = \{\alpha_1, \alpha_2, \cdots, \alpha_p\}$。

定理 4 的证明与定理 3 类似，证明略。

若把 F-区块元属性析取扩展(式(32))、$\bar{F}$-区块元属性析取收缩(式(33))分别扩展到 F-区块 $\bar{G}^F$ 属性析取扩展、$\bar{F}$-区块 $\bar{G}^{\bar{F}}$ 属性析取收缩，则分别得到式(32)、式(33)的一般形式，这些讨论略。

定理 5(α^*-矩阵等价类定理)　若 α_k^F 是 F-区块元矩阵 $\bar{\boldsymbol{B}}_k^F$ 的属性集合，满足：

$$\alpha^* = \bigcap_{k=1}^{n} \alpha_k^F \tag{34}$$

则 F-区块元矩阵 $\bar{\boldsymbol{B}}_k^F$ 构成关于 α^* 的矩阵等价类 $[\bar{\boldsymbol{B}}^*]$，或者：

$$[\bar{\boldsymbol{B}}^*] = \{\bar{\boldsymbol{B}}_k^F \mid k = 1, 2, \cdots, n\} \tag{35}$$

证明：给定 F-区块元 $(\bar{x})_i^F$，α_i^F 是 $(\bar{x})_i^F$ 的属性集合，$i = 1, 2, \cdots, n$；α_i^F 满足式(34)，显然，$\forall \bar{\boldsymbol{B}}_i^F$ 都具有属性集合 α^*。设 $\bar{\boldsymbol{B}}^F = \{\bar{\boldsymbol{B}}_i^F \mid i = 1, 2, \cdots, n\}$，定义 α^* 是 $\bar{\boldsymbol{B}}^F \times \bar{\boldsymbol{B}}^F$ 上的关系 R，$R = \alpha^*$，则有：1) $\forall \bar{\boldsymbol{B}}_i^F \in \bar{\boldsymbol{B}}^F$，$\bar{\boldsymbol{B}}_i^F \alpha^* \bar{\boldsymbol{B}}_i^F$；2) $\forall \bar{\boldsymbol{B}}_i^F, \bar{\boldsymbol{B}}_j^F \in \bar{\boldsymbol{B}}^F$，$\bar{\boldsymbol{B}}_i^F \alpha^* \bar{\boldsymbol{B}}_j^F \Rightarrow \bar{\boldsymbol{B}}_j^F \alpha^* \bar{\boldsymbol{B}}_i^F$；3) $\forall \bar{\boldsymbol{B}}_i^F, \bar{\boldsymbol{B}}_j^F, \bar{\boldsymbol{B}}_k^F \in \bar{\boldsymbol{B}}^F$，$\bar{\boldsymbol{B}}_i^F \alpha^* \bar{\boldsymbol{B}}_j^F, \bar{\boldsymbol{B}}_j^F \alpha^* \bar{\boldsymbol{B}}_k^F \Rightarrow \bar{\boldsymbol{B}}_i^F \alpha^* \bar{\boldsymbol{B}}_k^F$。$\bar{\boldsymbol{B}}_i^F, \bar{\boldsymbol{B}}_j^F, \bar{\boldsymbol{B}}_k^F$ 关于 α^* 满足自反性、对称性与传递性，$\bar{\boldsymbol{B}}_i^F, \bar{\boldsymbol{B}}_j^F, \bar{\boldsymbol{B}}_k^F$ 构成 $[\bar{\boldsymbol{B}}^*]$，得到式(35)。

定理 6(α^o-矩阵等价类定理)　若 $\alpha_k^{\bar{F}}$ 是 $\bar{F}$-区块元矩阵 $\bar{\boldsymbol{B}}_k^{\bar{F}}$ 的属性集合，满足：

$$\alpha^o = \bigcap_{k=1}^{m} \alpha_k^{\bar{F}} \tag{36}$$

则 $\bar{F}$-区块元矩阵 $\bar{\boldsymbol{B}}_k^{\bar{F}}$ 构成关于 α^o 的矩阵等价类 $[\bar{\boldsymbol{B}}^o]$，或者：

$$[\bar{\boldsymbol{B}}^o] = \{\bar{\boldsymbol{B}}_k^{\bar{F}} \mid k = 1, 2, \cdots, m\} \tag{37}$$

定理 6 的证明与定理 5 类似，证明略。

推论 1　具有属性集合 α^* 的 F-区块元 $(\bar{x})_k^F$ 构成关于 α^* 的区块元等价类 $[(\bar{x})^*]$，或者：

$$[(\bar{x})^*] = \{(\bar{x})_k^F \mid k = 1, 2, \cdots, n\} \tag{38}$$

证明：由定理 5 得到，$\forall \bar{\boldsymbol{B}}_i^F \in [\bar{\boldsymbol{B}}^*]$ 具有 α^*，$\bar{\boldsymbol{B}}_i^F$ 是 $(\bar{x})_k^F$ 的生成，由定理 5 直接得到：生成 $\bar{\boldsymbol{B}}_i^F$ 的 $(\bar{x})_k^F$ 构成 $[(\bar{x})^*] = \{(\bar{x})_k^F \mid k = 1, 2, \cdots, n\}$，得到式(38)。

推论 2　具有属性集合 α^o 的 $\bar{F}$-区块元 $(\bar{x})_k^{\bar{F}}$ 构成关于 α^o 的区块元等价类 $[(\bar{x})^o]$，或者：

$$[(\bar{x})^o] = \{(\bar{x})_k^{\bar{F}} \mid k = 1, 2, \cdots, m\} \tag{39}$$

推论 2 的证明与推论 1 类似，证明略。

由定理 5、定理 6 与推论 1、推论 2 得到：

定理 7(F-区块智能获取定理)　若 F-区块 $\bar{G}_k^F, \bar{G}_{k+1}^F$ 与 F-区块矩阵 $\bar{\boldsymbol{B}}_k^F, \bar{\boldsymbol{B}}_{k+1}^F$ 满足：

$$\text{if } \bar{\boldsymbol{B}}_k^F \Rightarrow \bar{\boldsymbol{B}}_{k+1}^F, \text{ then } \bar{G}_k^F \Rightarrow \bar{G}_{k+1}^F \tag{40}$$

则存在缺失区块 $\Delta\bar{G}^F$，$\Delta\bar{G}^F \cap \bar{G}_k^F = \emptyset$，$\bar{G}_{k+1}^F$ 在 $\bar{G}_k^F$ 外被智能获取 。

$$\bar{G}_{k+1}^F = \bar{G}_k^F \cup \Delta\bar{G}^F \tag{41}$$

证明：若 F-区块矩阵 $\bar{\boldsymbol{B}}_k^F, \bar{\boldsymbol{B}}_{k+1}^F$ 与 F-区块 $\bar{G}_k^F, \bar{G}_{k+1}^F$ 满足推理(40)，$\bar{G}_k^F$ 是 $\bar{G}_{k+1}^F$ 的子-区块，$\bar{G}_k^F \subseteq \bar{G}_{k+1}^F$，一定存在 $\Delta\bar{G}^F \neq \emptyset$，$\bar{G}_{k+1}^F = \bar{G}_k^F \cup \Delta\bar{G}^F$；在推理条件 $\bar{\boldsymbol{B}}_k^F \Rightarrow \bar{\boldsymbol{B}}_{k+1}^F$ 的限定下，$\bar{G}_{k+1}^F$ 在 $\bar{G}_k^F$ 外被智能找到，得到定理 7。

定理 8($\bar{F}$-区块智能获取定理)　若 $\bar{F}$-区块 $\bar{G}_k^{\bar{F}}, \bar{G}_{k+1}^{\bar{F}}$ 与 $\bar{F}$-区块矩阵 $\bar{\boldsymbol{B}}_k^{\bar{F}}, \bar{\boldsymbol{B}}_{k+1}^{\bar{F}}$ 满足：

$$\text{if } \bar{\boldsymbol{B}}_{k+1}^{\bar{F}} \Rightarrow \bar{\boldsymbol{B}}_k^{\bar{F}}, \text{ then } \bar{G}_{k+1}^{\bar{F}} \Rightarrow \bar{G}_k^{\bar{F}} \tag{42}$$

