Physica A 498 (2018) 102–115

Contents lists available at ScienceDirect

Physica A

journal homepage: www.elsevier.com/locate/physa

Measuring transferring similarity via local information

Likang Yin [a,b], Yong Deng [a,c,*]

[a] Institute of Fundamental and Frontier Science, University of Electronic Science and Technology of China, Chengdu, 610054, China
[b] School of Hanhong, Southwest University, Chongqing 400715, China
[c] School of Computer and Information Science, Southwest University, Chongqing, 400715, China

HIGHLIGHTS

- This paper presents a proof of the range of classical transferring similarity.
- This paper proposes a novel method to fuse local similarity.
- We compare the proposed method with some common link prediction methods on 9 well-known datasets.

ARTICLE INFO

Article history:
Received 10 April 2017
Received in revised form 22 November 2017
Available online 31 January 2018

Keywords:
Transferring similarity
Link prediction
Dempster–Shafer evidence theory
Belief function
Recommender systems

ABSTRACT

Recommender systems have developed along with the web science, and how to measure the similarity between users is crucial for processing collaborative filtering recommendation. Many efficient models have been proposed (i.g., the Pearson coefficient) to measure the direct correlation. However, the direct correlation measures are greatly affected by the sparsity of dataset. In other words, the direct correlation measures would present an inauthentic similarity if two users have a very few commonly selected objects. Transferring similarity overcomes this drawback by considering their common neighbors (i.e., the intermediates). Yet, the transferring similarity also has its drawback since it can only provide the interval of similarity. To break the limitations, we propose the Belief Transferring Similarity (BTS) model. The contributions of BTS model are: (1) BTS model addresses the issue of the sparsity of dataset by considering the high-order similarity. (2) BTS model transforms uncertain interval to a certain state based on fuzzy systems theory. (3) BTS model is able to combine the transferring similarity of different intermediates using information fusion method. Finally, we compare BTS models with nine different link prediction methods in nine different networks, and we also illustrate the convergence property and efficiency of the BTS model.

1. Introduction

Recent years have witnessed the exponential growth of the internet [1], the new challenge for modern science is the information overload. The personalized recommendations [2,3] is a efficient way to refine the massive data. Therefore, the recommender systems [4,5] come to a vogue in both the community of network science [6–9] and data mining field [10,11]. The recommender systems are a kind of complex entity that predicts the potential interesting by measuring the similarity between users' historical selections [12]. Generally, there are two classical approaches to implement the recommender

* Corresponding author at: Institute of Fundamental and Frontier Science, University of Electronic Science and Technology of China, Chengdu, 610054, China.
E-mail address: dengentropy@uestc.edu.cn (Y. Deng).

https://doi.org/10.1016/j.physa.2017.12.144
0378-4371/

systems [13]. The first one is the content-based method which makes the recommendation by finding the similar items according to the target user's historical selections [14]. The other one is called the user-based method [15,16]. Both of the two methods are based on the preferences of the users for a set of items [17]. Based on this inherent mechanism of recommender systems, how to properly measure the similarity between two users is of great significance. To address this issue, we can construct a bipartite network by regarding the users and items as different nodes and the accessing records as the links respectively [18,19]. Moreover, we can make commendations to users by predicting the missing links between unconnected user–item pairs in this bipartite network [20–22]. By this way, the issue of recommender systems has been transformed to the link prediction problems.

There are many efficient models in predicting the missing links in the networks. For example, the Common Neighbor (CN) index considers that the more common neighbors of two users represent the higher similarity they possess. Moreover, the Adamic–Adar (AA) index [23] and Resource Allocation (RA) [24] index have been proposed. AA and RA improve the accuracy by assigning a small-degree neighbor to a larger weight. Note that, RA and AA are called the similarity-based link prediction methods since both of them are based on the similarity between different nodes. There are some other kinds of prediction methods, such as the maximum likelihood methods [25], and the community-based models [26]. The link prediction is one of the hot topics in networks science, and there are some other valuable issues in network science, such as identifying the influential spreaders [27,28], supply chain networks [29], vulnerability analysis [30], network evolution [31,32], and self-similarity analysis [33].

By quantifying the similarity of users, some establishments and institutions have used the recommender systems in some practical applications, such as simulating spatial public goods games [34], analyzing the news consumption on Facebook [35], and video recommendation [36]. The similarity quantification has been widely used in both theoretical and practical fields, however, the similarity-based recommender systems still face challenges, such as the cold start problem [37], personalization problem [38,39], and measuring the effect of time [40,41].

However, how to handle the uncertainty of the high dimensional data is also an open issue in recommender systems [42–44]. For example, suppose there are three users, labeled as A, B, and C. There is a low similarity between A and C, but A and C are both very similar with B. In fact, the three users can share very similar potential preferences, and the low similarity between A and C may be caused by the sparsity of the data. That is to say, A and C share very few commonly selected items. The sparsity of dataset makes the direct similarity less accurate. To address this issue, the transferring similarity has been proposed [45]. However, the classical transferring similarity measure only presents an uncertain interval, and its range depends on the direct similarity degree. Therefore, the classical transferring similarity measure is hard to make accurate predictions.

The motivation of this study is to address both the issue of sparsity of datasets in the direct similarity measure and the issue of fuzzy similarity in classical models. The contributions of this paper are listed as follows. First, this paper proves the range of the interval of the classical transferring similarity. Second, the proposed method is able to fuse the transferring similarity of the intermediates (i.e., the common neighbors). Third, our model presents a specific state to denote the degree of similarity using fuzzy system theory. Moreover, to show the accuracy and efficiency, we compare the proposed method with nine common link prediction models on nine well-known datasets.

The structure of proposed method is listed as below. Given the raw user–item rating matrix, the first step is to calculate the direct similarity (i.e., the Pearson correlation) between two users. Second, based on the classical transferring similarity, we compute the upper bound and lower bound of the similarity interval. The third step is to construct the mass functions by discretizing the similarity interval. Next, we need to fuse the mass functions of all intermediates. Finally, we can make prediction according to the ultimate similarity state.

2. Preliminaries

2.1. Recommender systems

The recommender systems contain two main parts, one is the user and another is the object, and all users rated the each object in the target list. Hence, if there exists N users denotes as $U = u_1, u_2, \dots, u_N$ and M objects denotes as $O = o_1, o_2 \dots, o_M$, a rating $N \times M$ matrix can fully describe the input of the recommender systems. The common practice of recommender systems is to search for the preferences of a large group of people, and to find a small group of people who share the same taste as the target user. The recommender systems examine the preferences of these people, and combines them to create a ranking list of recommendations. There are various ways to help us determine who is close to our taste. Among them, there are two simple but very efficient methods, the first one is named Euclidean Distance Score (EDS) [46]. First, let features of the item, which are evaluated by different users, denote the multi-dimensional preference space, the coordinate of the user can be located according to the scores they give, and the distance between any two users reflects the degree of similarity between the two users. In general, the shorter the distance in the preference space, the higher the similarity between two users. Namely,

$$s_E(x, y) = \sqrt{\sum_{i=1}^{N}(x_i - y_i)^2} \tag{1}$$

where $s(x,y)$ denotes the similarity between user x and user y, i is the evaluated item, x_i denotes the score marked by one user and y_i denotes the score marked by another user.

Except the EDS measure, Pearson correlation coefficient [47] has been widely used. Pearson correlation coefficient is more complex than EDS, however, it presents better prediction result than EDS especially when the datasets have not been normalized (e.g., the pessimist tends to give a lower score than the optimists). The mathematical definition of Pearson correlation coefficient can be written as

$$s_P(x,y) = \frac{cov(x,y)}{\sigma_x \sigma_y} = \frac{E(xy) - E(x)E(y)}{\sqrt{E(x^2) - E^2(x)}\sqrt{E(y^2) - E^2(y)}} \tag{2}$$

where the $cov(x,y)$ is the covariance of variable x and y, σ_x, σ_y are standard deviation of x and y, respectively. $E(x)$ is the mean of variable x.

Due to the Pearson correlation considers the mean of the preference, it amends the grade inflation in EDS measure which will cause awful result. In other words, if user x always tends to mark a lower score than user y and the difference between those scores maintains the same, the similarity between user x and y will still be very high. In this paper, we consider all the similarity degree transits from -1 to 1. Plus, $s=-1$ denotes two objects are totally different, while $s=1$ denotes the two objects are the same, and $s=0$ denotes they are completely uncorrelated.

