Research Article

The Relationship between the Unicost Set Covering Problem and the Attribute Reduction Problem in Rough Set Theory

Qingyuan Xu,$^{1,2}$ and Jinjin Li$^3$

$^1$School of Computer Science, Minnan Normal University, Zhangzhou 363000, China
$^2$Key Laboratory of Granular Computing, Minnan Normal University, Zhangzhou 363000, China
$^3$School of Mathematics and Statistics, Minnan Normal University, Zhangzhou 363000, China

Correspondence should be addressed to Qingyuan Xu; xqyyuan871@163.com and Jinjin Li; jinjinlimnu@126.com

Received 24 February 2020; Revised 22 April 2020; Accepted 22 May 2020; Published 16 June 2020

Copyright © 2020 Qingyuan Xu and Jinjin Li. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The unicost set covering problem and the attribute reduction problem are NP-complete problems. In this paper, the relationship between these two problems are discussed. Based on the transformability between attribute reductions and minimal solutions in unicost set covering models, two methods are provided. One is to induce an information table from a given unicost set covering model. With no doubt, it shows that the unicost set covering problem can be investigated by rough set theory. The other is to induce a unicost set covering model from a given information table. Similarly, it shows that the attribute reduction problem can be studied by set covering theory. As an application of the proposed theoretical results, a rough set heuristic algorithm is presented for the unicost set covering problem.

1. Introduction

The set covering problem (SCP) is a classical problem in operational research for combination optimization. It is popularly applied in aviation personnel scheduling, circuit design, transport vehicle route arrangement, service location, job assignment in manufacturing, selection of operators, etc. [1–3]. The SCP can be described as follows: selecting several columns from a $m$–row $n$–column $0 \to 1$ matrix $M = (m_{ij})_{m \times n}$, such that these columns can cover all rows with the least cost. The set of selected columns is called a solution of the SCP [4]. Here, (1) choosing each row must pay a certain cost. (2) $m_{ij}$ is the $i$ row and the $j$ column element of $M$. $m_{ij} = 1$ means that the column $j$ covers the row $i$. When all rows have identical cost, for convenience, consider them as 1, the SCP is called the unicost set covering problem (USCP) [5]. This problem is also called the minimum cardinality set covering problem (MCSCP) [6] or the location set covering problem (LSCP) [7]. In this paper, we use the name of the unicost set covering problem. The SCP is a NP-complete problem [8]. Various methods have been proposed to solve it, such as branch and bound complete method [9], genetic method [10], ant colony method [11], and others [12]. Due to the increasing scale of USCP, the time complexity of these methods increases exponentially. Hence, these methods are all effective relatively and there is no efficient method absolutely to handle it, so far.

Rough set theory [13] has been proved to be a useful tool in handling inexact, uncertain, and fuzzy knowledge in information tables [14]. It provides a powerful basis in discovering important data structures and classifying complex objects. Up to now, it has become a hot research area of intelligent computing. As an important concept in rough set theory, the purpose of attribute reduction is to find the minimal attribute set with the same knowledge description ability as the whole attribute set [13]. Scholars have successfully used attribute reduction in data mining [15, 16], pattern recognition [17, 18], and artificial intelligence and classification [19–21] in the past twenty years. The attribute reduction problem (ARP) attracts many researchers to study it [18, 20, 22–25]. Skowron and Rauszer proposed a beautiful method, which is effective to obtain all attribute reductions...
based on discernibility matrices [20]. The method constructed a discernibility function and showed that the set of all attribute reductions is the set of prime implicants of the discernibility function. However, the time complexity of this method is NP-hard [26].

The cross research of set covering and rough set theory has attracted attention of some scholars. Covering-based rough sets [27] is an important way of set covering and rough sets. Zhu and Wang proposed three types of covering-based rough sets [28]. These models have the ability to implement other features such as multi-granularity or fuzziness [29]. Based on 0-1 integer programming, Xu et al. gave an attribute reduction method to deal with the dynamic data [30]. In the variable precision rough set model, Liu et al. provided an approach to calculate the attribute reduction using set covering concepts [31]. In order to deal with the test-cost-sensitive reduction problem [32], Tan et al. presented an optimization algorithm based on set covering theory [33]. Xu et al. proposed an algorithm based on multirelation granular computing model for the USCP [34]. In this paper, we mainly build the relationship between the USCP and the ARP. Firstly, by constructing an induced information table of a given USCP, we find that computing a minimal solution or a minimum solution in a USCP can be converted into computing an attribute reduction or a minimum attribute reduction in the constructed information table. Therefore, the USCP can be converted into the APR in rough set theory. Secondly, by constructing an induced uniset covering (USC) model according to an information table, we find that calculating an attribute reduction or a minimum attribute reduction of an information table is equal to calculating a minimal solution or a minimum solution of the constructed USCP. Therefore, the APR can also be converted into the USCP. Let the two problems to be characterized by each other, and bringing new methods mutually is the main aim of the paper. It must be noted that the ARP is a NP-complete problem [26] and the SCP is also a NP-complete problem [8], and using set covering theory to find the optimal attribute reduction with the minimum number of attribute is also NP-hard.

Here, we briefly introduce the contents of the following sections. Section 2 mainly introduces basic concepts of the SCP, the USCP, and the APR in rough set theory, respectively. Section 3 constructs an information table by a USC model and studies the relationship between the USCP and the ARP in the constructed information table. Section 4 constructs a USC model by an information table and considers the relationship between the ARP and the USCP in the constructed USC model. Section 5 provides a heuristic minimal solution method based on rough set theory for the USCP and followed by an example. Section 6 gives the conclusions.

### 2. Preliminaries

In this section, some basic concepts of SCP, USCP, information table, and ARP in rough sets are reviewed.

#### 2.1. The SCP and the USCP

In addition to what is described in Section 1, the mathematical formulation of the SCP model is usually described as follows [35]. Let $E = \{e_1, e_2, \ldots, e_m\}$ be a set of elements and $S = \{s_1, s_2, \ldots, s_n\}$ be a set of subsets of $E$. The goal is to find a minimal set covering $S'$ such that $S'$ covers all elements in $E$.

The SCP can be formulated as an integer programming problem:

$$
\text{Minimize} \quad \sum_{j=1}^{n} c_j s_j, \quad (1)
$$

s.t. \quad \sum_{j=1}^{n} m_{ij} s_j \geq 1, \quad i = 1, 2, \ldots, m, \quad (2)

$$
\quad s_j \in \{0, 1\}, \quad j = 1, 2, \ldots, n. \quad (3)
$$

Here, $m_{ij}$ is the $i$-th row and the $j$-th column element of the $0-1$ matrix $M$. $m_{ij} = 1$ means that the $j$-th column covers the $i$-th row; $s_j$ flags whether the $j$-th column is included in the solution. That is, when $s_j$ takes 1, it means column $j$ is included in the solution. $c_j$ represents the cost value of column $j$. Formula (1) denotes the minimal cost required for the solution; formula (2) implies that the solution must cover all lines. When formulas (1) and (2) are satisfied, let $S' = \{s_1', s_2', \ldots, s_k'\}$, where symbol $r$ denotes transposition; then, $S'$ is a solution of the SCP. When all subsets in $S$ have identical cost, for convenience, consider them as 1, and this problem is the USCP. That is, in a USCP $0-1$ integer programming, formula (1) is replaced by the following formula:

$$
\text{Minimize} \quad \sum_{j=1}^{n} s_j, \quad (4)
$$

Here, formulas (2) and (4) imply that a solution of the USCP must cover all lines and the number of the elements in the solution must be least.

Denote USCP by $\text{USCP} = (E, S)$, where $E = \{e_1, e_2, \ldots, e_m\}$ is the object set and $S = \{s_1, s_2, \ldots, s_n\}$ is a set covering of $E$ with the $0-1$ matrix $M = (m_{ij})_{m \times n}$.

Note. The theory of soft sets introduced by Molodtsov [36] is a relatively new approach to discuss vagueness and getting popularity among the researchers. N-soft set, first introduced by Fatimah et al. [37], can provide a finer granular structure with higher distinguishable power [38]. Given a set covering model $(E, S)$ with $E = \{h_1, h_2, \ldots, h_m\}$, $S = \{e_1, e_2, \ldots, e_n\}$. Treat the elements set
Let $A = \{h_1, h_2, \ldots, h_m\}$ be the universe $U$, the set covering $S = \{e_1, e_2, \ldots, e_n\}$ as the subset $A$, and $m_{ij} = c_j(1 \leq i \leq m, 1 \leq j \leq n)$ as $F(h_i, e_j)$ [39]. The set covering model $(E, S)$ is a special N-soft set, where $\forall e_j \in S, e_j \neq \emptyset$ and $\cup S = E$. Similarly, the USC model is a special soft set.

