F O U N D A T I O N S O F C O M P U T I N G A N D D E C I S I O N S C I E N C E S
Vol. 49 (2024) No. 4
ISSN 0867-6356
DOI: 10.2478/fcds-2024-0019 e-ISSN 2300-3405

Dynamic Interest Portrait of Community Users based on Three-way Fuzzy Concept Lattices

Wang Kai *

Abstract. Aiming at the multi-granularity knowledge discovery of users in online community, a method of dynamic user interest portrait based on Three-way Fuzzy Concept Lattice (3WFCL) is proposed to reveal the relationship between long-term and short-term interest drift, which improves the accuracy of the user portrait from a fuzzy perspective. Firstly, the label topic matrix of community users is constructed based on the topic analysis and the dynamic interest modelling is realized. Secondly, the environment migration index is calculated by extracting the topic and sentiment characteristics of the user environment. Meanwhile, interest inheritance is introduced into the calculation of interest characteristics, and interest attenuation vector is established to perceive the short-term interest label set dynamically. Finally, the membership degree of the three-way fuzzy interests is put forward into the multi-feature concept modelling of the long-term and short-term interest labels. Afterwards, the multi-granularity fuzzy association rules are generated to complete the multi-granularity characterization of differentiated interests by constructing the three fuzzy formal contexts. The effectiveness of the proposed method is validated on real data sets. The results show that it is of great significance to utilize the three-way fuzzy concept analysis into user portrait modelling, which is helpful to mine semantic association information and improve the analysis ability of fine-grained interest hierarchy.

Keywords: multi-granularity, interest inheritance, three-way neutrosophic fuzzy context, interest identification, portraits of interest.

1. Introduction

With the rapid development of social media, user-generated content has come into being gradually, when users socialize network resources according to their own knowledge context and interests. As a lightweight semantic fragment, social labels can reflect user interests, needs and emotional preferences. However, the characteristics of social labels such as fuzziness, complexity and weak correlation increase the difficulty of data mining greatly. In this context, interest-based user portrait method is introduced to realize the

* School of Health Management, Bengbu Medical University, 2600 Donghai Avenue, Bengbu City, Anhui Province, China.

feature abstraction by establishing interest label set, which has become the research focus of commercial marketing and personalized push. Therefore, it is of great practical significance to perceive and predict the recessive features of interest, by analysing the evolution law of dynamic interest and identifying the semantic correlation of long and short interest.

In recent years, the researches of user interest portrait are abundant, and the hotspots are concentrated in the fields of commercial information recommendation, scientific research user portrait, and public opinion analysis of government affairs. The core of user portrait is to build user interest feature set by collecting behaviour information such as browsing or attention, so as to realize short text description of user portrait. In the field of library and information, the main research results can be divided into two categories, namely user portraits based on explicit interest (UP-EI) and user portraits based on implicit interest (UP-II). UP-EI are mainly based on theories such as probability potential semantic analysis and text topic discovery to analyse the characteristics of user network behaviours and obtain semantic information of interests. For example, Xu H.L. [21] mined user behavioural attributes and demand characteristics based on concept lattice theory, which explored the internal correlation, and painted interest portraits of library group users. Liu X.X. [9] adopted personalized PageRank algorithm to identify twitter text content features and account social information, improving the interest mining accuracy of inactive Twitter users. UP-II mainly adopts the ontology or concept modelling theory, which utilizes the structural relationship associated in the concept system to dig the hidden theme and emotional features of the user text. For example, Li Yan [8] identified the user theme interests in multiple language environments based on the multilingual author theme model (JointAT), improving the generalization ability of the theme model. By establishing theme and emotion labels, Wang Shuai [16] subdivided user interests into multiple sub-attributes and realized differentiated emotional portraits of five community user groups. In addition, in the field of computer application, from the perspectives of dynamic interest modelling [19] and interest evolution [20], many scholars have adopted the theoretical model of multi-feature fusion to analyse the dynamic drift of user interest, and constructed multi-dimensional interest portraits using neural network [17], deep learning [11], double-layer Xgboost [4], and achieved fruitful results.

Although the above research realized the application exploration for specific fields to a certain extent, there is still room for improvement in the following aspects. Specifically, the lack of necessary semantic connections among user interest labels reduces the interpretability of interest portraits. In addition, the current methods have problems such as low situational sensitivity and weak situational perception of immediate interest, which cannot be applied to user personalized recommendation and interest community discovery. On the basis of previous studies, we focus on solving three problems. (1) How to use user-generated text label information to quantify the similarity between user topic and interest label. (2) How to recognize the evolution characteristics of long-term and short-term interest from the perspective of interest topic transfer and emotional change. (3) How to establish the short-term implicit interest set with granularity, and the application effect of the dynamic interest portrait in the open data set.

A Three-way Fuzzy Concept Lattice for Dynamic Interest Portrait (3WFCL-DIP) is presented. Specifically, at the theoretical level, the evolution characteristics of dynamic interest in time dimension are studied, and the dependency between interest scenarios and interest evolution is described by defining and calculating the environmental migration

index. On this basis, the significance and function of the above indexes in the classification of the long-term and short-term interests are clarified. In practice, the granulated representation of short-term interest based on three-way decision theory is realized by introducing the idea of interest inheritance. In addition, we explore the fuzzy correlation law of long-term and short-term interest, and verify the applicability and effectiveness of fuzzy interest discovery on the basis of building three-way fuzzy concept lattices for dynamic interest.

2. Related Work

2.1. Topic Model

Topic model is a generative probability model based on text content, which abstracts user text into multiple topics with their key words to realize the probability representation of text implicit information. From the perspective of application technology, topic models can be divided into probabilistic latent semantic analysis [3] (PLSA), latent Dirichlet analysis [2] (LDA) and author topic analysis [25] (ATM). PLSA utilizes probabilistic statistical model to construct the word bag space and establish the relationship between semantic dimension and document vocabulary. However, PLSA is unable to obtain probabilistic information at the document level, unable to solve the problem of overfitting the model. LDA is an unsupervised probabilistic topic model, which establishes document-topic, topic-feature word probability matrix by searching potential topics in text. Based on the classic topic model, the author information is introduced into the ATM model, and the author-topic probability distribution and topic-word probability distribution are generated through text modelling. In addition, due to the excellent generalization ability of LDA, many scholars proposed OLDA (Online LDA) [24], BTM (Biterm Topic Model) [22], which enriched the application field of topic model.

The core of 3WFCL-DIP is to build the label topic clustering for user dynamic interests. Considering that user labels are lightweight semantic units, BTM topic model is used to cluster short texts with label level. The topic probability distribution is constructed in the form of two-word pin-topic, and the topic division of user labels is carried out. Specifically, the indexes of interest popularity and interest stability are established based on the combination of annotated frequency and time factor. Finally, the dynamic interest model is built from the two measures of interest activity and change rate.

2.2. Three-way Fuzzy Concepts

Formal Concept Analysis [1] (FCA) is a mathematical modelling method based on the deterministic mapping relationship between objects and attributes. FCA realizes the hierarchical representation of concept specialization and generalization by establishing a concept poset with lattice structure. However, FCA relies on the binary relationship between objects and attributes completely, which makes the application scenario limited greatly. To this end, some scholars combined uncertainty theories such as Fuzzy and Rough

sets with FCA theory, and proposed generalized concept analysis methods such as Fuzzy concept lattice and Rough concept lattice, which improved the concept analysis ability of fuzzy data.

However, in the real scene, the acceptance or rejection decision of both deterministic decision model and uncertain decision model (based on fuzzy membership degree and attribute importance degree, etc.) is likely to pay a large decision cost. Under the above context, delayed decision making is not only beneficial to reduce decision cost and risk, but also to gain valuable decision time for in-depth research on problem solving methods and information processing modes. Yao [23] proposed a three-way decision theory model, which subdivides the knowledge domain into the accept-denial-delay domain, thus adding a non-commitment option to the two-way decision, which enriches the research perspective of uncertainty theory. Wei Ling et al. [18] combined the three-way decision theory with formal concept analysis, and derived the three-way concepts by defining objects and attributes, which quantified the fuzzy dependency relationship between objects and attributes.