In this paper, we focus on how to fuse the local information for making better decision instead of considering how to calculate the similarity in the fuzzy systems [48]. Moreover, we focus on how to take advantage of the transferring similarity especially when there exists several intermediaries between two target users. The next part we shall introduce the specific fusion method based on Dempster–Shafer evidence theory [49,50].

2.2. Dempster–Shafer theory

The real world is uncertain [51–54]. Theory of the evidence [55,56] is an efficient tool to handle uncertain information between many information sources such as uncertain decision making [57], D numbers modeling [58–63], information fusion problem [64,65], supplier selection [66]. In order to get better explanation of this method, some basic concepts are introduced follows.

The Frame Of Discernment (FOD) is used to represent the set of all observed events. Let ϕ be the set of mutually exclusive and collectively exhaustive events E_i, namely

$$\phi = \{E_1, E_2, \ldots, E_i, \ldots, E_n\}. \tag{3}$$

The power set of ϕ is denoted by 2^ϕ, and

$$2^\phi = \{\varnothing, \{E_1\}, \ldots, \{E_n\}, \{E_1, E_2\}, \ldots, \phi\} \tag{4}$$

where the $\varnothing$ is denoted empty set.

For the FOD $\phi = \{E_1, E_2, \ldots, E_\phi\}$, a mass function is a mapping m from 2^ϕ to $[0, 1]$, mass function is used to transform every event to probability, formally defined as

$$m: 2^\phi \to [0, 1] \tag{5}$$

which need to meet the succeeding properties

$$m(\varnothing) = 0 \quad and \quad \sum_{\theta \in 2^\phi} m(\theta) = 1,\ 0 \le m(\theta) \le 1. \tag{6}$$

Dempster–Shafer theory has many merits in uncertainty modeling due to its high efficiency and accuracy. In the theory of evidence, the Belief function (Bel) and Plausibility function (Pl) are defined as

$$Bel(A) = \sum_{B \subseteq A} m(B) \tag{7}$$

$$Pl(A) = \sum_{B \cap A \neq \varnothing} m(B). \tag{8}$$

The belief function $Bel(A)$ represents the justified specific support for the focal element (or proposition) A, while the plausibility function $Pl(A)$ represents the potential specific support for A. The length of the belief interval $[Bel(A), Pl(A)]$ is used to represent the degree of imprecision for A.

Given two independent Basic Probability Assignments (BPA) m_1 and m_2, Dempster's rule of combination, denoted by $m = m_1 \bigoplus m_2$, is used to combine them and it is defined as follows

$$m(A) = \begin{cases} \frac{1}{K} \sum\limits_{B \cap C = A} m_1(B) m_2(C), & A \neq \varnothing \\ 0, & A = \varnothing \end{cases} \tag{9}$$

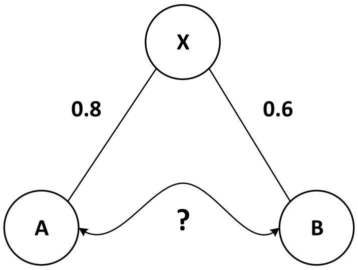


Fig. 1. The illustration for transferring similarity.

with

$$K=\sum_{B\cap C\neq\varnothing} m_1(B)m_2(C). \tag{10}$$

2.3. Pignistic probability transform

In the transferable belief model (TBM) [67,68], pignistic probabilities are used for decision making. Let m be a BPA on the frame of discernment Θ. Its associated pignistic probability function $BetP_m\colon \Theta \to [0, 1]$ is defined as

$$BetPm(\omega)=\sum_{A\subseteq P(\Theta),\omega\in A}\frac{1}{|A|}\frac{m(A)}{1-m(\varnothing)},\ m(\varnothing)\neq 1 \tag{11}$$

where $|A|$ is the cardinality of subset A. The process of pignistic probability transform (PPT) is that basic probability assignment transferred to probability distribution.

3. Proposed belief transferring similarity model

3.1. Transferring similarity

Transferring similarity is the crucial mathematical tool which is used to measure the similarity between two indirect users. However, due to the inherent uncertainty of transferring similarity, existing transferring similarity measures cannot precisely address the similarity induced by indirect relationship called the high-order correlations [69]. Notably, both EDS and Pearson correlation coefficient are the direct similarity between two indirect users. In order to make the concept of proposed model easier to understand, we draw an illustration in Fig. 1. Firstly, let us go back to the fundamental definition of transferring similarity. We simplify the problem as assuming there are three variables A, B and X, and provided that both A and B are relative to X with high correlation, can we surely say that A and B are highly correlated?

The answer is *uncertain*. Since both A and B are partially similar to X even with high correlation, we cannot make sure the relative components of A and B are still similar to X simultaneously. In fact, the similarity between A and B is an interval instead of a certain degree. The theorem is listed as follows.

Theorem. Let the ρ_A and ρ_B denote the correlation coefficient between variable *A and X*, *B and X* respectively. The correlation coefficient between variable A and B is an interval, namely

$$Corr(A,B)=[\ \rho_A\cdot\rho_B-\sqrt{1-\rho_A^2}\sqrt{1-\rho_B^2},\ \ \rho_A\cdot\rho_B+\sqrt{1-\rho_A^2}\sqrt{1-\rho_B^2}\ \ ]. \tag{12}$$

Proof. Let variable A, B and X be the random variable with theirs standard deviation equals to 1, then decompose both A and B into two parts that first part is completely correlated to the variable X and second part is completely uncorrelated to the X using linear regression method, namely

$$\begin{aligned} A&=\rho_A X+\epsilon_A\\ B&=\rho_B X+\epsilon_B \end{aligned} \tag{13}$$

where ϵ_A and ϵ_B are the second part of A, B respectively that uncorrelated to X, with their means equal to zero. Then we have

$$Var(\epsilon_A)=1-\rho_A^2\sigma_X^2=1-\rho_A^2 Var(\epsilon_B)=1-\rho_B^2\sigma_X^2=1-\rho_B^2 \tag{14}$$

where $Var(\epsilon_A)$ and $Var(\epsilon_B)$ are the variance of variable A and B, respectively. The σ_X denotes the standard deviation of X.

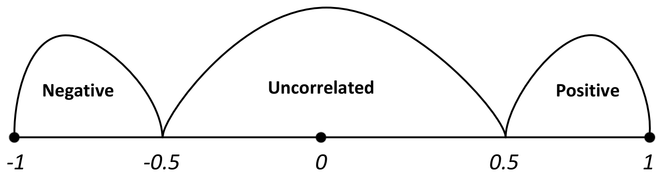


Fig. 2. The illustration of three basic states with respect to the correlation degree.

Now we need to follow the definition to calculate the correlation coefficient, it reads

$$\begin{aligned} Corr(A, B) &= Cov(A, B) \\ &= Cov(\rho_A X + \epsilon_A, \rho_B X + \epsilon_B) = \rho_A \rho_B Var(X) + Cov(\epsilon_A, \epsilon_B) \\ &= \rho_A \rho_B + Corr(\epsilon_A, \epsilon_B)\sigma(\epsilon_A)\sigma(\epsilon_B) \\ &= \rho_A \rho_B + Corr(\epsilon_A, \epsilon_B)\sqrt{1-\rho_A^2}\sqrt{1-\rho_B^2}. \end{aligned} \tag{15}$$

From Eq. (15) we can see that, given the ρ_A and ρ_B, the correlation coefficient between $Corr(A, B)$ only depends on the value of ϵ_A and ϵ_B. Meanwhile, $Corr(\epsilon_A, \epsilon_B)$ can be any real number in the interval $[-1, 1]$. Thus, we are able to derive Eq. (12). □

3.2. Belief transferring similarity model

As described in the above section, due to the inherent property of transferring similarity, we know that the transferring similarity is an uncertain interval instead of a real number. However, an interval is inconvenient for solving decision-making problem, such as predicting missing links, constructing supply chain and predicting stock price especially in the forecasting systems. In essence, this is due to classical transferring similarity model being unable to represent and handle the uncertainty of state description.

Facing this problem, we consider to proposed a novel method named Belief Transferring Similarity (BTS) to improve the accuracy of classical model using Dempster–Shafer theory. Dempster–Shafer theory is an efficient mathematical tool to handle uncertainty and information fusion problem [70–72]. Moreover, as we can see in the next section, proposed BTS model can deal with several intermediates which can also be used in both complex network and recommender systems. First, the definition of belief transferring similarity is given as below.