**Example 1.** In Ad-hoc sensor networks, the signal coverage is the main measure of network quality of service. Under the premise of ensuring the quality of signal coverage, how to save network construction cost as much as possible is particularly important. Let Ad-hoc sensor networks be as shown in Figure 1. The five circular areas $s_1, s_2, \ldots, s_5$ are the signal coverage of the same type of sensors, $e_1, e_2, \ldots, e_9$ are nine customers, who want to receive the networks service. How to finding the optimum scheme for setting up sensors?

Consider nine service customers as objects. Then, the object set $E = \{e_1, e_2, \ldots, e_9\}$. Sets $s_1, s_2, \ldots, s_5$ are formed by customers in the signal covering range of each sensor, i.e., $S = \{s_1, s_2, \ldots, s_5\}$ with $s_1 = \{e_1, e_2, e_4, e_5\}$, $s_2 = \{e_1, e_2, e_6, e_7, e_8, e_9\}$, $s_3 = \{e_1, e_3, e_5, e_9\}$, $s_4 = \{e_1, e_3, e_5, e_9\}$, and $s_5 = \{e_2, e_3, e_6, e_7, e_8\}$. Then, $S$ is a set covering of $E$, where $E^m = \{e_i\}_{1 \leq i \leq m}$ is the optimum scheme for setting up sensors.

Because five sensors are the same type, Example 1 is a USCP. A minimal solution of the USCP is an optimum sensors setting scheme for Example 1.

Example 1 can be described by the following 0-1 integer programming:

$$\begin{align*}
\text{Minimize} & \quad f = s_1 + s_2 + s_3 + s_4 + s_5, \\
\text{s.t.} & \quad \begin{align*}
& s_1 + s_2 + s_3 \geq 1, \\
& s_1 + s_2 + s_5 \geq 1, \\
& s_3 + s_4 + s_5 \geq 1, \\
& s_1 + s_3 \geq 1, \\
& s_3 + s_4 \geq 1, \\
& s_2 + s_5 \geq 1, \\
& s_2 + s_3 + s_5 \geq 1, \\
& s_1 + s_2 + s_4 \geq 1, \\
& s_1 + s_2 + s_3 + s_5 = 0 \text{ or } 1.
\end{align*}
\end{align*}$$

Use Matlab2016a to run the following procedures:

$$f = [1, 1, 1, 1, 1], \quad ic = [1, 2, 3, 4, 5]$$

$$A = [-1, -1, -1, -1, -1, 0, 0, -1, 0, 0, -1, -1, -1, -1, -1, -1; -1, -1, -1, -1, -1, 0, 0, -1, 0, 0, -1, -1, -1, -1, -1, -1; 0, -1, -1, -1, -1, 0, 0, -1, 0, 0, -1, -1, -1, -1, -1, -1]$$

$$b = [-1; -1; -1; -1; -1; -1; -1; -1; -1; -1]$$

$$lb = \text{zeros}(5, 1)$$

$$ub = \text{ones}(5, 1)$$

$$[X, f, flag] = \text{intlinprog}\,(f, ic, A, b, [], [], lb, ub)$$

We can obtain an optimum solution $X = (0, 1, 1, 0, 0)^T$.

That is, $\{s_3, s_5\}$ is the optimum scheme for setting up sensors in Example 1.

In this paper, we describe Example 1 as a USCP $=(E, S)$, where $E = \{e_1, e_2, \ldots, e_9\}$ is the object set and $S = \{s_1, s_2, s_3, s_4, s_5\}$ with $\emptyset \neq s_j \subseteq E$, $\cup S = E$, $1 \leq j \leq 5$ is a set covering of $E$ with the $0-1$ matrix:

$$M = \begin{bmatrix}
1 & 1 & 0 & 1 & 0 \\
1 & 1 & 0 & 1 & 0 \\
0 & 0 & 1 & 1 & 1 \\
1 & 0 & 1 & 0 & 0 \\
0 & 0 & 1 & 1 & 0 \\
0 & 1 & 0 & 0 & 1 \\
0 & 1 & 0 & 0 & 1 \\
0 & 1 & 1 & 0 & 1 \\
1 & 1 & 0 & 1 & 0
\end{bmatrix}, \quad (5)$$

$$M = \begin{bmatrix}
1 & 1 & 0 & 1 & 0 \\
1 & 1 & 0 & 1 & 0 \\
0 & 0 & 1 & 1 & 1 \\
1 & 0 & 1 & 0 & 0 \\
0 & 0 & 1 & 1 & 0 \\
0 & 1 & 0 & 0 & 1 \\
0 & 1 & 0 & 0 & 1 \\
0 & 1 & 1 & 0 & 1 \\
1 & 1 & 0 & 1 & 0
\end{bmatrix}, \quad (7)$$

2.2. ARP in Rough Set Theory. Rough set theory is one of the most successful tools for uncertainty management. In rough sets, information is presented by an information table or an information system. An information table is usually described as a pair $IT = (U, A)$ with a mapping $f$ from $U \times A$ to $V = \cup_{a \in A} V_a$, where $U$, called the universe, is a finite and nonempty set of objects, $A$ is a finite and nonempty set of
attributes, and $V_a$, called the $a'$ domains, is the value set of attribute $a$ [13].

For any nonempty attribute subset $B \subseteq A$,

$$R_B = \left\{ (x_i, x_j) \in U \times U \mid f(x_i, a) = f(x_j, a), \forall a \in B \right\},$$

(8)

where $R_B$ is called an indiscernibility relation on $U$ w.r.t. $B$. It is easy to see that $R_B$ is an equivalence relation on the universe $U$ with the partition $(U/R_B) = \left\{ [x_i]_B \mid x_i \in U \right\}$, where $[x_i]_B = \left\{ x_j \in U \mid (x_i, x_j) \in R_B \right\}$ is the equivalence class of $x_i$ w.r.t. $B$ and $\forall x_i, x_j \in U$. If $(x_i, x_j) \in R_B$, then we call $x_i$ and $x_j$ are indiscernible in the universe w.r.t. $B$.

Attribute reduction is one of the core problems in rough sets. It is a key process for discovering important data structures and classifying complex objects. Based on the indiscernibility relation discussed above, Pawlak proposed the concept of attribute reduction in an information table. The attribute reduction is a minimal subset of attributes, which has the same indiscernibility relation as $R_A$ [40].

Definition 1. Let $IT = (U, A)$ be an information table, $B \subseteq A$. If $R_B = R_A$, then we call $B$ an attribute consistent set of $IT$. If $B$ is an attribute consistent set, for $\forall C \subseteq B$, $C$ is not a consistent set; then, we call $B$ an attribute reduction of the $IT$.

Definition 1 means that an attribute reduction is a minimal subset of attributes with $R_B = R_A$.

Normally, there is more than one attribute reduction in an information table. The one with the minimum cardinal number is called the minimum attribute reduction. RED($IT$) denotes the set of all attribute reductions of an information table $IT$.

From the viewpoint of the importance of attribute, Skowron and Rauszer [20] define the notion of core:

Definition 2 (see [20]). Let $IT = (U, A)$ be an information table with $U = \{x_1, \ldots, x_n\}$. An attribute $a \in B \subseteq A$ is dispensable in $B$ if $R_B = R_{B \setminus \{a\}}$, otherwise $a$ is indispensable. The set of all indispensable attributes is called the core set of $IT$, denoted as CORE($IT$).

CORE($IT$) is also the intersection of all attribute reductions of $IT$.

Scholars provided many methods for attribute reduction, where the method based on the discernibility matrix and discernibility function can compute all the attribute reductions of an information table using logical operations [20].

Definition 3. Let $IT = (U, A)$ be an information table, $U = \{x_1, \ldots, x_n\}$. For any $x_i, x_j \in U$,

1. The discernibility set of object pair $(x_i, x_j)$ in $IT$ is denoted as $d(x_i, x_j) = \{a \in A \mid f(x_i, a) \neq f(x_j, a)\}$.

2. The discernibility matrix of $IT$ is denoted as

$$D = (d(x_i, x_j))_{n \times n}$$

3. The family of nonempty discernibility set of $IT$ is denoted as

$$\mathcal{D} = \{d(x_i, x_j) \mid (x_i, y_j) \in U \times U, d(x_i, x_j) \neq \emptyset\}$$

Definition 3 means that if $d(x_i, x_j) \neq \emptyset$, then object $x_i$ and object $x_j$ can be distinguished by any attribute in $d(x_i, x_j)$. Otherwise, object $x_i$ and object $x_j$ cannot be distinguished in $IT$. Obviously, $d(x_i, x_j) = d(x_j, x_i)$ and $d(x_i, x_i) = \emptyset$. The discernibility matrix $D$ is a symmetric matrix. It is usually to consider only the lower triangle or the upper triangle of the matrix. $\mathcal{D}$ is the simplified form of $D$.