则存在冗余区块 $\nabla\bar{G}^{\bar{F}}$，$\nabla\bar{G}^{\bar{F}} \cap \bar{G}_k^{\bar{F}} \neq \emptyset$，$\bar{G}_{k+1}^{\bar{F}}$ 在 $\bar{G}_k^{\bar{F}}$ 内被智能获取。

$$\bar{G}_{k+1}^{\bar{F}} = \bar{G}_k^{\bar{F}} - \nabla\bar{G}^{\bar{F}} \tag{43}$$

定理 8 的证明与定理 7 类似，证明略。

由定理 1—定理 8 与推论 1、推论 2 得到：

命题 6　F-区块 $\bar{G}_k^F, \bar{G}_{k+1}^F$ 满足：

$$card(\bar{G}_{k+1}^F) - card(\bar{G}_k^F) > 0 \tag{44}$$

其中，$\bar{G}_{k+1}^F = \{\bar{G}_1^F, \bar{G}_2^F, \cdots, \bar{G}_n^F\}$，$\bar{G}_k^F = \{\bar{G}_1^F, \bar{G}_2^F, \cdots, \bar{G}_m^F\}$，$m < n$。

命题 7　$\bar{F}$-区块 $\bar{G}_k^{\bar{F}}, \bar{G}_{k+1}^{\bar{F}}$ 满足：

$$card(\bar{G}_{k+1}^{\bar{F}}) - card(\bar{G}_k^{\bar{F}}) < 0 \tag{45}$$

其中，$\bar{G}_{k+1}^{\bar{F}} = \{\bar{G}_1^{\bar{F}}, \bar{G}_2^{\bar{F}}, \cdots, \bar{G}_m^{\bar{F}}\}$，$\bar{G}_k^{\bar{F}} = \{\bar{G}_1^{\bar{F}}, \bar{G}_2^{\bar{F}}, \cdots, \bar{G}_n^{\bar{F}}\}$，$m < n$。

命题 8　F-区块 $\bar{G}_{k+1}^F$ 在 $\bar{G}_k^F$ 外被智能获取，$\bar{G}_{k+1}^F$ 的属性集合 α_{k+1}^F 与 $\bar{G}_k^F$ 的属性集合 α_k^F 满足：

$$card(\alpha_{k+1}^F) - card(\alpha_k^F) > 0 \tag{46}$$

命题 9　$\bar{F}$-区块 $\bar{G}_{k+1}^{\bar{F}}$ 在 $\bar{G}_k^{\bar{F}}$ 内被智能获取，$\bar{G}_{k+1}^{\bar{F}}$ 的属性集合 $\alpha_{k+1}^{\bar{F}}$ 与 $\bar{G}_k^{\bar{F}}$ 的属性集合 $\alpha_k^{\bar{F}}$ 满足：

$$card(\alpha_{k+1}^{\bar{F}}) - card(\alpha_k^{\bar{F}}) < 0 \tag{47}$$

式(44)—式(47)中，$card$ 为 cardinal number。

命题 10　F-区块元 $(\bar{x})_k^F \in [(\bar{x})^*]$ 的属性集合 α^* 与基区块元 $(x)_k \in (x)$ 的属性集合 α_k 满足：

$$(\alpha_k \cup \Delta\alpha^F) - \alpha^* = \emptyset \tag{48}$$

其中，$\Delta\alpha^F \neq \emptyset$。

证明 ：由第 2 节的式(20) ，(x) 是基数据，定义 (x) 是基区块 $G = \{(x)_1, (x)_2, \cdots, (x)_q\}$，$\forall (x)_k \in G$ 是基区块 G 的区块元，设 α_k 是 $(x)_k \in G$ 的属性集合，定义 F-区块 $\bar{G}^F = \{(\bar{x})_1^F,$

$(\bar{x})_2^F,\cdots,(\bar{x})_n^F\}$，由第 2 节的式(1)－式(3)，$(\bar{x})_1^F,(\bar{x})_2^F,\cdots,(\bar{x})_n^F$ 的属性集合 $\alpha_1^F,\alpha_2^F,\cdots,\alpha_n^F$ 与 $(x)_k$ 的属性集合 α_k 满足：$\alpha_k\subseteq\alpha_1^F\subseteq\alpha_2^F\subseteq\cdots\subseteq\alpha_n^F$，或者 $\alpha_k\cap\alpha_1^F\cap\alpha_2^F\cap\cdots\cap\alpha_n^F=\alpha_k$，存在 $\Delta\alpha^F\neq\varnothing$，$\alpha_k\cup\Delta\alpha^F=\bigcap_{i=1}^{n}\alpha_i^F=\alpha^*$，或者$(\alpha_k\cup\Delta\alpha^F)-\alpha^*=\varnothing$，得到式(48)。

命题 11 $\bar{F}$-区块元$(\bar{x})_k^{\bar{F}}\in[(\bar{x})^o]$的属性集合 α^o 与基区块元$(x)_k\in(x)$的属性集合 α_k 满足：

$$(\alpha_k-\nabla\alpha^{\bar{F}})-\alpha^o=\varnothing \tag{49}$$

其中，$\nabla\alpha^{\bar{F}}\neq\varnothing$。

命题 11 的证明与命题 10 类似，证明略。

4 区块属性推理与区块元智能分离、区块智能检索-过滤

若 $\alpha_k^F,\alpha_{k+1}^F$ 与 $\boldsymbol{\bar{B}}_k^F,\boldsymbol{\bar{B}}_{k+1}^F$ 满足：

$$\text{if } \alpha_k^F\Rightarrow\alpha_{k+1}^F,\text{then } \boldsymbol{\bar{B}}_k^F\Rightarrow\boldsymbol{\bar{B}}_{k+1}^F \tag{50}$$

称式(50)是 F-区块元矩阵的属性生成的属性推理；$\alpha_k^F\Rightarrow\alpha_{k+1}^F$ 称作属性推理条件，$\boldsymbol{\bar{B}}_k^F\Rightarrow\boldsymbol{\bar{B}}_{k+1}^F$ 称作属性推理结论。

式(50)中，$\alpha_k^F,\alpha_{k+1}^F$ 分别是 F-区块元矩阵 $\boldsymbol{\bar{B}}_k^F,\boldsymbol{\bar{B}}_{k+1}^F$ 的属性集合；$\alpha_k^F\Rightarrow\alpha_{k+1}^F$ 与 $\alpha_k^F\subseteq\alpha_{k+1}^F$ 等价，$\boldsymbol{\bar{B}}_k^F\Rightarrow\boldsymbol{\bar{B}}_{k+1}^F$ 与 $\boldsymbol{\bar{B}}_k^F\subseteq\boldsymbol{\bar{B}}_{k+1}^F$ 等价。

若 $\alpha_{k+1}^{\bar{F}},\alpha_k^{\bar{F}}$ 与 $\boldsymbol{\bar{B}}_{k+1}^{\bar{F}},\boldsymbol{\bar{B}}_k^{\bar{F}}$ 满足：

$$\text{if } \alpha_{k+1}^{\bar{F}}\Rightarrow\alpha_k^{\bar{F}},\text{then } \boldsymbol{\bar{B}}_{k+1}^{\bar{F}}\Rightarrow\boldsymbol{\bar{B}}_k^{\bar{F}} \tag{51}$$

称式(51)是 $\bar{F}$-区块元矩阵的属性生成的属性推理；$\alpha_{k+1}^{\bar{F}}\Rightarrow\alpha_k^{\bar{F}}$ 称作属性推理条件，$\boldsymbol{\bar{B}}_{k+1}^{\bar{F}}\Rightarrow\boldsymbol{\bar{B}}_k^{\bar{F}}$ 称作属性推理结论。