Inspired by the above theory, a Three-way Fuzzy Concept Analysis based on neutrosophic sets is proposed, which realizes the dynamic interest recognition of community users. Aiming at the diversity and semantic fuzziness of user labels, the membership degree of three-way interest is defined. To enact this need, an intuitive fuzzy poset among short-term interests in the sense of granularity is established based on the hierarchical network structure of user interests to realize the online identification of user instant interest features.

3. Research Methods

3.1. Research Framework

The construction of user dynamic interest portraits mainly includes three steps, namely dynamic interest modelling, long- term and short-term interest classification, and three-way interest portraits. The main process is shown in Figure 1.

1) Dynamic interest modelling. First, the text generated by users in the network community is annotated. And the topic annotation of text labels is realized based on the topic model. Secondly, the label interest is calculated to complete the interest quantification of text labels combined with labelling frequency and time factor. Finally, the cosine similarity between the user topic and the label interest is calculated to obtain the user dynamic interest set.

2) Classification of long-term and short-term interests. The environmental transfer index is defined to quantify the degree of environmental change. On this basis, the idea of interest genetic interruption is introduced, and the long and short term interest classification is realized by calculating the interest decay vector.

3) Three-way interest portraits. Three-way decision making and formal concept analysis are introduced into the dynamic evolution of interests. By defining three-way interest memberships, three-way fuzzy interest concept lattice are constructed and fuzzy posets among short-term interests are established in the sense of granularity.

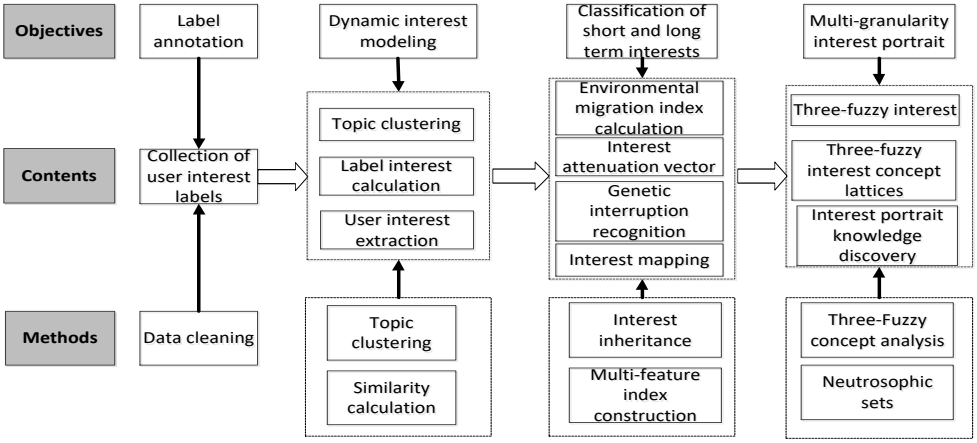

Figure 1. Multi-granularity user interest profile construction process

3.2. Dynamic interest modelling

In order to improve the accuracy and timeliness of user interest recognition, the user label interest weight is quantified in the form of label interest degree based on the evolution method of user interest level proposed in literature [20]. It is mainly divided into three steps. Firstly, the collection of user social labels is obtained based on the BTM topic model. Secondly, two secondary indexes of interest heat and interest stability are established by calculating the label interest. Finally, the similarity between label topics with interests is calculated to obtain user interest set.

The user label matrix $M_u(u,l)$ is constructed, and the matrix element is the frequency of label l_i by user u. Afterwards, the topic words are regarded as interest labels based on BTM, and the feature words under different topics are extracted with the co-occurrence word pairs using Python. The document-topic probability distribution and topic-feature word probability distribution are obtained to construct the label-topic probability distribution matrix $M_t(l,topic)$, the matrix element of which is the probability that label l_i belongs to topic t_j. Besides, the long-term interest tends to change slowly over time, which is affected by the frequency of annotation. By contrast, short-term interest fluctuates sharply in time dimension, with the characteristics of interest forgetting or interest strengthening. Therefore, the frequency and time factor of behavior labeling are introduced into label interest calculation by synthesizing interest popularity and interest stability to obtain the label -interest matrix $M_{int}(l,u)$. The definitions of interest heat $I_{\text{int}}^{t_i}$ and interest stability $I_{sta}^{t_i}$ are shown in formula 1 and 2 respectively. In formula 1, $f(u,l_i)$ is the frequency of labeling l_i by user u. $\sum_{i=1}^{n} f(u,l_i)$ is the total labeling frequency of all labels for user u. In formula 2, $I_{sta}^{t_i}$ is the interest stability of user u at t_i time. When the label l_k does not appear at t_j time, it indicates that the interest of label l_k attenuates. At this time, the decay of label interest is described based on the forgetting function [26], where k is the forgetting rate

assigned the value of *ln* 2/10. When the label l_k appears again at t_j time, it indicates that the interest of the label l_k is strengthened. $\Delta I_{sta}^{t_i}$ is the increment of the interest of the label l_k at t_i time, and the value is the labeling frequency increased by the new labeling behavior of the label l_k at this time.

$$I_{\text{int}}^{t_i} = \frac{f(u,l_i)}{\sum_{i=1}^{n} f(u,l_i)} \tag{1}$$

$$I_{sta}^{t_i} = \begin{cases} I_{sta}^{t_j} \cdot e^{-k(t_i-t_j)} t_j \,(l_k \text{ dose not appear at } t_j) \\ \Delta I_{sta}^{t_i} + I_{sta}^{t_j} \cdot e^{-k(t_i-t_j)} \,(_k \text{ appears at } t_j) \end{cases} \tag{2}$$

The quantification of user interest drift features in time dimension is realized by synthesizing formula 1 and 2. The label interest $W(u,l_j)$ is calculated as shown in formula 3, where θ is the interest weight factor. The larger the value of θ is, the greater the influence of the annotation frequency on the interest degree tends to be. On the contrary, it shows that time factor has a higher influence on interest degree.

$$W(u,l_j) = \theta I_{\text{int}}^{t_i} + (1-\theta) I_{sta}^{t_i} \tag{3}$$

The label-interest matrix $M_{int}(l, u)$ and label-topic matrix $M_t(l, topic)$ are carried out to obtain the label-interest matrix $M^t_{int}(u, topic)$, and the matrix element represents user interest in the topic. Finally, the cosine similarity is used to calculate the interest similarity between *u* and *topic*. When the similarity is greater than the interest threshold, *a* is added to its interest set to complete the user dynamic interest modeling.

3.3. Classification of long-term and short-term interests

The core of dynamic user interest recognition is to calculate the user text label interest in time dimension. Since user interest is dynamic and fuzzy, the label interest will also change dynamically with the migration of user interest, which is manifested in its differential distribution in the life cycle. However, the label interest of the long-term interest has inertia in the distribution of eigenvalues [27], which means that the changes in the personalized needs are not immediately affecting their own long-term label interest. For example, at the moment *t*, the user is affected by the surrounding environmental factors and loses focus temporarily, resulting in a sudden decline in the intensity of interest. However, its long-term label interest as a cyclical concept will not decrease immediately. On the contrary, some high-intensity short-term interests generated by external hot spots or their own needs will not evolve directly into stable long-term interests because of the sudden increase in their interest degree, and the inertial characteristics of this label interest degree in the time dimension is called interest inheritance. In the absence of strong stimulation from the external environment, the long-term interest of users is time-smooth. And the label interest at time *t* is affected by time *t*-1. Because of the relatively short life cycle of short-term interest, it is easy to be affected by changes in the external environment and the interest will decline. Therefore, when it is disturbed by changes in the context environment, it rarely has

the characteristics of interest inheritance on some continuous time snapshots, which is called the interest inheritance interruption.

Based on the analysis above, the change degree of context environment is quantified by defining environmental migration index. When the growth rate of environmental migration index is greater than the environmental threshold λ, it indicates that environmental mutation occurs at this time. In than case, the interest decay vector under the environment mutation window is calculated, from which *topK* interest labels are screened and the short-term interest sets of users are added. In the process of interest inheritance, the growth rate of environmental migration index is an important quantitative index to distinguish the long-term and short-term interest of users. The positive growth rate indicates the expansion of the overall interest of users, which means the interests become more extensive and the interest heat rises. The negative growth rate indicates that the overall interest shrinks, and the interest becomes more concentrated and the interest heat decreases. When the heat of interest declines, short-term interest is easily disturbed by exogenous factors. Therefore, the classification of the long and short term interests is realized by capturing the above characteristics of environmental changes. Assuming that the environmental migration index at time t is I_t and the exponential increment is ΔI_t, the environmental migration index growth rate μ_t at time t is $\Delta I_t / I_t$.

