Research Article
Preference Mining Using Neighborhood Rough Set Model on Two Universes

Kai Zeng

Faculty of Information Engineering, Guizhou Institute of Technology, No. 1 Caiguan Road, Guiyang 550003, China

Correspondence should be addressed to Kai Zeng; zengkailink@sina.com

Preference mining plays an important role in e-commerce and video websites for enhancing user satisfaction and loyalty. Some classical methods are not available for the cold-start problem when the user or the item is new. In this paper, we propose a new model, called parametric neighborhood rough set on two universes (NRSTU), to describe the user and item data structures. Furthermore, the neighborhood lower approximation operator is used for defining the preference rules. Then, we provide the means for recommending items to users by using these rules. Finally, we give an experimental example to show the details of NRSTU-based preference mining for cold-start problem. The parameters of the model are also discussed. The experimental results show that the proposed method presents an effective solution for preference mining. In particular, NRSTU improves the recommendation accuracy by about 19% compared to the traditional method.

1. Introduction

In recent years, electronic retailers and content providers offer a huge selection of products. It makes the users confused in making decisions among various kinds of products (called items in this paper), such as books, music, and movies. Matching consumers with most appropriate items is important in enhancing user satisfaction and loyalty.

Many researchers have proposed different solutions to discover the preference relations between user and items. The most popular such technology is collaborative filtering (CF) [1]. There are two primary disciplines of CF: the neighborhood approach and latent factor models. A user-based neighborhood [2] approach evaluates the preference of a user by analyzing the historical rating data of neighbors who have similar taste to this user. The similar taste is usually interpreted that these users have rated most of the same items. Items that the neighbors like are then recommended to this user, as he/she will probably also like them. Latent factor models [3] built a two-dimensional matrix to describe the relation about users and items. The elements in the matrix

are the ratings given by users. Thus the issue about preference mining translates to the issue about matrix completion. However, the classical CF-based method cannot work when it meets a new user who never rates any item [4]. As a matter of fact, there even are some long-time users who have never rated any of items. It is difficult to recommend an item since there is not any historical rating information about these users to facilitate personalized recommendation. The key reason is that the CF method only relies on the historical rating data. This problem can be called cold-start which is the intrinsic limitation of CF [4].

To deal with the cold-start problem [5], Schein et al. [6] developed a combined method for recommending items that no one has yet rated. Bobadilla et al. [7] proposed improved collaborative filtering to mitigate the new user cold-start problem. The content-based filtering methods [8, 9] are also studied for the cold-start problem. However, most of the researchers have addressed the cold-start problem where either the user or the item is new. The situation with both new user and new item has seldom been considered [4]. To cope with cold-start problem when the users and items are new,

Min and Zhu [4] pioneered a cold-start recommendation approach for the new situation based rough set on two-universe model (RSTU). First, the authors considered the features of users' and items', such as age, gender, and price, to form the equivalence-based information granules. Then they use the definition of approximation operator in rough set theory for generating association rules between users and items. This granule-based recommendation model does not only rely on the rating data. It can work on better ways for dealing with the cold-start problem. However, some problems in [4] still need to be studied. For example, the equivalence-based information granule is only suitable for dealing with nominal data, such as male or female, good or bad. If the attributes are numerical, such as the price of the items, we have to adopt discretization technique to transform the nonnominal to the nominal, which would bring loss of information inevitably. It is obviously unreasonable to measure similarity or dissimilarity with Euclidean distance as to categorical attributes in numerical methods. Furthermore, the usual scoring method contains 5 scales, while the authors. [4] only consider whether or not a user has rated a movie. That is, a user will be labeled "like" if he has rated an item. It means that the user, who gives an item low mark, could be regarded as an admirer of the item. In fact, he takes dislike to the item because of the low-scoring. The key reason is that there is no evaluation about the user's rating baseline in [4]. This is also indistinguishable between positive rules and negative rules as the results.

Thus, so far, little work in this field has succeeded in holding the promise of personalized recommendation because of the following problems:

(1) cold-start problems;

(2) loss of information by discretization;

(3) no distinction between positive rules and negative rules.

Aiming at above problems, the contribution of this paper includes the following: (1) we construct the parametric neighborhood rough sets model on two universes. Users and items are described through neighborhood granules. It can overcome the loss of information by the discretization of the data. (2) The rating table is divided into positive mapping and negative mapping based by using the rating baseline. The neighborhood lower approximation operator is used for defining the preference rules. Then, we can get the positive preference rules which means "like." Some negative preference rules are also mined to be seen as "dislike." (3) The method based on the neighborhood preference rules is proposed. The cold-start problem can be solved well. The experimental results show that our model presents an effective solution for preference mining.

The paper is organized as follows: Section 2 briefly reviews the related work about cold-start issue and rough set theory on two universes. In Section 3, some basic concepts about neighborhood rough set model on single universe and granular computing-based preference mining are briefly reviewed. In Section 4, the data model and the method about baseline evaluation are investigated. In Section 5, we propose the parametric neighborhood rough set model on two universes (NRSTU). Section 6 shows the applications of NRSTU for preference rules mining and recommendation. Numeric experiments are reported in Section 7. Finally, Section 8 concludes the paper.

2. Related Work

2.1. The Cold-Start Issue. In a real recommender system, users usually present their interest by five numeric scores [10]. CF-based approach is based on the assumption that the users who have similar rating-behaviors are grouped together to help each other make a choice among the potential items. However, CF-based method is unavailable because of the cold-start problem. It occurs when the recommender system is short of ratings. We can distinguish three kinds of cold-start problems: only new user, only new item, and both of new user and new item [7].

2.1.1. New User [11–13]. New user means we face the user who never rates any item. It is hard for a new user to obtain the preferred items since no available usage information can be employed in personalized recommendation [10]. The improved CF-based method such as CF-content or CF-demographic is the common strategy to tackle the new user problem [7]. For example, Braunhofer aimed at solving cold-start problem by using various contextually tagged rating datasets [14]. Leung et al. proposed a new content-based hybrid approach that makes use of cross-level association rules to integrate content information about domains items [15]. Ahn presented a new heuristic similarity measure named PIP that focuses on improving recommendation performance under new user condition [16]. All the former approaches are based on the assumption that the users are new. As a matter of fact, there could also be lots of new items in a real recommender system.

2.1.2. New Item [17, 18]. The new item problem arises due to the fact that the new items have no initial ratings. New items are not likely to be recommended since there is not any rating information to facilitate personalized recommendation. Simplest solution is to encourage the motivated users who are responsible for rating each new item in the system [7]. In addition, the authors [19] investigated the value of even a few ratings in regard to predictive power. Cremonesi et al. presented two different approaches for building hybrid collaborative content recommender systems to overcome the new item issue [5]. Elbadrawy and Karypis proposed a feature-based similarity model for top-n recommendation of new items [20]. In a word, the new item problem [21] is more often addressed.

2.1.3. New User and New Item. The new user or new item problem is addressed individually. However, the situation with both new user and new item has seldom been considered. It is very difficult to recommend a new item to a new user since the historical data of the current user and item is unknown. Rough set model so far is the best way to deal with

this situation. Min and Zhu studied the cold-start problem by using the rough set model on two universes. They provide a means for describing users and items through information granules, a means for generating association rules between users and items, and a means for recommending items to users using these rules [4]. Min and Zhu explained the concept about granular association rule in detail [22]. Then, some rough set based methods are proposed to solve the new user and new item problem by extending Min work [23, 24].

2.2. Rough Set Model on Two Universes. The rough sets theory [25], proposed by Pawlak in 1982, is a powerful mathematical method for the study of incomplete or imprecise information. This theory has been successfully applied to many fields, such as data mining, decision making, pattern recognizing, machine learning, and intelligent controlling [26–29].