For $B \subseteq A$, then if $B \cap d(x_i, x_j) \neq \emptyset$, it means that the attribute set $B$ can discern the object pair $(x_i, x_j)$ [18].

Theorem 1 (see [20]). Let $IT = (U, A)$ be an information table, $U = \{x_1, \ldots, x_n\}$, and $D = (d(x_i, x_j))_{n \times n}$ be the discernibility matrix of $IT$. For any attribute set $B \subseteq A$,

1. $B$ is a consistent set iff $\forall (x_i, x_j) \in U \times U$, when $d(x_i, x_j) \neq \emptyset$, $B \cap d(x_i, x_j) \neq \emptyset$.

2. $B$ is an attribute reduction iff $B$ is an attribute consistent set and $\forall a \in B$, $\exists (x_i, x_j) \in U \times U$, and $d(x_i, x_j) \neq \emptyset \land ((B \setminus \{a\}) \cap d(x_i, x_j) = \emptyset$.

Theorem 2 (see [20]). Let $IT = (U, A)$ be an information table, $U = \{x_1, \ldots, x_n\}$, and $D = (d(x_i, x_j))_{n \times n}$ be the discernibility matrix of $IT$. CORE($IT$) = $\{a \in A \mid d(x_i, x_j) = [a]\}$ for some $x_i, x_j \in U$.

Skowron and Rauszer [20] proposed a method to compute all the attribute reductions of an information table, which is based on notions of discernibility function and logical operation. A discernibility function $f_k$ for $IT$ is a Boolean function, which has $m$ Boolean variables $a_1, a_2, \ldots, a_m$ corresponding to $m$ attributes $a_1, a_2, \ldots, a_m$, respectively.

Definition 4. Let $IT = (U, A)$ be an information table, $U = \{x_1, \ldots, x_n\}$:

$$f_k(a_1, a_2, \ldots, a_m) = \land \left\{ \lor d(x_i, x_j) \mid d(x_i, x_j) \in \mathcal{D} \right\},$$

(9)

where $\lor d(x_i, x_j)$ is the disjunction of all attributes $a \in d(x_i, x_j)$ and $f_k(a_1, a_2, \ldots, a_m)$ is called the discernibility function of $IT$.

Theorem 3 (see [20]). Let $IT = (U, A)$ be an information table and $f_k(a_1, a_2, \ldots, a_m)$ be the discernibility function. An attribute subset $B \subseteq A$ is an attribute reduction of $IT$ iff $\land_{a \in B} a_r$ is a prime implicant of $f_k$.

Theorem 3 means that a prime implicant of discernibility function $f_k(a_1, a_2, \ldots, a_m)$ is an attribute reduction of $IT$, and all the prime implicants of discernibility function $f_k(a_1, a_2, \ldots, a_m)$ are all the attribute reductions of $IT$. So, we can use the disjunction $\lor$ and conjunction $\land$ operations to compute all the attribute reductions of an information table.

If $f_k(a_1, a_2, \ldots, a_m) = \land_{t=1}^{m} (\lor_{s=1}^{t} a_{s,t})$, where $\lor_{s=1}^{t} a_{s,t}$ are all the prime implicants of the discernibility function $f_k$, then $B_r = [a_{s,t} \mid s \leq s, r \leq t]$, and $r \leq t$ is an attribute reduction of $IT$. Denote all the attribute reductions of $IT$ by $\mathcal{B} = \{B_r \mid r \leq t\}$.
From the above conclusion, one can obtain all the attribute reductions of an information table by computing the discernibility function using Boolean operations.

**Example 2.** Let IT = (U, A) be an information table, as shown in Table 1, with $U = \{x_1, x_2, x_3, x_4, x_5, x_6\}$ and $A = \{a_1, a_2, a_3, a_4, a_5\}$.

Table 2 represents the discernibility matrix $D$ of IT, where we only consider the upper triangle matrix. The discernibility function $f_S$ of IT is

$$f_S(a_1, a_2, a_3, a_4, a_5) = \bigwedge \{ \forall d(x_i, x_j) \mid d(x_i, x_j) \in D \}$$

$$= (a_1 \lor a_2 \lor a_3) \land (a_1 \lor a_2 \lor a_3)$$

$$\land (a_1 \lor a_2) \land (a_2 \lor a_3)$$

$$\land (a_1 \lor a_2 \lor a_4 \lor a_5) \land a_2$$

$$\land (a_1 \lor a_2 \lor a_4 \lor a_5)$$

$$\land (a_1 \lor a_2 \lor a_4 \lor a_5)$$

$$\land (a_1 \lor a_2 \lor a_4 \lor a_5)$$

$$\land (a_1 \lor a_2 \lor a_4 \lor a_5)$$

(10)

Using Boolean logical operations, we can obtain

$$f_S(a_1, a_2, a_3, a_4, a_5) = (a_1 \land a_2 \land a_4 \land a_5) \land (a_1 \land a_3 \land a_4 \land a_5).$$

(11)

According to Theorem 3, IT has two attribute reductions: $(a_1, a_2, a_4, a_5)$ and $(a_2, a_3, a_4, a_5)$. CORE (IT) = $\{a_1, a_2, a_3, a_4, a_5\}$.

**3. Inducing an Information Table from a USC Model**

In this section, we induce an information table from a USC model and give an example to explain the inducing process. Then, we discuss the relationship between the USCP and the ARP in the induced information table.

**Definition 5.** Given a USCP = $(E, S)$ with $E = \{e_1, e_2, \ldots, e_m\}$, $S = \{s_1, s_2, \ldots, s_n \mid \emptyset \neq s \subseteq E, U_{S_j} = E\}$, and $M = (m_{ij})_{i,j \leq m}$, we call the pair IT = (U, A) an induced information table of the USCP, where $U = E \cup \{e_{m+1}\} = \{e_1, e_2, \ldots, e_m, e_{m+1}\}$, $A = \{s_1, s_2, \ldots, s_n\}$, and the mapping $f$ is constructed as follows:

(a) $f(e_i, a_j) = i \times m_{ij}, 1 \leq i \leq m, 1 \leq j \leq n$

(b) $f(e_{m+1}, a_j) = 0, 1 \leq j \leq n$

**Note.** The induced information table IT = (U, A), constructed by a USC model as above, is a special N-soft set, where $\forall e_i \in U, F(e_i, s_j) = 0 \lor i(1 \leq i \leq m, 1 \leq j \leq n)$, and $F(e_{m+1}, s_j) = 0(1 \leq j \leq n)$.

**Example 3.** The induced IT of the USC in Example 1 is shown in Table 3.

**Theorem 4.** Given a USCP = $(E, S)$ with $E = \{e_1, e_2, \ldots, e_m\}$, $S = \{s_1, s_2, \ldots, s_n \mid \emptyset \neq s \subseteq E, U_{S_j} = E\}$, $M = (m_{ij})_{i,j \leq m}$, IT = (U, A) is the induced information table. For $\forall e_i, e_k \in U$, $i \neq k$, $d(e_i, e_k) = \{s_j \mid m_{ij} = 1 \lor m_{kj} = 1, 1 \leq j \leq n\}$ and $d(e_i, e_k) \neq \emptyset$.

**Proof** of Theorem 4. (1) We prove that, for $\forall e_i, e_k \in U, i \neq k$, \{s_j | m_{ij} = 1 or m_{kj} = 1, 1 \leq j \leq n\} \subseteq d(e_i, e_k). (a) $\forall 1 \leq j \leq n$, if $m_{ij} = 1$ and $m_{kj} = 0$, according to Definition 5, then $1 \leq i \leq m$ and $f(e_i, s_j) = i \times m_{ij} = 1, i \leq k \leq m + 1$, $f(e_k, s_j) = k \times m_{kj} = 0(\text{when } k = m + 1)$, so $f(e_i, s_j) \neq f(e_k, s_j), s_j \in d(e_i, e_k)$; (b) similarly, $\forall 1 \leq j \leq n$, if $m_{ij} = 0$ and $m_{kj} = 1$, then $1 \leq i \leq m$ and $f(e_i, s_j) = i \times m_{ij} = 0(\text{when } i = m + 1)$, $1 \leq k \leq m, f(e_k, s_j) = k \times m_{kj} = k$, so $f(e_i, s_j) \neq f(e_k, s_j), s_j \in d(e_i, e_k)$. (c) Assume that, for $\forall e_i, e_k \in U$ and $i \neq k$, $\forall 1 \leq j \leq n$, if $m_{ij} = 1$ or $m_{kj} = 1$, then $s_j \in d(e_i, e_k)$, \{s_j | m_{ij} = 1 or m_{kj} = 1, 1 \leq j \leq n\} \subseteq d(e_i, e_k).

