Inverse P-data Models and Data Intelligent Separation

Ling Zhang, Xuefang Ren and Kaiquan Shi

ABSTRACT

Firstly by improving inverse P-sets, inverse P-data model is proposed, it consists of internal and external inverse P-data models, and demonstrates dynamic characteristics; then the expansion-contraction generations of attribute disjunctive normal form for inverse P-data model are shown, and inverse P-data reasoning is proposed; finally intelligent data separation-discovery is given, and verified by an application.

INTRODUCTION

Inverse P-sets and its theoretical models are put forward in Ref. [1,2], it has been applied in the research of dealing dynamic data or information [1-5]. Inverse P-set is a set pair consisting of the internal inverse P-set \( \overline{X}^p \) and the external inverse P-set \( \overline{X}^r \), namely \( (\overline{X}^p, \overline{X}^r) \), and it has dynamic characteristics as shown in Ref. [1,2].

The dynamic characteristics of inverse P-sets are the same as the dynamic characteristics of the processes that data unit \( y_i \) outside \( y \) invades into \( y \), or the redundancy data unit \( y_i \) inside data \( y \) is deleted from \( y \). Data unit \( y_i \in y \) invades into

Ling Zhang, Xuefang Ren, School of Information Engineering, Longyan University, 1 Dongxiao Beilu, Longyan, Fujian, 364012, China
Kaiquan Shi, School of Mathematics, Shandong University, 27 Shanda Nanlu, Jinan, Shandong, 250100, China
y can be found in the experiment of resisting interfere to keeping the system output stable when system is interfered by unknown data. Data unit \( y_i \in F \) is deleted from \( y \) can be found in the experiment of obtaining selective preference data from data \( y \), if a data unit hasn’t selected, it will be deleted from \( y \). By improve inverse P-set to research inverse P-data is the main topic of this paper. For convenience of discussion, the main results of inverse P-sets are given in section 2[1,2,6].

**INVERSE P-SETS THEORY [1, 2]**

Given finite ordinary element set \( X = \{x_1, x_2, \cdots, x_q\} \subseteq U \), and its attribute set \( \alpha = \{\alpha_1, \alpha_2, \cdots, \alpha_q\} \subseteq V \).

Suppose \( F = \{f_1, f_2, \cdots, f_n\} \) and \( \bar{F} = \{\bar{f}_1, \bar{f}_2, \cdots, \bar{f}_n\} \) are element or attribute transfer families, \( f \in F \) and \( \bar{f} \in \bar{F} \) are element or attribute transfers. \( \bar{X}^F \) is called the internal inverse P-set generated by \( X \), moreover,

\[
\bar{X}^F = X \cup X^*
\]

Where \( X^\ast \) is the \( F \)-element supplemented set of \( X \), and \( X^\ast = \{u_i \mid u_i \in U, u_i \in X \} \). The attribute set \( \alpha^F \) of \( \bar{X}^F \) satisfies

\[
\alpha^F = \alpha \cup \{\alpha_i' \mid f(\beta) = \alpha_i' \in \alpha, \beta \in \alpha, f \in F\}. \quad (1)
\]

Where \( \bar{X}^F = \{x_1, x_2, \cdots, x_r\} \), \( q \leq r, q, r \in \mathbb{N}^+ \).

\( \bar{X}^F \) is the external inverse packet set generated by \( X \), moreover

\[
\bar{X}^F = X - X^\ast,
\]

Where \( X^\ast \) is \( F \)-element deleted set of \( X \), and \( X^\ast = \{x_i \mid x_i \in X, \bar{f}(x_i) = u_i \subseteq X, \bar{f} \in \bar{F}\} \). The attribute set \( \alpha^F \) of \( \bar{X}^F \) satisfies

\[
\alpha^F = \alpha - \{\beta_i \mid \bar{f}(\alpha_i) = \beta_i \in \alpha, \alpha_i \in \alpha, \bar{f} \in \bar{F}\}. \quad (2)
\]

Where \( \bar{X}^F = \{x_1, x_2, \cdots, x_q\} \), \( p \leq q, p, q \in \mathbb{N}^+ \).

The set pair \( (\bar{X}^F, \bar{X}^\ast) \) is called the inverse P-set generated by \( X \).