The Pawlak approximation operators are defined by an equivalence relation on the universe. The equivalence relation forms a partition of the universe. Partition or equivalent relation is still restrictive for many applications. To address this problem, Lin firstly proposed the theory of neighborhood system [30–32]. The idea about neighborhoods of tolerance is pointed out by Professor Lin. Neighborhood theory can be seen as a breakthrough of rough set. Professor Zhu studied the covering-based rough set [33, 34]. He gives the definition of smallest covering. It is a key issue to reduce the redundant information in data mining. Zhu investigated the relationship among basic concepts in covering-based rough sets. This excellent research helps us have a better understanding of covering-based rough set. Neighborhood rough set and cover-based rough set are all meaningful extensions to equivalent relation.

Most of the researches have been conducted on the assumption of the same universe [26–29]. However, two-universe model is more appropriate for the preference mining problem. In general, for a certain user, he or she may grade many items. Meantime, it also could include many items rated by some users in a real recommender system. Then an effective method to describe this problem is using two different universes. One of the universes is the set of all users. Another is the set of all the items. Some results have been generated in the rough sets theory on two universes. In [35], the authors gave a general framework of the two-universe based rough sets model. Shen and Wang defined a variable precision rough sets model based on classical Pawlak model, in which the lower approximation and the upper approximation were generalized to two universes [36]. The concept of the probabilistic rough sets on two universes was firstly defined by Gong and Sun [37]. In [38], the authors focus on the properties of probabilistic rough sets on two universes.

Previous studies are all the extension of Pawlak rough set model on two universes. The equivalence-based information granules in Pawlak rough set are not suitable for dealing with numerical data. It means we have to adopt discretization technique to transform the nonnominal to the nominal for the numerical attributes. The strength of the approach presented in this paper lies in the ability of neighborhood granules to avoid the discretization of the data for the new user and new items problem.

3. Preliminary

In this section, we first review the model preference mining in the view of granular computing which was proposed in [4, 39]. Then, the classical neighborhood rough set model on single universe is also reviewed.

3.1. Preference Mining in the View of Granular Computing

Definition 1 (see [40]). Knowledge representation is realized via the information system (IS) which is a tabular form, similar to databases. An information system is pair representation realized via the information system (IS) which is a tabular form, similar to databases. An information system is pair IS $= (U, A)$, where $U = \{x_1, x_2, \ldots, x_n\}$ is a nonempty finite set of objects and A is a nonempty finite set of attributes.

Definition 2 (see [40]). An equivalence granule can be defined as follows:

$$E(x) = \{(x, z) \in U \mid \forall a \in A, \ a(x) = a(z)\}. \tag{1}$$

In an information system, a granule coincides with a concept, which is a basic unit of human thought understood as a pair of intension and extension [41].

The rough sets theory [25], proposed by Pawlak in 1982, is a powerful mathematical method for the study of incomplete or imprecise information. This theory has been successfully applied to many fields, such as data mining, decision making, pattern recognizing, machine learning, and intelligent controlling. The following definitions, which had been defined in the theory of rough set, were employed by Min and Zhu [39].

Definition 3 (see [42]). Let $U = \{x_1, x_2, \ldots, x_n\}$ and $W = \{y_1, y_2, \ldots, y_n\}$ be two nonempty set of objects. $\mathfrak{R} \subseteq U \times W$ is the set-valued mapping from universe U to W. Then, $(U, A, W, B, \mathfrak{R})$ is called the general approximation spaces, where (U, A) and (W, B) are two information systems.

This data model can be called two-universe model. In the preference mining method [39], U can be consider as the set of users and W is the set of items where the two entities are connected by the relation $\mathfrak{R}$.

Definition 4. Given the general approximation space $(U, A, W, B, \mathfrak{R})$, for a subset $Y \subseteq W$, we define the lower and upper approximations of Y on the space $(U, A, W, B, \mathfrak{R})$ as follows, respectively.

$$\underline{apr}(Y) = \{x \in U \mid \mathfrak{R}(x) \subseteq Y\},$$
$$\overline{apr}(Y) = \{x \in U \mid \mathfrak{R}(x) \cap Y \neq \varnothing\}. \tag{2}$$

The above granular computing model is called classical rough set model on two universes.

Definition 5 (see [39]). A granular preference rule is an implication of the form

$$E(x) \Longrightarrow E(y) \tag{3}$$

$$\text{iff } \mathfrak{R}(x) \cap E(y) \neq \varnothing, \ x \in U, \ y \in W.$$

In the view of rough set [39], the definition of upper approximation is used for describing the preference rules where $E(y)$ is substituted for the subset $Y \subseteq W$. The granular preference rule can be interpreted that the users in $E(x)$ like the items in $E(y)$.

3.2. Neighborhood Rough Set on the Single Universe. As the previous sections described, the classical rough set model is unavailable when the datasets are numerical. Hu et al. introduced a neighborhood rough set model in [43] for the heterogeneous data to avoid discretization. The authors considered that the data model is on the single universe. Then, we will give the basic concepts about the neighborhood rough set model on single universe.

Definition 6 (see [43]). U is a nonempty finite set of objects and Δ is a given distance function. We say (U, A, δ) is a neighborhood approximation space where

(1) $\Delta(x_i, x_j) \geq 0$, if and only if $x_i = x_j$, $\Delta(x_i, x_j) = 0$;

(2) $\Delta(x_i, x_j) = \Delta(x_j, x_i)$;

(3) $\Delta(x_i, x_k) \leq \Delta(x_i, x_j) + \Delta(x_j, x_k)$.

Definition 7 (see [43]). Given a neighborhood approximation space (U, A, δ), $\forall x \in U$, $\delta \geq 0$, we say $\delta(x)$ is a δ neighborhood of x whose centre is x and radius is δ, where $\delta(x) = \{z \mid \Delta(x, y) \leq \delta, \ z \in U\}$.

Here, $\delta(x)$ can be seen as the neighborhood granule.

Remark 8 (see [40]). Given two points $x_i = \{x_i^1, x_i^2, \ldots, x_i^n\}$ and $x_j = \{x_j^1, x_j^2, \ldots, x_j^n\}$ in N-dimensional Euclidean space, the distance of them can be computed as

$$\Delta(x_i, x_j) = \sqrt{\sum_{l=1}^{N} d_{a_l}(x_i, x_j)^2},$$

$$d_{a_l}(x_i, x_j) \tag{4}$$

$$= \begin{cases} \text{nom\_diff}_{a_l}(x_i, x_j) & \text{if } a_l \text{ is a nominal attribute} \\ \text{num\_diff}_{a_l}(x_i, x_j) & \text{if } a_l \text{ is a numerical attribute,} \end{cases}$$

where

$$\text{nom\_diff}_{a_l}(x_i, x_j) = \begin{cases} 0 & \text{if } x_i^l = x_j^l \\ 1 & \text{if } x_i^l \neq x_j^l, \end{cases} \tag{5}$$

$$\text{num\_diff}_{a_l(x_i, x_j)} = \frac{\left| x_i^l - x_j^l \right|}{\max\left(V_{a_l}\right) - \min\left(V_{a_l}\right)}.$$

TABLE 1: Rating table and baseline evaluation.

Users	Movies	Ratings	Baseline
u_230	m_280	4.0	2.8
u_13	m_219	1.0	2.3

TABLE 2: Rating table and baseline evaluation.

Users	Movies	Ref [4]	Our study
u_230	m_280	1	1
u_13	m_219	1	0

Definition 9 (see [43]). Given a neighborhood approximation space (U, A, δ), for any subset $X \subseteq U$, we define the lower and upper approximations of X on the space (U, A, δ) as follows, respectively.

$$\underline{\text{apr}}(X) = \{x \in U \mid \delta(x) \subseteq X\},$$

$$\overline{\text{apr}}(X) = \{x \in U \mid \delta(x) \cap X \neq \varnothing\}. \tag{6}$$

Obviously, classical neighborhood rough set model on single universe is not suitable for the user-item data mode structure.

4. Data Model and Baseline Evaluation

4.1. Data Model. In this study, we illustrate the movie recommendation by making use of MovieLens [44] which is widely used in many preference mining researches (e.g., [4, 10, 39]). We use the version with 943 users and 1,682 movies. Thus, the items are movies in this study. The database schema is follows.

(i) Users: $A = \{$u_id, age, gender, occupation$\}$.

(ii) Movies: $B = \{$m_id, release − year, genre$\}$.