3.3.1. Construction of environment migration index

Social label is an important manifestation of the behaviour labelling. The interpretability of the interest behaviour cycle can be improved by analysing the topic semantics and emotion fluctuations, so as to provide the necessary data support for the accurate identification of long-term and short-term interest. User-generated texts, as an information carrier reflecting user behaviour characteristics, contain a variety of topic characteristics and emotional characteristics. As a quantitative index to monitor the change of user interest, the environmental migration index can perceive the interest inflection point dynamically and reflect the expansion or contraction of the overall interest. Therefore, the environment transfer index includes the dynamic change of text content and the transfer change of text emotion, which can be subdivided into two secondary indexes: topic transfer index and emotion transfer index.

The TF-IDF algorithm is utilized to calculate the weight of topic feature words at t-1 and t moments, and the word frequency matrix of adjacent moments is obtained, which is viewed as the change quantity of topic feature words at t moments. The topic migration index is shown in formula 4.

$$I_t = \frac{W_{s_1}^{t_1} - W_{s_2}^{t_2}}{t_1 - t_2} \tag{4}$$

In formula 4, $W_{s_1}^{t_1}$ and $W_{s_2}^{t_2}$ represent the topic weights of interest labels at time t_1 and t_2 respectively shown in formula 5. In formula 5, $tfidf_{ij}$ represents the value of the j_{th} feature word contained in topic I at time t. s is the number of topics and p is the number of feature words contained in the topic. s_1 and s_2 are the number of topics contained in each context.

$$W_s^t = \sum_{j=1}^{s} \sum_{i=1}^{p} tfidf_{ij} \quad (5)$$

The sentiment shift index identifies changes on topics of interest to the user. Based on WordNet sentiment dictionary [7], sentiment feature words under different interest labels are identified, and positive and negative sentiment word sets are constructed respectively. Afterwards, negative words and degree adverbs are extracted to determine the degree of sentiment modification [15]. The sentiment value of the interest label is determined by the comprehensive sentiment words it contains. Through linear summation of the sentiment weights of different categories, the sentiment value of different interest labels at time t is determined shown in formula 6.

$$S_t(w) = \sum_{i=1}^{k} \frac{S_t(w_i)}{|w|} \quad (6)$$

In formula 6, $S_t(w_i)$ is the sentiment intensity of the i_{th} interest label of the user at time t (formula 7). k represents the total number of sentiment feature words with the same emotional tendency contained in an interest label, and $|w|$ represents the total number of interest labels of the user. In formula 7, pw_i and nw_i are the positive and negative sentiment words respectively. M and N represent the total number of sentiment words respectively. The sentiment shift index (formula 8) can be obtained by combining formula 6 and 7. $S_{t_1}(w)$ and $S_{t_1}(w)$ represent the sentiment weight of interest labels at t_1 and t_2 respectively.

$$S_t(w_i) = \frac{\sum_{i=1}^{M} sim(w, pw_i)}{M} - \frac{\sum_{j=1}^{N} sim(w, nw_i)}{N} \quad (7)$$

$$I_s = \frac{S_{t_1}(w) - S_{t_1}(w)}{t_1 - t_2} \quad (8)$$

As a quantitative index of dynamic perception of the interest migration, environment migration index can provide important topic and emotional information for quantifying the difference of the long-term and short-term interests. The environmental shift index I_e (formula 9) is obtained by integrating topic shift index and sentiment shift index. In formula 9, ε is the shift weight factor. On the contrary, it shows that the sentiment intensity of interest label has less influence on the change of interest.

$$I_e = \varepsilon I_t + (1-\varepsilon) I_s \quad (9)$$

3.3.2. Calculation of interest attenuation vector

By introducing interest inheritance into the calculation of interest features, the interest drift path is drawn using the environmental migration index. The mutation of interest environment is perceived on condition that the growth rate of environmental migration index is greater than the environmental threshold λ. Meanwhile, the interest decay vector at the same time is calculated to capture the short-term interest label set dynamically, where

the interest inheritance is interrupted with the change of environment. For the sake of explanation, this paper describes the time interval between the recognition of users' short-term interests from t_0 when users' interests are differentiated in the long and short term to t. During the period of interest window, the attenuation mark of interest degree cannot be detected because the attenuation of environmental migration index has not reached the set threshold, which makes it difficult to perceive the real starting point of short-term interest accurately. To enact this need, the latency of short-term interest recognition is reduced by extending the sampling time range of interest attenuation vector. To be specific, the first n time intervals equal to t_1-t_2 are selected to calculate the decay rate of label interest associated with their interest (topic), assuming that the growth rate of environmental migration index at t_1 and t_2 is greater than the environmental threshold. In that case, the interest decay feature matrix is established through normalization processing. The feature term represents the average decay rate of interest degree of interest label l_j at time t_i.

According to the interest decay rate of labels, the interest label set can be mapped to the long-term and short-term interest set. The specific process is as follows. First, the label set with the top N interest decay rate is selected and mapped to the short-term interest set $S_t\{topic_1, topic_2, \ldots, topic_s\}$. The remaining labels are classified as long-term interest sets $L_t\{topic_1, topic_2, \ldots, topic_r\}$. Secondly, interest labels with high interest decay rate are repeated in multiple time intervals. Considering that the interest labels still have the characteristic of periodic change of interest inheritance, they are classified as long-term interest sets.

3.4. Multi-granularity interest portrait

Short-term interest is the temporary focus of attention shown by users in a specific context, which is susceptible to interference and gradually evolves into two stable states, namely either disappearing with the end of its life cycle or transforming into a relatively stable medium and long-term interest. In view of the complexity and uncertainty of short-term interest evolution, the interest genetic interruption is regarded as the basis of short-term interest recognition. In that case, three-way interest membership degrees are defined, to establish intuitionistic fuzzy posets among short-term interests in the sense of granularity. To be specific, the users are mapped to a set of objects and the topics contained in their collection of short-term interests are mapped to a set of attributes. Afterwards, the three-way fuzzy context is established by assigning three-way interest membership degrees. In addition, three-way fuzzy concept lattice is constructed combined with the interest confidence threshold, and multi-granularity association knowledge is induced through the lattice operation, which realizes the fuzzy association representation of interest labels to complete the knowledge extraction of user interest in dynamic scenarios.

3.4.1. Three-way fuzzy representations of interest labels

In order to quantify the fuzzy dependency of different interest levels, the definition of three-way interest membership degrees is given inspired by literature [6], and the binary fuzzy relationship between users and their interest labels is extended to the three-way hesitate neutrosophic sets to characterize the granular information of the interests.

Definition 1. If $K = (U, A, R)$ is the three-way neutrosophic fuzzy context, where U and A are the user object set and the attribute label set respectively. R is a three-way fuzzy

relation on $U \times A$, on condition that there are three mappings expressed as formula 10-12. The constraint conditions shown in formula 13 are satisfied. On this condition $\mu_T(u,a)$, $\mu_I(u,a)$, $\mu_F(u,a)$ are the membership degree, hesitant membership degree and non-membership degree belonging to the three-way fuzzy relation on (u,a) respectively.