Definition 1. Assume A and B denote the random variables. Let $\{\chi_n : n > 0\}$ be the sequence of the intermediates between A and B. The similarity matrix $\Re$ is defined as

$$\Re = \begin{bmatrix} \rho_{A1} & \rho_{A2} & \cdots & \rho_{Ai} & \cdots & \rho_{An} \\ \rho_{B1} & \rho_{B2} & \cdots & \rho_{Bi} & \cdots & \rho_{Bn} \end{bmatrix} \tag{16}$$

where ρ_{Ai} denotes the correlation between variable A and the ith intermediate between A and B. The illustration of similarity matrix is in Fig. 3. The network is a belief transferring similarity model if

$$\rho_{mn} \in [-1, 1]\ \forall \rho_{mn} \in \Re \tag{17}$$

where the $\rho_{m,n}$ denote the correlation between element m and the nth intermediate.

Definition 2. In the BTS model, we use states to describe the frame of discernment Θ is defined as

$$\Theta = \{Positive\text{-}correlation(P),\ Uncorrelated(U),\ Negative\text{-}correlation(N)\}. \tag{18}$$

Meanwhile, the three focal elements (P, U, N) are defined according to the correlation degree. The boundaries among the three focal elements are 0.5 and −0.5 as listed in Table 1 and graphical shown in Fig. 2. Theoretically, the predicting result will be more accurate and stable if more evidences are obtained. Notably, in the BTS model, it is a consistent one-to-one match between intermediate and mass function. Thus, the correlation information from each intermediate regards as an evidence, which supports users to measure the similarity between two objects. These following properties are valid in the BTS model.

Property 1. *If there is no intermediate (or common neighbor) between two users, thus the frame of discernment $\Theta = \{\varnothing\}$. In this sense, it is reasonable to treat the users as two uncorrelated users since there is no evidence to support they are similar or dissimilar. Namely,*

$$S(A, B) = 0 \text{ if } A \cup B = \varnothing. \tag{19}$$

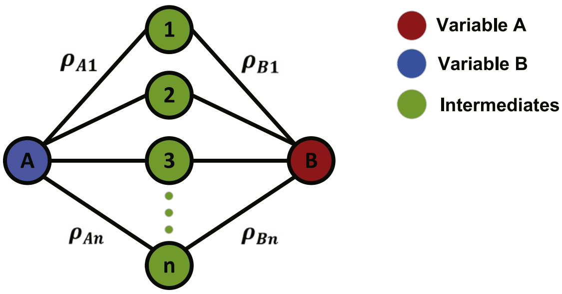


Fig. 3. The illustration of similarity matrix $\Re$.

Table 1
The classification of three basic states with respect to the correlation degree.

The correlation degree	[1, 0.5)	[0.5, −0.5]	(−0.5, −1]
State	Positive-correlation	Uncorrelated	Negative-correlation
Abbreviation	*P*	*U*	*N*

Property 2. *The evidence fusion process follow the commutative law and associative law. It reads*

$$m_1 \oplus m_2 = m_2 \oplus m_1 \tag{20}$$

$$m_1 \oplus m_2 \oplus m_3 = (m_1 \oplus m_2) \oplus m_3 = m_1 \oplus (m_2 \oplus m_3). \tag{21}$$

The proposed BTS model inherits the desirable properties of the Dempster–Shafer theory, and it can totally be transformed to network frame when we treat all users and similarity degree as nodes and weight of the links, respectively. In this scenario, mass functions can be constructed by the adjacent matrix of the networks, and the new model can be treated as a generalized tool to handle the link prediction problem in the network science [73]. The construction and application of the BMC model are represented in the following steps, graphically represented by Fig. 4. The pseudo code is also listed as below, and it is shown that the complexity of the proposed method is $O(n^3)$ since it needs to run two $for-loops$ and one iterative information fusion process. However, the calculation process can be greatly accelerated by taking advantages of matrix operations in MATLAB and other similar software platforms.

Step 1 Given the raw user–item rating matrix, we can derive the direct similarity between two users using either the linear correlation measure (i.e., the Pearson correlation) or the distance measure (i.e., the EDS). At this step, some useful technologies, such as normalization and aggregation, can be used for preprocessing if necessary.

Step 2 According to Eq. (12), we can calculate the interval of transferring similarity of each pair of users based on the previous direct similarity.

Step 3 Define the mass function on the frame of discernment Θ according to the practical case. Thus, the BPA of each intermediate can be constructed by discretizing the interval. The BPA of a particular intermediate denotes the membership degree to each state (i.e., similar or dissimilar).

Step 4 Fusing information is the most important part of the proposed BTS model. Since the BPAs generated by different intermediates regard as the different evidences, fusing function is used to combine two evidence based on Dempster–Shafer theory. Generally, we need to fuse all the BPAs and finally get the ultimate BPA m' using Eq. (9).

Step 5 Make the prediction and recommendation according to the BPA m'.

Notably, the ultimate m' denotes the distribution of the degree of belief of each proposition instead of the probability distribution of the basic states. Various methods has been proposed to aid decision making according to the BPAs. In the frame of D–S theory, a common mathematical tool is the pignistic probability transformation (PPT) [67,68], which is introduced in Section 2. Within the PPT method, the mass function can be transformed to a specific probability distribution, and then the decision can be made based on the existing probability distribution in a more reasonable way.

4. Case study

Given a simple binary user–item rating matrix over 20 different items (the ratings are restricted to −1 and 1). Notably, the proposed BTS model focuses on handling how to measure the transferring similarity, but not how to calculate the initial direct similarity between users and intermediates. However, to make our logic clear, we still start at the user–item rating matrix here. This example is the simple but effective case to make recommendations to users based on transferring similarity. Firstly, let us first define two target users A and B and their intermediates I_1 and I_2 as

$$A = [1]_{16} + [-1, -1, 1, 1],$$

$$B = [1]_{16} + [-1, 1, -1, 1],$$

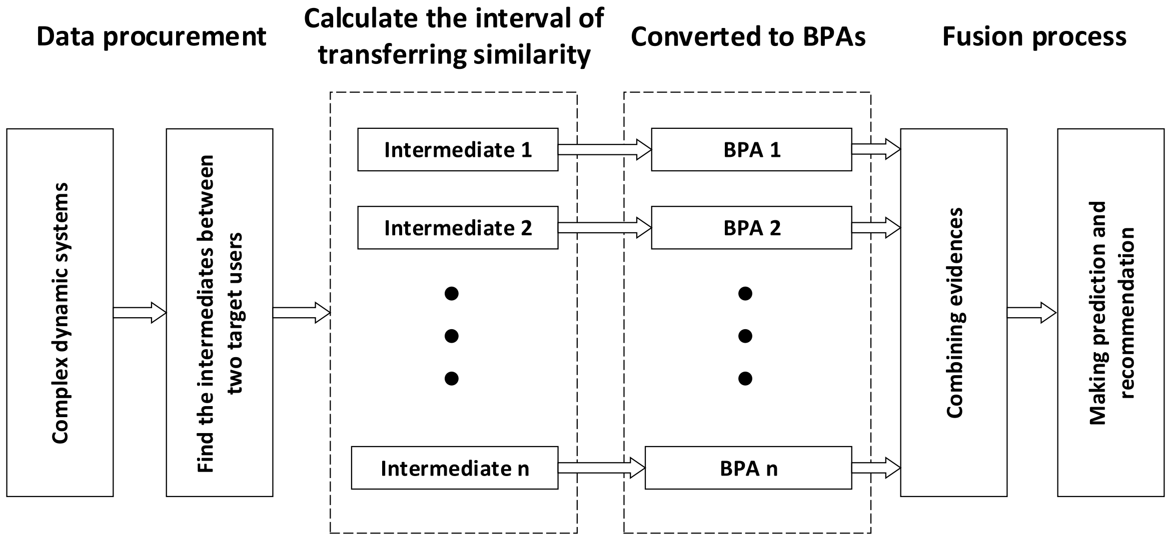


Fig. 4. The flow of constructing belief transferring similarity model.

Algorithm 1 The algorithm of belief transferring similarity.