若$(\alpha_k^F,\alpha_{k+1}^{\bar{F}})$，$(\alpha_{k+1}^F,\alpha_k^{\bar{F}})$与$(\boldsymbol{\bar{B}}_k^F,\boldsymbol{\bar{B}}_{k+1}^{\bar{F}})$，$(\boldsymbol{\bar{B}}_{k+1}^F,\boldsymbol{\bar{B}}_k^{\bar{F}})$满足：

$$\text{if } (\alpha_k^F,\alpha_{k+1}^{\bar{F}})\Rightarrow(\alpha_{k+1}^F,\alpha_k^{\bar{F}}),\text{then } (\boldsymbol{\bar{B}}_k^F,\boldsymbol{\bar{B}}_{k+1}^{\bar{F}})\Rightarrow(\boldsymbol{\bar{B}}_{k+1}^F,\boldsymbol{\bar{B}}_k^{\bar{F}}) \tag{52}$$

称式(52)是$(\boldsymbol{\bar{B}}_k^F,\boldsymbol{\bar{B}}_{k+1}^{\bar{F}})$，$(\boldsymbol{\bar{B}}_{k+1}^F,\boldsymbol{\bar{B}}_k^{\bar{F}})$的属性生成的属性推理；$(\alpha_k^F,\alpha_{k+1}^{\bar{F}})\Rightarrow(\alpha_{k+1}^F,\alpha_k^{\bar{F}})$称作属性推理条件，$(\boldsymbol{\bar{B}}_k^F,\boldsymbol{\bar{B}}_{k+1}^{\bar{F}})\Rightarrow(\boldsymbol{\bar{B}}_{k+1}^F,\boldsymbol{\bar{B}}_k^{\bar{F}})$称作属性推理结论。

式(52)中，$(\alpha_k^F,\alpha_{k+1}^{\bar{F}})$是 F-区块元 $(\bar{x})_k^F$ 与 $\bar{F}$-区块元 $(\bar{x})_{k+1}^{\bar{F}}$ 构成的$((\bar{x})_k^F,(\bar{x})_{k+1}^{\bar{F}})$的属性集合；$(\alpha_k^F,\alpha_{k+1}^{\bar{F}})\Rightarrow(\alpha_{k+1}^F,\alpha_k^{\bar{F}})$表示：$\alpha_k^F\Rightarrow\alpha_{k+1}^F$，$\alpha_{k+1}^{\bar{F}}\Rightarrow\alpha_k^{\bar{F}}$。

若把式(50)－式(52)中的属性推理结论定义成推理条件，则得到定理 9。

定理 9（F-区块元智能分离定理） 若 $\boldsymbol{\bar{B}}_k^F,\boldsymbol{\bar{B}}_{k+1}^F$ 与$(\bar{x})_k^F$，$(\bar{x})_{k+1}^F$满足：

$$\text{if } \boldsymbol{\bar{B}}_k^F\Rightarrow\boldsymbol{\bar{B}}_{k+1}^F,\text{then } (\bar{x})_k^F\Rightarrow(\bar{x})_{k+1}^F \tag{53}$$

则：

1）F-区块元$(\bar{x})_{k+1}^F$在$(\bar{x})_k^F$ 之外被智能分离；

2）F-区块元$(\bar{x})_{k+1}^F$与$(\bar{x})_k^F$满足：

$$(\bar{x})_{k+1}^F\cap(\bar{x})_k^F\neq\varnothing \tag{54}$$

证明：给定区块 $\bar{G}_\lambda^F=\{(\bar{x})_1^F,(\bar{x})_2^F,\cdots,(\bar{x})_k^F,(\bar{x})_{k+1}^F,\cdots,(\bar{x})_n^F\}$，$(\bar{x})_i^F$ 是 $\bar{G}_\lambda^F$ 的区块元，$(\bar{x})_k^F=\{x_1,x_2,\cdots,x_p\}$，$(\bar{x})_{k+1}^F=\{x_1,x_2,\cdots,x_q\}$，$p\leqslant q$；$x_i\in(\bar{x})_k^F$，$x_j\in(\bar{x})_{k+1}^F$ 分别是$(\bar{x})_k^F$，$(\bar{x})_{k+1}^F$ 的数据元；区块元矩阵 $\boldsymbol{\bar{B}}_k^F,\boldsymbol{\bar{B}}_{k+1}^F$ 分别是$(\bar{x})_k^F$，$(\bar{x})_{k+1}^F$的生成，满足式(53)；则$(\bar{x})_k^F$ 内被补充数据元 x_{p+1}，$x_{p+2},\cdots,x_q$，在 $\boldsymbol{\bar{B}}_k^F\Rightarrow\boldsymbol{\bar{B}}_{k+1}^F$ 条件下，$(\bar{x})_k^F$ 智能生成$(\bar{x})_{k+1}^F=(\bar{x})_k^F\cup\{x_{p+1},x_{p+2},\cdots,x_q\}$，$(\bar{x})_{k+1}^F$依赖$(\bar{x})_k^F$，在$(\bar{x})_k^F$ 外被智能分离。显然，$(\bar{x})_{k+1}^F\cap(\bar{x})_k^F\neq\varnothing$，得到定理 9。

定理 10（$\bar{F}$-区块元智能分离定理） 若 $\boldsymbol{\bar{B}}_{k+1}^{\bar{F}},\boldsymbol{\bar{B}}_k^{\bar{F}}$ 与$(\bar{x})_{k+1}^{\bar{F}}$，$(\bar{x})_k^{\bar{F}}$ 满足：

$$\text{if } \boldsymbol{\bar{B}}_{k+1}^{\bar{F}}\Rightarrow\boldsymbol{\bar{B}}_k^{\bar{F}},\text{then } (\bar{x})_{k+1}^{\bar{F}}\Rightarrow(\bar{x})_k^{\bar{F}} \tag{55}$$

则：

1）$\bar{F}$-区块元$(\bar{x})_{k+1}^{\bar{F}}$在$(\bar{x})_k^{\bar{F}}$ 之内被智能分离；

2）$\bar{F}$-区块元$(\bar{x})_{k+1}^{\bar{F}}$与$(\bar{x})_k^{\bar{F}}$满足：

$$(\bar{x})_{k+1}^{\bar{F}}\cap(\bar{x})_k^{\bar{F}}\neq\varnothing \tag{56}$$

定理 10 的证明与定理 9 类似，证明略。

把定理 9、定理 10 的推理条件 $\boldsymbol{\bar{B}}_k^F\Rightarrow\boldsymbol{\bar{B}}_{k+1}^F$，$\boldsymbol{\bar{B}}_{k+1}^{\bar{F}}\Rightarrow\boldsymbol{\bar{B}}_k^{\bar{F}}$ 分别用推理条件 $\alpha_k^F\Rightarrow\alpha_{k+1}^F$，$\alpha_{k+1}^{\bar{F}}\Rightarrow\alpha_k^{\bar{F}}$ 代替，推广定理 9、定理 10 得到定理 11、定理 12。

定理 11（F-区块智能检索定理） 若 F-区块 $\bar{G}_k^F,\bar{G}_{k+1}^F$ 与它们的属性集合 $\alpha_k^F,\alpha_{k+1}^F$ 满足：

$$\text{if } \alpha_k^F\Rightarrow\alpha_{k+1}^F,\text{then } \bar{G}_k^F\Rightarrow\bar{G}_{k+1}^F \tag{57}$$

则：

1）F-区块 $\bar{G}_{k+1}^F$ 在 $\bar{G}_k^F$ 外被智能检索发现；

2）$\bar{G}_\lambda^F$ 是 $\bar{G}_k^F$ 的智能检索-过滤剩余，$\lambda=n+1,n+2,\cdots,m$。

这里，$\bar{G}_k^F=\{\bar{G}_1^F,\bar{G}_2^F,\cdots,\bar{G}_n^F\}$，$\bar{G}_{k+1}^F=\{\bar{G}_1^F,\bar{G}_2^F,\cdots,\bar{G}_m^F\}$，$n<m$；$\alpha_i^F$ 是 $\bar{G}_i^F\in\bar{G}_k^F$ 的属性集合，$\alpha_k^F=\bigcup_{i=1}^{n}\alpha_i^F$ 是 $\bar{G}_k^F$ 的属性集合；$\alpha_{k+1}^F=\bigcup_{i=1}^{m}\alpha_i^F$ 是 $\bar{G}_{k+1}^F$ 的属性集合。