$$\mu_T : U \times A \to]0^-, 1^+[\quad (10)$$

$$\mu_I : U \times A \to]0^-, 1^+[\quad (11)$$

$$\mu_F : U \times A \to]0^-, 1^+[\quad (12)$$

$$0^- \le \mu_T(u,a) + \mu_I(u,a) + \mu_F(u,a) \le 3^+ \quad (13)$$

Liu Rui et al. [10] analyzed that short-term interest has the characteristics of low interest rating and large change of attention frequency within a certain time range, verifying the low-frequency change trend of the numerical intensity of long-term interest with small fluctuations. Based on this, the three-way fuzzy relation sets are regarded as the core of the fuzzy context. By analyzing the characteristics of the labeling behavior in the time dimension, the characteristics of the long-term and short-term interests are summarized as follows. Firstly, the long-term interest of the user has the characteristics of interest inheritance in its life cycle, and always retains a certain degree of interest in the long-term interest label. Meanwhile, the short-term interest has a greater volatility, and there is the possibility of converting into long-term interest. Based on the analysis above, it is assumed that $topic_i$ is the short-term interest of user u, and the assignment process of the membership degree of the three-way interests is as follows. Firstly, the membership degree $\mu_T(u,a)$ is the possibility that attribute a ($topic_i$) is transformed into long-term interest, so the membership degree $\mu_T(u,a)$ from the long-term interest set of $L_t\{topic_1, topic_2, \ldots, topic_r\}$, is selected successively to obtain the long-term interest similarity set $D_l = \{topic_1, topic_2, \ldots, topic_p\}$. Afterwards, the membership degree is the mean value of label interest (formula 14), where p is the number of topics contained in long-term interest similarity set D_l, and m is the number of labels contained in the topic set. $N(m)$ and $N(m)$ is the maximum number of labels contained respectively. Secondly, the hesitancy membership $\mu_I(u,a)$ is the mean value of label interest in $topic_i$ (formula 15), where n is the number of labels in $topic_i$. Besides, the degree of non-membership $\mu_F(u,a)$ is the mean value of label interest contained in the topic in the long-term interest dissimilar set $D - D_l = \{t_1, t_2, \ldots, t_q\}$.

$$\mu_T(u,a) = \frac{\sum_{i=1}^{p}\sum_{j=1}^{m} w(u,l_j)}{N(m)} \quad (14)$$

$$\mu_I(u,a) = \frac{\sum_{j=1}^{n} w(u,l_j)}{N(n)} \quad (15)$$

In this paper, on the basis of the fuzzy concept lattice proposed in literature [14], three-way interest membership degrees are mapped to the hesitate neutrosophic set on the discourse domain. The dynamic evolution characteristics of short-term interest are described from the perspective of three-way concept lattice, among which three-way fuzzy concepts and three-way fuzzy concept lattice are shown in Definitions 2 and 3.

Definition 2. If $K=(U,A,R)$ is the three-way neutrosophic fuzzy context on the confidence threshold (α , β , $\gamma \in [0,1]$). For any $X \subseteq U$, $Y \subseteq A$, there are three-way fuzzy operators shown in formula 16 and 17, with the constraints of formula 18 and 19, where $\mu_T^R(u,a)$, $\mu_I^R(u,a)$ and $\mu_F^R(u,a)$ represent the membership degree, hesitant membership degree and non-membership degree respectively. The binary group (X,Y) is called a three-way fuzzy concept on *K*, where $X^{\uparrow}$ indicates the object set whose membership degree is not less than the confidence threshold (α,β,γ). Meanwhile, $Y^{\downarrow}$ indicates the object set whose membership degree is also not less than the confidence threshold (α,β,γ).

$$X^{\uparrow}=\{a\in A\,|\,\forall u\in X,\mu_T^{x_i}(u,a)\geq\alpha,\mu_I^{x_i}(u,a)\leq\beta,\mu_F^{x_i}(u,a)\leq\gamma\} \text{ and } X^{\uparrow}=\{Y,(\mu_T^{x_i}(u,a),\mu_I^{x_i}(u,a),\mu_F^{x_i}(u,a))\} \tag{16}$$

$$Y^{\downarrow}=\{u\in U\,|\,\forall a\in Y,\mu_T^{y_j}(u,a)\geq\alpha,\mu_I^{y_j}(u,a)\leq\beta,\mu_F^{y_j}(u,a)\leq\gamma\} \text{ and}$$

$$Y^{\downarrow}=\{X,(\mu_T^{y_j}(u,a),\mu_I^{y_j}(u,a),\mu_F^{y_j}(u,a))\} \tag{17}$$

$$\begin{cases}\mu_T^{x_i}(u,a)=\min_{y_j\in A}\mu_T^R(u,a)\\ \mu_I^{x_i}(u,a)=\max_{y_j\in A}\mu_I^R(u,a)\\ \mu_F^{x_i}(u,a)=\max_{y_j\in A}\mu_F^R(u,a)\end{cases} \tag{18}$$

$$\begin{cases}\mu_T^{y_j}(u,a),=\min_{x_i\in U}\mu_T^R(u,a)\\ \mu_I^{y_j}(u,a)=\max_{x_i\in U}\mu_I^R(u,a)\\ \mu_F^{y_j}(u,a)=\max_{x_i\in U}\mu_F^R(u,a)\end{cases} \tag{19}$$

It should be noted that in definition 2, $\mu_T^X(u,a)$, $\mu_I^X(u,a)$, $\mu_F^X(u,a)$ represents the membership, hesitant membership, and non-membership of the definition derived from the object *X* respectively, whose superindices are intended to distinguish it from the membership derived from attribute *Y*. In application scenarios, the number of three-way fuzzy concepts can be adjusted by the confidence threshold to achieve the purpose of controlling the granularity level.

Definition 3. If $C_1=(X_1,Y_1)$ and $C_2=(X_2,Y_2)$ are two three-way fuzzy concepts on $K=(U,A,R)$, C_1 is called a subconcept of C_2 if it satisfies the following constraints shown in formula 20 - 23. It is easy to prove that if the three-way fuzzy concept sets

derived from K are represented $CF(K)$, then the relation set $\leq$ is a set of partial order relations on $CF(K)$, and $(CF(K),\leq)$ constitutes a complete lattice CL^F which is called the three-way fuzzy concept lattice represented as $(CF(K),\wedge,\vee,(1,0,0),(0,1,1))$.

$$(X_1,Y_1)\leq(X_2,Y_2)\Leftrightarrow X_1\subseteq X_2(\Leftrightarrow Y_2\subseteq Y_1) \tag{20}$$

$$\mu_T^X(X_1,Y_1)\leq\mu_T^X(X_2,Y_2) \tag{21}$$

$$\mu_r^X(X_1,Y_1)\geq\mu_r^X(X_2,Y_2) \tag{22}$$

$$\mu_c^X(X_1,Y_1)\geq\mu_c^X(X_2,Y_2) \tag{23}$$

3.4.2. Interest discovery based on Three-way fuzzy concepts

At the theoretical level, the evolutionary characteristics of dynamic interest in time dimension are analysed, and the correlation between environmental migration index in interest scene characterization and interest inheritance is clarified. Besides, fuzzy recognition of long and short term interest is realized. At the practical level, the practical application direction of personalized interest discovery is explored and verified by real data from online communities. In general, the core contributions are as follows. Firstly, the intensity of user interest in the dimension of label topic is quantified, and the correlation degree between the topic and label interest is analysed. Secondly, the interest inheritance is introduced into the label interest feature calculation to perceive the short-term interest label set dynamically. Finally, the multi-granularity characterization of dynamic interest is completed by defining from the perspective of fuzzy interest.

In order to explain the utilize of three-way fuzzy concept lattice to identify the dynamic interest in detail, more specific research steps are given to facilitate readers to grasp the technical process of the model. The key process of community user interest discovery based on three-way concept lattice and its application examples are expressed in the following steps.

The label topic matrix of community users is constructed based on topic analysis. Specifically, it collects data about user behaviour in the community, such as browsing history, likes, comments, etc., as well as the profiles and labels that users fill in. The label interest is calculated by formula 3. Secondly, formula 9 is used to calculate the environment migration index by extracting the topic and sentiment characteristics of the user environment. Meanwhile, the interest inheritance is introduced into the interest feature calculation, and the interest decay vector under different time windows is calculated according to the interest decay rate of the label, while the user interest label set is mapped to the long and short interest set.

Three-way interest membership degrees are introduced into the multi-feature concept modelling of user short-term interest labels from the perspective of fuzzy interest. By constructing three-way fuzzy contexts, multi-granularity association knowledge is generated. Specifically, definition 1 is utilized to map users to a set of objects and topics contained in their collection of short-term interests to a set of attributes. On this basis, three-way interest membership degrees are obtained by using formula 10 and 11. In addition, three-way fuzzy context are established, and combined with the interest confidence threshold, three-way fuzzy concepts can be calculated to represent the user

long-term interest and short-term interest, based on definition 2. Finally, three-way fuzzy concept lattices are constructed to induce fuzzy associations based on interest labels to complete knowledge extraction of user interests based on definition 3.