Input: The $n \times n$ adjacency matrix of training set T;
Output: The $n \times n$ transferring similarity matrix S;

```
Initialize the similarity matrix S to null matrix;
for node i from i = 1 to i = n do
    for node j from j = i + 1 to j = n do
        Obtain the set of intermediates I with k common neighbors of node i and j.
        Calculate the interval of the transferring similarity of the node i and node j.
        Discretize the interval into the BPA of each intermediate m_k in I.
        Get the final fusing result based on evidence combination method;
        S_ij ← (m_1 ⊕ m_2 ··· ⊕ m_k)
        j ← j + 1
    end for
    i ← i + 1;
end for
```

$$I_1 = [1]_{16} + [1, -1, -1, -1],$$

$$I_2 = [-1]_{16} + [1, -1, 1, 1].$$

Given the user–item rating matrix above, we now use the simplest strategy to measure the direct similarity since the key point of this study is measuring the transferring similarity. Namely,

$$S_{i,j} = \sum_{N}^{k=1} i[k] * j[k] \tag{22}$$

where i and j denote the rating matrix of two different users.

The result of direct similarity measure is listed as follows. The illustration of the direct similarity is shown in Fig. 5.

$$S_{A,I_1} = 0.8, S_{A,I_2} = -0.7$$

$$S_{B,I_1} = 0.7, S_{B,I_2} = -0.9.$$

At the first step, the frame of discernment Θ has been generated according to the given data. It still assumes there exists three basic states, they are negative correlation (N), uncorrelated (U) and positive correlation (P), respectively. The three basic states form the frame of discernment, which denotes as $\Theta = \{N, U, P\}$.

The second stage in our five-stage strategy is essential, the mass function will be defined on the frame of discernment Θ according to the practical similarity distribution. Assume the mass function is defined as follows

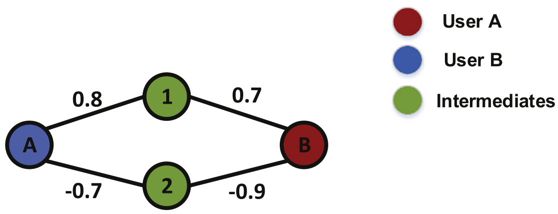


Fig. 5. The illustration of the case study. The blue node and the red node denote two different users, and the green nodes represent the intermediates between the two nodes. The value of similarity of two users is shown in the link.

$$m(\{N\}) = \begin{cases} 1, & \rho \in [-1, -0.65) \\ (-0.5-\rho)/0.15, & \rho \in [-0.65, -0.5] \\ 0, & \rho \in (-0.5, 1] \end{cases} \tag{23}$$

$$m(\{U\}) = \begin{cases} 0, & \rho \in [-1, -0.5) \\ (\rho+0.5)/0.15, & \rho \in [-0.5, 0.35) \\ 1, & \rho \in [-0.35, 0.35] \\ (0.5-\rho)/0.15, & \rho \in (-0.35, 0.5] \\ 0, & \rho \in (0.5, 1] \end{cases} \tag{24}$$

$$m(\{P\}) = \begin{cases} 0, & \rho \in [-1, 0.5) \\ (\rho-0.5)/0.15, & \rho \in [0.5, 0.65] \\ 1, & \rho \in (0.65, 1] \end{cases} \tag{25}$$

$$m(\{N, U\}) = \begin{cases} 0, & \rho \in [-1, -0.65) \\ (\rho+0.65)/0.15, & \rho \in [-0.65, -0.5] \\ (-0.35-\rho)/0.15, & \rho \in (-0.5, -0.35] \\ 0, & \rho \in (-0.35, 1] \end{cases} \tag{26}$$

$$m(\{U, P\}) = \begin{cases} 0, & \rho \in [-1, 0.35) \\ (\rho-0.35)/0.15, & \rho \in [0.35, 0.5] \\ (0.65-\rho)/0.15, & \rho \in (0.5, 0.65] \\ 0, & \rho \in (0.65, 1] \end{cases} \tag{27}$$

$$m(\{N, P\}) = 0 \tag{28}$$

$$m(\{N, U, P\}) = 0 \tag{29}$$

$$m(\{\varnothing\}) = 0. \tag{30}$$

Here, the above mass function is used to discretize the continuous correlation scope, graphically shown in Fig. 6. After discretizing the correlation scope, the initial state space $\{N, U, P\}$ has been transformed to the belief space $\{N, \{N, U\}, U, \{U, P\}, P\}$ which contains 5 propositions. It is worth noting that the belief space contains 2 special propositions $\{N, U\}$ and $\{U, P\}$. Moreover, through the defined mass function, it describes each state in a smoother way, which helps the fusing result converge to an optimized value.

Notably, the mixture of two basic states (e.g., the $\{N, U\}$) implies the state can be either negative correlation or uncorrelated correlation. However, it cannot be distinguished or separated simply. The existence of mixtures enhances the ability of BTS model to express uncertain information inside the correlation value, which is entirely different from the concept of fuzzy states. According to the given mass function, the BPA of each intermediate can be obtained, and the BPAs of these intermediates are represented in Table 2. As shown in Table 2, both ρ_1 and ρ_2 are the correlation between I_1 and I_2, respectively. The left point denotes the left end of similarity interval and the right point denotes the right end of similarity interval. The belief of proposition is assigned according to the corresponding area's proportion of total area under the piecewise function.

At the third step, the interval of transferring similarity of each intermediate is obtained according to Eq. (12). It is obvious that the interval radius depends on the absolute value of correlation degree, and the larger the correlation degree the smaller the radius.

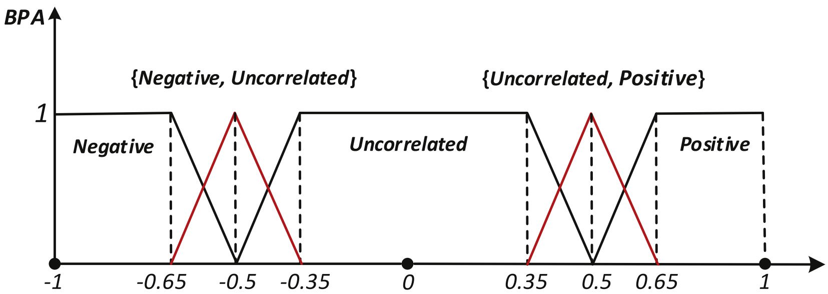


Fig. 6. The illustration of three basic states with respect to the correlation degree.

Table 2
The procedure of constructing BPAs according to the correlation.

User	ρ_1	ρ_2	Left-point	Mid-point	Right-point
A	0.8	−0.7	0.1315	0.5600	0.9885
B	0.7	−0.9	0.3187	0.6300	0.9413
BPA	$m(\{N\})$	$m(\{N, U\})$	$m(\{U\})$	$m(\{U, P\})$	$m(\{P\})$
m_A	0	0	0.3425	0.1750	0.4825
m_B	0	0	0.1707	0.2409	0.5884

At the fourth step, the BPAs of all intermediates are derived according to the mass function. Then, the BPAs are combined to get the ultimate BPA m' according to Eq. (9). Notably, the belief of proposition is assigned according to the corresponding area's proportion of total area under the piecewise function in Fig. 6. Namely,

$$m'(\{P\}) = 0.7025,$$

$$m'(\{U, P\}) = 0.0589,$$

$$m'(\{U\}) = 0.2386,$$

$$m'(\{N, U\}) = 0,$$

$$m'(\{N\}) = 0.$$

Then, using PPT approach, a probability distribution of states is obtained:

$$m'(\{P\}) = 0.7320,$$

$$m'(\{U\}) = 0.2680,$$

$$m'(\{N\}) = 0.$$

From the fusing result we can see that the most probable relation between user A and user B is the *Positive correlation*. It is consistent with our intuition since both user A and user B have similar correlation with same intermediate. Moreover, Table 2 indicates that the I_1 represents an evidence, which supports the similarity should be in the interval 0.1315 to 0.9885 with the midpoint equals to 0.56, and the I_2 supports the similarity should be in the interval from 0.3187 to 0.9413 with its midpoint equals 0.63. The wide range of similarity interval can contain more than one predicted state (e.g., uncorrelated correlation and positive correlation in this case), while this recommendation makes no sense to users since it can only imply that the correlation between user A and user B is not negative.

However, on the one hand, as the most probable state, state P has a certainty of 0.7320, which is larger than both midpoints of similarity interval. On the other hand, the jumping phenomenon does not occur in predicting process of proposed BTS model. Thus, those results imply that the predicted results are consistent, and the belief transferring similarity model effectively overcomes the drawbacks of the classical transferring similarity model.