It is obtained from formula (1) that \( \alpha^F_1 \subseteq \alpha^F_2 \subseteq \cdots \subseteq \alpha^F_n \subseteq \alpha^F_n \), then the corresponding internal inverse P-sets satisfying

\[
\bar{X}^F_1 \subseteq \bar{X}^F_2 \subseteq \cdots \subseteq \bar{X}^F_{n-1} \subseteq \bar{X}^F_n. \quad (1)
\]
It is obtained from formula (2) that \( \alpha_n^F \subseteq \alpha_{n-1}^F \subseteq \cdots \subseteq \alpha_2^F \subseteq \alpha_1^F \), then the corresponding outer inverse P-sets satisfying
\[
\bar{X}_n^F \subseteq \bar{X}_{n-1}^F \subseteq \cdots \subseteq \bar{X}_2^F \subseteq \bar{X}_1^F .
\]  

Inverse P-set family is obtained from Eq. (3) and Eq.(4) as follow
\[
\{(\bar{X}_i^F, \bar{X}_j^F) | i \in I, j \in J\}.
\]  

Eq. (5) is the inverse P-set family generated by \( X \).

For \( \forall x_i \in X \), the attribute \( \alpha_i \) of \( x_i \) satisfies attribute disjunctive normal form, namely
\[
\alpha_i = \alpha_1 \lor \alpha_2 \lor \cdots \lor \alpha_q ,
\]  

Given \( \bar{X}_n^F = \{x_1, \ldots, x_q, x_{q+1}, \ldots, x_r\} \), and its attribute set \( \alpha_n^F = \{\alpha_1, \ldots, \alpha_q, \alpha_{q+1}, \ldots, \alpha_r\} \), for \( \forall x_j \in \bar{X}_n^F \), its attribute \( \alpha_j \) satisfies attribute disjunctive normal form expansion, namely
\[
\alpha_j = (\alpha_1 \lor \alpha_2 \lor \cdots \lor \alpha_q) \lor \alpha_{q+1} \lor \cdots \lor \alpha_r \\
= \alpha_1 \lor \alpha_2 \lor \cdots \lor \alpha_q \lor \alpha_{q+1} \lor \cdots \lor \alpha_r
\]  

Given \( \bar{X}_n^F = \{x_1, x_2, \ldots, x_p\} \), and its attribute set \( \alpha_n^F = \{\alpha_1, \ldots, \alpha_p\} \), for \( \forall x_k \in \bar{X}_n^F \), the attribute \( \alpha_k \) of \( x_k \) satisfies attribute disjunctive normal form contraction, namely
\[
\alpha_i = (\alpha_1 \lor \alpha_2 \lor \cdots \lor \alpha_p \lor \alpha_{p+1} \lor \cdots \lor \alpha_q) - (\alpha_p \lor \alpha_{p+1} \lor \cdots \lor \alpha_q) \\
= \alpha_1 \lor \alpha_2 \lor \cdots \lor \alpha_p
\]  

Given inverse P-set \( (\bar{X}_n^F, \bar{X}_n^F) \), and its attribute set \( (\alpha_n^F, \alpha_n^F) \), for \( \forall x_i \in \bar{X}_n^F \) and \( \forall x_j \in \bar{X}_n^F \), the attribute set \( \alpha_i \) of \( x_i \) and the attribute set \( \alpha_j \) of \( x_j \) are satisfies attribute disjunctive normal form expansion-contraction, namely
\[
(\alpha_i, \alpha_j) = (\alpha_1 \lor \alpha_2 \lor \cdots \lor \alpha_p, \alpha_1 \lor \alpha_2 \lor \cdots \lor \alpha_p)
\]
Where Eq.(9) indicates $\alpha_i = \alpha_1 \wedge \alpha_2 \wedge \cdots \wedge \alpha_j$, $\alpha_j = \alpha_1 \wedge \alpha_2 \wedge \cdots \wedge \alpha_p$; $p, q, r$ in Eqs. (9) and (6) satisfy $p < q < r$; $p, q, r \in \mathbb{N}^+$. 