(2) We prove that, for $\forall e_i, e_k \in U, i \neq k$, $d(e_i, e_k) \subseteq \{s_j | m_{ij} = 1 or m_{kj} = 1, 1 \leq j \leq n\}$. Assume $e_i, e_k \in U$, $i \neq k$, $\forall s \in d(e_i, e_k) \land \forall 1 \leq j \leq n$, (a) assume $1 \leq i \leq m$, $1 \leq k \leq m$, $f(e_i, s) = i \times m_{ij}$ and $f(e_k, s) = k \times m_{kj}$ because $f(e_i, s) \neq f(e_k, s)$, $s_j \in d(e_i, e_k)$. (b) Assume $1 \leq i \leq m$, $k = m + 1$, according to Definition 5, $f(e_i, s) = i \times m_{ij}$ and $f(e_k, s) = 0$ because $f(e_i, s) \neq f(e_k, s), s_j \in d(e_i, e_k)$. (c) Assume $1 \leq i \leq m$, $1 \leq k \leq m$, according to Definition 5, $f(e_i, s) = 0$, and $f(e_k, s) = k \times m_{kj}$ because $f(e_i, s) \neq f(e_k, s), s_j \in d(e_i, e_k)$. Combining (a) and (b) with (c), we obtain that, for $\forall e_i, e_k \in U, i \neq k, \forall s \in d(e_i, e_k), s \subseteq \{s_j | m_{ij} = 1 or m_{kj} = 1, 1 \leq j \leq n\}$. By (1) and (2), we get that, for $\forall e_i, e_k \in U, i \neq k$, $d(e_i, e_k) \subseteq \{s_j | m_{ij} = 1 or m_{kj} = 1, 1 \leq j \leq n\}$.

(3) Here, we prove that $d(e_i, e_k) \neq \emptyset$. (a) Assume $1 \leq i \leq m, k = m + 1$, then $s_j \in S$ and $m_{ij} = 1$, so $s_j \in d(e_i, e_k)$ and $d(e_i, e_k) \neq \emptyset$. (b) Similarly, assume $1 \leq k \leq m, i = m + 1$, then $s_j \in S$ and $m_{kj} = 1$, so $s_j \in d(e_i, e_k)$ and $d(e_i, e_k) \neq \emptyset$. (c) Assume $1 \leq i \leq m, 1 \leq k \leq m$, then $s_j \in S$ and $m_{ij} = 1$, $1 \leq j \leq n$, so $s_j \in d(e_i, e_k)$, $s_j \in d(e_i, e_k)$, and $d(e_i, e_k) \neq \emptyset$.

**Theorem 5.** Given a USCP = $(E, S)$ with $E = \{e_1, e_2, \ldots, e_m\}$, $S = \{s_1, s_2, \ldots, s_n \mid \emptyset \neq s \subseteq E, U_{S_j} = E\}$, $M = (m_{ij})_{i,j \leq m}$, the induced information table IT = (U, A). For $S \subseteq S$, $S$ is a
set covering of the USCP iff $S'$ is an attribute consist set of the induced IT.

**Proof of Theorem 5.** (1) $\implies$ Suppose $S' \subseteq S$ is a set covering of the USCP. Then, for $(e_i, e_k) \in U \times U, i \neq k$, (a) when $1 \leq i \leq m, 1 \leq k \leq m, \exists s_j \in S', s_j \in S'$ and $m_{ij} = 1, m_{ki} = 1$. Because $d(e_i, e_k) = \{s_j | m_{ij} = 1 \text{ or } m_{kj} = 1, 1 \leq j \leq n\}$, so $s_j \in d(e_i, e_k)$ and $s_j \in d(e_k, e_i)$, hence $S' \cap d(e_i, e_k) \neq \emptyset$; (b) when $m = 1 + 1, 1 \leq k \leq m, \exists s_j \in S', m_{kj} = 1$, so $s_j \in d(e_i, e_k)$, thus $S' \cap d(e_i, e_k) \neq \emptyset$. Combining (a) and (b) with (c), we have that $S' \cap d(e_i, e_k) \neq \emptyset$. By Theorem 1, one sees that $S'$ is an attribute consist set of the induced IT.

(2) $\implies$ Suppose $S' \subseteq S$ is an attribute consist set of the induced IT. Then, for $(e_i, e_k) \in U \times U, i \neq k$, $d(e_i, e_k) \cap (S' - \{s_j\}) = \emptyset$. Thus, for $\forall s_i \in (S' - \{s_j\})$, $m_{ij} = 0$ and $m_{ki} = 0$, i.e., $(S' - \{s_j\})$ does not cover $e_i$ and $e_k$. $(S' - \{s_j\})$ is no longer a set covering of the USCP. Hence, $S'$ is a minimal solution of the USCP.

Combining (1) with (2), it completes the proof.

**Theorem 7.** Given a USCP $= (E, S)$ with $E = \{e_1, e_2, \ldots, e_m\}$, $S = \{s_1, s_2, \ldots, s_n\} \setminus \emptyset \neq s_i \subseteq E, \forall s_i = E\}$, and $M = (m_{ij})_{n \times n}$ the induced information table $IT = (U, A)$. For $S' \subseteq S$, $S'$ is a minimal solution of the USCP iff $S'$ is a minimum attribute reduction of the induced IT.

**Proof of Theorem 7.** It can be directly concluded by Theorem 6.

**Theorem 8.** Given a USCP $= (E, S)$ with $E = \{e_1, e_2, \ldots, e_m\}$, $S = \{s_1, s_2, \ldots, s_n\} \setminus \emptyset \neq s_i \subseteq E, \forall s_i = E\}$, and $M = (m_{ij})_{n \times n}$ the induced information table $IT = (U, A)$. For $S' \subseteq S$, $S'$ is a minimal solution of the USCP iff $S'$ is a minimum attribute reduction of the induced IT.

**Proof of Theorem 8.** (1) $\implies$ Suppose $s_j \in S$ is a core of the induced IT. By Theorem 2, $\exists e_i, e_k \in U, i \neq k, d(e_i, e_k) = \{s_j\}$. Thus, by Theorem 4, (a) when $1 \leq i \leq m, 1 \leq k \leq m, m_{ij} = 1, m_{ki} = 1$, $m_{kj} = 0, m_{ji} = 0$. (b) When $1 \leq k \leq m, \forall s_i = E, m_{ij} = 1, 1 \leq j \leq n$, $m_{ij} = 0$. (c) When $1 \leq i \leq m, \forall s_i = E, m_{ij} = 1$. Hence, combining (a) and (b) with (c), we have $s_j$ is a core of the induced IT. $\exists e_i, e_k \in U, m_{ij} = 1$ and $\forall s_i \in S, s_i \neq s_j, m_{ij} = 0$.

(2) $\implies$ Suppose $s_j \in S, \forall s_i \in E, m_{ij} = 1$ and $\forall s_i \in S, s_i \neq s_j, m_{ij} = 0$. Because $\forall s_i \in S, m_{ij+1} = 0$, by Theorem 4, $d(e_i, e_{m+1}) = s_j$. According to Theorem 1, $s_j \in \text{CORE}(IT), i.e., s_j$ is a core of the induced IT.

Combining (1) with (2), it completes the proof.

From Theorem 8, if $1 \leq i \leq m, \exists 1 \leq j \leq n, m_{ij} = 1, \forall 1 \leq i \leq m, t \neq j, m_{it} = 0, then s_j is a core of the induced S.$

**Definition 5** and **Theorems 6 and 7** make it possible to use attribute reduction methods to compute a minimal solution or the minimum solution of a USCP. Thus, the USCP can be solved by using rough set theory.

In the following example, we illustrate how to compute a minimal solution or the minimum solution of a USCP by calculating an attribute reduction or the minimum attribute reduction of the constructed information table.
Example 4. The discernibility matrix $D$ of IT constructed by the USCP in Example 1 can be represented in Table 4, where we only consider the upper triangle matrix.