5. Empirical experiments

In this section, we first compare the result of proposed BTS model with the famous Pearson correlation coefficient over the well-known dataset in recommender systems. Second, the comparison of proposed BTS model with other two well-known similarity-based link prediction methods and the Pearson correlation over three recommender systems datasets is shown as Fig. 7. Note that, we generate the Pearson correlation over the three dataset (i.e., the MovieLens (100k & 1M) [17] and the FilmTrust [74]) in the first place. Then, we implement the proposed model over the direct similarity matrix (i.e., the Pearson

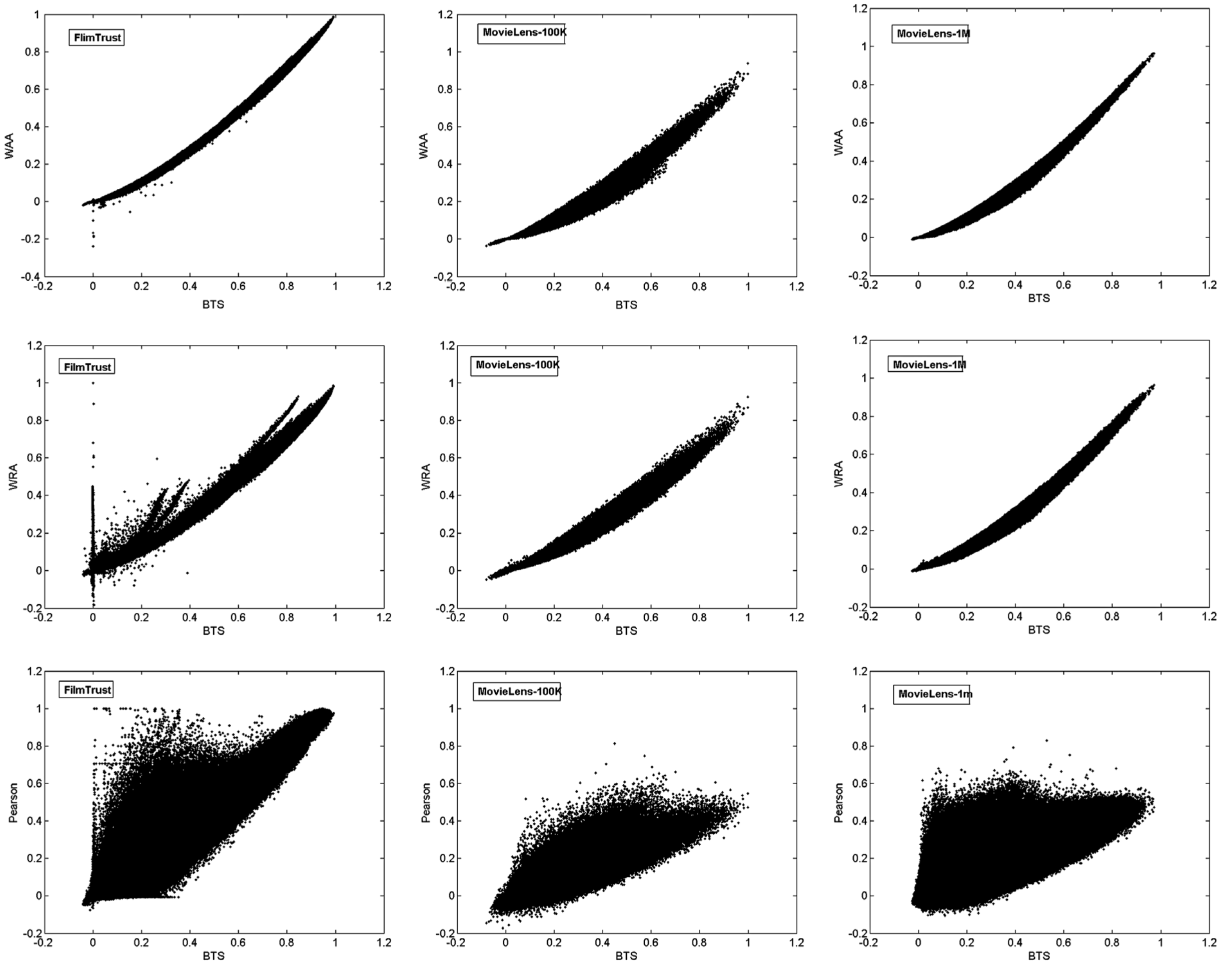


Fig. 7. We compare relative similarity of nodes measured by the proposed BTS model with two well-known link prediction indexes and Pearson correlation. In order to eliminate the effect caused by the difference of absolute value, we set the percentage values in x-axis and y-axis as the top percentage ranking values of the three measures, respectively. As shown in this figure, the ranking dots of BTS model are highly consistent with both RA index and AA index.

correlation). The result shows that the proposed method act in a manner to lower the degree of similarity, specially compare to the original Pearson correlation. Plus, there are some negative dots as shown in the figure. The possible reason for the two phenomenon is the fact that the Pearson correlation is in the interval $[-1, 1]$, and therefore, the transferring similarity measure results in lower the degree of similarity.

Moreover, to compare with existing link prediction method, we also test the accuracy of the proposed BTS model based on some common datasets. The BTS model has some drawbacks under the precision evaluation approach, the possible reason is that the transferring similarity enlarge the uncertain interval of similarity and it also tends to lower all the similarity degree. However, the performance of BTS model exceeds the other nine indices in link prediction based on AUC over some datasets. The highest accuracy of all nine networks in each line is emphasized in bold. From Table 3, we can see that the similarity-based link prediction method (i.e., the RA and AA) BTS model are well-matched in both the weighted and unweighted network, especially on the AUC measure.

It should be pointed out that this paper does not focus on the accuracy of link prediction, since the BTS model is used to handle the higher-order similarity. However, the input of the dataset is not the similarity between nodes but the connection between nodes. Theoretically, the experiment over some link prediction dataset is not meaningful for BTS model, since the links do not equal to the direct similarity (i.e., the Pearson correlation), but stand for other connection (i.e., the frequency of US air). In particular, if the link holds for a negative relations, such as "dislike" in Tweets and Facebook, we may get a totally reversed answer as directly implementing the BTS model before getting the direct similarity of the networks.

6. View from Dempster–Shafer

In the present study, we investigated the property of transferring similarity and proposed a novel model named belief transferring similarity. Moreover, our experimental results demonstrated that the traditional transferring similarity

Table 3
The accuracy of BTS compare with other nine measures under AUC and Precision (Precision values are in brackets). The mean of AUC and Precision values are obtained by the mean of 100 independent realizations. The entries corresponding to the highest value among these measures are emphasized in bold.

Network	CN	Salton	Jaccard	Sorens	HPI	LHN	AA	RA	PAI	BTS
C.elegans	0.8478 (0.1211)	0.7979 (0.0134)	0.7904 (0.0126)	0.7911 (0.0126)	0.8058 (0.0111)	0.7234 (0.0106)	0.8663 (**0.1358**)	**0.8701** (0.1266)	0.7613 (0.0688)	0.8503 (0.0267)
Jazz	0.9566 (0.8140)	0.9663 (0.7658)	0.9625 (0.7544)	0.9627 (0.7544)	0.9478 (0.0126)	0.9036 (0.060)	0.9639 (**0.8404**)	**0.9723** (0.8226)	0.7707 (0.1876)	0.9547 (0.0373)
Metabolic	0.9241 (0.1786)	0.8158 (0.0834)	0.7764 (0.0642)	0.7765 (0.0642)	0.9161 (0.1633)	0.7398 (0.0440)	0.9556 (0.2494)	**0.9606** (**0.3067**)	0.8232 (0.1514)	0.9221 (0.0112)
NS	0.9911 (0.8707)	0.9914 (0.4922)	0.9912 (0.0842)	0.9914 (0.0842)	0.9912 (0.0126)	0.9908 (0.2376)	0.9914 (**0.9722**)	0.9915 (0.9694)	0.7298 (0.0031)	**0.9922** (0.0271)
PB	0.9223 (**0.3988**)	0.8786 (0.0021)	0.8759 (0.0021)	0.8757 (0.0003)	0.8538 (0.1633)	0.7617 (0.0002)	0.9262 (0.3718)	0.9265 (0.2733)	0.9093 (0.1167)	**0.9285** (0.0676)
Power	0.6269 (0.1064)	0.6272 (0.0267)	0.6267 (0.0400)	0.6264 (0.0400)	0.6267 (0.0033)	0.6266 (0.0267)	0.6268 (0.1033)	0.6269 (0.0741)	0.5788 (0.0433)	**0.6284** (0.0036)
Router	0.6522 (0.0910)	0.6514 (0.0004)	0.6518 (0.0021)	0.6522 (0.0003)	0.6512 (0.0005)	0.6511 (**0.1633**)	0.6526 (0.1052)	0.6530 (0.0846)	**0.9550** (0.0196)	0.6530 (0.0072)
USAir	0.9532 (0.5820)	0.9249 (0.0004)	0.9144 (0.0102)	0.9142 (0.0101)	0.8813 (0.0037)	0.7789 (0.0067)	0.9655 (0.6232)	**0.9721** (**0.6348**)	0.9087 (0.4782)	0.9527 (0.0233)
Yeast	0.7348 (0.1746)	0.7323 (0.0044)	0.7328 (0.0033)	0.7329 (0.0033)	0.7329 (0.0037)	0.7318 (0.0012)	0.7359 (**0.2244**)	0.7358 (0.1826)	**0.8872** (0.0333)	0.7345 (0.0192)