**INVERSE P-DATA AND ITS EXPANSION-CONTRACTION GENERATION OF ATTRIBUTE DISJUNCTIVE NORMAL FORM**

**Definition 1.** \(Y\) is data set generated by the finite ordinary element set \(X = \{x_1, x_2, \cdots, x_q\}\), \(\alpha = \{\alpha_1, \alpha_2, \cdots, \alpha_q\}\) is its attribute set, moreover

\[
y = \{y_1, y_2, \cdots, y_n\} = \left\{\sum_{i=1}^{q} y_{i,1}, \sum_{i=1}^{q} y_{i,2}, \cdots, \sum_{i=1}^{q} y_{i,n}\right\}. \quad (10)
\]

Where for any \(y_i\) of \(Y\), \(y_i = (y_{i,1}, y_{i,2}, \cdots, y_{i,n})\), \(y_i\) is the \(n\) dimensional data of element \(x_i \in X\), \(i = 1, 2, \cdots, q\); \(\forall y_{k,i} \in y_i, y_{k,i} \in \mathbb{R}^+\).

**Definition 2.** \(\bar{Y}^F\) is called internal inverse P-data generated by internal inverse P-set \(\bar{X}^F = \{x_1, x_2, \cdots, x_q, x_{q+1}, \cdots, x_r\}\), where \(\alpha^F = \{\alpha_1, \cdots, \alpha_r\}\) is its attribute set, moreover

\[
\bar{y} = \{y_1', y_2', \cdots, y_n'\} = \left\{\sum_{i=1}^{r} y_{i,1}', \sum_{i=1}^{r} y_{i,2}', \cdots, \sum_{i=1}^{r} y_{i,n}'\right\}. \quad (11)
\]

**Definition 3.** \(\bar{Y}^F\) is called external inverse P-data generated by external inverse P-set \(\bar{X}^F = \{x_1, x_2, \cdots, x_p\}\), \(\alpha^F = \{\alpha_1, \alpha_2, \cdots, \alpha_p\}\) is its attribute set, moreover

\[
\bar{y} = \{y_1', y_2', \cdots, y_n'\} = \left\{\sum_{i=1}^{p} y_{i,1}', \sum_{i=1}^{p} y_{i,2}', \cdots, \sum_{i=1}^{p} y_{i,n}'\right\}. \quad (12)
\]

In definitions 1-3, \(p, q\) and \(r\) satisfy \(p < q < r\); \(p, q, r \in \mathbb{N}^+\). \(\bar{Y}^F\) is obtained by supplemented \(r-q\) data into \(y\). \(\bar{Y}^F\) is obtained by deleted \(q-p\) data from \(y\).

**Definition 4.** The set pair composed by \(\bar{Y}^F\) and \(\bar{Y}^F\) is called the inverse P-data generated by inverse P-set \((\bar{X}^F, \bar{X}^F)\), denoted by

\[
(\bar{Y}^F, \bar{Y}^F). \quad (13)
\]

**Theorem 3.** The sufficient and necessary condition for the internal inverse P-data \(\bar{Y}^F\) generated by \(y\) is that the attribute \(\alpha_i \in \alpha^F\) of \(y_i' \in \bar{Y}^F\) satisfies attribute disjunctive normal form of \(y_i \in Y\), namely
\[\alpha_i = \left( \bigvee_{i=1}^q \alpha_i \right) \bigvee_{i+q+1}^r \alpha_i\]

Where \(\alpha_i = \bigvee_{i=1}^q \alpha_i\) is attribute disjunctive normal form of \(y_i \in y\).

**Theorem 4.** The sufficient and necessary condition for external inverse P-data \(\bar{y}^F\) generated by data \(y\) is that attribute \(\alpha_j \in \alpha^F\) of \(y_j \in \bar{y}^F\) satisfies the attribute disjunctive normal form contraction of \(y_j \in y\), namely

\[\alpha_j = \left( \bigvee_{i=1}^q \alpha_i \right) - \bigvee_{i+p}^q \alpha_i\]

**Corollary 1.** If \((\bar{y}^F_i, \bar{y}^F_j)\) is the inverse P-data generated by data \(y\), then the attribute \(\alpha_i\) of \(\forall y_i \in \bar{y}^F\) and the attribute \(\alpha_j\) of \(\forall y_j \in \bar{y}^F\) are the attribute disjunctive normal form expansion and attribute disjunctive normal form contraction of \(y_i \in y\) respectively, namely

\[(\alpha_i, \alpha_j) = \left( \left( \bigvee_{i=1}^q \alpha_i \right) \bigvee_{i+q+1}^r \alpha_i, \left( \bigvee_{i=1}^q \alpha_i \right) - \bigvee_{i+p}^q \alpha_i \right)\]

Where \(\alpha_i = \left( \bigvee_{i=1}^q \alpha_i \right) \bigvee_{i+q+1}^r \alpha_i\) and \(\alpha_j = \left( \bigvee_{i=1}^q \alpha_i \right) - \bigvee_{i+p}^q \alpha_i\).