By Table 4, the discernibility function $f_5$ of the induced IT is as follows:

$$f_5(s_1, s_2, s_3, s_4) = \bigwedge \{ d(e_1, e_2) \mid d(e_1, e_2) \in \mathcal{D} \}$$

\begin{align*}
= & (s_1 \vee s_2 \vee s_3 \vee s_4) \wedge (s_1 \vee s_2 \vee s_3 \vee s_5) \wedge (s_1 \vee s_2 \vee s_1 \vee s_4) \\
\wedge & (s_1 \vee s_2 \vee s_3) \wedge (s_1 \vee s_2 \vee s_3 \vee s_5) \wedge (s_1 \vee s_2 \vee s_5) \\
\wedge & (s_1 \vee s_2 \vee s_3 \vee s_5) \\
\wedge & (s_1 \vee s_2 \vee s_5) \\
\wedge & (s_1 \vee s_2) \wedge (s_1 \vee s_3) \vee (s_1 \vee s_2) \\
\wedge & (s_2 \vee s_3) \vee (s_2 \vee s_5) \vee (s_2 \vee s_3) \vee (s_2 \vee s_4) \\
\end{align*}

Using Boolean logical operations, we can obtain

$$f_5(s_1, s_2, s_3, s_4) = (s_1 \wedge s_2 \wedge s_3) \vee (s_1 \wedge s_2 \wedge s_4) \vee (s_1 \wedge s_2 \wedge s_5) \vee (s_2 \wedge s_4) \vee (s_2 \wedge s_5).$$

(12)

According to Theorem 3, the induced IT has five attribute reductions: \{s_1, s_2, a_3\}, \{s_1, s_4, s_5\}, \{s_1, s_2, s_4\}, \{s_3, s_4, s_5\}, and \{s_2, s_3\}, where \{s_2, s_3\} is the unique minimum attribute reduction. By Theorem 2, there is no core in the induced IT.

By Theorem 6, we get that \{s_1, s_3, a_3\}, \{s_1, s_4, s_5\}, \{s_1, s_2, a_3\}, \{s_1, s_2, s_4\}, and \{s_2, s_3\} are all minimal solutions of the USCP in Example 1, and \{s_2, s_3\} is the unique minimum solution.

4. Inducing a USC Model from an Information Table

From the analysis above, we know that a USC model can induce an information table, and the USCP can be transformed into the ARP in the induced information table. One can derive a minimal solution or the minimum solution of a USCP by calculating an attribute reduction or the minimum attribute reduction of the constructed information table. Conversely, in this section, we are interested in whether a given information table can induce a USC model, and the ARP is corresponding to the USCP.

In what follows, with the help of the simplified discernibility set proposed by Yao and Zhao [18], we construct a USC model from a given information table, and the ARP is corresponding to the USCP. Thus, we give an example to explain the constructing process. Then, we discuss the connection between the ARP and the constructed USCP.

Definition 6 (see [18]). For a matrix element $d(x_i', x_j') \neq \emptyset$, if $\emptyset \neq d(x_i', x_i') \subset d(x_i, x_j)$, let $d(x_i', x_j')$ absorb and replace another element $d(x_i, x_j)$, and we call the process element absorption.

Definition 7 (see [18]). For a discernibility matrix $D = (d(x_i, x_j))_{n \times n}$, if we use element absorption to process all the elements in $D$, and then it can be converted into another simplified discernibility matrix $D' = (d'(x_i, x_j))_{n \times n}$, we call the process matrix absorption and call $D'$ the simplified discernibility matrix of $D$.

In Definition 7, we can see that no elements in $D'$ can be absorbed by each other. Yao and Zhao point out that the discernibility matrix $D'$ has the same set of attribute reductions as the discernibility matrix $D$ [18], so do $\mathcal{D}'$ and $\mathcal{D}$.

Example 5. The simplified discernibility matrix $D'$ of $D$ in Example 2 can be represented in Table 5 as follows, where we only consider the upper triangle matrix.

Using the simplified discernibility matrix $D'$, we can obtain

$$f_5(a_1, a_2, a_3, a_4, a_5) = \bigwedge \{ d(x_i, x_j) \mid d(x_i, x_j) \in \mathcal{D}' \}$$

$$= a_1 \wedge a_2 \wedge a_3 \wedge (a_1 \vee a_2).$$

(14)

Using Boolean logical operations, we have

$$f_5(a_1, a_2, a_3, a_4, a_5) = (a_1 \wedge a_2 \wedge a_3 \vee a_2 \wedge a_3 \wedge a_4 \wedge a_5).$$

(15)

The result is the same as Example 2.

Inspired by the result proposed by [18], in the following, we construct an induced USC model from a given information table by using the simplified discernibility matrix.

Definition 8. Given an information table $IT = (U, A)$, $U = \{x_1, x_2, \ldots, x_m\}$, $A = \{a_1, a_2, \ldots, a_n\}$, and $\mathcal{D}' = \{d_1, d_2, \ldots, d_t\}$:

(1) Assign $\mathcal{D}'$ as an object set $E$, that is, $E = \{d_1, d_2, \ldots, d_t\}$

(2) Assign a covering set $S = \{s_{a_1}, s_{a_2}, \ldots, s_{a_n}\}$ with $s_{a_j} = \{d_i \mid a_j \in d_i, 1 \leq i \leq t \}$, $1 \leq j \leq n$

The object set $E$ and the covering set $S$ constitute a USC model with the $t$–row $n$–column $0 \times 1$ matrix $M = (m_{ij})_{n \times t}$ as follows:

\( a \) $m_{ij} = 1$ if $a_j \in d_i$, $1 \leq j \leq n$, $1 \leq i \leq t$

\( b \) $m_{ij} = 0$ if $a_j \notin d_i$, $1 \leq j \leq n$, $1 \leq i \leq t$

The USC model $E = (E, S)$ is called the induced USC model of the given IT.

In Definition 8, for $\forall d_i \in E$, $d_i \neq \emptyset$, there is at least $\exists a_j \in A, a_j \in d_i$, such that $d_i \in S_{a_j}$ and $d_i \in \cup_{j=1}^{n} S$. Thus, we have $\cup_{j=1}^{n} S = E$, i.e., $S$ is exactly a set covering of $E$. Hence, the pair $(E, S)$ is a USC.

Example 6. Construct an induced USC model from the IT in Example 2.

According to Table 5, $\mathcal{D}' = \{a_1, a_2, a_3, a_4\}$, and assign the object set $E = \{e_1, e_2, e_3, e_4\}$. $e_1 = \{a_1\}$, $e_2 = \{a_1\}$, $e_3 = \{a_3\}$, and $e_4 = \{a_1, a_3\}$. Assign a covering set
Table 4: The discernibility matrix $D$ of the IS induced by the USCP in Example 1.

|      | $e_1$       | $e_2$       | $e_3$       | $e_4$       | $e_5$       | $e_6$       | $e_7$       | $e_8$       | $e_9$       | $e_{10}$    |
|------|-------------|-------------|-------------|-------------|-------------|-------------|-------------|-------------|-------------|-------------|
| $e_1$| $\emptyset$ | $\{s_1, s_4, s_5\}$ | $\{s_1, s_2, s_3, s_5\}$ | $\{s_1, s_2, s_3, s_4\}$ | $\{s_1, s_2, s_3, s_5\}$ | $\{s_1, s_2, s_4, s_5\}$ | $\{s_1, s_2, s_3, s_5\}$ | $\{s_1, s_2, s_3, s_5\}$ | $\{s_1, s_2, s_3, s_5\}$ | $\{s_1, s_2, s_3, s_5\}$ |
| $e_2$| $\emptyset$ | $\emptyset$ | $\{s_1, s_2, s_3, s_5\}$ | $\{s_1, s_2, s_3, s_4\}$ | $\{s_1, s_2, s_3, s_5\}$ | $\{s_1, s_2, s_3, s_5\}$ | $\{s_1, s_2, s_3, s_5\}$ | $\{s_1, s_2, s_3, s_5\}$ | $\{s_1, s_2, s_3, s_5\}$ | $\{s_1, s_2, s_3, s_5\}$ |
| $e_3$| $\emptyset$ | $\emptyset$ | $\emptyset$ | $\{s_1, s_2, s_3, s_4\}$ | $\{s_1, s_2, s_3, s_4\}$ | $\{s_1, s_2, s_3, s_4\}$ | $\{s_1, s_2, s_3, s_4\}$ | $\{s_1, s_2, s_3, s_4\}$ | $\{s_1, s_2, s_3, s_4\}$ | $\{s_1, s_2, s_3, s_4\}$ |
| $e_4$| $\emptyset$ | $\emptyset$ | $\emptyset$ | $\emptyset$ | $\{s_1, s_2, s_3, s_4\}$ | $\{s_1, s_2, s_3, s_4\}$ | $\{s_1, s_2, s_3, s_4\}$ | $\{s_1, s_2, s_3, s_4\}$ | $\{s_1, s_2, s_3, s_4\}$ | $\{s_1, s_2, s_3, s_4\}$ |
| $e_5$| $\emptyset$ | $\emptyset$ | $\emptyset$ | $\emptyset$ | $\emptyset$ | $\{s_1, s_2, s_3, s_4\}$ | $\{s_1, s_2, s_3, s_4\}$ | $\{s_1, s_2, s_3, s_4\}$ | $\{s_1, s_2, s_3, s_4\}$ | $\{s_1, s_2, s_3, s_4\}$ |
| $e_6$| $\emptyset$ | $\emptyset$ | $\emptyset$ | $\emptyset$ | $\emptyset$ | $\emptyset$ | $\{s_1, s_2, s_3, s_4\}$ | $\{s_1, s_2, s_3, s_4\}$ | $\{s_1, s_2, s_3, s_4\}$ | $\{s_1, s_2, s_3, s_4\}$ |
| $e_7$| $\emptyset$ | $\emptyset$ | $\emptyset$ | $\emptyset$ | $\emptyset$ | $\emptyset$ | $\emptyset$ | $\{s_1, s_2, s_3, s_4\}$ | $\{s_1, s_2, s_3, s_4\}$ | $\{s_1, s_2, s_3, s_4\}$ |
| $e_8$| $\emptyset$ | $\emptyset$ | $\emptyset$ | $\emptyset$ | $\emptyset$ | $\emptyset$ | $\emptyset$ | $\emptyset$ | $\{s_1, s_2, s_3, s_4\}$ | $\{s_1, s_2, s_3, s_4\}$ |
| $e_9$| $\emptyset$ | $\emptyset$ | $\emptyset$ | $\emptyset$ | $\emptyset$ | $\emptyset$ | $\emptyset$ | $\emptyset$ | $\emptyset$ | $\{s_1, s_2, s_3, s_4\}$ |
| $e_{10}$| $\emptyset$ | $\emptyset$ | $\emptyset$ | $\emptyset$ | $\emptyset$ | $\emptyset$ | $\emptyset$ | $\emptyset$ | $\emptyset$ | $\emptyset$ |
Table 5: The simplified discernibility matrix $D_I$ for the IS in Example 2.