generally obtain an interval instead of a certain degree. Currently, we shall view the simulation results from the point view of Dempster–Shafer and see that the proposed method is indeed a better solution for decision making. As mentioned above, the main drawbacks are twofold. First, the boundaries between states of the traditional transferring similarity are overly sensitive so that unreasonable phenomenon appears in simulation results (e.g., $\rho = 0.51$ regards as positive correlation but $\rho = 0.49$ is treated as uncorrelated relation). Second, the traditional model can hardly obtain a precise number to represent the transferring similarity. As a result, traditional transferring similarity model is difficult to handle decision making problem. Moreover, the former drawback has been settled by discretizing the continuous correlation interval. Let us now divert our attention to the latter one.

Recall that the usual practice for fusing information is taking the average value of various results. However, the proposed BTS model transfers the network structure to the framework of Dempster–Shafer by generating BPAs according to the similarity interval, then the evidence fusion function is used to combine multiple evidence instead of taking average values. The BTS model, luckily, could benefit from the polarizability of evidence fusion function.

Recall the given example in this paper, the three basic states N, U and P and the power set of the frame of discernment is $2^{\Theta} = \{N, \{N, U\}, U, \{U, P\}, P\}$. Here, the example of case study is used again to show the polarizability and efficiency of the fusing function.

Example. For a belief transferring similarity model that consists two intermediates A and B which has already been transferred to the mass functions named m_A and m_B

$$m_A(N) = 0, m_A(N, U) = 0, m_A(U) = 0.3425, m_A(U, P) = 0.1750, m_A(P) = 0.4825.$$

$$m_B(N) = 0, m_B(N, U) = 0, m_B(U) = 0.1707, m_B(U, P) = 0.2409, m_B(P) = 0.5884.$$

According to Eq. (9), we can obtain the final mass function m' by combining two mass functions. Namely, $m' = m_A \oplus m_B$. The fusing result is shown as follows.

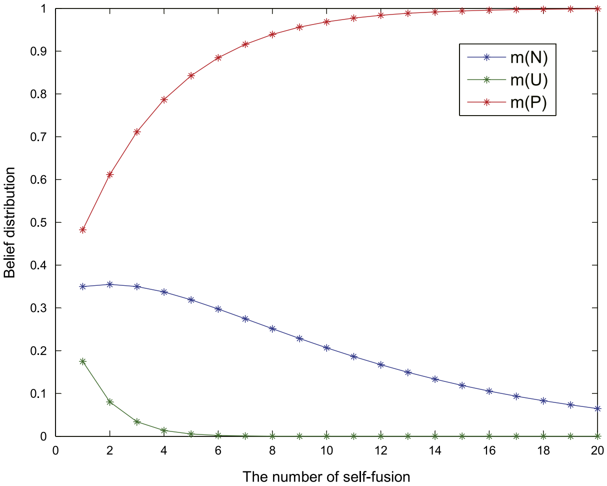
$$m'(\{P\}) = 0.7320,$$

$$m'(\{U\}) = 0.2680,$$

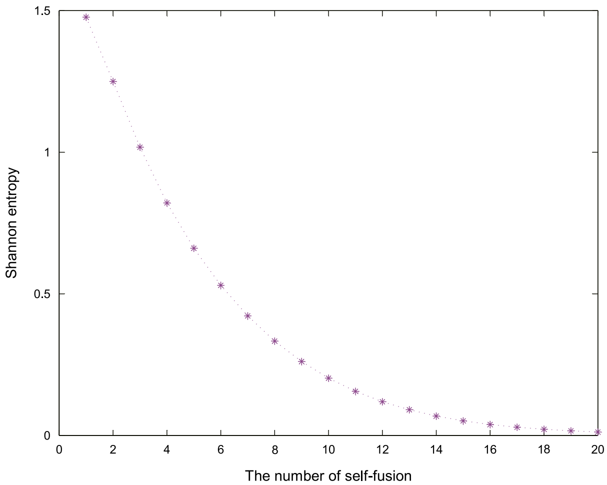
$$m'(\{N\}) = 0.$$

From above fusing result, on the one hand, we can see that the final mass function has larger belief to proposition U than m_A even than m_B. A possible reason for this phenomenon, as we shall subsequently see, is that the fusing function is essentially a special form to polarize belief distribution. That is to say, given a mass function m_i, if m_i supports proposition X, then we fuse m_i with itself, the fusing result $m' = m_i \oplus m_i$ will assign more belief on proposition X.

Let us regard m_A as an example to show the characteristics of self-fusion. As shown in Fig. 8(a), the mass function m_A is fusion with itself over and over again. Even though the initial belief distribution is slightly different. As we can see, the belief of proposition P keeps increasing till converging to 1 since proposition P has higher initial belief, while the belief of other two propositions are declining quickly. This phenomenon implies that the evidence fusing function has the character



(a) The polarized change of belief distribution after doing continuous iterations.



(b) The declining of Shannon entropy after doing continuous iterations.

Fig. 8. The illustration of the property of self-fusing.

that tends to amplify the congenital advantage of the proposition with higher initial belief (e.g., proposition P in this case). Namely,

$$m'(X) \geq m_i(X). \tag{31}$$

On the other hand, from the point view of information science, evidence fusion function tends to decrease the uncertainty of mass function. As we know, Shannon entropy [75] is widely used to measure the uncertainty of an complex system or a simple information with the Bayesian structure. Here, Shannon entropy is used again to illustrate the efficiency of BTS model. As shown in Fig. 8(b), the uncertainty of the mass function m_A is dramatically decreasing with the repeated self-fusing process. Due to this property of evidence fusion, BTS model is more effective than traditional transferring similarity especially when applied to decision making problem. However, the proposed BTS model also has drawbacks which inherents from the frame of D–S theory. It may get a irregular and unreasonable fusing result m' or we cannot even combine the evidences when some evidences are highly conflicting. For a instance, given two evidence m_1 and m_2 where

$$m_1(\{A\}) = 1,\ m_1(\{B\}) = 0,$$

$$m_2(\{B\}) = 0,\ m_2(\{B\}) = 1.$$

When we try to combine m_1 and m_2, we will find the coefficient $\frac{1}{k}$ cannot be obtained since the $K = 0$. In this case, some necessary precautions need to be carried out to handle the conflicts before fusing process. Some previous studies have fully discussed and proposed some models to eliminate the conflicts between evidences.

7. Conclusion

In this paper, the limitations of the classical transferring similarity have been pointed out. A novel model, which is called Belief Transferring Similarity (BTS) model, is proposed to deal with the uncertainty of fuzzy boundaries. The BTS model, as the generalization of the classical transferring similarity, addresses the issue of the sparsity of dataset by considering the high-order similarity. The BTS model transforms uncertain interval to a certain state using fuzzy systems theory. Moreover, an illustrative example is presented to show the new model's procedure, and the BTS model is implemented to compare the accuracy with nine link prediction methods in nine networks. Due to the strong adaptability of BTS model, it is very promising to be applied to many other potential applications.

Acknowledgments

The authors greatly appreciate the anonymous reviews' suggestions and the editor's encouragement. The work is partially supported by National Natural Science Foundation of China (Grant Nos. 61573290, 61503237).