4. Experiment

4.1. Sources of experimental data

In order to verify the validity of 3WFCL-DIP on different data sets, the experimental design is divided into two parts. The first part is to select the user information of *Tumblr* (https://github.com/tumblr) for user interest identification. *Tumblr* is an open site that combines blogging and social networking. The *Stuff picks official blog* (SPOB) is selected with more user participation by using the *octopus collector* to climb the posting data of website users for a natural year since April 10, 2021. In that case, a total of 2,063,548 valid comments are obtained, involving 103,653 users. After data cleaning such as text segmentation and removal of stop words, two data sets of *userid* and *labname* are constructed. In the second part, the open source data set *MovieLens* is selected as the research object, and the accuracy and rationality of 3WFCL-DIP is verified by comparison and analysis with the classical algorithms.

4.2. Parameter Analysis

The dependency between interest scene and interest evolution is described by analyzing the evolution characteristics of dynamic interest in time dimension. Besides, the fuzzy correlation information in dynamic interest evolution is mined by constructing three-way fuzzy concept lattices oriented to dynamic interests. The above calculation process involves the discussion of several threshold parameters. In order to facilitate the explanation, the parameters are divided into three categories according to the nature of the parameters, namely weight parameters, similarity parameters and threshold parameters. The relevant descriptions are shown in Table 1.

Table 1. Parameter classification description

Classification	Parameter	Description
weight parameter	θ	Adjust the weight of interest intensity and interest stability
	ε	Adjust the weight of topic shift index and sentiment shift index
similarity parameter	a	Determine if the topic is of user interest
	σ	Determine if there are thematic similarities between the topics of long-term interest
threshold parameter	λ	Determine if the environmental migration index has environmental abrupt change
	(α, β, γ)	Adjust the generation scale of three-way fuzzy concepts

The weight parameters are mainly to adjust the proportion of different components in the corresponding index. In the absence of environmental mutation, both long-term and short-term interests have time smoothness in their life cycle. The interest at time t is

affected by the characteristic value at time *t*-1. Similarly, topic migration and sentiment change also maintain time inertia [5], showing a relatively strong correlation under adjacent time windows. To this end, the optimal parameter assignment is determined by calculating the correlation coefficient of the degree of interest between *t*-1 and *t*. Taking the parameter θ as an example, different parameter values are selected in formula 3, and the calculation results are shown in Figure 2. As can be seen from Figure 2, when θ=0.45, the correlation coefficient reaches the maximum, and the assignment for adjusting interest intensity and interest stability is the best. According to the similar method, the optimal weight for ε is determined.

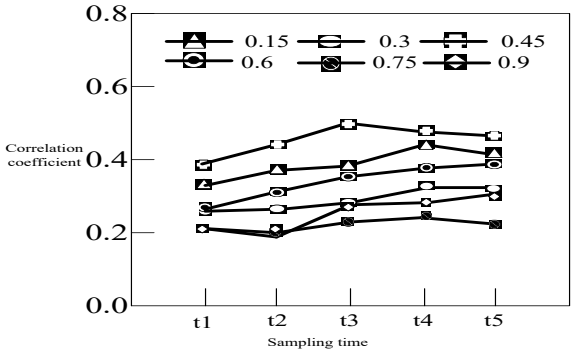

Figure 2. Correlation analysis of interest weight factors

The similarity parameter is mainly to cluster the interest in the label (or the similarity between topics). The accuracy rate of the interests (or similar topics) on the prediction label is divided by comparing different parameters. Taking the threshold *a* as an example, the calculation is shown in formula 24, where *Acc* is the ratio of the frequency of the labels contained in the topics marked as their interest to the frequency of their total label (the prediction accuracy rate). N_n is the total number of times the labels contained in the topic marked as interest at *t*+1 time. N_{all} indicates the total number of times that users annotate labels at *t*+1. The experimental results are shown in Figure 3. When the interest threshold *a*=0.4, the prediction accuracy reaches the maximum, and the division of interest labels is the best at this time.

$$Acc = \frac{N_n}{N_{all}} \tag{24}$$

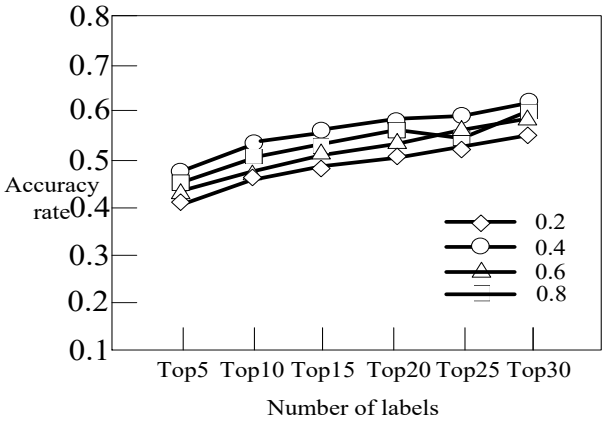

Figure 3. Prediction accuracy analysis of interest labels

Threshold parameters include the environment threshold λ and confidence threshold (α, β, γ). Considering that the environmental migration index mainly reflects the degree of contextual environment change of dynamic interest, and the environmental threshold is utilized to perceive the time window information of environmental sudden change. Therefore, it is assumed that environmental mutation is the direct cause of the interruption of interest inheritance, so different environmental thresholds are selected to calculate the user label interest at the time *t* of negative growth rate under the window of environmental mutation. The decline of label interests is regarded as the decline of interests, and the corresponding environmental threshold set $\{\lambda_1, \lambda_2, \ldots \lambda_k\}$ is screened out and its expectation is taken as the final environmental threshold. The confidence threshold reflects the graininess degree of the fuzzy relationship between the three-way fuzzy concepts, which controls the scale of concept generation and screening association rules. To this end, the experiment compared the number of three-way fuzzy concepts obtained from 27 groups with different values (from 0.3 to 0.9 with step size of 0.3), and the results were shown in figure 4. As can be seen from figure 4, when the hesitation membership threshold and non-membership threshold remain unchanged, the number of three-way fuzzy concepts decreases with the increase of membership threshold. Under the condition of constant membership threshold, the number of three-way fuzzy concepts increases gradually with the increase of hesitant membership threshold and non-membership threshold.

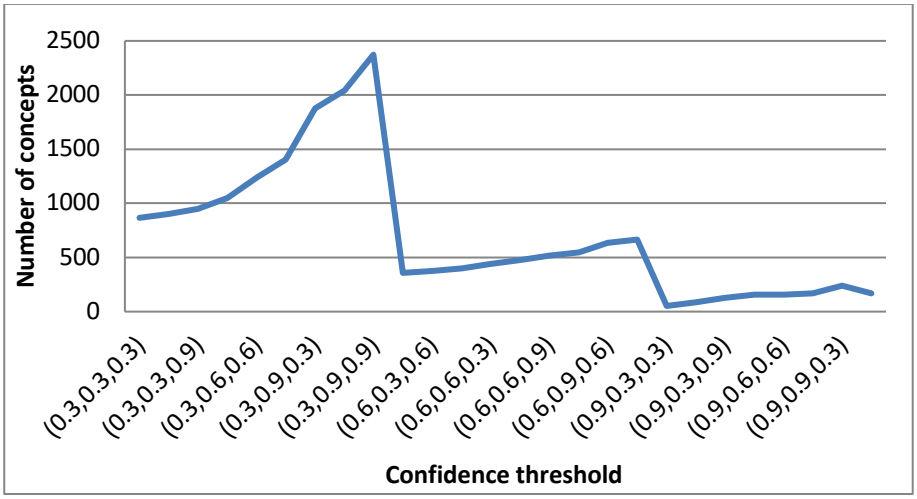

Figure 4. Comparison of three fuzzy concepts under different confidence thresholds

4.3. Analysis of experimental results

The *Gensim toolkit* based on Python realized the topic annotation of user text based on the BTM model. After 2000 iterations, the confusion degree is the lowest when the number of topic K=85. Partial clustering results are shown in Table 2.