|   | $x_1$ | $x_2$ | $x_3$ | $x_4$ | $x_5$ |
|---|---|---|---|---|---|
| $x_1$ | $\emptyset$ | $\{a_3\}$ | $\{a_2\}$ | $\{a_2\}$ | $\{a_1, a_3\}$ |
| $x_2$ | $\emptyset$ | $\{a_3\}$ | $\{a_2\}$ | $\{a_2\}$ | $\{a_1\}$ |
| $x_3$ | $\emptyset$ | $\{a_3\}$ | $\{a_2\}$ | $\{a_2\}$ | $\{a_1\}$ |
| $x_4$ | $\emptyset$ | $\{a_3\}$ | $\{a_2\}$ | $\{a_2\}$ | $\{a_1\}$ |
| $x_5$ | $\emptyset$ | $\{a_3\}$ | $\{a_2\}$ | $\{a_2\}$ | $\{a_1\}$ |
| $x_6$ | $\emptyset$ | $\{a_3\}$ | $\{a_2\}$ | $\{a_2\}$ | $\{a_1\}$ |

Note: The simplified discernibility matrix $D_I$ is not the unique one simplified discernibility matrix of $D$. The ordering in which different elements are absorbed can obtain different simplified discernibility matrix. Nevertheless, all the absorbed matrices have the same set of attribute reductions as the discernibility matrix $D$ [18].

$S = \{\{e_4\}, \{e_1\}, \{e_2\}, \{e_3\}\}$. The object set $E$ and the covering set $S$ constitute the induced USC model of the IT in Example 2, where

$$M = (m_{ij})_{4 \times 5} = \begin{bmatrix}
0 & 1 & 0 & 0 & 0 \\
0 & 0 & 1 & 0 & 0 \\
0 & 0 & 0 & 1 & 1 \\
1 & 0 & 1 & 0 & 0
\end{bmatrix}.$$  \hspace{1cm} (16)

Theorem 9. Given an information table $IT = (U, A)$, $U = \{x_1, x_2, \ldots, x_m\}$, $A = \{a_1, a_2, \ldots, a_n\}$, and $\mathcal{D}' = \{d_1, d_2, \ldots, d_{\ell}\}$, the induced USC model of $U$ with $E = \{d_1, d_2, \ldots, d_{\ell}\}$, $S = \{s_1, s_2, \ldots, s_{\ell}\}$, is defined as $M = (m_{ij})_{\ell \times \ell}$ with $m_{ij}$ being the object set $S$ in the induced USC.

Proof of Theorem 9. (1)$\Rightarrow$ Suppose $B$ is a consistent set of the given IT. Then, for all $d_i \in \mathcal{D}'$, $B \cap \mathcal{D}' \neq \emptyset$, that is, there is at least $\exists a_j \in B$ and $d_i \in \mathcal{D}'$, according to Definition 8, so $d_i \in s_{a_j}$, $d_i$ is covered by $s_{a_j}$. Hence, $S' = \{s_{a_j} \mid a_j \in B\}$ covers $d_i$. Thus, $S' = \{s_{a_j} \mid a_j \in B\}$ is a set covering of the induced USC.

(2)$\Leftarrow$ Suppose $S' = \{s_{a_j} \mid a_j \in B\}$ is a set covering of the induced USC. Then, for all $d_i \in \mathcal{D}'$, $\exists \mathcal{S}' \subseteq S'$, $d_i \in \mathcal{S}'$, according to Definition 8, $a_j \in \mathcal{D}'$, so $a_j \cap d_i \neq \emptyset$. Hence, $B \cap d_i \neq \emptyset$, $d_i$ is an attribute consistent set of the given IT.

Combining (1) with (2), it completes the proof.

Similar to Section 3, we can easily obtain some results.

Theorem 10. Given an information table $IT = (U, A)$, $U = \{x_1, x_2, \ldots, x_m\}$, $A = \{a_1, a_2, \ldots, a_n\}$, and $\mathcal{D}' = \{d_1, d_2, \ldots, d_{\ell}\}$, the induced USC model of $U$ with $E = \{d_1, d_2, \ldots, d_{\ell}\}$, $S = \{s_1, s_2, \ldots, s_{\ell}\}$, is defined as $M = (m_{ij})_{\ell \times \ell}$ with $m_{ij}$ being the object set $S$ in the induced USC.

Proof of Theorem 10. By Theorems 1 and 9, it can easily be proved.

Theorem 11. Given an information table $IT = (U, A)$, $U = \{x_1, x_2, \ldots, x_m\}$, $A = \{a_1, a_2, \ldots, a_n\}$, and $\mathcal{D}' = \{d_1, d_2, \ldots, d_{\ell}\}$, the induced USC model of $U$ with $E = \{d_1, d_2, \ldots, d_{\ell}\}$, $S = \{s_1, s_2, \ldots, s_{\ell}\}$, and the $0 - 1$ matrix $M = (m_{ij})_{\ell \times \ell}$ with $m_{ij}$ being the object set $S$ in the induced USC.

Proof of Theorem 11. By Theorem 10, it is easy to prove.

Theorem 12. Given an information table $IT = (U, A)$, $U = \{x_1, x_2, \ldots, x_m\}$, $A = \{a_1, a_2, \ldots, a_n\}$, and $\mathcal{D}' = \{d_1, d_2, \ldots, d_{\ell}\}$, the induced USC model of $U$ with $E = \{d_1, d_2, \ldots, d_{\ell}\}$, $S = \{s_1, s_2, \ldots, s_{\ell}\}$, and the $0 - 1$ matrix $M = (m_{ij})_{\ell \times \ell}$ with $m_{ij}$ being the object set $S$ in the induced USC.

Proof of Theorem 12. (1)$\Rightarrow$ Suppose $a_j \in A$, $a_j$ is a core of the given IT. Then, $\exists d_i \in \mathcal{D}'$, $d_i = \{a_j\}$, According to Definition 8, $d_i \in E$, $\exists s_{a_j}$, $s_{a_j} \in B$. By Theorem 2, we have that $a_j$ is a core of the given IT.

Combining (1) and (2), it completes the proof.

From Theorems 10 and 11, we can see that an attribute reduction or a minimum attribute reduction of a given information table is a solution or the minimum solution of the constructed USC. Thus, the ARP can be converted into the USC.

We can obtain all of the minimal solution of the constructed USC in Example 6 by using Matlab: $X = \{1, 1, 0, 1, 1\}^t$ and $0, 1, 1, 1, 1\}^t$. That is, $\{a_1, a_2, a_3, a_5\}$ and $\{a_1, a_3, a_4, a_5\}$ are all attribute reductions of the information table IT in Example 2. And by Theorem 12, $\{a_1, a_3, a_4, a_5\}$ is the core set.