References

[1] I. Mashal, O. Alsaryrah, T.-Y. Chung, Performance evaluation of recommendation algorithms on Internet of Things services, Physica A 451 (2016) 646–656. http://dx.doi.org/10.1016/j.physa.2016.01.051.
[2] L.-J. Chen, Z.-K. Zhang, J.-H. Liu, J. Gao, T. Zhou, A vertex similarity index for better personalized recommendation, Physica A 466 (2017) 607–615. http://dx.doi.org/10.1016/j.physa.2016.09.057.

[3] G. Chen, T. Gao, X. Zhu, H. Tian, Z. Yang, Personalized recommendation based on preferential bidirectional mass diffusion, Physica A 469 (2017) 397–404. http://dx.doi.org/10.1016/j.physa.2016.11.091.

[4] L. Hou, K. Liu, J. Liu, R. Zhang, Solving the stability-accuracy-diversity dilemma of recommender systems, Physica A 468 (2017) 415–424. http://dx.doi.org/10.1016/j.physa.2016.10.083.

[5] J. Bao, Y. Zheng, D. Wilkie, M. Mokbel, Recommendations in location-based social networks: a survey, Geoinformatica 19 (3) (2015) 525–565. http://dx.doi.org/10.1007/s10707-014-0220-8.

[6] M. Ramezani, J. Haddad, N. Geroliminis, Dynamics of heterogeneity in urban networks: aggregated traffic modeling and hierarchical control, Transp. Res. B 74 (2015) 1–19. http://dx.doi.org/10.1016/j.trb.2014.12.010.

[7] D. Chen, L. Lü, M.-S. Shang, Y.-C. Zhang, T. Zhou, Identifying influential nodes in complex networks, Physica A 391 (4) (2012) 1777–1787. http://dx.doi.org/10.1016/j.physa.2011.09.017.

[8] R. Zhang, B. Ashuri, Y. Deng, A novel method for forecasting time series based on fuzzy logic and visibility graph, Adv. Data Anal. Classif. 11 (4) (2018) 759–783. http://dx.doi.org/10.1007/s11634-017-0300-3.

[9] R.A. Barrio, T. Govezensky, Ifego Ruiz-Gutirrez, K.K. Kaski, Modelling trading networks and the role of trust, Physica A 471 (2017) 68–79. http://dx.doi.org/10.1016/j.physa.2016.11.144.

[10] M. Ramezani, P. Moradi, F. Akhlaghian, A pattern mining approach to enhance the accuracy of collaborative filtering in sparse data domains, Physica A 408 (2014) 72–84. http://dx.doi.org/10.1016/j.physa.2014.04.002.

[11] B. Moradabadi, H. Beigy, A new real-coded Bayesian optimization algorithm based on a team of learning automata for continuous optimization, Genet. Program. Evolvable Mach. 15 (2) (2014) 169–193.

[12] X. Han, S. Cao, Z. Shen, B. Zhang, W.-X. Wang, R. Cressman, H.E. Stanley, Emergence of communities and diversity in social networks, Proc. Natl. Acad. Sci. USA 114 (11) (2017) 2887–2891. http://dx.doi.org/10.1073/pnas.1608164114.

[13] P. Moradi, S. Ahmadian, F. Akhlaghian, An effective trust-based recommendation method using a novel graph clustering algorithm, Physica A 436 (2015) 462–481. http://dx.doi.org/10.1016/j.physa.2015.05.008.

[14] J. Zhang, Q. Peng, S. Sun, C. Liu, Collaborative filtering recommendation algorithm based on user preference derived from item domain features, Physica A 396 (2014) 66–76. http://dx.doi.org/10.1016/j.physa.2013.11.013.

[15] X. Shi, M.-S. Shang, X. Luo, A. Khushnood, J. Li, Long-term effects of user preference-oriented recommendation method on the evolution of online system, Physica A 467 (2017) 490–498. http://dx.doi.org/10.1016/j.physa.2016.10.033.

[16] Z.-K. Zhang, T. Zhou, Y.-C. Zhang, Personalized recommendation via integrated diffusion on usercitemctag tripartite graphs, Physica A 389 (1) (2010) 179–186. http://dx.doi.org/10.1016/j.physa.2009.08.036.

[17] J. Bobadilla, F. Ortega, A. Hernando, A. Gutirrez, Recommender systems survey, Knowl.-Based Syst. 46 (Supplement C) (2013) 109–132. http://dx.doi.org/10.1016/j.knosys.2013.03.012.

[18] L. Lü, T. Zhou, Link prediction in complex networks: A survey, Physica A 390 (6) (2011) 1150–1170. http://dx.doi.org/10.1016/j.physa.2010.11.027.

[19] Q. Zhang, M. Li, Y. Deng, Measure the structure similarity of nodes in complex networks based on relative entropy, Physica A 491 (2018) 749–763.

[20] K. ke Shang, M. Small, W. sheng Yan, Link direction for link prediction, Physica A 469 (2017) 767–776.

[21] L. Lü, T. Zhou, Link prediction in weighted networks: The role of weak ties, Europhys. Lett. 89 (1) (2010) 18001.

[22] C. Ma, T. Zhou, H.-F. Zhang, Playing the role of weak clique property in link prediction: A friend recommendation model, Sci. Rep. 6 (2016). http://dx.doi.org/10.1038/srep30098.

[23] L.A. Adamic, E. Adar, Friends and neighbors on the web, Soc. Networks 25 (3) (2003) 211–230.

[24] T. Zhou, L. Lü, Y.-C. Zhang, Predicting missing links via local information, Eur. Phys. J. B 71 (4) (2009) 623–630. http://dx.doi.org/10.1140/epjb/e2009-00335-8.

[25] S. Redner, Networks - Teasing out the missing links, Nature 453 (7191) (2008) 47–48. http://dx.doi.org/10.1038/453047a.

[26] A. Biswas, B. Biswas, Community-based link prediction, Multimedia Tools Appl. 76 (18) (2017) 18619–18639. http://dx.doi.org/10.1007/s11042-016-4270-9.

[27] M.A. Al-garadi, K.D. Varathan, S.D. Ravana, Identification of influential spreaders in online social networks using interaction weighted k-core decomposition method, Physica A 468 (2017) 278–288. http://dx.doi.org/10.1016/j.physa.2016.11.002.

[28] L. Lü, D. Chen, X. Ren, Q. Zhang, Y. Zhang, T. Zhou, Vital nodes identification in complex networks, Phys. Rep. 650 (2016) 1–63. http://dx.doi.org/10.1016/j.physrep.2016.06.007. Vital nodes identification in complex networks.

[29] L. Tang, K. Jing, J. He, H.E. Stanley, Robustness of assembly supply chain networks by considering risk propagation and cascading failure, Physica A 459 (2016) 129–139. http://dx.doi.org/10.1016/j.physa.2016.04.030.

[30] W.-X. Jin, P. Song, G.-Z. Liu, H.E. Stanley, The cascading vulnerability of the directed and weighted network, Physica A 427 (2015) 302–325. http://dx.doi.org/10.1016/j.physa.2015.02.035.

[31] J. Toruniewska, K. Suchecki, J.A. Holyst, Unstable network fragmentation in co-evolution of potts spins and system topology, Physica A 460 (2016) 1–15. http://dx.doi.org/10.1016/j.physa.2016.04.029.

[32] R. Karan, B. Biswal, A model for evolution of overlapping community networks, Physica A 474 (2017) 380–390. http://dx.doi.org/10.1016/j.physa.2017.01.083.

[33] C. Song, S. Havlin, H. Makse, Self-similarity of complex networks, Nature 433 (7024) (2005) 392–395. http://dx.doi.org/10.1038/nature03248.

[34] Z. Yang, Z. Li, T. Wu, L. Wang, Role of recommendation in spatial public goods games, Physica A 392 (9) (2013) 2038–2045. http://dx.doi.org/10.1016/j.physa.2012.11.024.

[35] A.L. Schmidt, F. Zollo, M. Del Vicario, A. Bessi, A. Scala, G. Caldarelli, H.E. Stanley, W. Quattrociocchi, Anatomy of news consumption on facebook, Proc. Natl. Acad. Sci. USA 114 (12) (2017) 3035–3039. http://dx.doi.org/10.1073/pnas.1617052114.

[36] M. Ramezani, F. Yaghmaee, A novel video recommendation system based on efficient retrieval of human actions, Physica A 457 (2016) 607–623. http://dx.doi.org/10.1016/j.physa.2016.03.101.

[37] B. Lika, K. Kolomvatsos, S. Hadjiefthymiades, Facing the cold start problem in recommender systems, Expert Syst. Appl. 41 (4, Part 2) (2014) 2065–2073. http://dx.doi.org/10.1016/j.eswa.2013.09.005.