证明：$\alpha_k^F=\bigcup_{i=1}^{n}\alpha_i^F$，$\alpha_{k+1}^F=\bigcup_{i=1}^{m}\alpha_i^F$ 分别是 $\bar{G}_k^F,\bar{G}_{k+1}^F$ 的属性集合，由式(57)的 if $\alpha_k^F\Rightarrow\alpha_{k+1}^F$，then $\bar{G}_k^F\Rightarrow\bar{G}_{k+1}^F$ 得到：if $\bigcup_{i=1}^{n}\alpha_i^F\Rightarrow\bigcup_{i=1}^{m}\alpha_i^F$，then $\bar{G}_k^F\Rightarrow\bar{G}_{k+1}^F$；在 $\bigcup_{i=1}^{n}\alpha_i^F\Rightarrow\bigcup_{i=1}^{m}\alpha_i^F$ 的条件下，$\bar{G}_k^F=\{\bar{G}_1^F,\bar{G}_2^F,\cdots,\bar{G}_n^F\}$内补充 $\bar{G}_{n+1}^F,\bar{G}_{n+2}^F,\cdots,\bar{G}_m^F$；$\bar{G}_k^F$ 智能生成 $\bar{G}_{k+1}^F=\bar{G}_k^F\cup\{\bar{G}_{n+1}^F,\bar{G}_{n+2}^F,\cdots,\bar{G}_m^F\}$，$\bar{G}_{k+1}^F$ 在 $\bar{G}_k^F$ 之外被智能检索发现-获取，$\bar{G}_k^F\subseteq\bar{G}_{k+1}^F$。定义 $GRD(G)=card(G)/card(G)$是基数(x)生成的基区块的颗粒度，$\bar{G}_k^F,G,\bar{G}_{k+1}^F$ 满足 $\bar{G}_k^F\subseteq G\subseteq\bar{G}_{k+1}^F$，$GRD(G)$是基区块的颗粒度(粒度)；显然，$GRD(\bar{G}_k^F)=card(\bar{G}_k^F)/card(G)\leqslant GRD(G)=card(G)/card(G)\leqslant GRD(\bar{G}_{k+1}^F)=card(\bar{G}_{k+1}^F)/card(G)$，或者 $GRD(\bar{G}_k^F)\leqslant GRD(G)\leqslant GRD(\bar{G}_{k+1}^F)$。定义 $GRD(G)$是过滤筛子 Ω 的孔径(孔的尺寸)，则 $\bar{G}_i^F$ 从 Ω 内被过滤剩余到 Ω 之外被检索，$\lambda=1,2,\cdots,n$；$\bar{G}_{n+1}^F,\bar{G}_{n+2}^F,\cdots,\bar{G}_m^F$ 被留在 Ω 内；或者 $\bar{G}_{n+1}^F,\bar{G}_{n+2}^F,\cdots,\bar{G}_m^F$ 是 $\bar{G}_i^F$ 被智能过滤的剩余。GRD=granulation degree，*card*=cardinal number。

定理 12（$\bar{F}$-区块智能检索定理） 若 $\bar{F}$-区块 $\bar{G}_k^{\bar{F}},\bar{G}_{k+1}^{\bar{F}}$ 与它们的属性集合 $\alpha_k^{\bar{F}},\alpha_{k+1}^{\bar{F}}$ 满足：

$$\text{if } \alpha_{k+1}^{\bar{F}}\Rightarrow\alpha_k^{\bar{F}},\text{then } \bar{G}_{k+1}^{\bar{F}}\Rightarrow\bar{G}_k^{\bar{F}} \tag{58}$$

则：

1）$\bar{F}$-区块 $\bar{G}_{k+1}^{\bar{F}}$ 在 $\bar{G}_k^{\bar{F}}$ 内被智能检索发现；

2）$\bar{G}_j^{\bar{F}}$ 是 $\bar{G}_{k+1}^{\bar{F}}$ 的智能检索-过滤剩余，$\lambda=m+1,m+2,\cdots,n$。

这里，$\bar{G}_k^{\bar{F}}=\{\bar{G}_1^{\bar{F}},\bar{G}_2^{\bar{F}},\cdots,\bar{G}_n^{\bar{F}}\}$，$\bar{G}_{k+1}^{\bar{F}}=\{\bar{G}_1^{\bar{F}},\bar{G}_2^{\bar{F}},\cdots,\bar{G}_m^{\bar{F}}\}$，$m<n$；$\alpha_i^{\bar{F}}$ 是 $\bar{G}_i^{\bar{F}}\in\bar{G}_k^{\bar{F}}$ 的属性集合，$\alpha_{k+1}^{\bar{F}}=\bigcap_{i=1}^{m}\alpha_i^{\bar{F}}$ 是 $\bar{G}_{k+1}^{\bar{F}}$ 的属性集合；$\alpha_k^{\bar{F}}=\bigcap_{i=1}^{n}\alpha_i^{\bar{F}}$ 是 $\bar{G}_k^{\bar{F}}$ 的属性集合。

定理 12 的证明与定理 11 类似，证明略。

4.1 区块元智能分离-区块智能检索准则

(1) 区块元智能分离准则

给定区块 $\bar{G}_k^F=\{(\bar{x})_1^F,(\bar{x})_2^F,\cdots,(\bar{x})_n^F\}$，$\alpha_k^F=\{\alpha_1^F,\alpha_2^F,\cdots,\alpha_n^F\}$是 $\bar{G}_k^F$ 的属性集合，区块元$(\bar{x})_{\lambda+1}^F$在区块元$(\bar{x})_\lambda^F$ 外被智能分离，$(\bar{x})_{\lambda+1}^F$的属性集合 $\alpha_{\lambda+1}^F$与$(\bar{x})_\lambda^F$ 的属性集合 α_λ^F 满足推理条件：$\alpha_\lambda^F=\bigcup_{i=1}^{\lambda}\alpha_i^F\Rightarrow\alpha_{\lambda+1}^F$；$(\bar{x})_{\lambda+1}^F$，$(\bar{x})_\lambda^F\in\bar{G}_k^F$。

给定区块 $\bar{G}_k^{\bar F}=\{(\bar{x})_1^{\bar F},(\bar{x})_2^{\bar F},\cdots,(\bar{x})_n^{\bar F}\}$，$\alpha_k^{\bar F}=\{\alpha_1^{\bar F},\alpha_2^{\bar F},\cdots,\alpha_n^{\bar F}\}$是 $\bar{G}_k^{\bar F}$ 的属性集合，区块元$(\bar{x})_{\lambda+1}^{\bar F}$在区块元$(\bar{x})_\lambda^{\bar F}$ 内被智能分离，$(\bar{x})_{\lambda+1}^{\bar F}$的属性集合 $\alpha_{\lambda+1}^{\bar F}$与$(\bar{x})_\lambda^{\bar F}$ 的属性集合 $\alpha_\lambda^{\bar F}$ 满足推理条件：$\alpha_{\lambda+1}^{\bar F}\Rightarrow\bigcap_{i=1}^{\lambda}\alpha_i^{\bar F}$；$(\bar{x})_{\lambda+1}^{\bar F}$，$(\bar{x})_\lambda^{\bar F}\in\bar{G}_k^{\bar F}$。

(2) 区块智能检索准则

给定区块 $\bar{G}_k^F=\{\bar{G}_1^F,\bar{G}_2^F,\cdots,\bar{G}_n^F\}$，$\bar{G}_{\lambda+1}^F$在 $\bar{G}_\lambda^F$ 外被智能检索，$\bar{G}_{\lambda+1}^F$生成的区块矩阵 $\bar{\boldsymbol{B}}_{\lambda+1}^F$与 $\bar{G}_\lambda^F$ 生成的区块矩阵 $\bar{\boldsymbol{B}}_\lambda^F$ 满足推理条件：$\bar{\boldsymbol{B}}_\lambda^F\Rightarrow\bar{\boldsymbol{B}}_{\lambda+1}^F$；$\bar{G}_{\lambda+1}^F,\bar{G}_\lambda^F\in\bar{G}_k^F$，$\bar{G}_\lambda^F\subseteq\bar{G}_{\lambda+1}^F$。