5. A Rough Set Heuristic Algorithm for USC

Based on the discussion above, the relationship between the USC and the ARP is established. One can investigate the USC by rough set theory and discuss the ARP of an information table via set covering theory.

In this section, based on rough set theory, we propose a heuristic algorithm based on rough set theory for USC as an application of the proposed theoretical results.

There is no doubt that the method based on discernibility function can compute all attribute reductions or an minimum attribute reduction, but the time complexity of it is NP-hard [26]. However, finding an attribute reduction based on heuristic algorithms can dramatically reduce the time complexity. By eliminating irrelevant states or unlikely possibilities, heuristic algorithms can lessen the computational efficiency. The classical heuristic algorithms are discernibility matrix methods [41–43], positive-region methods...
where the importance of knowledge. Usually, they make the entropy methods [23, 47, 48].

Example 7. The notion of entropy is introduced by Shannon, and Szekely applied Shannon’s entropy to compute an attribute reduction.

Definition 9. Given an information table as $\mathcal{IT} = (U, A)$. The entropy of $B$, where $A \subseteq U$ and $B \subseteq A$ and

$$H(B) = -\sum_{i=1}^{r} \frac{|X_i|}{|U|} \log_2 \left( \frac{|X_i|}{|U|} \right),$$

where $H(B)$ is the entropy of $B$.

Here, $X_i \in (U/R_p) = \{X_1, X_2, \ldots, X_r\}$ and $\cdot \cdot \cdot$ means the number of elements in a set. If $H(B) = H(A)$, then $B$ is an attribute reduction of $IT$ [23, 47]. The significance of attribute set $B$ can be measured by $H(B)$.

Definition 10. Given an information table $\mathcal{IT} = (U, A)$, $B \subseteq A$, $\forall a \notin B$, and

$$\text{Sig}(a, B) = H(B \cup \{a\}) - H(B),$$

where $\text{Sig}(a, B)$ is called the significance of attribute $a$ w.r.t. $B$.

Now, we construct a rough set heuristic algorithm for USCP via the entropy.

By Algorithm 1, we can compute a minimal solution of a given USCP by selecting the most important attribute one by one according to its entropy.

In Algorithm 1, step 2 selects all cores according to Theorem 8 and its time complexity is $O(mm)$. The time complexity of step 3 is $O(mmn)$. Step 4 is the most important step. It selects the most important attribute in $\mathcal{IT} - \text{Red}$ w.r.t. the Red. Its time complexity is $O((|U|^2)|A|)$ [46], that is, $O(m^2n)$. So, the time complexity of Algorithm 1 is $O(m^2n)$.

Example 7. Compute a minimal solution of the given USCP in Example 1 by using Algorithm 1.

**Input:** the USCP $= (E, S)$ in Example 1.

1. Let $\text{Red} = \emptyset$.
2. For all $1 \leq i \leq m$, there is not a $1 \leq j \leq n$, $m_{ij} = 1$ and $\forall 1 \leq t \leq n, t \neq j$, $m_{it} = 0$, then let $\text{Red} = \text{Red} \cup \{s_j\}$.
3. Induce an information set $\mathcal{IS} = (U, S)$ from the given USCP.
4. While $(H(\text{Red}) \neq H(S))$ do {
   (i) For all $a_i \in S - \text{Red}$, compute Sig($s_j$, Red), $H(\{s_j\} \cup \{\text{Red}\}) = H(S)$, Red = $\text{Red} \cup \{s_j\}$ and go to Step 5;
   (ii) Let Sig($s_j$, Red) = max[Sig($s_j$, Red) | $a_i \in S - \text{Red}$];
   (iii) Let $\text{Red} = \text{Red} \cup \{s_j\}$;
5. Return $\text{Red}$, that is, $\text{Red}$ is a minimal solution of the USCP.

**Algorithm 1:** A rough set heuristic algorithm for USCP.

**Input:** A USCP = $(E, S)$ with $E = \{e_1, e_2, \ldots, e_m\}$, $S = \{s_1, s_2, \ldots, s_n\}$, $M = (m_{ij})_{mn}$

**Output:** A minimal solution.

(1) Let $\text{Red} = \emptyset$;
(2) For all $1 \leq i \leq m$, if $\exists 1 \leq j \leq n$, $m_{ij} = 1$ and $\forall 1 \leq t \leq n, t \neq j$, $m_{it} = 0$, then let $\text{Red} = \text{Red} \cup \{s_j\}$;
(3) Induce an information system $\mathcal{IS} = (U, S)$ from the given USCP;
(4) While $(H(\text{Red}) \neq H(S))$ do {
   (i) For all $a_i \in S - \text{Red}$, compute Sig($s_j$, Red), if $H(\{s_j\} \cup \{\text{Red}\}) = H(S)$, Red = $\text{Red} \cup \{s_j\}$ and go to Step 5;
   (ii) Let Sig($s_j$, Red) = max[Sig($s_j$, Red) | $a_i \in S - \text{Red}$];
   (iii) Let $\text{Red} = \text{Red} \cup \{s_j\}$;
5. Return $\text{Red}$, that is, $\text{Red}$ is a minimal solution of the USCP.

6. Conclusions

This paper has established the relationship between the USCP and the ARP. The results have shown that a
minimal or a minimum solution of a USC model is equivalent to finding an attribute reduction or a minimum attribute reduction in the induced information table. Also, finding an attribute reduction or a minimum attribute reduction of an information table is equivalent to finding a minimal or a minimum solution in the induced USC model. The results have provided the two problems to be characterized by each other and bring new methods mutually. As an application, a rough set heuristic algorithm for USCP has been given.

In practical applications, data is always not free. In view of this situation, Min et al. formally defined test-cost-sensitive decision system [49]. In test-cost-sensitive decision system, there is a test cost for each data item. The test-cost-sensitive ARP is to find a reduction which has the minimal cost [32]. In addition, in the SCP, each subset \( s_j \) associates a cost \( c_j \geq 0 \). The SCP is to find a minimal cost set covering where the cost is minimal. Our future work will plan to discuss the relationship between the test-cost-sensitive ARP and the SCP. In this paper, the induced information table constructed by a USC model is an N-soft set. Parameter reduction is an active area of research in soft set theory as well [50]. Discussing the relationship between the parameter reduction problem and the SCP is also a future work of us. In rough set theory and granular computing, one normally considers measures of granularity, as discussed in Section 5. It might be interesting to look at other measures, for example, complexity measures [51]. In the future research, we will study the attribute reduction problem based on complexity measure.

Data Availability

All data, models, and code generated or used during the study appear in the article and all reference data has been annotation references.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This work was supported by grants from National Natural Science Foundation of China (nos. 11871259 and 11701258), National Natural Science Youth Foundation of China (no. 61603173), and Natural Science Foundation of Fujian Province, China (nos. 2019J01748 and 2019J01749).