Table 2. Results of community theme-word clustering

ID	Keywords with their probabilities	Content summary
Topic1	mixing/0.0625 banjo/0.0612 trance/0.0537 intro/0.0526 mooc/0.0510 drone/0.0501 standard/0.0455 may/0.0418	music
Topic2	right/0.0726 citizen/0.0703 power/0.0659 justice/0.0574 public/0.0536 country /0.0522 party/0.0508 rule/0.0475	politics
Topic3	guide/0.0863 track/0.0842 race/0.0751 court/0.0728 long/0.0663 fence/0.0639 diet/0.0631 flush/0.0612	sports
Topic4	cognition/0.0764 logic/0.0646 novel /0.0631 reality/0.0452 ideal/0.0546 youth/0.0529 biography/0.0512 history/0.0503	reading
Topic5	party/0.0465 food/0.0432 club/0.0432 traveling/0.0413 sex/0.0312 show/0.0366 pal/0.0440 community/0.0362	lifestyle

The user groups with high activities were selected from Table 2. When θ is 0.45, the label interests were calculated respectively with formula 3, and the results were shown in Table 3. The cosine similarity between the label interest and label topic value at different timing was calculated to extract interest topics shown in Table 4. The highest 20 similarity topics for interest labels were selected based on Tables 3 and 4, when the migration weight factor is 0.42. The results were divided into 12 subsets in the unit of months. Afterwards, the topic migration index and sentiment migration index of u_i under the continuous time snapshot were calculated respectively. The environmental migration index was calculated shown in Table 5. It can be seen that the growth rate of the environmental migration index

changes greatly. Specially, the exponential growth rate is -3.343 at windows 2 to 3, indicating that his interest may become more concentrated and his interest enthusiasm decreases. Besides, a total of 4 mutation windows of interest environment were screened when the environmental threshold is 0.125, and the time interval is determined to be 2 months. The decay rates of label interest associated with their topics were calculated respectively, and the average decay rates of label interest were obtained through normalization processing as shown in Table 6. Finally, the short-term interest set $S_t\{topic_{33},topic_{31},topic_{15},\ldots\}$ was obtained, where $topic_4$ is the enhanced short-term interest point of u_i.

Table 3. Label interest of *ui* at *ti*

ID	Interest intensity	Interest stability	Interest degree
l_{13}	0.068	0.123	0.098
l_{51}	0.041	0.116	0.082
l_{46}	0.053	0.053	0.053
l_{19}	0.061	0.035	0.047

Table 4. Cosine similarity between labels and topics for u_i

ID	Similarity	ID	Similarity
$topic_{38}$	0.768	$topic_{31}$	0.354
$topic_{25}$	0.529	$topic_4$	0.306
$topic_{16}$	0.475	$topic_{15}$	0.224
$topic_{22}$	0.396	$topic_{40}$	0.189

Table 5. Environment migration index of ui under multiple time windows

ID	Date span	Topic shift index	Sentiment shift index	Environment migration index	Growth rate
1	2021.4.10-5.10	0.018	0.158	0.099	-------
2	2021. 5.10-6.10	0.013	0.184	0.112	0.131
3	2021. 6.10-7.10	-0.014	-0.443	-0.263	-3.343
4	2021. 7.10-8.10	-0.016	0.092	0.047	-1.177
5	2021. 8.10-9.10	0.023	0.116	0.077	0.650
6	2021. 9.10-10.10	0.085	-0.067	-0.003	-1.041

Table 6. Average decay rate of interest of labels under the environment mutation window

ID	Interest ID	Interest decay rate	ID	Interest ID	Interest decay rate
	$topic_4$	0.352		$topic_{31}$	0.366
2	$topic_{15}$	0.347	7	$topic_7$	0.273
	……	……		……	……
	$topic_4$	0.414		$topic_{38}$	0.439
5	$topic_{22}$	0.268	8	$topic_4$	0.331
	……	……		……	……

Three-way fuzzy concept lattice for dynamic interest was constructed, which can be divided into three-way fuzzy context construction and three-way fuzzy concept lattice

construction. The specific process is described as follows. First, the target user was selected to build a finite set of objects $\{u_1,u_2,\ldots,u_m\}$. Secondly, the short-term interests with their labels in matrix $M_t(l,topic)$ were extracted to build a finite set of attributes $\{a_1,a_2,\ldots,a_n\}$ based on $S_t\{topic_1, topic_2,\ldots,topic_s\}$. Afterwards, membership degree, hesitant membership degree and non-membership degree on (u, a) were calculated according to formula 9 and 10 respectively, and three-way fuzzy relation sets were constructed to complete the fuzzy dependence relationship labeling between user labeling behavior and short-term interest shown in Table 7. As can be seen from definition 1, some users in Table 7 have strong short-term interest characteristics, which are reflected in their high degree of interest in labels of short-term interest. For example, u_4 has a high degree of hesitancy membership on multiple topics, which reflects the universality of the interests. In addition, the high membership degree or non-membership degree of some users indicates that their short-term interest has a high probability of evolving into long-term interest or dying out, reflecting the dynamic evolution of interests. For example, the membership degree to the topic a_3 is 0.73 for u_5, and the topic is more likely to be transformed into the long-term interest. The non-membership degree of u_5 for topic a_1 is 0.74, which reflects the high probability of extinction of this topic.

Table 7. Three-way fuzzy context of the short-term interest

	$a_1(topic_4)$	$a_2(topic_{21})$	$a_3(topic_{16})$	$a_4(topic_{35})$	$a_5(topic_{40})$
u_1	(0.46,0.60,0.36)	(0.34,0.28,0.55)	(0.41,0.37,0.69)	(0.42,0.39,0.34)	(0.52,0.41,0.67)
u_2	(0.55,0.37,0.41)	(0.46,0.61,0.36)	(0.59,0.32,0.42)	(0.43,0.59,0.57)	(0.41,0.68,0.56)
u_3	(0.62,0.46,0.38)	(0.52,0.47,0.58)	(0.64,0.62,0.39)	(0.55,0.64,0.49)	(0.53,0.74,0.75)
u_4	(0.53,0.57,0.72)	(0.36,0.60,0.39)	(0.52,0.78,0.41)	(0.38,0.82,0.62)	(0.46,0.80,0.35)
u_5	(0.37,0.65,0.74)	(0.30,0.60,0.33)	(0.73,0.64,0.54)	(0.63,0.47,0.67)	(0.57,0.76,0.48)

The common characteristics of user interest attributes were mined and the internal correlation of user personalized subdivision interest was analyzed. HNCL[6] progressive concept lattice construction algorithm was used to add three-way interest membership degree labels to the attributes of concept nodes, and three-way fuzzy operators were calculated on the three-way neutrosophic fuzzy context according to confidence thresholds (0.6,0.9,0.9). Finally, three-way fuzzy concept lattice with partial three-way fuzzy concepts satisfying the partial order relationship were obtained, as shown in Figure 5 and Table 8.