给定区块 $\bar{G}_k^{\bar F}=\{\bar{G}_1^{\bar F},\bar{G}_2^{\bar F},\cdots,\bar{G}_n^{\bar F}\}$，$\bar{G}_{\lambda+1}^{\bar F}$在 $\bar{G}_\lambda^{\bar F}$ 内被智能检索，$\bar{G}_{\lambda+1}^{\bar F}$生成的区块矩阵 $\bar{\boldsymbol{B}}_{\lambda+1}^{\bar F}$与 $\bar{G}_\lambda^{\bar F}$ 生成的区块矩阵 $\bar{\boldsymbol{B}}_\lambda^{\bar F}$ 满足推理条件：$\bar{\boldsymbol{B}}_{\lambda+1}^{\bar F}\Rightarrow\bar{\boldsymbol{B}}_\lambda^{\bar F}$；$\bar{G}_{\lambda+1}^{\bar F},\bar{G}_\lambda^{\bar F}\in\bar{G}_k^{\bar F}$，$\bar{G}_{\lambda+1}^{\bar F}\subseteq\bar{G}_\lambda^{\bar F}$。

在 $\bar{\boldsymbol{B}}_\lambda^F\Rightarrow\bar{\boldsymbol{B}}_{\lambda+1}^F$ 条件下，$\bar{G}_1^F,\bar{G}_2^F,\cdots,\bar{G}_\lambda^F$ 被智能过滤；$\bar{G}_{\lambda+1}^F,\bar{G}_{\lambda+2}^F,\cdots,\bar{G}_n^F$ 是过滤剩余；$\bar{G}_\lambda^F=\bigcup_{i=1}^{\lambda}\bar{G}_i^F$。

在 $\bar{\boldsymbol{B}}_{\lambda+1}^{\bar F}\Rightarrow\bar{\boldsymbol{B}}_\lambda^{\bar F}$ 条件下，$\bar{G}_\lambda^{\bar F},\bar{G}_{\lambda+1}^{\bar F},\cdots,\bar{G}_n^{\bar F}$ 被智能过滤；$\bar{G}_{\lambda-1}^{\bar F},\bar{G}_{\lambda-2}^{\bar F},\cdots,\bar{G}_1^{\bar F}$ 是过滤剩余；$\bar{G}_\lambda^{\bar F}=\bigcup_{i=\lambda}^{n}\bar{G}_i^{\bar F}$。

4.2 $\bar{F}$-数据元智能分离-区块智能检索算法

为了简单又不失理论与应用的一般性，这里只给出 $\bar{F}$-数据元的智能分离与智能检索的算法，如图 1 所示。F-数据元的智能分离与智能检索算法、$\bar{F}$-区块的智能检索(F-区块的智能检索)算法，略。事实上，把 $\bar{F}$-数据元智能分离与智能检索算法稍加改进，容易得到 F-数据元智能分离与智能检索算法、$\bar{F}$-区块智能检索(F-区块智能检索)算法。

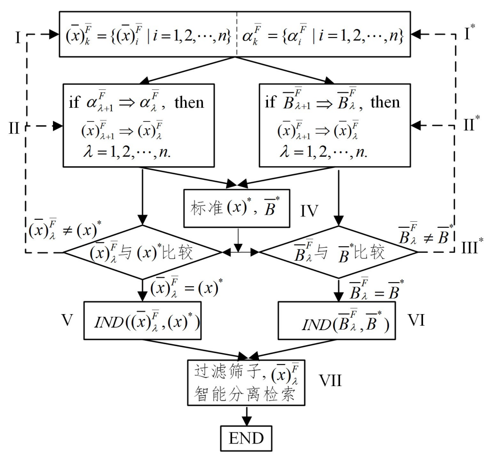

图 1 $\bar{F}$-数据元智能分离-检索算法框图

Fig. 1 Block diagram of $\bar{F}$-data element intelligent separation-retrieval algorithm

图 1 中，$\bar{G}_k^{\bar F}$ 是 $\bar{F}$-大数据的一个区块，$(\bar{x})_i^{\bar F}$ 是 $\bar{G}_k^{\bar F}$ 的区块元，$\bar{G}_k^{\bar F}=\{(\bar{x})_1^{\bar F},(\bar{x})_2^{\bar F},\cdots,(\bar{x})_n^{\bar F}\}$，$\forall(\bar{x})_i^{\bar F}\in\bar{G}_k^{\bar F}$，$(\bar{x})_i^{\bar F}$ 由数据元 x_i 构成，$(\bar{x})_i^{\bar F}=\{x_1,x_2,\cdots,x_m\}$，$(\bar{x})_i^{\bar F}$ 与$(\bar{x})_{i+1}^{\bar F}$满足$(\bar{x})_{i+1}^{\bar F}\subseteq(\bar{x})_i^{\bar F}$，$\alpha_k^{\bar F}$ 是$\bar{G}_k^{\bar F}$ 的属性集合，$\alpha_i^{\bar F}\in\alpha_k^{\bar F}$ 是$(\bar{x})_i^{\bar F}\in\bar{G}_k^{\bar F}$ 的属性集合；$\alpha_i^{\bar F}$ 与 $\alpha_{i+1}^{\bar F}$满足 $\alpha_{i+1}^{\bar F}\subseteq\alpha_i^{\bar F}$。I 与 I* 是算法准备，是算法的基础模块；II 与 II* 分别是属性推理族，是矩阵推理族生成模块，它们分别是属性递推推理、矩阵递推推理库；IV 是被定义的$(\bar{x})_k^{\bar F}$ 的识别标准$(x)^*$ 以及 $\bar{\boldsymbol{B}}_k^{\bar F}$ 的识别标准 $\bar{\boldsymbol{B}}^*$，$(x)^*$ 与 $\bar{\boldsymbol{B}}^*$构成标准模块，它是 II，II* 区块元智能分离-检索的生成；III 与 III* 分别是$(\bar{x})_k^{\bar F}$，$\bar{\boldsymbol{B}}_k^{\bar F}$ 的比较模块；V，VI 分别是$(\bar{x})_k^{\bar F}$ 与$(x)^*$，$\bar{\boldsymbol{B}}_k^{\bar F}$ 与 $\bar{\boldsymbol{B}}^*$ 的分辨模块；VII 是智能检索-过滤模块，IND=indiscernible。

算法的过程如下：

Step 1 给定 $\bar{F}$-区块元$(\bar{x})_k^{\bar F}=\{x_1,x_2,\cdots,x_n\}$，$(\bar{x})_k^{\bar F}$ 的属性集合 $\alpha_k^{\bar F}=\{\alpha_1,\alpha_2,\cdots,\alpha_n\}$；$\bar{F}$-区块元$(\bar{x})_\lambda^{\bar F}$ 生成的 $\bar{F}$-区块矩阵 $\bar{\boldsymbol{B}}_\lambda^{\bar F}$；I 进入 II，II*。

Step 2 在 $\alpha_k^{\bar F}$ 内删除属性的条件下，$(\bar{x})_\lambda^{\bar F}$ 生成$(\bar{x})_{\lambda+1}^{\bar F}$，$(\bar{x})_{\lambda+1}^{\bar F}\subseteq(\bar{x})_\lambda^{\bar F}$，$(\bar{x})_\lambda^{\bar F}$ 的属性集合 $\alpha_\lambda^{\bar F}$ 与$(\bar{x})_{\lambda+1}^{\bar F}$的属性集合 $\alpha_{\lambda+1}^{\bar F}$满足 $\alpha_{\lambda+1}^{\bar F}\subseteq\alpha_\lambda^{\bar F}$；属性推理族 if $\alpha_{\lambda+1}^{\bar F}\Rightarrow\alpha_\lambda^{\bar F}$，then $(\bar{x})_{\lambda+1}^{\bar F}\Rightarrow(\bar{x})_\lambda^{\bar F}$ 生成，$\lambda=1,2,\cdots,n$。$\bar{F}$-区块矩阵 $\bar{\boldsymbol{B}}_\lambda^{\bar F}$，$\bar{\boldsymbol{B}}_{\lambda+1}^{\bar F}$的生成与区块元的生成类似，$\bar{\boldsymbol{B}}_{\lambda+1}^{\bar F}\subseteq\bar{\boldsymbol{B}}_\lambda^{\bar F}$，矩阵推理族 if $\bar{\boldsymbol{B}}_{\lambda+1}^{\bar F}\Rightarrow\bar{\boldsymbol{B}}_\lambda^{\bar F}$，then $(\bar{x})_{\lambda+1}^{\bar F}\Rightarrow(\bar{x})_\lambda^{\bar F}$ 生成，II，II* 进入 III，III*。