According to Definitions 1 and 3, U and W are the set of users and movies. The set A and B are the features of users and movies, respectively. The specific information can be found in Tables 3 and 4 in Section 7.

The original rating data contains 5 scales showed in Table 1. Reference [4] only considers whether or not a user has rated a movie. That is, it can be established as the mapping relationship $\mathfrak{R}$ from u_13 to m_219 because that u_13 has rated m_219. An example of set-valued mapping $\mathfrak{R}$ is given in Table 2. Then we can mine some preference rules from this type of mapping relationship.

As a matter of fact, 1.0 point means u_13 in all probability dislike m_219. In other words, some of such rules are negative rules. The implication of the word "negative" is "dislike" in this study. Obviously, this type of mapping relationship is unreasonable. Therefore, how to establish the appropriate mapping relationship from users to movies is a key issue for preference mining.

4.2. Baseline Evaluation. The rating baseline tends to capture the basic emotions of a user for one item [3]. In this study,

TABLE 3: Information of movies.

ID	Release year	Genre
m_192	1980	Drama
m_861	1986	Horror
m_1140	1994	Comedy
m_1165	1996	Comedy, drama
m_1641	1995	Documentary
m_1680	1998	Drama, romance
m_280	1996	Drama, romance
m_219	1984	Horror

TABLE 4: Information of users.

ID	Age	Gender	Occupation
u_13	47	Male	Educator
u_32	28	Female	Student
u_44	26	Male	Technician
u_936	26	Male	Other
u_937	48	Male	Educator
u_940	32	Male	Administrator
u_230	28	Female	Student

it can be considered that a user likes a movie if he (she) gives a higher score than his (her) raging baseline. The simplest computational method of baseline is the overall average rating of movies. For example, the average rating over all movies is 3.7 stars. User Joe can be regarded as an admirer of the movie Avatar if he has given Avatar 4.0 points. However, the authors [3, 45] proposed that the user and item bias exist in the real rating system. Some users, for instance, frequently give higher ratings than others. Similarly, some items always receive higher ratings than others. That is to say, the rating baseline should be personalized. Therefore, we use the baseline evaluation method in [45]. A user's rating baseline for a movie b_{ui} accounts for the user and item bias effects.

$$b_{ui} = \mu + b_u + b_i. \tag{7}$$

μ is the overall average rating. The parameters b_u and b_i indicate the raging bias of user u and item i, respectively. We also use the example about Avatar to explain the meaning of b_{ui}. Avatar is more popular than an average movie, so it tends to be rated 0.5 stars above the average. On the other hand, Joe is a critical user, who tends to rate 0.1 stars lower than the average. Thus, the baseline estimate for Avatar's rating by Joe would be 4.1 stars by calculating $3.7 - 0.1 + 0.5$. It means Joe might not be interested in Avatar if he has given Avatar 4.0 points.

In order to estimate b_u and b_i one can solve the least squares problem [1, 3, 45]:

$$C(\kappa) = \min_{b_*} \sum_{(u,i) \in \kappa} (r_{ui} - \mu - b_u - b_i)^2 + \lambda \left(\sum_u b_u^2 + \sum_i b_i^2 \right). \tag{8}$$

Here, κ is all the training samples. The first term $\sum_{(u,i) \in \kappa} (r_{ui} - \mu - b_u - b_i)^2$ strives to find b_u's and b_i's that fit the given ratings r_{ui}. The regularizing term $\lambda(\sum_u b_u^2 + \sum_i b_i^2)$ avoids overfitting by penalizing the magnitudes of the parameters. And then, we can take the derivative with respect to b_u and b_i.

$$\frac{\partial C}{\partial b_u} = -2 \cdot b_i + 2\lambda b_u,$$
$$\frac{\partial C}{\partial b_i} = -2 \cdot b_u + 2\lambda b_i. \tag{9}$$

Thus, gradient descent was adopted as learning algorithm [1, 3, 45].

$$b_u \longleftarrow b_u + \alpha (b_i - \lambda b_u),$$
$$b_i \longleftarrow b_i + \alpha (b_u - \lambda b_i). \tag{10}$$

The parameter α is set to 0.003 in this study. Thus, μ, b_u, and b_i can be calculated with the above formulas.

We use this baseline evaluation method for building the mapping relationship between the two universes. A user can be considered as an admirer of a movie if he gives a higher rating than the baseline b_{ui} of the movie. The fourth column in Table 1 shows the results by the baseline estimate. Then the set-valued mapping $\Re$ is rebuilt in the fourth column in Table 2.

5. Neighborhood Rough Set Model on Two Universes

In this section, we build the neighborhood rough set model on two universes (NRSTU).

Definition 10. Let U and W be two nonempty finite universes. $\delta(x)$ is a δ neighborhood of x whose centre is $x \in U$. $\Re$ is the set-valued mapping from universe U to W where $\Re(x) \subseteq W$. Then, we have the object set $\Re(\delta(x)) \subseteq W$ if $\delta(x) \subseteq U$. Hence, $(U, W, \Re, \delta)$ is called neighborhood approximation space on two universes.

$\Re(\delta(x))$ is the set of the elements in W which are mapped from $\delta(x) \subset U$.

Definition 11. Given a neighborhood approximation space on two universes $(U, W, \Re, \delta)$ and $Y \subseteq W$, the lower and upper approximation about Y can be defined as

$$\underline{apr}(Y) = \{x \in U \mid \Re(\delta(x)) \subseteq Y\},$$
$$\overline{apr}(Y) = \{x \in U \mid \Re(\delta(x)) \cap Y \neq \varnothing\}. \tag{11}$$

The boundary region, positive region, and negative region can be defined as

$$BN(Y) = \overline{apr}(Y) - \underline{apr}(Y),$$
$$POS(Y) = \underline{apr}(Y),$$
$$NEG(Y) = U - \overline{apr}(Y). \tag{12}$$

Theorem 12. *Let $(U, W, \mathfrak{R}, \delta)$ be neighborhood rough sets on two universes. For any $Y \subseteq W$, given two positive numbers δ_1 and δ_2, if $\delta_1 \geq \delta_2$, one has*

$$\underline{apr}_{\delta_1}(Y) \subseteq \underline{apr}_{\delta_2}(Y),$$

$$\overline{apr}_{\delta_1}(Y) \supseteq \overline{apr}_{\delta_2}(Y). \tag{13}$$

Proof. (1) $\forall x \in U$ and $\delta_1 \geq \delta_2$: we have $\mathfrak{R}(\delta_1(x)) \supseteq \mathfrak{R}(\delta_2(x))$. Assuming $R(\delta_1(x)) \subseteq Y$, we have $\mathfrak{R}(\delta_2(x)) \subseteq Y$. Therefore, we must have $x \in \underline{apr}_{\delta_2}(Y)$ if $x \in \underline{apr}_{\delta_1}(Y)$. However, x is not sure in $\underline{apr}_{\delta_1}(Y)$ if we have $x \in \underline{apr}_{\delta_2}(Y)$. Hence, $\underline{apr}_{\delta_1}(Y) \subseteq \underline{apr}_{\delta_2}(Y)$.

(2) $\forall x \in U$ and $\delta_1 \geq \delta_2$: we have $\mathfrak{R}(\delta_1(x)) \supseteq \mathfrak{R}(\delta_2(x))$. Assuming $\mathfrak{R}(\delta_2(x)) \cap Y \neq \phi$, we have $\mathfrak{R}(\delta_1(x)) \cap Y \neq \phi$. Therefore, we must have $x \in \underline{apr}_{\delta_1}(Y)$ if $x \in \underline{apr}_{\delta_2}(Y)$. However, x is not sure in $\underline{apr}_{\delta_2}(Y)$ if we have $x \in \underline{apr}_{\delta_1}(Y)$. Hence, $\overline{apr}_{\delta_1}(Y) \supseteq \overline{apr}_{\delta_2}(Y)$. □

In the practical environment, neighborhood upper approximation is inadequate for describing user preference. For example, we need to know how many items in the granule users like. Here, we proposed a parametric neighborhood rough sets model on two universes where $|\cdot|$ is the cardinality of the set.