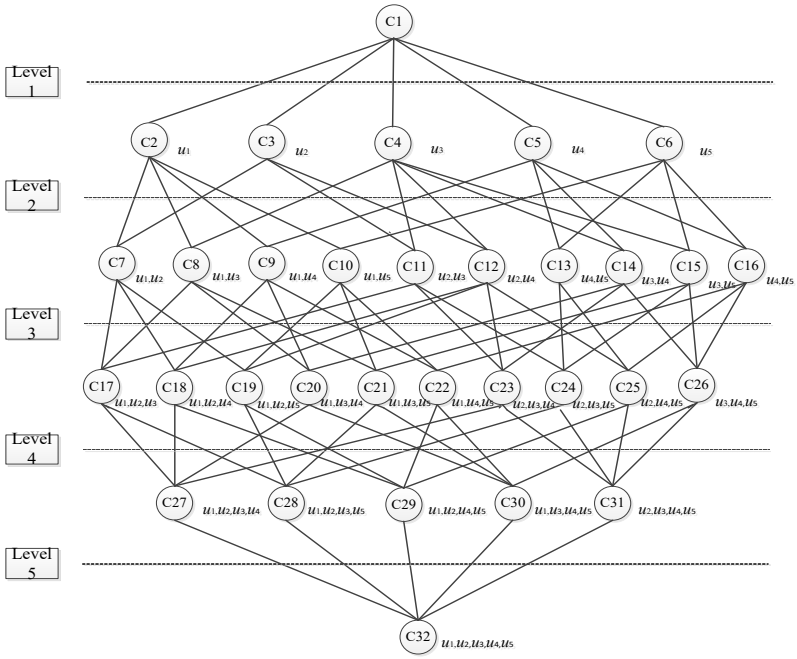

Figure 5. Three-way fuzzy concept lattice for user interests

Table 8. Three-way fuzzy concepts oriented to user interests

Hierarchy	No	Three-way fuzzy concepts with their membership degrees
L1	*C*1	{u_1(0.0,1.0,1.0)+u_2(0.0,1.0,1.0)+u_3(0.0,1.0,1.0)+u_4(0.0,1.0,1.0)+u_5(0.0,1.0,1.0); a_1(0.37,0.65,0.74)+a_2(0.30,0.61,0.58)+a_3(0.41,0.78,0.69)+a_4(0.38,0.82,0.67)+a_5(0.41,0.80,0.75)}
L2	*C*2	{u_1(1.0,0.0,0.0)+u_2(0.41,0.68,0.57)+u_3(0.52,0.74,0.75)+u_4(0.36,0.82,0.72)+u_5(0.30,0.76,0.74); a_1(0.46,0.60,0.36)+a_2(0.34,0.28,0.55)+a_3(0.41,0.37,0.69)+a_4(0.42,0.39,0.34)+a_5(0.52,0.41,0.67)}
	*C*3	{u_1(0.34,0.60,0.69)+u_2(1.0,0.0,0.0)+u_3(0.52,0.74,0.75)+u_4(0.36,0.82,0.72)+u_5(0.30,0.76,0.74); a_1(0.55,0.37,0.41)+a_2(0.46,0.61,0.36)+a_3(0.59,0.32,0.42)+a_4(0.43,0.59,0.57)+a_5(0.41,0.68,0.56)}
	…	…
L3	*C*7	{u_1(1.0,0.0,0.0)+u_2(1.0,0.0,0.0)+u_3(0.52,0.74,0.75)+u_4(0.36,0.82,0.72)+u_5(0.30,0.76,0.74); a_1(0.46,0.60,0.41)+a_2(0.34,0.61,0.55)+a_3(0.41,0.37,0.69)+a_4(0.42,0.59,0.57)+a_5(0.41,0.74,0.75)}
	*C*8	{u_1(1.0,0.0,0.0)+u_2(0.41,0.68,0.57)+u_3(1.0,0.0,0.0+u_4(0.36,0.82,0.72)+u_5(0.30,0.76,0.74); a_1(0.46,0.60,0.38)+a_2(0.34,0.47,0.58)+a_3(0.41,0.62,0.69)+a_4(0.42,0.64,0.49)+a_5(0.52,0.74,0.75)}
	…	…
L4	*C*17	{u_1(1.0,0.0,0.0)+u_2(1.0,0.0,0.0)+u_3(1.0,0.0,0.0)+u_4(0.36,0.82,0.72)+u_5(0.30,0.76,0.74); a_1(0.46,0.60,0.41)+a_2(0.34,0.61,0.58)+a_3(0.41,0.62,0.69)+a_4(0.42,0.64,0.57)+a_5(0.41,0.74,0.75)}
	*C*18	{u_1(1.0,0.0,0.0)+u_2(1.0,0.0,0.0)+u_3(0.52,0.74,0.75)+u_4(1.0,0.0,0.0)+u_5(0.30,0.76,0.74

		); a_1(0.46,0.60,0.72)+a_2(0.34,0.61,0.55)+a_3(0.41,0.78,0.69)+a_4(0.38,0.82,0.62)+a_5(0.41,0.80,0.67)}
	…	…
L5	*C*27	{u_1(1.0,0.0,0.0)+u_2(1.0,0.0,0.0)+u_3(1.0,0.0,0.0)+u_4(1.0,0.0,0.0)+u_5(0.30,0.76,0.74); a_1(0.46,0.60,0.72)+a_2(0.34,0.61,0.58)+a_3(0.41,0.78,0.69)+a_4(0.38,0.82,0.62)+a_5(0.41,0.80,0.75)}
	*C*28	{u_1(1.0,0.0,0.0)+u_2(1.0,0.0,0.0)+u_3(1.0,0.0,0.0)+u_4(0.36,0.82,0.72)+u_5(1.0,0.0,0.0); a_1(0.37,0.65,0.74)+a_2(0.30,0.61,0.58)+a_3(0.41,0.64,0.69)+a_4(0.42,0.64,0.67)+a_5(0.41,0.76,0.75)}
	…	…

Note: The membership degree of objects shaded (1.0,0.0,0.0) in the table is 1, indicating that they are most likely to represent the corresponding user object, which is called the classical three-way object [12].

4.4. Model Verification

To verify the validity of 3WFCL-DIP, the open source dataset *MovieLens* was selected as the research object, which contains 25,000,095 user ratings and 1,093,360 labels, involving 62,423 movies and 162,541 users. And each user has rated at least 20 movies. The rating time in *ratings.dat* was taken as the basis for division, and the latest rating data of each user was selected as the test set while the remaining data as the training set. In comparison with classical interest discovery methods in recommendation, IDCL method proposed in literature [21], UI-LDA method proposed in literature [20] and CIBF method proposed in literature [13] were presented respectively. The names and explanations of each method are shown in Table 9. In addition, *accuracy rate*, *recall rate*, *F-value* and *Mean Reciprocal Rank* (MRR) were used as evaluation indicators, where MRR is the average reciprocal rank used to analyze the ranking position of correct values in the interest recommendation list. The calculation is shown in formula 25, where N is the number of the interest annotations on resources. $rank_i$ is the location of the i_{th} interest resource in the recommended list.

The movie rating value can be regarded as the intensity of the interest in different categories of movies, according to the user rating characteristics of the open source dataset in *MovieLens*. The rating time is viewed as the interest tendency in different time periods. The computational performance of 3WFCL-DIP in short-term interest recommendation applications was analyzed. The assignment principles of relevant parameters are determined by referring to the analysis method in Section 4.2. The experimental results are shown in Figure 6. Compared with IDCL, UI-LDA and TIMA-TI, 3WFCL-DIP increased the accuracy rate by 12.3%, 10.4% and 2.5%, the recall rate by 17.1%, 13% and 11%, and the F-value by 14.4%, 11.5% and 6.1% respectively.

$$MRR = \frac{1}{N}\sum_{i=1}^{N}\frac{1}{rank_i} \tag{25}$$

Table 9. The meanings of the comparison method with its explanation

Name	Explanation
IDCL[21]	Interest discovery method based on concept lattice mining user behavior attributes and demand characteristics.
UI-LDA[20]	The method of acquiring label features based on user interest topics and combining the evolution rule of interest hierarchy structure.
TIMA-TI[13]	An approach based on user interest drift and social relationship discovery.
3WFCL-DIP	The proposed method.

Table 10. MRR values of different methods

Name	Top10	Top20
IDCL[21]	0.2197	0.1743
UI-LDA[20]	0.2487	0.1997
TIMA-TI[13]	0.2525	0.2056
3WFCL-DIP	0.2784	0.2132