Step 3 定义$(\bar{x})_\lambda^{\bar F}$ 的识别标准$(x)^*$，$\bar{\boldsymbol{B}}_\lambda^{\bar F}$ 的识别标准 $\bar{\boldsymbol{B}}^*$，生成 IV。

Step 4 属性递推推理生成$(\bar{x})_r^{\bar F}$ 与$(x)^*$的比较，矩阵递推推理生成 $\bar{\boldsymbol{B}}_r^{\bar F}$ 与 $\bar{\boldsymbol{B}}^*$ 的比较，若$(\bar{x})_r^{\bar F}\neq(x)^*$，III 返回到 I，重复 Step1－Step4；若 $\bar{\boldsymbol{B}}_r^{\bar F}\neq\bar{\boldsymbol{B}}^*$，III* 返回到 I*，重复 Step1－Step4。

Step 5 若$(\bar{x})_r^{\bar F}=(x)^*$，$\bar{\boldsymbol{B}}_r^{\bar F}=\bar{\boldsymbol{B}}^*$ 满足 $IND((\bar{x})_r^{\bar F},(x)^*)$，$IND(\bar{\boldsymbol{B}}_r^{\bar F},\bar{\boldsymbol{B}}^*)$，V 与 VI 进入 VII；$(\bar{x})_r^{\bar F}$ 从$(\bar{x})_\lambda^{\bar F}$ 内被分离，被过滤筛子检索获取；$\bar{\boldsymbol{B}}_r^{\bar F}$ 对$(\bar{x})_r^{\bar F}$ 的正确性给出验证-确认，进入 Step 6。

Step 6 END。

利用第 2 节中的预备概念以及第 3 节、第 4 节中的理论结果，给出第 5 节的内容。

5 大数据智能检索与大数据区块元智能分离-获取的应用

本节的例子取自公共安全系统大数据中的 $\bar{F}$-区块元，因为一些原因，例子中的属性名称略。例子中的 $\bar{F}$-区块元矩阵中的数值是原始数值经过技术方法后得到的，技术方法处理后的数值不影响例子的分析。取 $\bar{F}$-区块元$(\bar{x})_t^{\bar F}\in\bar{G}_k^{\bar F}$，$\alpha_t^{\bar F}$ 是$(\bar{x})_t^{\bar F}$ 的属性集合。

$$(\bar{x})_t^{\bar F}=\{x_1,x_2,x_3,x_4,x_5,x_6,x_7\} \tag{59}$$

$$\alpha_t^{\bar F}=\{\alpha_1,\alpha_2,\alpha_3,\alpha_4,\alpha_5,\alpha_6,\alpha_7\} \tag{60}$$

式(59)中，$\forall x_i\in(\bar{x})_t^{\bar F}$ 是$(\bar{x})_t^{\bar F}$ 的数据元，$\alpha_i\in\alpha_t^{\bar F}$ 是 x_i 的属性，$\bar{\boldsymbol{B}}_t^{\bar F}$ 是 $\bar{F}$-区块元$(\bar{x})_t^{\bar F}$ 生成的 $\bar{F}$-区块元矩阵。

$$\bar{\boldsymbol{B}}_t^{\bar F}=\begin{pmatrix}1.68 & 1.77 & 1.59 & 1.70 & 1.63 & 1.78 & 1.60\\ 1.10 & 1.26 & 1.03 & 0.98 & 1.04 & 1.26 & 0.83\\ 0.73 & 0.80 & 0.60 & 0.70 & 0.67 & 0.79 & 0.59\end{pmatrix} \tag{61}$$

在 $t+1\in T$ 时刻，案例现场对 $t\in T$ 时刻获取的属性(证据) $\alpha_4,\alpha_6,\alpha_7$ 不被采信，$\alpha_4,\alpha_6,\alpha_7$ 从 $\alpha_t^{\bar F}$ 内被删除，式(59)、式(60)生成$(\bar{x})_{t+1}^{\bar F}$，$\alpha_{t+1}^{\bar F}$。

$$(\overline{x})_{t+1}^{\overline{F}}=\{x_1,x_2,x_3,x_5\} \tag{62}$$

$$\alpha_{t+1}^{\overline{F}}=\{\alpha_1,\alpha_2,\alpha_3,\alpha_5\} \tag{63}$$

$(\overline{x})_{t+1}^{\overline{F}}$生成$\overline{F}$-区块元矩阵$\overline{\boldsymbol{B}}_{t+1}^{\overline{F}}$。

$$\overline{\boldsymbol{B}}_{t+1}^{\overline{F}}=\begin{pmatrix}1.68 & 1.77 & 1.59 & 1.63\\ 1.10 & 1.26 & 1.03 & 1.04\\ 0.73 & 0.80 & 0.60 & 0.67\end{pmatrix} \tag{64}$$

式(59)－式(64)生成算法中的属性推理、矩阵推理。

$$\text{if } \alpha_{t+1}^{\overline{F}}\Rightarrow\alpha_t^{\overline{F}}, \text{ then } (\overline{x})_{t+1}^{\overline{F}}\Rightarrow(\overline{x})_t^{\overline{F}} \tag{65}$$

$$\text{if } \overline{\boldsymbol{B}}_{t+1}^{\overline{F}}\Rightarrow\overline{\boldsymbol{B}}_t^{\overline{F}}, \text{ then } (\overline{x})_{t+1}^{\overline{F}}\Rightarrow(\overline{x})_t^{\overline{F}} \tag{66}$$

在$t+k\in T$时刻，以案例现场目击者提供的证据，属性α_2不被采信，式(62)、式(63)生成$(\overline{x})_{t+k}^{\overline{F}}$，$\alpha_{t+k}^{\overline{F}}$。

$$(\overline{x})_{t+k}^{\overline{F}}=\{x_1,x_3,x_5\} \tag{67}$$

$$\alpha_{t+k}^{\overline{F}}=\{\alpha_1,\alpha_3,\alpha_5\} \tag{68}$$

$(\overline{x})_{t+k}^{\overline{F}}$生成$\overline{F}$-区块元矩阵$\overline{\boldsymbol{B}}_{t+k}^{\overline{F}}$。

$$\overline{\boldsymbol{B}}_{t+k}^{\overline{F}}=\begin{pmatrix}1.68 & 1.59 & 1.63\\ 1.10 & 1.03 & 1.04\\ 0.73 & 0.60 & 0.67\end{pmatrix} \tag{69}$$

类似得到：

$$\text{if } \alpha_{t+k}^{\overline{F}}\Rightarrow\alpha_{t+1}^{\overline{F}}, \text{ then } (\overline{x})_{t+k}^{\overline{F}}\Rightarrow(\overline{x})_{t+1}^{\overline{F}} \tag{70}$$

$$\text{if } \overline{\boldsymbol{B}}_{t+k}^{\overline{F}}\Rightarrow\overline{\boldsymbol{B}}_{t+1}^{\overline{F}}, \text{ then } (\overline{x})_{t+k}^{\overline{F}}\Rightarrow(\overline{x})_{t+1}^{\overline{F}} \tag{71}$$

$(\overline{x})_{t+k}^{\overline{F}}$从$(\overline{x})_{t+1}^{\overline{F}}$内被智能分离-检索，$(\overline{x})_{t+k}^{\overline{F}}$被智能筛选-获取。