Definition 13. Let $(U, W, \mathfrak{R}, \delta, p)$ be parametric neighborhood approximation space on two universes. For any $0 \leq \beta < \omega \leq 1$, $Y \subseteq W$. Then, the lower and upper approximations of Y with ω and β are as follows.

$$\underline{apr}_p^\omega(Y) = \left\{ x \in U \mid \frac{|\mathfrak{R}(\delta(x)) \cap Y|}{|Y|} \geq \omega \right\},$$

$$\overline{apr}_p^\beta(Y) = \left\{ x \in U \mid \frac{|\mathfrak{R}(\delta(x)) \cap Y|}{|Y|} > \beta \right\}. \tag{14}$$

We have proposed a variable precision neighborhood rough sets model in the previous research [46]. In this study, ω is the coverage degree of $\mathfrak{R}(\delta(x))$ to Y. It does not mean the precision of an approximation. This is why we call this model parametric neighborhood rough sets instead of variable precision neighborhood rough sets.

Definition 14. Let $(U, W, \mathfrak{R}, \delta, p)$ be variable precision neighborhood approximation space on two universes. For any $x_i \in U$, the neighborhood uncertainty of x is defined as

$$\text{NH}^{x_i}(U, W) = -\log \frac{|\mathfrak{R}(\delta(x_i))|}{|W|}. \tag{15}$$

The uncertainty of approximation space is computed as

$$\text{NH}(U, W) = -\frac{1}{|U|} \sum_{x_i} \log \frac{|\mathfrak{R}(\delta(x_i))|}{|W|}. \tag{16}$$

6. Preference Rules and Recommendation

In this section, we first define the preference rules by using the neighborhood rough set models on two universes. Then the method of recommendation is also discussed in detail.

6.1. Preference Rules. Min has given the formulation of the preference rule on the view of equivalence granule [4]. In our study, we define a neighborhood granular preference rule by extending Min's work.

Definition 15. Given a neighborhood approximation space on two universes $(U, W, \mathfrak{R}, \delta)$ and $x_i \in U$, $y_j \in W$. A neighborhood granular preference rule can be described as follows.

$$\delta(x_i) \Longrightarrow \delta(y_j) \quad \text{iff } \mathfrak{R}(\delta(x_i)) \cap \delta(y_j) \neq \varnothing. \tag{17}$$

The formulation is consistent with the upper approximation as Definition 11. In an information system, a granule is a basic unit of human thought [41]. In a regular recommended system, such as CF method, the meaning of neighborhood is that the similar people have similar interests [45]. Therefore, the definition of neighborhood upper approximation can be interpreted that some similar users in $\delta(x_i)$ like the similar items in $\delta(y_j)$.

The rules are usually evaluated through two measures, namely, support degree and confidence degree, which are well defined in [47] for the single universe model. In the two-universe model, there are user and item support degree of the rules, respectively.

$$\sup_{\text{user}} = \frac{|\delta(x_i)|}{|U|} \geq \eta,$$

$$\sup_{\text{item}} = \frac{|\delta(y_j)|}{|W|} \geq \gamma. \tag{18}$$

On the other hand, the higher proportion of items in $\delta(y_j)$ users like also indicates that the preference rule is stronger. Hence, we also use the third measure called confidence degree for mining the stronger rules.

$$\text{conf} = \frac{|\mathfrak{R}(\delta(x_i)) \cap \delta(y_j)|}{|\delta(y_j)|} \geq \omega. \tag{19}$$

Thus, the parameter ω in the parametric neighborhood rough sets model can serve as the confidence degree of the rules. This is why we propose the parametric neighborhood rough sets model in Definition 13. The formulation of preference rules is redefined by using the lower approximation in the parametric neighborhood rough sets model.

$$\delta(x_i) \Longrightarrow \delta(y_j) \quad \text{iff } \frac{\mathfrak{R}(\delta(x_i)) \cap \delta(y_j)}{\delta(y_j)} \geq \omega. \tag{20}$$

Here, the neighborhood lower approximation is employed for the definition of the preference rule. This type of rule can be read as "the users in $\delta(x_i)$ like at least $\omega\%$ of items in $\delta(y_j)$". In our study, we only consider the condition that the users and items are described by all of the features.

A straightforward algorithm for preference rules mining is given by Algorithm 1 which has two steps.

```
Input $(U, W, \mathfrak{R}, \delta), \eta, \gamma, \omega$
Output All rules satisfying given constraints
(1) find all user neighborhood granules $G_{\text{user}}$ where $|\delta(x_i)|/|U| \geq \eta$
(2) find all item neighborhood granules $G_{\text{item}}$ where $|\delta(y_j)|/|U| \geq \gamma$
(3) for each $\delta(x_i)$ in $G_{\text{user}}$
(4)     for each $\delta(y_j)$ in $G_{\text{item}}$
(5)         if conf = $|\mathfrak{R}(\delta(x_i)) \cap \delta(y_j)|/|\delta(y_j)| \geq \omega$
(6)             output rule $\delta(x_i) \Rightarrow \delta(y_j)$
(7)         end if
(8)     end for
(9) end for
```

ALGORITHM 1: Preference rules mining.

Step 1. Search all neighborhood granules meeting the minimal support threshold of user granule η and γ. This step corresponds to Lines 1 and 2 of the algorithm, where G_{user} and G_{item} stand for user granules and item granules, respectively.

Step 2. Check all possible rules regarding G_{user} and G_{item} and output valid ones. This step corresponds to Line 3 through Line 9 of the algorithm.

6.2. Preference Rules for Recommendation. The recommendation method is derived from the idea that the neighbors have the same taste. Given a preference rule $\delta(x_i) \Rightarrow \delta(y_j)$ and a new user u, we recommend items in $\delta(y_j)$ if u is a neighbor of x_i.

The performance of a recommender is evaluated mainly by the recommendation accuracy [4]. Formally, let the number of recommended items be M and the number of appropriate recommendations N; the accuracy is N/M [4]. In our study, appropriate recommendation can be interpreted that the real score of the recommended item is higher than the rating baseline of the user. We will elaborate it in detail by the next experiments.

7. Experiments

In this section, we will evaluate our model through experimentation. In this study, the method of preference mining has improved by two aspects as follows.

(1) NRSTU is proposed to overcome the cold-start problem.

(2) NRSTU enhanced the effectiveness of preference mining by avoiding the discretization and the dividing positive rules and negative rules.

We design three experiments to verify the two points above. First of all, we give an experimental example to show the details of NRSTU-based preference mining. It elaborates how to solve the cold-start problem by using NRSTU. Then, we will discuss the parameters of our model through experimental analysis. To compare the effectiveness of NRSTU, we choose the classical rough set model on two universes (RSTU) as the benchmark. RSTU are employed in

[4, 39] for preference mining. We download the discretization samples from [48] which are preprocessed by Min as in [39]. This experiment is called NRSTU versus RSTU.

7.1. The Meaningfulness of Neighborhood Preference Rules and Recommendation. Firstly, we look at some rules mined from MovieLens dataset by Algorithm 1. The setting is as follows: $\eta = \gamma = 0.02$ and $\omega = 0.3$. The training samples percentage and test samples percentage are 90% and 10%, respectively. We can get lots of positive rules form the users who are labeled "like the movies." Similarly, some negative rules are also obtained from the users who dislike those movies. Some of them are listed below as the examples.

Some positive preference rules are as follows:

(1) $\delta_{\text{user}}(13) \Rightarrow \delta_{\text{movie}}(1641)$.

(2) $\delta_{\text{user}}(32) \Rightarrow \delta_{\text{movie}}(1680)$.

(3) $\delta_{\text{user}}(44) \Rightarrow \delta_{\text{movie}}(1165)$.

(4) $\delta_{\text{user}}(940) \Rightarrow \delta_{\text{movie}}(861)$.

Some negative preference rules are as follows:

(5) $\delta_{\text{user}}(937) \nRightarrow \delta_{\text{movie}}(861)$.

(6) $\delta_{\text{user}}(936) \nRightarrow \delta_{\text{movie}}(1140)$.