[38] S.M. McNee, J. Riedl, J.A. Konstan, Being accurate is not enough: How accuracy metrics have hurt recommender systems, in: CHI'06 Extended Abstracts on Human Factors in Computing Systems, CHI EA'06, ACM, New York, NY, USA, 2006, pp. 1097–1101. http://dx.doi.org/10.1145/1125451.1125659.

[39] N. Hurley, M. Zhang, Novelty and diversity in Top-N recommendation - analysis and evaluation, ACM Trans. Internet Technol. 10 (4) (2011). http://dx.doi.org/10.1145/1944339.1944341.

[40] Q. Guo, W.-J. Song, L. Hou, Y.-L. Zhang, J.-G. Liu, Effect of the time window on the heat-conduction information filtering model, Physica A 401 (2014) 15–21. http://dx.doi.org/10.1016/j.physa.2014.01.012.

[41] L. Hou, X. Pan, Q. Guo, J.-G. Liu, Memory effect of the online user preference, Sci. Rep. 4 (2014). http://dx.doi.org/10.1038/srep06560.

[42] J. Gao, T. Zhou, Evaluating user reputation in online rating systems via an iterative group-based ranking method, Physica A 473 (Supplement C) (2017) 546–560. http://dx.doi.org/10.1016/j.physa.2017.01.055.

[43] R. Katarya, O.P. Verma, Recent developments in affective recommender systems, Physica A 461 (2016) 182–190. http://dx.doi.org/10.1016/j.physa.2016.05.046.

[44] D. Matesanz, G.J. Ortega, On business cycles synchronization in Europe: A note on network analysis, Physica A 462 (2016) 287–296. http://dx.doi.org/10.1016/j.physa.2016.06.097.
[45] D. Sun, T. Zhou, J.-G. Liu, R.-R. Liu, C.-X. Jia, B.-H. Wang, Information filtering based on transferring similarity, Phys. Rev. E 80 (1, 2) (2009). http://dx.doi.org/10.1103/PhysRevE.80.017101.
[46] M.M. Deza, E. Deza, Encyclopedia of distances, in: Encyclopedia of Distances, Springer, 2009, pp. 1–583.
[47] J. Benesty, J. Chen, Y. Huang, I. Cohen, Pearson correlation coefficient, in: Noise Reduction in Speech Processing, Springer, 2009, pp. 1–4.
[48] L. Fei, H. Wang, L. Chen, Y. Deng, A new vector valued similarity measure for intuitionistic fuzzy sets based on OWA operators, Iran. J. Fuzzy Syst. 15 (5) (2018) 31–49.
[49] L. Yin, H. Zheng, T. Bian, Y. Deng, An evidential link prediction method and link predictability based on Shannon entropy, Physica A 482 (2017) 699–712. http://dx.doi.org/10.1016/j.physa.2017.04.106.
[50] F. Li, Y. Qian, J. Wang, J. Liang, Multigranulation information fusion: A Dempster-Shafer evidence theory-based clustering ensemble method, Inf. Sci. 378 (2017) 389–409. http://dx.doi.org/10.1016/j.ins.2016.10.008.
[51] B. Kang, G. Chhipi-Shrestha, Y. Deng, K. Hewage, R. Sadiq, Stable strategies analysis based on the utility of z-number in the evolutionary gam, Appl. Math. Comput. (2017). http://dx.doi.org/10.1016/j.amc.2017.12.006.
[52] F. Xiao, M. Aritsugi, Q. Wang, R. Zhang, Efficient processing of multiple nested event pattern queries over multi-dimensional event streams based on a triaxial hierarchical model, Artif. Intell. Med. 72 (2016) 56–71.
[53] F. Xiao, C. Zhan, H. Lai, L. Tao, Z. Qu, New parallel processing strategies in complex event processing systems with data streams, Int. J. Distrib. Sens. Netw. 13 (8) (2017) 1–15.
[54] X. Zheng, Y. Deng, Dependence assessment in human reliability analysis based on evidence credibility decay model and iowa operator, Ann. Nucl. Energy 112 (2018) 673–684.
[55] A.P. Dempster, Upper and lower probabilities induced by a multivalued mapping, Ann. Math. Statist. (1967) 325–339.
[56] G. Shafer, et al., A Mathematical Theory of Evidence, Vol. 1, Princeton University Press, Princeton, 1976.
[57] B. Kang, G. Chhipi-Shrestha, Y. Deng, et al., Development of a predictive model for clostridium difficile infection incidence in hospitals using Gaussian mixture model and Dempster-Shafer theroy, Stoch. Environ. Res. Risk Assess. (2017). http://dx.doi.org/10.1007/s00477-017-1459-z.
[58] H. Mo, Y. Deng, A new aggregating operator in linguistic decision making based on D numbers, Internat. J. Uncertain. Fuzziness Knowledge-Based Systems 24 (6) (2016) 831–846.
[59] X. Deng, Y. Deng, D-AHP method with different credibility of information, Soft Comput. (2018). http://dx.doi.org/10.1007/s00500-017-2993-9. Published online.
[60] F. Xiao, An intelligent complex event processing with Dnumbers under fuzzy environment, Math. Probl. Eng. 2016 (1) (2016) 1–10.
[61] T. Bian, H. Zheng, L. Yin, Y. Deng, Failure mode and effects analysis based on Dnumbers and topsis, Qual. Reliab. Eng. Int. (2018).. http://dx.doi.org/10.1002/qre.2268. Article ID: QRE2268.
[62] B. Liu, Y. Hu, Y. Deng, New failure mode and effects analysis based on d numbers downscaling method, Int. J. Comput. Commun. Control 13 (2) (2018) 42–49.
[63] X. Zhou, X. Deng, Y. Deng, S. Mahadevan, Dependence assessment in human reliability analysis based on D numbers and AHP, Nucl. Eng. Des. 313 (2017) 243–252.
[64] H. Zheng, Y. Deng, Y. Hu, Fuzzy evidential influence diagram and its evaluation algorithm, Knowl.-Based Syst. 131 (2017) 28–45.
[65] Y. Dong, J. Wang, F. Chen, Y. Hu, Y. Deng, Location of facility based on simulated annealing and "ZKW" algorithms, Math. Probl. Eng. 2017 (2017). http://dx.doi.org/10.1155/2017/4628501. Article ID 4628501.
[66] T. Liu, Y. Deng, F. Chan, Evidential supplier selection based on DEMATEL and game theory, Int. J. Fuzzy Syst. (2017). http://dx.doi.org/10.1007/s40815-017-0400-4.
[67] P. Smets, Decision making in the tbm: the necessity of the pignistic transformation, Internat. J. Approx. Reason. 38 (2) (2005) 133–147. http://dx.doi.org/10.1016/j.ijar.2004.05.003.
[68] P. Smets, R. Kennes, The transferable belief model, Artficial Intell. 66 (2) (1994) 191–234. http://dx.doi.org/10.1016/0004-3702(94)90026-4.
[69] L. Montangie, F. Montani, Higher-order correlations in common input shapes the output spiking activity of a neural population, Physica A 471 (2017) 845–861. http://dx.doi.org/10.1016/j.physa.2016.12.002.
[70] H. Zheng, Y. Deng, Evaluation method based on fuzzy relations between Dempster-Shafer belief structure, Int. J. Intell. Syst. (2017). http://dx.doi.org/10.1002/int.21956.
[71] H. Xu, Y. Deng, Dependent evidence combination based on shearman coefficient and pearson coefficient, IEEE Access (2018). http://dx.doi.org/10.1109/ACCESS.2017.2783320.
[72] V.-D. Nguyen, V.-N. Huynh, Two-probabilities focused combination in recommender systems, Int. J. Approximate Reasoning 80 (2017) 225–238. http://dx.doi.org/10.1016/j.ijar.2016.09.005.
[73] B. Moradabadi, M.R. Meybodi, Link prediction based on temporal similarity metrics using continuous action set learning automata, Physica A 460 (2016) 361–373.
[74] G. Guo, J. Zhang, N. Yorke-Smith, A novel bayesian similarity measure for recommender systems, in: Proceedings of the 23rd International Joint Conference on Artificial Intelligence, IJCAI, 2013, pp. 2619–2625.
[75] C.E. Shannon, A mathematical theory of communication, ACM SIGMOBILE Mob. Comput. Commun. Rev. 5 (1) (2001) 3–55.