**Corollary 2.** Given inverse P-data \((\bar{y}^F_i, \bar{y}^F_j)\) generated by data \(y\), \((\bar{y}^F_i, \bar{y}^F_j)\) constitutes inverse P-data family, denoted by \(\{(\bar{y}^F_i, \bar{y}^F_j)\} \mid i \in I, j \in J\} \).
\[
\text{if } \alpha_{k+1}^F \Rightarrow \alpha_k^F, \text{ then } \bar{y}_{k+1}^F \Rightarrow \bar{y}_k^F, \quad (14)
\]

it is called the external inverse P-data reasoning generated by external inverse P-data, \( \alpha_{k+1}^F \Rightarrow \alpha_k^F \) is the antecedent, \( \bar{y}_{k+1}^F \Rightarrow \bar{y}_k^F \) is the consequent.

**Definition 7.** Suppose \((\alpha_k^F, \alpha_{k+1}^F)\) and \((\alpha_k^F, \alpha_{k+1}^F)\) are the attribute set of inverse P-data \( (\bar{y}_k^F, \bar{y}_{k+1}^F) \) and \( (\bar{y}_k^F, \bar{y}_{k+1}^F) \) respectively. Given reasoning

\[
\text{if } (\alpha_k^F, \alpha_{k+1}^F) \Rightarrow (\alpha_{k+1}^F, \alpha_{k+1}^F), \text{ then } (\bar{y}_k^F, \bar{y}_{k+1}^F) \Rightarrow (\bar{y}_{k+1}^F, \bar{y}_k^F),
\]

It is called the inverse P-data reasoning generated by inverse P-data, \( (\alpha_k^F, \alpha_{k+1}^F) \Rightarrow (\alpha_{k+1}^F, \alpha_{k+1}^F) \) is the antecedent, \( (\bar{y}_k^F, \bar{y}_{k+1}^F) \Rightarrow (\bar{y}_{k+1}^F, \bar{y}_k^F) \) is the consequent.

Where \( (\alpha_k^F, \alpha_{k+1}^F) \Rightarrow (\alpha_{k+1}^F, \alpha_{k+1}^F) \) denotes \( \alpha_k^F \Rightarrow \alpha_{k+1}^F \) and \( \alpha_{k+1}^F \Rightarrow \alpha_{k+1}^F \); \( (\bar{y}_k^F, \bar{y}_{k+1}^F) \Rightarrow (\bar{y}_{k+1}^F, \bar{y}_k^F) \) denotes \( \bar{y}_k^F \Rightarrow \bar{y}_{k+1}^F \) and \( \bar{y}_{k+1}^F \Rightarrow \bar{y}_k^F \).

**Conclusion 1.** Given attribute sets \( \alpha_k^F \) and \( \alpha_{k+1}^F \), and internal inverse P-data \( \bar{y}_k^F \), then the unknown internal inverse P-data \( \bar{y}_{k+1}^F \), whose attribute satisfies \( \alpha_k^F \Rightarrow \alpha_{k+1}^F \), is intelligent separated or discovered during the process of internal inverse P-data reasoning, moreover \( \bar{y}_k^F \subseteq \bar{y}_{k+1}^F \).

**Conclusion 2.** Given attribute sets \( \alpha_k^F \) and \( \alpha_{k+1}^F \), and external inverse P-data \( \bar{y}_k^F \), then the unknown external inverse P-data \( \bar{y}_{k+1}^F \), whose attribute satisfies \( \alpha_k^F \Rightarrow \alpha_{k+1}^F \), is intelligent separated or discovered by external inverse P-data reasoning, moreover \( \bar{y}_k^F \subseteq \bar{y}_{k+1}^F \).