(7) $\delta_{\text{user}}(940) \nRightarrow \delta_{\text{movie}}(192)$.

As in Tables 3 and 4, we use 8 movies and 7 users to expound the meaningfulness of neighborhood preference rules. For example, $\delta_{\text{user}}(32) \Rightarrow \delta_{\text{movie}}(1680)$ means that the users in $\delta_{\text{user}}(32)$ like the movies in $\delta_{\text{movie}}(1680)$. It can be interpreted that the female students who are about 28 years old like the movies which are labeled drama and romance. Then, we can recommend movies to other users by using the neighborhood preference rule as the Algorithm 2. Suppose u_230 is a new user who has not rated any of the movies. The classical CF-based method cannot work if we do not know any of the historical rating data of u_230. We can get her registration information. Then we will know she is a female student. Hence, we can recommend the movie m_280 to her because of m_280 $\in \delta_{\text{movie}}(1680)$. As a matter of fact, m_280 is scored 4.0 points by user u_230 where the baseline of u_230 on movie 280 is 2.8 (see Table 1). It means that user u_230

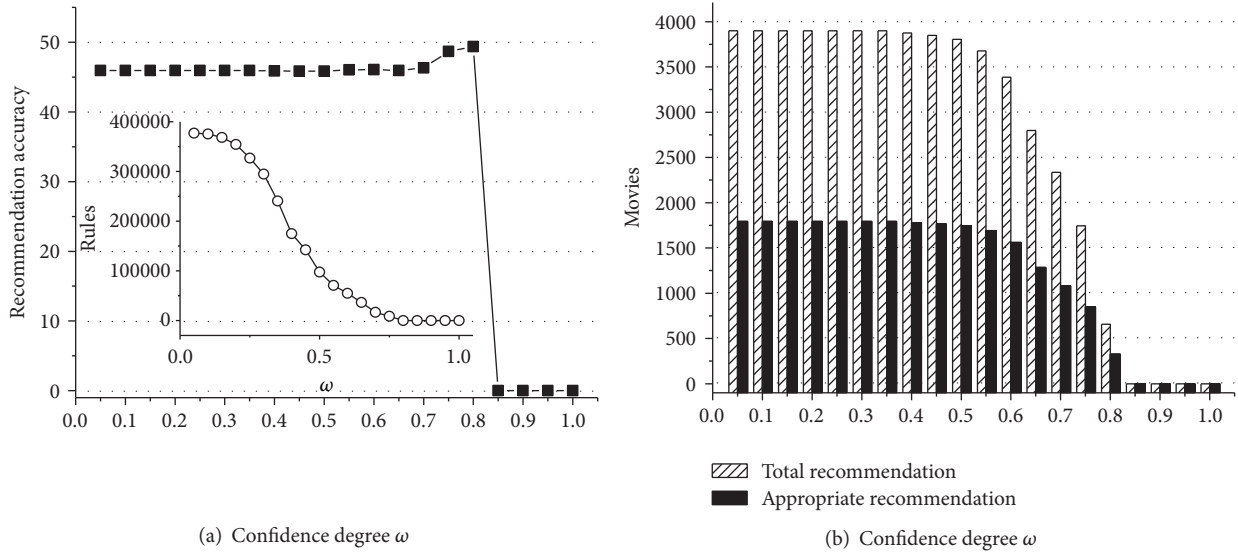(a) Confidence degree ω

(b) Confidence degree ω

FIGURE 1: The results of recommendation where the support degree is $\eta = \gamma = 0.02$. (a) Recommendation accuracy (%) and amount of the rules with the confidence degree ω; (b) total and appropriate recommendations with the confidence degree ω.

```
Input Test users, All rules
Output The recommendations for the test users
(1) for each test users u
(2)    for each rule
(3)       if u is a neighbor of x_i
(4)          recommend δ(y_j) for u.
(5)       end if
(6)    end for
(7) end for
```

ALGORITHM 2: Recommendation method.

really enjoys the movie m_280. Here we say it is an appropriate recommendation.

On the other hand, our method also mines some negative preference rules which can solve the problem in [4, 39]. The authors only consider whether or not a user has rated a movie. A user will be labeled "like" if he has rated a movie. It means that the user, who gives a movie low mark, could be regarded as an admirer of the movie. That is to say, the negative preference rules (5), (6), and (7) will be regarded as the positive preference rules in the view of [4, 39]. For example, as Table 1 in Section 3, suppose u_13 is a new user who is a neighbor of u_937. Movie m_219 will be recommended to u_13 because of m_219 $\in \delta_{\text{movie}}(861)$. As a matter of fact, m_219 is scored 1.0 point by user u_13 where the baseline of u_13 on movie 219 is 2.3. It means that user u_13, who is a male educator and 47 years old, does not like the horror movie at all. Apparently it is an improper recommendation. This experiment shows the baseline evaluation in an important step in preference mining

7.2. Parameters Discussion. There are four parameters in our model. They are neighborhood metric δ, confidence degree

ω, support degree η, and γ, respectively. Literature [43] has explained that the result is optimal if threshold δ is set between 0.1 and 0.2 in the neighborhood system. In our study, threshold δ is set to 0.15. Then, the selection of ω, η, and γ is discussed through a series of experiments. We set $\eta = \gamma$ from 0.02 to 0.12 with step 0.02 because we cannot get any of rules when $\eta = \gamma > 0.12$. We try confidence degree ω from 0.05 to 1.0 with step 0.05. For the MovieLens datasets, we randomly divide the samples into 10 subsets and use nine of them as training set and the rest one as the test set. After 10 rounds, we compute the average recommendation accuracy as the final performance.

According to experiment data above, we can obtain some useful conclusions. First of all, Figures 1–6 show that the highest recommendation accuracy always keeps from 45% to 50% if $0 < \omega \leq 0.55$ in most cases. When we increase the numerical value of ω from 0.45 to 0.55, it concomitantly reduces the number of the rules. However, total and appropriate recommendations do not significantly reduce in this interval. The amount of the rules will drop to 0% if $\omega \geq 0.65$. Then there are no longer any of recommendations. That is, we can obtain the highest recommendation accuracy by fewer rules when $0.45 < \omega \leq 0.55$. On the other hand, the amount of total and appropriate recommendations did not significantly reduce when $0.45 < \omega \leq 0.55$ in most cases. It can be concluded that some preference rules are redundant if ω is set in $(0, 0.45]$. As a matter of fact, we just need more recommendations and the highest recommendation accuracy by using the minimum amount of rules. In this point of view, $0.45 < \omega \leq 0.55$ is a cost-effective choice for confidence degree ω.

For the support degree η and γ, they only impact on the amount of the rules and recommendations showed in Figures 1(b), 2(b), 3(b), 4(b), 5(b), and 6(b). The recommendation accuracy has nothing to do with η and γ. It is obvious that the higher the support degree we set the less rules and