References

[1] A. Caprara, M. Fischetti, and P. Toth, “A heuristic method for the set covering problem,” Operations Research, vol. 47, no. 5, pp. 730–743, 1999.
[2] E. Marchiori and A. Steenbeek, “An evolutionary algorithm for large scale set covering problems with application to airline crew scheduling,” Lecture Notes in Computer Science, vol. 1803, pp. 367–381, Springer, Berlin, Germany, 2000.
[3] G. Lan, G. W. Depuy, and G. E. Whitehouse, “An effective and simple heuristic for the set covering problem,” European Journal of Operational Research, vol. 176, no. 3, pp. 1387–1403, 2007.
[4] J. Bautista and J. Pereira, “A GRASP algorithm to solve the unicost set covering problem,” Computers & Operations Research, vol. 34, no. 10, pp. 3162–3173, 2007.
[5] C. Toregas, R. Swain, C. Revelle, and L. Bergman, “The location of emergency service facilities,” Operations Research, vol. 19, pp. 1363–1373, 1971.
[6] F. J. Vasko and G. R. Wilson, “Hybrid heuristics for minimum cardinality set covering problems,” Naval Research Logistics Quarterly, vol. 33, no. 2, pp. 241–249, 1986.
[7] A. T. Murray, D. Tong, and K. Kim, “Enhancing classic coverage location models,” International Regional Science Review, vol. 33, no. 2, pp. 115–133, 2010.
[8] C. E. Lemke, H. M. Salkin, and K. Spielberg, “Set covering by single-branch enumeration with linear-programming sub-problems,” Operations Research, vol. 19, no. 4, pp. 998–1022, 1971.
[9] C. Mannino and A. Sassano, “Solving hard set covering problems,” Operations Research Letters, vol. 18, no. 1, pp. 1–5, 1995.
[10] M. Solar, V. Parada, and R. Urrutia, “A parallel genetic algorithm to solve the set-covering problem,” Computers & Operations Research, vol. 29, no. 9, pp. 1221–1235, 2002.
[11] L. Lessing, I. Dumitrescu, and T. Stützle, “A comparison between ACO algorithms for the set covering problem,” Ant Colony Optimization and Swarm Intelligence, vol. 3172, pp. 1–12, 2004.
[12] M. Ohlsson, C. Peterson, and B. Söderberg, “An efficient mean field approach to the set covering problem,” European Journal of Operational Research, vol. 133, no. 3, pp. 583–595, 2001.
[13] Z. a. Pawlak, “Rough sets,” International Journal of Computer & Information Sciences, vol. 11, no. 5, pp. 341–356, 1982.
[14] W. Pedrycz and W. Homenda, “Building the fundamentals of granular computing: a principle of justifiable granularity,” Applied Soft Computing, vol. 13, no. 10, pp. 4209–4218, 2013.
[15] F. Feng, X. Liu, V. Leoreanu-Fotea, and Y. B. Jun, “Soft sets and soft rough sets,” Information Sciences, vol. 181, no. 6, pp. 1125–1137, 2011.
[16] W. Ziarko, “Introduction to the special issue on rough sets and knowledge discovery,” Computational Intelligence, vol. 11, no. 2, pp. 223–226, 1995.
[17] Q. Hu, D. Yu, and M. Guo, “Fuzzy preference based rough sets,” Information Sciences, vol. 180, no. 10, pp. 2003–2022, 2010.
[18] Y. Yao and Y. Zhao, “Discernibility matrix simplification for constructing attribute reducts,” Information Sciences, vol. 179, no. 7, pp. 867–882, 2009.
[19] J. F. Peters, A. Skowron, and Z. Suraj, “An application of rough set methods in control design,” Fundamenta Informaticae, vol. 43, no. 1–4, pp. 269–290, 2000.
[20] A. Skowron and C. Rauszer, “The discernibility matrices and functions in information systems,” in Intelligent Decision Support; Handbook of Applications and Advances of the Rough Sets Theory, pp. 331–362, Kluwer, Dordrecht, Netherlands, 1992.
[21] P. Synak, J. G. Bazan, A. Skowron, and J. F. Peters, “Spatio-temporal approximate reasoning over complex objects,” Fundamenta Informaticae, vol. 67, pp. 249–269, 2005.
[22] G. Lang, Q. Li, M. Cai, T. Yang, and Q. Xiao, “Incremental approaches to knowledge reduction based on characteristic matrices,” International Journal of Machine Learning and Cybernetics, vol. 8, no. 1, pp. 203–222, 2017.
[23] D. Slezak, "Approximate entropy reducts," *Fundamenta Informaticae*, vol. 53, pp. 365–390, 2002.

[24] W.-Z. Wu, M. Zhang, H.-Z. Li, and J.-S. Mi, "Knowledge reduction in random information systems via Dempster-Shafer theory of evidence," *Information Sciences*, vol. 174, no. 3-4, pp. 143–164, 2005.

[25] M. Li, C. Shang, S. Feng, and J. Fan, "Quick attribute reduction in inconsistent decision tables," *Information Sciences*, vol. 254, pp. 155–180, 2014.

[26] S. K. M. Wong and W. Ziarko, "On optimal decision rules in decision tables," *Bulletin of the Polish Academy of Sciences Mathematics*, vol. 33, pp. 693–696, 1985.

[27] W. Zakowski, "Approximations in the space (U, Π)," *Demonstratio Mathematica*, vol. 16, pp. 761–769, 1983.

[28] W. Zhu and F.-Y. Wang, "On three types of covering-based rough sets," *IEEE Transactions on Knowledge and Data Engineering*, vol. 19, no. 8, pp. 1131–1144, 2007.

[29] L. Zhang, J. Zhan, Z. Xu, and J. C. R. Alcantud, "Covering-based general multigranulation intuitionistic fuzzy rough sets and corresponding applications to multi-attribute group decision-making," *Information Sciences*, vol. 494, pp. 114–140, 2019.

[30] Y. Xu, L. Wang, and R. Zhang, "A dynamic attribute reduction algorithm based on 0-1 integer programming," *Knowledge-Based Systems*, vol. 24, no. 8, pp. 1341–1347, 2011.

[31] J. N. K. Liu, Y. Hu, and Y. He, "A set covering based approach to find the reduct of variable precision rough set," *Information Sciences*, vol. 275, pp. 83–100, 2014.

[32] F. Min, H. He, Y. Qian, and W. Zhu, "Test-cost-sensitive attribute reduction," *Information Sciences*, vol. 181, no. 22, pp. 4928–4942, 2011.

[33] A. Tan, W. Wu, and Y. Tao, "A set-cover-based approach for the test-cost-sensitive attribute reduction problem," *Soft Computing*, vol. 21, no. 20, pp. 6159–6173, 2017.

[34] Q. Xu, A. Tan, and Y. Lin, "A rough set method for the unicast set covering problem," *International Journal of Machine Learning and Cybernetics*, vol. 8, no. 3, pp. 781–792, 2017.

[35] N. Bilal, P. Galinier, and F. Guibault, "A new formulation of the set covering problem for metaheuristic approaches," *ISRN Operations Research*, vol. 2013, p. 10, 2013.

[36] D. Molodstov, "Soft set theory-first results," *Computers & Mathematics with Applications*, vol. 37, pp. 19–31, 1999.

[37] F. Fatimah, D. Rosadi, R. B. F. Hakim, and J. C. R. Alcantud, "N-soft sets and their decision making algorithms," *Soft Computing*, vol. 22, no. 12, pp. 3829–3842, 2018.

[38] J. C. R. Alcantud, F. Feng, and R. R. Yager, "An N-soft set approach to rough sets," *IEEE Transactions on Fuzzy Systems*, p. 1, 2019.

[39] M. Irfan Ali, "A note on soft sets, rough soft sets and fuzzy soft sets," *Applied Soft Computing*, vol. 11, no. 4, pp. 3329–3332, 2011.

[40] Z. Pawlak, "Theoretical aspects of reasoning about data," in *Rough Sets*, Kluwer Academic Publishers, Dordrecht, Netherlands, 1991.

[41] D. Chen, S. Zhao, L. Zhang, Y. Yang, and X. Zhang, "Sample pair selection for attribute reduction with rough set," *IEEE Transactions on Knowledge and Data Engineering*, vol. 24, no. 11, pp. 2080–2093, 2012.

[42] G. Y. Wang, J. Zhao, J. J. An, and Y. Wu, "A comparative study of algebra viewpoint and information viewpoint in attribute reduction," *Fundamenta Informaticae*, vol. 68, pp. 289–301, 2005.

[43] J. Wang and J. Wang, "Reduction algorithms based on discernibility matrix: the ordered attributes method," *Journal of Computer Science and Technology*, vol. 16, no. 6, pp. 489–504, 2001.

[44] J. W. Grzymala-Busse, "Handbook of applications and advances of the rough set theory," in *Intelligent Decision Support*, Kluwer Academic Publishers, Dordrecht, Netherlands, 1992.

[45] X. Hu and N. Cercone, "Learning in relational databases: a rough set approach," *Computational Intelligence*, vol. 11, no. 2, pp. 323–338, 1995.

[46] Y. Qian, J. Liang, W. Pedrycz, and C. Dang, "Positive approximation: an accelerator for attribute reduction in rough set theory," *Artificial Intelligence*, vol. 174, no. 9-10, pp. 597–618, 2010.

[47] J. Liang and Z. Shi, "The information entropy, rough entropy and knowledge granulation in rough set theory," *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, vol. 12, no. 1, pp. 37–46, 2004.

[48] Y. Qian and J. Liang, "Combination entropy and combination granulation in rough set theory," *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, vol. 16, no. 2, pp. 179–193, 2008.

[49] F. Min and Q. Liu, "A hierarchical model for test-cost-sensitive decision systems," *Information Sciences*, vol. 179, no. 14, pp. 2442–2452, 2009.

[50] G. Zhang, N. Xie, and Z. Li, "Parameter reductions of soft equivalence relations," *International Journal of Machine Learning and Cybernetics*, vol. 8, no. 2, pp. 711–720, 2017.

[51] M. X. Yao, "Granularity measures and complexity measures of partition-based granular structures," *Knowledge-Based Systems*, vol. 163, pp. 885–897, 2019.