**Conclusion 3.** If given attribute sets \( (\alpha_k^F, \alpha_{k+1}^F) \) and \( (\alpha_k^F, \alpha_{k+1}^F) \), and inverse P-data \( (\bar{y}_k^F, \bar{y}_k^F) \), then the unknown inverse P-data \( (\bar{y}_{k+1}^F, \bar{y}_{k+1}^F) \), whose attribute set satisfies \( (\alpha_k^F, \alpha_{k+1}^F) \Rightarrow (\alpha_{k+1}^F, \alpha_{k+1}^F) \), is intelligent separated or discovered by inverse P-data reasoning, moreover \( \bar{y}_k^F \subseteq \bar{y}_{k+1}^F \), \( \bar{y}_{k+1}^F \subseteq \bar{y}_k^F \).

These conclusions have been applied in data intelligent separation-discovery. Inverse P-data reasoning is obtained by improving inverse P-reasoning \([3,5]\).

**The Applications of External Inverse P-data Model and Data Intelligent Separation-Discovery**

**Assumption:** For the sake of simplicity, only the application of external inverse P-data models in data intelligent separation-discovery are given. The following example comes from the intelligent classification-screening recognition for students.

Longyan University is a comprehensive university possessing science and engineering etc., it is located in the west of Fujian province of China, consisting of 12 schools, and about 11780 students. Every autumn, about 3000 new students enter into. Every new students must reach the fractional admission score line \( \lambda_0 \), \( \lambda_0 = 450 \)
for example. Suppose set \( X = \{x_1,x_2,x_3,x_4,x_5,x_6,x_7,x_8\} \) is a part of new students entering in autumn of 2014.

Suppose \( y_i \) is the admission score of student \( x_i \), \( i = 1,2, \cdots , 8 \); \( y \) is the random distribution set consisted of \( y_1,y_2, \cdots , y_8 \), namely
\[
y = \{y_1,y_2,y_3,y_4,y_5,y_6,y_7,y_8\}.
\]
\( \alpha \) is the attribute set of \( X \) (or \( y \)), which consists of students’ native place, namely
\[
\alpha = \{\alpha_1,\alpha_2,\alpha_3,\alpha_4,\alpha_5,\alpha_6,\alpha_7,\alpha_8\}. \tag{16}
\]

The new student set \( X \), the random distribution set \( y \) consisting of admission score of new students, and attribute set \( \alpha \) of new student set \( X \) are listed in TABLE I.

<table>
<thead>
<tr>
<th>( X )</th>
<th>( x_1 )</th>
<th>( x_2 )</th>
<th>( x_3 )</th>
<th>( x_4 )</th>
<th>( x_5 )</th>
<th>( x_6 )</th>
<th>( x_7 )</th>
<th>( x_8 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( y )</td>
<td>384</td>
<td>405</td>
<td>396</td>
<td>390</td>
<td>412</td>
<td>506</td>
<td>460</td>
<td>448</td>
</tr>
<tr>
<td>( \alpha )</td>
<td>( \alpha_1 )</td>
<td>( \alpha_2 )</td>
<td>( \alpha_3 )</td>
<td>( \alpha_4 )</td>
<td>( \alpha_5 )</td>
<td>( \alpha_6 )</td>
<td>( \alpha_7 )</td>
<td>( \alpha_8 )</td>
</tr>
</tbody>
</table>

In TABLE I, for \( \forall x_i, x_j \in X \), \( x_i \neq x_j \) and \( y_i \neq y_j \) and \( \alpha_i \neq \alpha_j \); moreover \( x_i, x_j \) and \( \alpha_i, \alpha_j \in \alpha \) satisfy attribute disjunctive relationship as Eq. (9); or for \( \forall x_i \in X \), the attribute \( \alpha_i \) of \( x_i \) satisfies \( \alpha_i = \alpha_1 \lor \alpha_2 \lor \alpha_3 \lor \alpha_4 \lor \alpha_5 \lor \alpha_6 \lor \alpha_7 \lor \alpha_8 \).

It should be pointed out that, 1. People know the score \( y_i \) of new student \( x_i \) who entered this university in autumn of 2014, moreover \( y_i \geq \lambda_0 \), \( i = 1,2, \cdots , 8 \); 2. In Table 1, people only know \( x_1, x_2, \cdots , x_8 \) coming from one of the 8 different provinces \( \alpha_1, \alpha_2, \cdots , \alpha_8 \) of China. 3. People don’t know \( x_6 \) (or other \( x_i \neq 6 \) ) comes from which province, and they also don’t know the score \( y_6 \) of \( x_6 \) (or other \( x_i \neq 6 \) ).