(a) Confidence degree ω

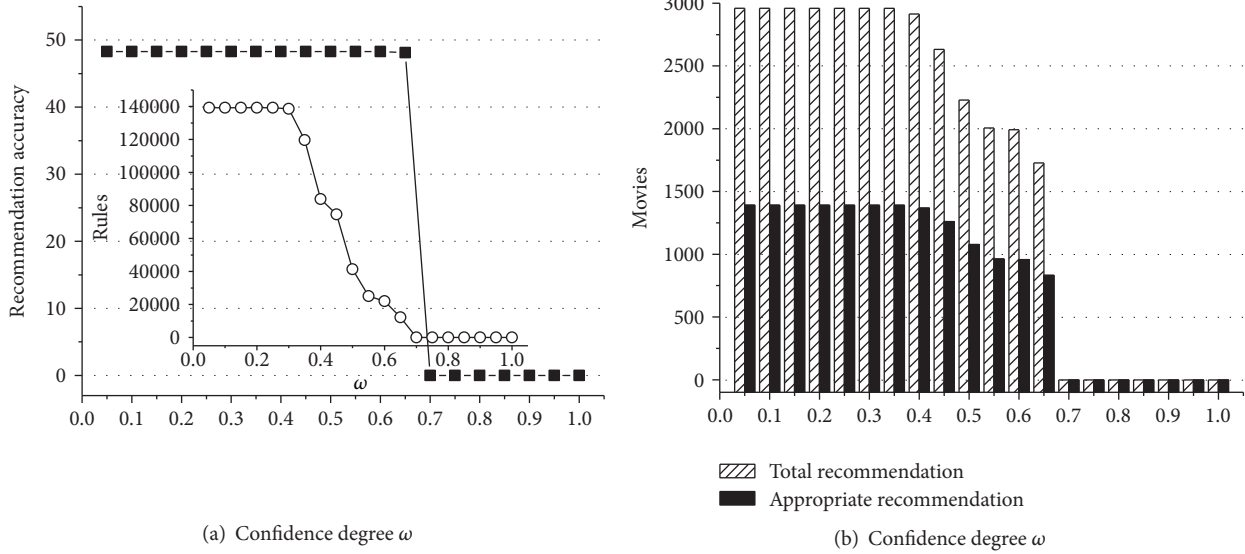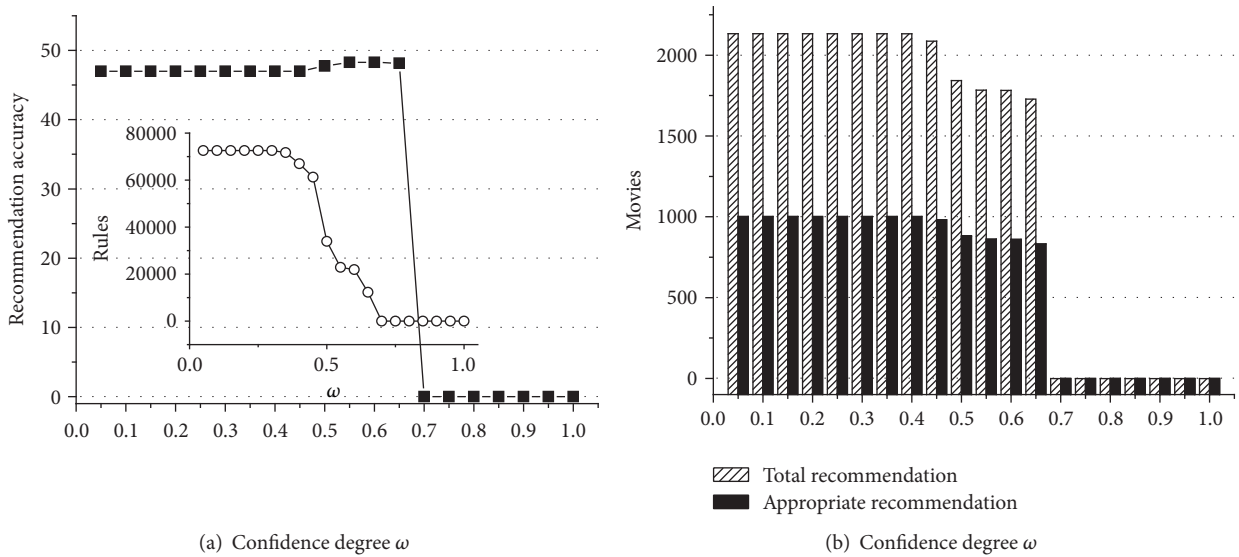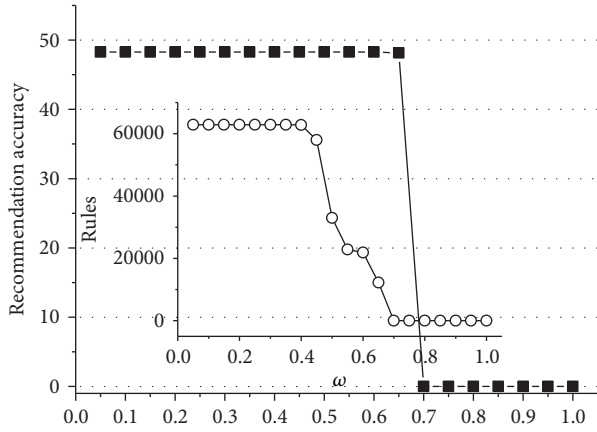(b) Confidence degree ω

FIGURE 2: The results of recommendation where the support degree is $\eta = \gamma = 0.04$. (a) Recommendation accuracy (%) and amount of the rules with the confidence degree ω; (b) total and appropriate recommendations with the confidence degree ω.

(a) Confidence degree ω

(b) Confidence degree ω

FIGURE 3: The results of recommendation where the support degree is $\eta = \gamma = 0.06$. (a) Recommendation accuracy (%) and amount of the rules with the confidence degree ω; (b) total and appropriate recommendations with the confidence degree ω.

recommendations we can get. Nonetheless, the high support degree means the strong preference rules. Consequently, it is more reasonable to set the support degree based on actual demand. It depends on which one is more needful between stronger preference rules and more recommendations for a real recommended system. We can even set different values for η and γ on a per-destination basis. Hence, it becomes an open-ended question.
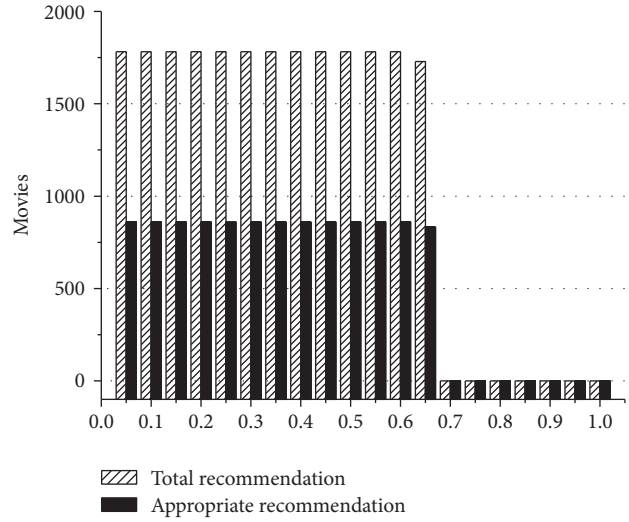
7.3. NRSTU versus RSTU. In this section, we choose RSTU [4] as the benchmark. In [4], the authors proposed a cold-start recommendation approach by using the classical parametric rough set model. The main difference between

NRSTU and RSTU is the granules that are structured by the equivalence relation in [4] rather than neighborhood relation. It means that the users (or items) can be sorted into one granule if their features are identical. Actually, the features are rarely exactly alike because of the numerical features such as age and release year. Therefore, it has to adopt discretization technique to transform the nonnominal to the nominal, which would bring loss of information inevitably. In this experiment, the setting is as follows: $\eta = \gamma = 0.02$ and $\omega = 0.5$.

The random recommender, which has an accuracy close to 6.2%, is illustrated for comparison in [4]. As is shown in Table 5, NRSTU yield better performance than RSTU
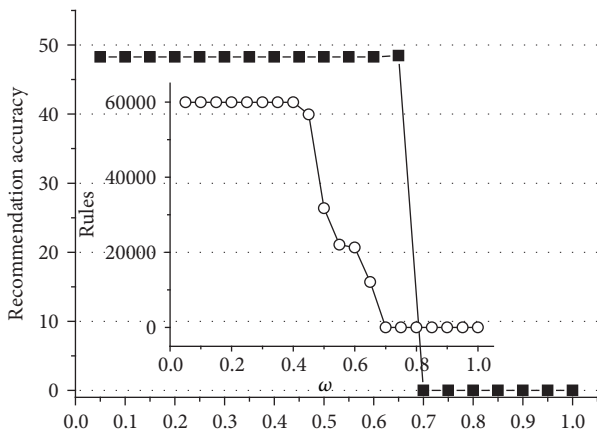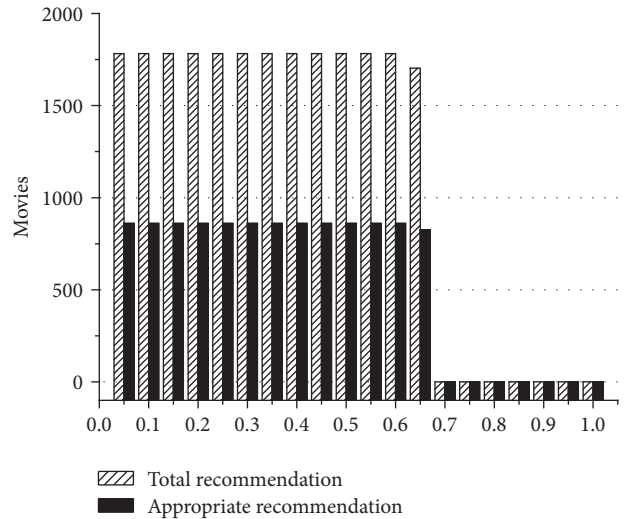
(a) Confidence degree ω

(b) Confidence degree ω

FIGURE 4: The results of recommendation where the support degree is $\eta = \gamma = 0.08$. (a) Recommendation accuracy (%) and amount of the rules with the confidence degree ω; (b) total and appropriate recommendations with the confidence degree ω.