结果分析与结果认证如下：例子给出的分析式(59)－式(71)和案例的最终结论式(70)、式(71)与实际结论相符并被确认。

例子的背景与结果的说明如下。

1)例子取自公共安全系统大数据的重大案例，案例的特征为：案例涉及多地，多人合伙交叉集团作案，案情复杂，案例侦破与结案历时多年。利用逆P-集合作为数学方法，模型和推理相结合形成$\overline{F}$-区块中的数据元(作案者)智能分离-检索算法，利用大数据的特征与逆P-集合的动态结构完成案例侦破的全过程。2)例子是整个案例侦破、结果获取的一个简化，一些原因的限制，案例侦破中的多个重要过程、案例属性删除的原因与属性的取证细节都被略去。3)x_1，x_3，x_5构成$\overline{F}$-区块元$(\overline{x})_{t+k}^{\overline{F}}$，$x_1$，$x_3$，$x_5$是案例的主犯；$x_1$，$x_3$，$x_5$从$x_1\sim x_7$中被分离-检索获取。

结束语 “大数据”是诸多应用系统研究中的一个名词，文献[53-68]给出了多个实际应用的案例，给出了多种可被应用的方法。简言之，大数据已渗透到多个应用研究中，为应用研究提供了新的思想方法，大数据的研究引起了人们的兴趣。大数据具有什么样的结构？大数据中潜藏着什么样的不为人知的特征？大数据是否可以分类？分类的依据是什么？在已有文献中，这些问题没有给出讨论。文献[1]利用具有动态特征的数学模型，给出大数据结构的讨论，给出几个具有重要应用的特征，这些特征存在于大数据内，其给出了两类大数据的结构与它们具有的逻辑特征。利用文献[1]中的结果，本文给出Ⅴ型大数据新的研究，并给出Ⅴ型大数据新的特征；把推理方法交叉到Ⅴ型大数据中，给出数据智能分离与智能检索的理论与应用研究，并给出应用实例。或许本文给出的研究与结果、文献[1]给出的研究与结果为从事大数据基础理论与应用的研究者提供了一个研究思路与必要的数学方法。大数据是一个具有数学概念内涵的理论与应用研究分支；把数学概念渗透到大数据与大数据规律[69-72]研究中，或许使得人们对大数据获得新的理论认识，从单纯的大数据应用研究过渡到大数据的理论研究成为一个必然过程。

参考文献

[1] SHI K Q. Big data structure-logic characteristics and big data law [J]. Journal of Shandong University (Natural Science), 2019,54(2):1-29.

[2] SHI K Q. P-sets[J]. Journal of Shandong University (Natural Science),2008,43(11):77-84.

[3] SHI K Q. P-sets and its applications [J]. Advances in Systems Science and Applications,2009,9(2):209-219.

[4] SHI K Q. P-sets and its applied characteristics [J]. Computer Science,2010,37(8):1-8.

[5] SHI K Q. P-reasoning and P-reasoning discovery- identification of information [J]. Computer Science,2011,38(7):1-9.

[6] SHI K Q. P-sets,inverse P-sets and the intelligent fusion-filter identification of information [J]. Computer Science, 2012, 39(4):1-13.

[7] FAN C X,LIN H K. P-sets and the reasoning- identification of disaster information [J]. International Journal of Convergence Information Technology,2012,7(1):337-345.

[8] LIN H K,FAN C X. The dual form of P-reasoning and identification of unknown attribute [J]. International Journal of Digital Content Technology and its Applications,2012,6(1):121-131.

[9] ZHANG L,CUI Y Q,SHI K Q. Outer P-sets and data internal recovery [J]. Systems Engineering and Electronics, 2010, 32(6):1233-1238.

[10] WANG Y,GENG H Q,SHI K Q. P-sets and dependence discovery of dynamic information [J]. Systems Engineering and Electronics,2011,33(9):2035-2038.

[11] LI Y Y,LIN H K,SHI K Q. Characteristics of data discrete interval and data discovery-application[J]. Systems Engineering and Electronics,2011,33(10):2258-2262.

[12] ZHANG L,TANG J H,SHI K Q. The fusion of internal packet information and its feature of attribute conjunction [J]. Journal of Shandong University(Natural Science),2014,49(2):93-97.

[13] SHI K Q,ZHANG L. Internal P-set and data outer-recovery [J]. Journal of Shandong University (Natural Science),2009, 44(4):8-14.

[14] ZHANG G Y,ZHOU H Y,SHI K Q. P-sets and the recovery-identification double [J]. Systems Engineering and Electronics, 2010,32(9):1919-1924.

[15] ZHANG L,REN X F. P-sets and $(f,\overline{f})$-heredity[J]. Quantitative Logic and Soft Computing,2010,2(1):735-743.

[16] ZHANG L,XIU M,SHI K Q. P-sets and application of internal-outer data circle [J]. Quantitative Logic and Soft Computing, 2010,2(1):581-591.

[17] QIU Y F,CHEN B H. f-model generated by P-sets [J]. Quantitative Logic and Soft Computing,2010,2(1):613-620.

[18] LI Y Y,ZHANG L,SHI K Q. Generation and recovery of compressed data and redundant data[J]. Quantitative Logic and Soft Computing,2010,2(1):661-671.

[19] XIU M,SHI K Q,ZHANG L. P-sets and $\bar{F}$-data selection-discovery [J]. Quantitative Logic and Soft Computing,2010,2(1): 791-799.

[20] ZHAO S L,FAN C X,SHI K Q. Outer P-information generation and its reasoning-searching discovery [J]. Journal of Shandong University (Natural Science),2012,47(1):99-104.

[21] ZHAO S L,WU S L,SHI K Q. Internal P- reasoning information recovery and attribute hiding searching discovery[J]. Computer Science,2013,40(4):209-213.

[22] SHI K Q. Function P-sets[J]. Journal of Shandong University (Natural Science),2011,46(2):62-69.

[23] SHI K Q. Function P-sets[J]. International Journal of Machine Learning and Cybernetics,2011,2(4):281-288.

[24] SHI K Q. P-information law intelligent fusion and soft information image intelligent generation[J]. Journal of Shandong University (Natural Science),2014,49(4):1-17.

[25] TANG J H,ZHANG L,SHI K Q. Intelligent fusion of information law and its inner separation[J]. Computer Science,2015, 42(2):204-209.

[26] LIN R,FAN C X. Function P-sets and dynamic characteristics of information law[J]. Journal of Shandong University (Natural Science),2012,47(1):121-126.

[27] REN X F,ZHANG L,SHI K Q. Two types of dynamic information law models and their applications in information camouflage and risk identification [J]. Computer Science, 2018, 45(9):230-236.

[28] CHEN B H,ZHANG L,SHI K Q. Intelligent dynamic fusion of packet information and the intelligent state recognition of information law[J]. Journal of Shandong University (Natural Science),2018,53(2):83-87.

[29] TANG J H,ZHANG L,SHI K Q,et al. Outer P-information law reasoning andits application in intelligent fusion and separating of information law [J]. Microsystem Technologies, 2018, 24(10):4389-4398.

[30] SHI K Q. P-augmented matrix and dynamic intelligent discovery identification of information [J]. Journal of Shandong University (Natural Science),2015,50(10):1-12.

[31] ZHANG L, REN X F, SHI K Q. The dynamic segmentation characteristics of P-augmented matrix and the dynamic intelligent acquisition of P-information [J]. International Journal of Applied Decision Sciences,2016,9(4):413-425.

[32] ZHANG L,REN X F,SHI K Q. Intelligent switch- camouflage of information laws and P-law augmented matrices [J]. Journal of Shandong University (Natural Science),2016,51(8):90-97.

[33] TANG J H,CHEN B H,SHI K Q. P-augmented matrix reasoning and intelligent decomposition mining of information [J]. Journal of Shandong University (Natural Science),2016,51(12): 61-66.

[34] LIU J Q,ZHANG H Y. Information P-dependence and P-dependence mining-sieving [J]. Computer Science, 2018, 45(7): 202-206.