By using Eq. (2) in section 2, and deleting attribute \( \alpha_3, \alpha_4, \alpha_5, \alpha_6, \alpha_7, \alpha_8 \) from \( \alpha \), \( \alpha \) becomes \( \alpha^F \), namely
\[
\alpha^F = \alpha - \{\alpha_3,\alpha_4,\alpha_5,\alpha_6,\alpha_7,\alpha_8\} = \{\alpha_1,\alpha_2\}. \tag{17}
\]
Due to the external inverse P-reasoning [5], and by using the attribute set \( \alpha \) in Eq. (16), the attribute set \( \alpha^F \) in Eq. (17) and the set \( X \) in Eq. (15), it is obtained that
\[
\text{if } \alpha^F \Rightarrow \alpha, \text{ then } \overline{X^F} \Rightarrow X. \tag{18}
\]

Where \( X^F \) is intelligent separated or discovered from \( X \) by using Eq. (18); \( x_i \) and \( x_2 \) have attribute set \( \alpha_i \) and \( \alpha_2 \) respectively, \( x_1 \) comes from Hebei province, \( x_2 \) comes from Zhejiang Province for example. For \( \forall x_i \in \overline{X^F} \), the attribute \( \alpha_i \) for \( \forall x_i \in \overline{X^F} \) satisfies \( \alpha_i = \alpha_1 \lor \alpha_2 \).
Due to Eq. (14), Eq.(16)-Eq.(17), and the data set y in Eq. (15), it can be ob-
tained that 

\[
if \alpha^F \Rightarrow \alpha, then \ \bar{y}^F \Rightarrow y.
\]

Where external inverse P-data \(\bar{y}^F = \{y_1, y_2\} = \{384, 405\}\). \(\bar{y}^F\) is intelligent separat-
ed or discovered from data y.

It is obtained easily from this example as follows:

1. If \(\alpha^F \Rightarrow \alpha\), then due to external inverse P-data reasoning, data \(\bar{y}^F\) is intelli-
gent screened out from data y by internal inverse P-data reasoning, and \(\bar{y}^F \subseteq y\), meanwhile \(\bar{x}^F\) is mined out from X, and \(\bar{x}^F \subseteq X\).

2. If there is the consequent \(\bar{y}^F \subseteq y\),due to external inverse P-data reasoning, then \(\bar{y}^F\) has attribute set \(\alpha^F = \{\alpha_1, \alpha_2\}\); or \(\bar{x}^F\) has attribute set \(\alpha^F = \{\alpha_1, \alpha_2\}\).

CONCLUSIONS

Inverse P-sets is the dual form of P-sets [7-17], they are dynamic models with
different characteristics. Inverse P-sets and P-sets offered new theories and methods
for the research of dynamic information recognition-discovery and dynamic infor-
modation mining-separation. Inverse P-sets has dynamic characteristics, but it doesn’t
have law or function characteristics. By introducing dynamic characteristics into
finite ordinary function set S and improved it, Function inverse P-sets is proposed
[18]; Function inverse P-sets has dynamic characteristics and law or function charac-
teristics; it has been applied in the research of dynamic information law. Function
inverse P-sets is the dual form of function P-sets [19, 20]. The attribute of function
in function inverse P-set satisfies attribute disjunctive normal form. The attribute set
of function in function P-set satisfies attribute conjunctive normal form. Function
inverse P-set and function P-set are two kinds of dynamic information law models
with different characteristics.

This paper gives new researches on inverse P-set as follows: by improving in-
verse P-set, inverse P-data models and inverse P-data reasoning are proposed, applica-
tions of data intelligent separation-discovery are given. Inverse P-data models are
new methods in research of dynamic data recognition or extracting.

ACKNOWLEDGEMENTS

This work is supported by the Program of Longyan Science and Technology Bu-
reau(No. 2011LY20), the Fujian Provincial Education and Scientific Research Pro-
gram for Middle and Young Teachers (No. JA15495).
REFERENCES