(a) Confidence degree ω

(b) Confidence degree ω

FIGURE 5: The results of recommendation where the support degree is $\eta = \gamma = 0.10$. (a) Recommendation accuracy (%) and amount of the rules with the confidence degree ω; (b) total and appropriate recommendations with the confidence degree ω.

and random recommender obviously. In particular, NRSTU improves the recommendation accuracy by about 19%. There are two reasons for the explanation. The first reason is that RSTU is unreasonable to measure similarity or dissimilarity with Euclidean distance as to categorical attributes in numerical methods. Hu et al. [43] pointed out that there are at least two categories of structures lost in discretization: neighborhood structure and order structure in real spaces. For example, we know the distances between samples and we can get how the samples are close to each other in real spaces. NRSTU addresses this deficiency by the neighborhood granule structure. Furthermore, there is no distinction between

positive rules and negative rules in [4]. It is consequent to give some unreasonable recommendations by RSTU which has been discussed in the first experiment. Thus it can be seen that NRSTU is more effective for dealing with the problem of preference mining.

8. Conclusion and Future Work

In this study, we proposed the parametric neighborhood rough set model on two universes for dealing the problem of preference mining. Firstly, baseline evaluation is investigated for dividing the positive and negative mapping. Furthermore,

(a) Confidence degree ω
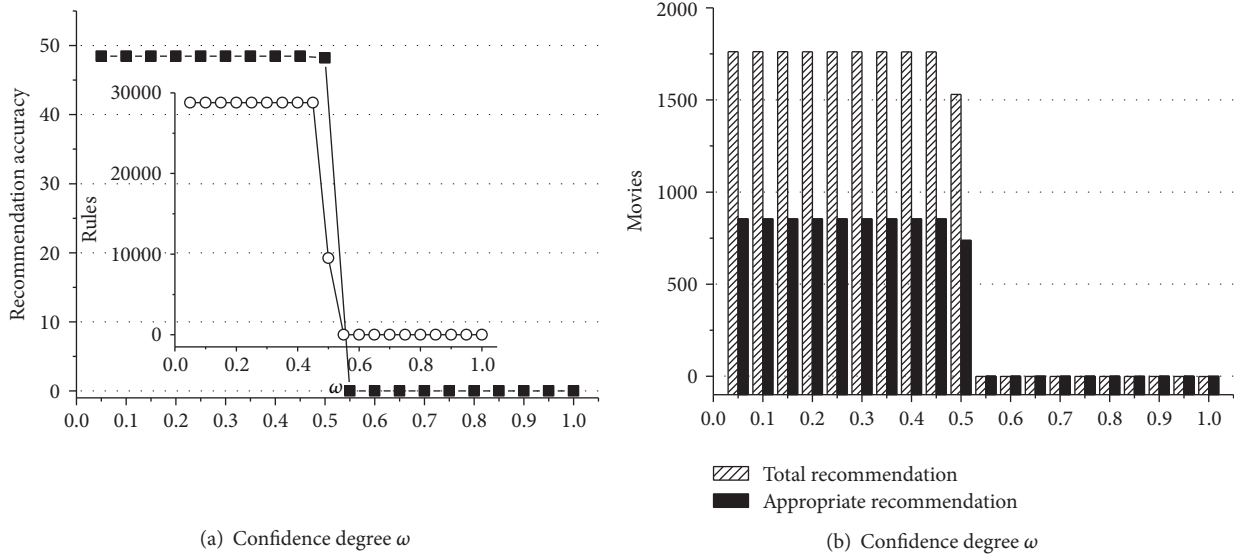

(b) Confidence degree ω

FIGURE 6: The results of recommendation where the support degree is $\eta = \gamma = 0.12$. (a) Recommendation accuracy (%) and amount of the rules with the confidence degree ω; (b) total and appropriate recommendations with the confidence degree ω.

TABLE 5: The comparison between NRSTU and RSTU.

	NRSTU	RSTU
Rules	97508.6	14260.2
Total recommendation	3804.7	854.1
Appropriate recommendation	1744.6	223.6
Recommendation accuracy	45.85%	26.17%

the definition of neighborhood lower approximation is proposed to define the user preference rules. The algorithms about preference rules mining and item recommendation are also given. Experiment 1 elaborates how NRSTU can overcome the cold-start problem. It simultaneously shows that baseline evaluation can avoid some unreasonable recommendations. The parameters of NRSTU are discussed in detail in Experiment 2. The last experiment shows that NRSTU improves the recommendation accuracy by about 19%. It can be concluded that NRSTU is more effective for dealing with the problem of preference mining.

The future work could move along two directions. First, many other rough set models also can be used to describe the user preference rules, such as kernel rough set and fuzzy rough set. Comparative analysis about the effectiveness of these rough set modes for preference mining is an important issue. Second, the application of our model for dealing with big data is necessary. Consequently, the version of NRSTU within distribute framework requires further attention.

Competing Interests

The author declared that there is no conflict of interests regarding the publication of this paper.

References

[1] Y. Koren, "Collaborative filtering with temporal dynamics," *Communications of the ACM*, vol. 53, no. 4, pp. 89–97, 2010.

[2] J. Wang, A. P. D. Vries, and M. J. Reinders, "Unifying user-based and item-based collaborative filtering approaches by similarity fusion," in *Proceedings of the 29th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval, (SIGIR '06)*, pp. 501–508, Seattle, Wash, USA, August 2006.

[3] Y. Koren, B. Robert, and C. Volinsky, "Matrix factorization techniques for recommender systems," *Computer*, vol. 42, no. 8, pp. 30–37, 2009.

[4] F. Min and W. Zhu, "Cold-start recommendation through granular association rules," https://arxiv.org/abs/1305.1372.

[5] P. Cremonesi, R. Turrin, and F. Airoldi, "Hybrid algorithms for recommending new items," in *Proceedings of the 2nd International Workshop on Information Hetero-Geneity and Fusion in Recommender Systems (HetRec '11)*, pp. 33–40, Chicago, Ill, USA, October 2011.

[6] A. I. Schein, A. Popescul, L. H. Ungar, and D. M. Pennock, "Methods and metrics for cold-start recommendations," in *Proceedings of the 25th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR '02)*, pp. 253–260, ACM, Tampere, Finland, August 2002.

[7] J. Bobadilla, F. Ortega, A. Hernando, and J. Bernal, "A collaborative filtering approach to mitigate the new user cold start problem," *Knowledge-Based Systems*, vol. 26, pp. 225–238, 2012.

[8] R. J. Mooney and L. Roy, "Content-based book recommending using learning for text categorization," in *Proceedings of the 5th ACM Conference on Digital Libraries (DL '00)*, pp. 195–204, ACM, 2000.

[9] M. J. Pazzani and D. Billsus, "Content-based recommendation systems," in *The Adaptive Web*, P. Brusilovsky, A. Kobsa, and W. Nejdl, Eds., vol. 4321 of *Lecture Notes in Computer Science*, pp. 325–341, Springer, Berlin, Germany, 2007.