[35] SHI K Q. Inverse P-sets [J]. Journal of Shandong University (Natural Science),2012,47(1):98-109.

[36] SHI K Q, TANG J H, ZHANG L. Intelligent fusion of inverse packet information and recessive transmission of information's intelligent hiding [J]. Systems Engineering and Electronics, 2015,37(3):599-605.

[37] SHI K Q. Function inverse P-sets and information law fusion [J]. Journal of Shandong University: Natural Science, 2012, 47(8):73-80.

[38] LIN K K,FAN C X. Embedding camouflage of inverse P-information and application[J]. International Journal of Convergence Information and Technology,2012,7(20):471-480.

[39] FAN C X,HUANG S L. Inverse P-reasoning discovery identification of inverse P-information[J]. International Journal of Digital Content Technology and its Applications,2012,6(20):735-744.

[40] SHI K Q. Function inverse P-sets and the hiding information generated by function inverse P-information law fusion[C]// Proceedings of the 13th IFIP WG 6.11 Conference on e-Business,e-Services,and e-Society. Sanya,China,2014:224-237.

[41] TANG J H,CHEN B H,ZHANG L,BAI X R. Function inverse P-sets and the dynamic separation of inverse P- information laws [J]. Journal of Shandong University (Natural Science), 2013, 48(8):104-110.

[42] GUO H L,REN X F,ZHANG L. Relationships between dynamic data mining and P-augmented matrix[J]. Journal of Shandong University (Natural Science),2016,51(8):105-110.

[43] GUO H L,CHEN B H,TANG J H. Inverse P-sets and intelligent fusion mining-discovery of information [J]. Journal of Shandong University (Natural Science), 2013, 48(8): 97-103, 110.

[44] ZHANG L, REN X F, SHI K Q, et al. Inverse packet matrix reasoning model-based the intelligent dynamic separation and acquisition of educational information[J]. Microsystem Technologies,2018,24(10):4415-4421.

[45] REN X F,ZHANG L,SHI K Q,et al. Inverse P-augmented matrix method-based the dynamic findings of unknown information [J]. Microsystem Technologies,2018,24(10):4187-4192.

[46] ZHANG L,REN X F,SHI K Q. Inverse P-information law models and the reality-camouflage intelligent transformations of information image[C]// Proceedings of the 2016 International Conference on Network and Information Systems for Computers. Washington:IEEE,2016:337-341.

[47] ZHANG L,REN X F,SHI K Q. Inverse P-data models and data intelligent separation[C]// Proceedings of the 2016 International Conference on Electronic Information Technology and Intellectualization. 2016.

[48] REN X F,ZHANG L,SHI K Q. Surplus- deficiency of cardinal number and inverse P-augmented matrices[J]. Journal of Shandong University (Natural Science),2015,50(10):13-18,26.

[49] ZHANG L,REN X F. Surplus-deficient theorem of cardinal number and data internal-outer mining- separation [J]. Journal of Shandong University (Natural Science),2015,50(8):90-94.

[50] GUO H L,ZHANG L. Data separation and its attribute state characteristics [J]. Journal of Shandong University (Natural Science),2017,52(12):89-94.

[51] REN X F,ZHANG L,SHI K Q. Boundary Characteristics of Inverse P-sets and System Condition Monitoring[J]. Computer Science,2016,43(10):211-213,255.

[52] REN X F,ZHANG L. Perturbation theorems of inverse P-sets and perturbation-based data mining[J]. Journal of Shandong University (Natural Science),2016,51(12):54-60.
[53] WANG F Y,CARLEY K M,ZENG D,et al. Social computing: From social informatics to social intelligence [J]. IEEE Intelligent systems,2007,22(2):79-83.
[54] ZHANG N,WANG F Y,ZHU F,et al. DynaCAS:Computational experiments and decision support for ITS [J]. IEEE Intelligent Systems,2008,23(6):19-23.
[55] WANG F Y. Parallel control and management for intelligent transportation systems:Concepts,architectures,and applications [J]. IEEE Transactions on Intelligent Transportation Systems, 2010,11(3):630-638.
[56] WANG F Y. Scanning the issue and beyond:Parallel driving with software vehicular robots for safety and smartness [J]. IEEE Transactions on Intelligent Transportation Systems, 2014,15(4):1381-1387.
[57] LV Y,DUAN Y,KANG W,et al. Traffic flow prediction with big data:A deep learning approach[J]. IEEE Trans. Intelligent Transportation Systems,2015,16(2):865-873.
[58] LV Y,ZHANG X,KANG W,et al. Managing Emergency Traffic Evacuation With a Partially Random Destination Allocation Strategy:A Computational-Experiment-Based Optimization Approach[J]. IEEE Trans. Intelligent Transportation Systems, 2015,16(4):2182-2191.
[59] YUAN Y,WANG F Y,ZENG D. Developing a cooperative bidding framework for sponsored search markets-An evolutionary perspective [J]. Information Sciences,2016,369:674-689.
[60] WANG F Y. The emergence of intelligent enterprises:From CPS to CPSS [J]. IEEE Intelligent Systems,2010,25(4):85-88.
[61] GOODHOPE K,KOSHY J,KREOS J,et al. Building LinkedIn's Real-time Activity Data Pipeline [J]. IEEE Data Eng. Bull. 2012,35(2):33-45.
[62] MCKUSICK K,QUINLAN S. GFS:evolution on fast-forward [J]. Communications of the ACM,2010,53(3):42-49.
[63] MELNIK S,GUBAREV A,LONG J J,et al. Dremel:interactive analysis of web-scale datasets [J]. Proceedings of the VLDB Endowment,2010,3(1/2):330-339.
[64] BU Y,HOWE B,BALAZINSKA M,et al. HaLoop:efficient iterative data processing on large clusters [J]. Proceedings of the VLDB Endowment,2010,3(1/2):285-296.
[65] CHEN S. Cheetah:a high performance,custom data warehouse on top of MapReduce [J]. Proceedings of the VLDB Endowment,2010,3(1/2):1459-1468.
[66] LUO T,LEE R,MESNIER M,et al. Storage-DB:heterogeneity-aware data management to exploit the full capability of hybrid storage systems [J]. Proceedings of the VLDB Endowment, 2012,5(10):1076-1087.
[67] WONG P C,SHEN H W,JOHNSON C R,et al. The top 10 challenges in extreme-scale visual analytics [J]. IEEE computer graphics and applications,2012,32(4):63-67.
[68] PIKE R,DORWARD S,GRIESEMER R,et al. Interpreting the data:Parallel analysis with Sawzall[J]. Scientific Programming, 2005,13(4):277-298.
[69] SHI K Q,YAO B X. Function S-rough sets and law identification [J]. Science in China E:Information Science,2008,38(4): 553-564.
[70] SHI K Q,ZHAO J L. Function S-rough sets and security-authentication of hiding law[J]. Science in China E:Information Science,2008,38(8):1234-1243.
[71] SHI K Q,YAO B X. Function S-rough sets and law identification [J]. Science in China F:Information Science,2008,51(5): 499-510.
[72] SHI K Q,ZHAO J L. Function S-rough sets and security-authentication of hiding law[J]. Science in China F:information Sciences,2008,51(7):924-935.

HAO Xiu-mei, born in 1965, Ph.D, professor, Ph.D supervisor. Her main research interests include risk data identification, big data analysis and application and rough system theory and application.

SHI Kai-quan, born in 1945, professor, Ph.D supervisor. His main research interests include big data theory and application, information intelligent system. He proposed S-rough sets, function S-rough sets, P-sets, inverse P-sets, function P-sets and function inverse P-sets. More papers were published in *Science in China E*, *Science in China F* and other journals. In 2010—2016, a number of papers were selected as "100 most influential domestic academic papers in China". The two papers published in *Computer Science*, "Function S-rough sets, function rough sets and separation-composition of law for information systems" and "P-sets, inverse P-sets and the intelligent fusion-filter identification of information" were listed among them. A number of papers were selected in *Frontrunner* 5000. A number of monographs have been published.