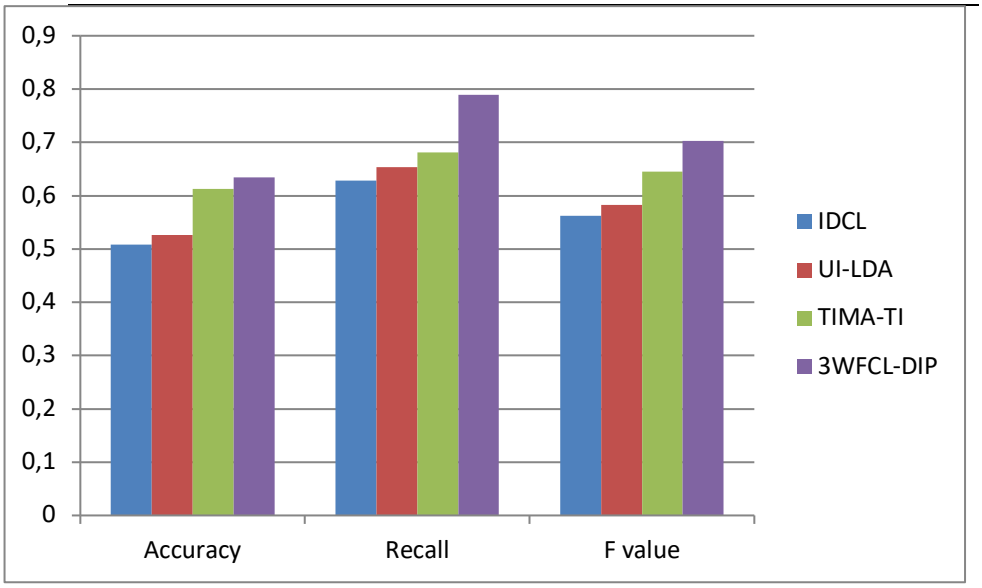

Figure 6. Accuracy, recall rate and F-value of different methods

In summary, 3WFCL-DIP obtains the real interest preference of users in the process of identifying the dynamic interests, compared with the other three methods. The reasons are as follows. Firstly, IDCL only considers the behavioral attribute characteristics of users, ignoring the semantic correlation between different attributes, which fails to consider the time sensitivity of the interests. UI-LDA analyzes the characteristics of interest hierarchy, which makes it improve to a certain extent in the above indicators, while without considering the attenuation characteristics of long-term and short-term interest drift. The problem of weak perception of scenes of short-term interest is not solved effectively. In addition, TIMA-TI combines the interest drift with social relations, making reasonable use of historical interaction data to improve the identification accuracy of interest discovery.

However, TIMA-TI relies on prior knowledge in the specific calculation process. The problem of data sparsity in the process of interest mining is more prominent, and the introduction of a large number of noisy data reduces the computational efficiency of large-scale data identification. By contrast, 3WFCL-DIP introduces interest inheritance into long and short term interest discovery, establishing three-way fuzzy interest concepts with time smoothing relationship, which integrates the important interest information and improves the computational accuracy of interest recognition.

5. Conclusion

The evolution characteristics of the dynamic interest are explored to propose a user dynamic interest portrait based on three-way fuzzy concept lattices. In specific, it introduces interest inheritance into the calculation of interest features to capture the short-term interest label set. In addition, the dependency of the multidimensional interest set attributes on group users is determined by defining three-way interest membership degrees. The fuzzy relation mapped to the fuzzy context improves the generalization and example path recognition ability of the user's short-term interest to a certain extent.

Availability of data and materials: The datasets used or analyzed during the current study are available from the corresponding author on reasonable request. Data are available from the author email at wangkai0552@126.com.

Competing interests: The authors declare that they have no competing interests.

Funding: This work is supported by the Key Project of Humanities and Social Science in Anhui Education Department(Grant Number 2022AH051412 and 2022AH051405), and the Key Project of Humanities and Social Sciences of Bengbu Medical College(Grant Number 2020byzd223sk).

Authors' contributions: Kai Wang conducted model building, experiment and result analysis, and paper writing.

References

[1] Bikram P. B., Ravi T, Amar R. C. Formal Concept Analysis and Its Validity in Power Reliability Engineering. *ELECTRICA*, **23**, 2, 2023, 696-712.

[2] Blei D. M., Ng A. Y., Jordan M I. Latent Dirichlet Allocation. *Journal of Machine Learning Research*, **3**, 2003, 993-1022.

[3] Chen L., Zhang H. Z, Jose J. M, et al. Topic Detection and Tracking on Heterogeneous Information. *Journal of Intelligent Information Systems*, **51**, 1, 2018, 115-137.

[4] Chen T, Guestrin C. XGBoost: A Scalabel Tree Boosting System. *Proceedings of the* 22*nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. ACM, 2016: 785-794.

[5] Huang L.H., Shi Y.X. Research on Monitoring Network Public Opinion events for Video barrage. *Journal of Information*, **41**, 2, 2022, 146-154.

[6] KAI. W, Cao Y. Algorithms for Constructing Hesitant Neutrosophic Concept Lattices and Computing their Similarity Distances in Medical Decision-Making. *Journal of Computing Science and Engineering*, **13**, 4, 2019, 151-162.

[7] Khan F. H., Qamar U., Bashir S. SentiMI: Introducing point-wise mutual information with SentiWordNet to improve sentiment polarity detection. *Applied Soft Computing*, **39**, 3, 2016, 140-153.

[8] Li Yan, LIU Z. H., GAO Y. F. Research on Multilingual author subject Model for interest portrait of researchers. *Journal of Information Science*, **39**, 6, 2020, 601-608.

[9] Liu X.X., LV Xiao Q., Wang X. L., et al. Twitter User Interest Mining Based on Wikipedia Category Graph, *Computer Science*, **46**, 9, 2019, 79-84.

[10] Liu R., Chen Ye. Research on Personalized movie Dynamic recommendation considering long-term and short-term interest and its evolution. *Data analysis and knowledge discovery*, **7**, 4, 2023, 1-18.

[11] Liu Y.Y.. Deep Learning Recommendation Model Based on Reader Interest Mining. *Journal of Southwest University*, **45**, 4, 2023, 201-209.

[12] Mao H., Cheng Y.L., Liu X.Q. The Relationship between Fuzzy Three Branches Concept and Fuzzy Concept Lattice. *Fuzzy Systems and Mathematics*, **34**, 6, 2020, 167-174.

[13] PIAO G. Y, BRESLIN J. G. Leveraging follow list memberships for inferring user interests for passive users on twitter. 28*th ACM International Conference on Hypertext and Social Media*, Prague, Czech Republic: ACM, 2017,155-164.

[14] Singh, Kumar P. Complex neutrosophic concept lattice and its applications to air quality analysis. Chaos Solitons & Fractals, **109**, 2018, 206-213.

[15] Wang H., Ji X.H., Ma J. Automatically Collect the Adjectives and Calculate their Sentiment Intensity in the Sentiment Analysis. *Applied Mechanics and Materials*, 2014, 1138-1142.

[16] Wang S., Ji X.M. A study on emotion expression characteristics based on user portraits in online health community. *Information theory and practice*, **45**, 6, 2022, 179-187.

[17] Wang H.B., YAN Jia. Group Intelligent Task Recommendation Based on Dynamic Preference and Competitiveness. *Journal of Software*, **34**, 4, 2023, 1666-1694.

[18] Wei L., ZHAO S. Particle and Knowledge Structure in Three-way Concept Analysis. *Journal of Northwest University*, **50**, 4, 2020, 537-545.

[19] Wu J.Y., Xu M.Z. Personalized Recommendation based on User profile and Video interest labels. *Information Science*, **39**, 1, 2021, 128-134.

[20] Xia L. X., Zeng J.Y., Bi C.W. Research on the evolution of user interest level based on LDA Topic Model. *Data Analysis and Knowledge Discovery*, **3**, 7, 2019, 1-13.

[21] Xu H.L., ZHANG H.T., ZHANG X.H.I, et al. Research on Interest Portrait of University Library Group Users Based on Concept Lattice, *Information Science*, **37**, 9, 2019, 153-158+176.

[22] Yan X., Guo J., Lan Y., et al. A Biterm Topic Model for Short Texts. *Proceedings of the* 22*nd International Conference on World Wide Web*. ACM, 2013, 1445-1456.

[23] YAO Y. Y. An Outline of a Theory of Three-Way Decisions. *Proceedings of the* 8*th International Conference on Rough Sets and Current Trends in Computing*. Berlin, Germany: Springer, 2012, 1-17.

[24] Yayık Apdullah, Kutlu Yakup. Online LDA based brain-computer interface system to aid disabled people. *Natural and Engineering Sciences*, **2**, 2, 2017, 383-395.

[25] Yuehua Z., Si C. Z., Jie Wu, et al. Detecting the research diversity of researchers in library and information science: An exploratory study. *Data Science and Informetrics*, **2**, 3, 2022, 22-36.

[26] Yu H.T., CUI R.F., Dong Qinqin. Microblogging user interest model based on forgetting curve. *Computer Engineering and Design*, **35**, 10, 2014, 3367-3372 + 3379.

[27] ZHU M., LIN X., LU T., et al. Identification of opinion leaders in social networks based on sentiment analysis: evidence from an automotive forum. International Conference on Modeling, *Simulation and Optimization Technologies and Applications*, 2017.

Received 24.11.2023, Accepted 11.07.2024