[10] J.-H. Su, B.-W. Wang, C.-Y. Hsiao, and V. S. Tseng, "Personalized rough-set-based recommendation by integrating multiple contents and collaborative information," *Information Sciences*, vol. 180, no. 1, pp. 113–131, 2010.

[11] A. M. Rashid, I. Albert, D. Cosley et al., "Getting to know you: learning new user preferences in recommender systems," in *Proceedings of the International Conference on Intelligent Users Interfaces (IUI '02)*, pp. 127–134, San Francisco, Calif, USA, January 2002.

[12] A. M. Rashid, G. Karypis, and J. Riedl, "Learning preferences of new users in recommender systems: an information theoretic approach," *ACM SIGKDD Explorations Newsletter*, vol. 10, no. 2, pp. 90–100, 2008.

[13] P. du Boucher-Ryan and D. Bridge, "Collaborative recommending using formal concept analysis," *Knowledge-Based Systems*, vol. 19, no. 5, pp. 309–315, 2006.

[14] M. Braunhofer, "Hybridisation techniques for cold-starting context-aware recommender systems," in *Proceedings of the 8th ACM Conference on Recommender Systems (RecSys '14)*, pp. 405–408, ACM, Foster City, Calif, USA, October 2014.

[15] C. W.-K. Leung, S. C.-F. Chan, and F.-L. Chung, "An empirical study of a cross-level association rule mining approach to cold-start recommendations," *Knowledge-Based Systems*, vol. 21, no. 7, pp. 515–529, 2008.

[16] H. J. Ahn, "A new similarity measure for collaborative filtering to alleviate the new user cold-starting problem," *Information Sciences*, vol. 178, no. 1, pp. 37–51, 2008.

[17] S.-T. Park and W. Chu, "Pairwise preference regression for cold-start recommendation," in *Proceedings of the 3rd ACM Conference on Recommender Systems (RecSys '09)*, pp. 21–28, ACM, 2009.

[18] Y.-J. Park and A. Tuzhilin, "The long tail of recommender systems and how to leverage it," in *Proceedings of the 2nd ACM International Conference on Recommender Systems (RecSys '08)*, pp. 11–18, October 2008.

[19] I. Pilászy and D. Tikk, "Recommending new movies: even a few ratings are more valuable than metadata," in *Proceedings of the 3rd ACM Conference on Recommender Systems (RecSys '09)*, pp. 93–100, ACM, New York, NY, USA, October 2009.

[20] A. Elbadrawy and G. Karypis, *Feature-Based Similarity Models for Top-n Recommendation of New Items*, Department of Computer Science, University of Minnesota, Minneapolis, Minn, USA, 2013.

[21] D. Sun, Z. Luo, and F. Zhang, "A novel approach for collaborative filtering to alleviate the new item cold-start problem," in *Proceedings of the 11th International Symposium on Communications and Information Technologies (ISCIT '11)*, pp. 402–406, October 2011.

[22] F. Min and W. Zhu, "Granular association rule mining through parametric rough sets for cold start recommendation," https://arxiv.org/abs/1210.0065.

[23] X. He, F. Min, and W. Zhu, "Parametric rough sets with application to granular association rule mining," *Mathematical Problems in Engineering*, vol. 2013, Article ID 461363, 13 pages, 2013.

[24] H.-R. Zhang, F. Min, X. He, and Y.-Y. Xu, "A hybrid recommender system based on user-recommender interaction," *Mathematical Problems in Engineering*, vol. 2015, Article ID 145636, 11 pages, 2015.

[25] Z. Pawlak, "Rough sets," *International Journal of Computer and Information Sciences*, vol. 11, no. 5, pp. 341–356, 1982.

[26] Q. H. Hu, D. R. Yu, and Z. X. Xie, "Neighborhood classifiers," *Expert Systems with Applications*, vol. 34, no. 2, pp. 866–876, 2008.

[27] J. Tang, K. She, F. Min, and W. Zhu, "A matroidal approach to rough set theory," *Theoretical Computer Science*, vol. 471, pp. 1–11, 2013.

[28] S.-Y. Jing, "A hybrid genetic algorithm for feature subset selection in rough set theory," *Soft Computing*, vol. 18, no. 7, pp. 1373–1382, 2014.

[29] S. Y. Jing and K. She, "A rough sets approach to user preference modeling," in *Rough Set and Knowledge Technology: 5th International Conference, RSKT 2010, Beijing, China, October 15–17, 2010. Proceedings*, vol. 6401 of *Lecture Notes in Computer Science*, pp. 344–352, Springer, Berlin, Germany, 2010.

[30] T. Y. Lin, "Granular computing: from rough sets and neighborhood systems to information granulation and computing in words," in *Proceedings of the European Congress on Intelligent Techniques and Soft Computing*, pp. 1602–1606, Aachen, Germany, September 1997.

[31] T. Y. Lin, "Neighborhood systems and approximation in database and knowledge base systems," in *Proceedings of the 4th International Symposium on Methodologies of Intelligent Systems (ISMIS '89)*, pp. 75–86, October 1989.

[32] T. Y. Lin, *Neighborhood Systems: A Qualitative Theory for Fuzzy and Rough Sets*, University of California, Berkeley, Berkeley, Calif, USA, 2007.

[33] W. Zhu, "Relationship among basic concepts in covering-based rough sets," *Information Sciences*, vol. 179, no. 14, pp. 2478–2486, 2009.

[34] W. Zhu, "Relationship between generalized rough sets based on binary relation and covering," *Information Sciences*, vol. 179, no. 3, pp. 210–225, 2009.

[35] D. W. Pei and Z.-B. Xu, "Rough set models on two universes," *International Journal of General Systems*, vol. 33, no. 5, pp. 569–581, 2004.

[36] Y. Shen and F. Wang, "Variable precision rough set model over two universes and its properties," *Soft Computing*, vol. 15, no. 3, pp. 557–567, 2011.

[37] Z.-T. Gong and B.-Z. Sun, "Probability rough sets model between different universes and its applications," in *Proceedings of the 7th International Conference on Machine Learning and Cybernetics (ICMLC '08)*, pp. 561–565, July 2008.

[38] W. M. Ma and B. Sun, "Probabilistic rough set over two universes and rough entropy," *International Journal of Approximate Reasoning*, vol. 53, no. 4, pp. 608–619, 2012.

[39] F. Min and W. Zhu, "Granular association rule ing through parametric rough sets," *Brain Informatics*, vol. 7670, pp. 320–331, 2012.

[40] S. Jing, K. She, and S. Ali, "A Universal neighbourhood rough sets model for knowledge discovering from incomplete heterogeneous data," *Expert Systems*, vol. 30, no. 1, pp. 89–96, 2013.

[41] Y. Yao and X. Deng, "A granular computing paradigm for concept learning," *Smart Innovation, Systems and Technologies*, vol. 13, pp. 307–326, 2013.

[42] W. Ma and B. Sun, "Probabilistic rough set over two universes and rough entropy," *International Journal of Approximate Reasoning*, vol. 53, no. 4, pp. 608–619, 2012.

[43] Q. Hu, D. Yu, J. Liu, and C. Wu, "Neighborhood rough set based heterogeneous feature subset selection," *Information Sciences*, vol. 178, no. 18, pp. 3577–3594, 2008.

[44] GroupLens, http://movielens.umn.edu.

[45] Y. Koren, "Factorization meets the neighborhood: a multi-faceted collaborative filtering model," in *Proceedings of the 14th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD '08)*, pp. 426–434, Las Vegas, Nev, USA, August 2008.

[46] K. Zeng and K. She, "Variable precision neighborhood rough sets on two universes," in *Proceedings of the IEEE International Conference on Granular Computing (GrC '13)*, pp. 418–422, IEEE, Beijing, China, December 2013.

[47] R. Srikant and R. Agrawal, "Mining quantitative association rules in large relational tables," *SIGMOD Record (ACM Special Interest Group on Management of Data)*, vol. 25, no. 2, pp. 1–12, 1996.

[48] F. Min, http://grc.fjzs.edu.cn/~fmin/index.html.