Journal of Computer Science and Cybernetics, V.38, N.2 (2022), 147–163
DOI no. 10.15625/1813-9663/38/2/16749

MODELING COMPUTATIONAL TRUST BASED ON INTERACTION EXPERIENCE AND REPUTATION WITH USER INTERESTS IN SOCIAL NETWORK

DINH QUE TRAN[1], PHUONG THANH PHAM[2,*]

[1] *Post and Telecommunications Institute of Technology, Hanoi, Vietnam*
[2] *Department of Mathematics and Informatics, ThangLong University, Hanoi, Vietnam*

Abstract. Computational trust is a reliability among peers that plays a crucial role in sharing information, decision making, searching or attracting recommendations in intelligent systems and social networks. Several trust models have been proposed in literature and most of them focus on investigating interaction forms rather than analyzing contexts such as comments, posts being dispatched by users. This paper is to present a novel model of estimating trustworthiness of a truster on a trustee based on experience trust and reputation trust from some community within the context of user's topic interests. Firstly, we construct a measure of experience topic-aware trust which is defined as a function of degrees of interaction from a truster to some trustee and a degree of trustee's interests in topics. Secondly, we construct a measure of reliability degree of community on some trustee by means of a function which is computed via degrees of reliability of truster on members of the community and similarity of these members with the trustee. Thirdly, we propose a composition function for estimating an overal topic-aware trust based on experience topic-aware trust and the reputation topic-aware trust. Our experimental results show that the degree of experience topic-aware trust depends on interaction degree among truster and trustee more than on trustee's interest degree. They also indicate that the overall topic-aware trust estimation depends on reputation from community more than user's own experience evaluation.

Keywords. Computational trust; Context; Intelligent systems; Interaction; Interests; Social network; Reputation.

1. INTRODUCTION

Trust is a reliability which a user (truster) has on his own partners (trustees) in his interaction process. It has become a crucial factor to share knowledge or to coordinate in actions with each others in systems such as recommender and decision making or search engines. Trust has been considered from research fields including sociology, psychology, economics and computer science [10]. There are various models of computational trust that have been proposed in literature [2, 3, 10–15, 21, 22]. In social networks, peers utilize their own tags, comments, post, etc., to annotate and organize items for searching or sharing viewpoints and opinions as well. Such text entries are types of meta-data composed of

*Corresponding author.
E-mail addresses: tdque@yahoo.com (D.Q.Tran); ppthanh216@gmail.com (P.T.Thanh)

keywords or terms to introduce bookmarks, article titles, comments of items or digital images etc. They have contributed to discovering user interests for various real world applications. These issues have attracted a large number of researchers from academy as well as application development [4–8, 13, 18, 19].

Computational trust models in literature can be categorized into three groups:

(i) Models utilizing the past interaction experience to estimate trustworthiness of peers in distributed systems. The interaction based approach has been utilized widely in multi-agent systems, P2P systems [6–8, 10, 11];

(ii) Models exploiting contexts of intereraction among peers such as tags, comments or user's profiles on social networks to estimate trustworthiness among users. These data resources are utilized to determine uer's interests, similarity as well as relationship between peers [4, 15];

(iii) Hybrid models combining interaction scores and degrees of interests or similarity of users [4, 6].

Along with the hybrid approach, we develop it furthermore by constructing a computational function which is a combination of two factors: (i) experience topic-aware trust; (ii) reputation topic-aware trust. In order to construct experience topic-aware trust, we analyse messages, named *entries*, dispatched by users to determine their interest in topics and utilize the score of interaction among users. And in turn, we share with other work [10], LoTrust [4], TidalTrust [5], SWTrust [8], TrustWalker [7] in computing the interaction score by relying on the assumption of frequency of interaction of closest users. In our work, the experience topic-aware trust is estimated by means of a composition function of interaction scores of a truster with trustee and trustee's interest degrees. However, in contrast with other ones such as [4] in which user's interest is extracted from his profiles via SPARQL Query Language, we make use of the semantic extension of words by means of wikipedia proposed by Kang et al., [3] and Gabrilovich et al. [9]. We analyse entries into words by the technique of tf-idf [1, 15] to compute the weight of word in a document for representing vectors of entries and topics. Based on such a vector model, we define similarity measures and interest degrees. And then computational function of estimating trustworthiness of users is defined by means of the degrees of interest, interaction scores.

Reputation trust is defined as reliability which is resulted from some community. Some work makes use of the propagation of trust estimation via the graph structure of network such as TidalTrust [5], SWTrust [8], TrustWalker [7] to construct the reputation trust. Their approach selects some path for computation to avoid computational complexity. For example, selecting the shortest path connecting the truster and trustee. The problem of this approach is that there is no basics in theory for such a path selection. Our approach is completely diffferent compared with these studies. We first construct a hierarchy structure of peers based on interaction layers with a truster and then define a common community of both a truster and some trustee. And the reputation topic-aware trust is estimated by means of average of all experience topic-aware trusts of the truster and similarity of truster with them. The overall trust, called topic-aware trust, is determined as a composition function of reputation and experience topic-aware trust. In this paper, we revise, upgrade and develop our previous studies [15–17] and our research results are interpreted in the contributions as follows:

- We propose a similarity measure among users which is defined as a composition function of profile and interest similarities. The interest similarities of users are computed by means of Pearson correlation of interest measures, named Max, Cor and Sum, which have been proposed previously by ourselves. The profile similarity is defined by the traditional cosin similairity of entries dispatched by users. In order to construct entry and interest vectors for such a computation, we perform a semantics extension with wikipedia of entries and then make use of tf-idf to compute word weights. We perform experimental evaluations to compare affects of three above interest measures on the similarity with respect to the mean deviation. Our experimental results show that the selected interest measure Max gets the lowest deviation.

- We upgrade the function of topic-aware experience trust among truster and trustee, which has been proposed by ourselves [16], and construct a reputation function of estimating a reliability degree of community on some trustee. The function of topic-aware experience trust shows the closeness degree of two users and the interest degree interpreted by user's expert degree on some topic. The reputation function is to estimate degrees of reliability of members in some community on some trustee. It is computed via degrees of reliability of truster on members of the community and similarity of these members with the trustee. We perform experiments to evaluate how affects of two factors on trustworthiness. Our experimental results show that the trust measure depends on interaction degree more than on interest degree. Furthremore, with the same interaction degree, the higher an interest degree is, the higher the corresponding experience trust measure obtains.
- We propose a overal topic-aware trust, which is a combination function of two factors: topic-aware experience trust and the reputation topic-aware trust. We conduct experiments to consider how affections of two factors. The experimental results show that topic-aware trust estimation depends on reputation more than user's own experience evaluation.

The remainder of this paper is structured as follows. Section 2 presents a model of social network and vectorial representation of entries and topics. Section 3 describes interest degrees, user's profile and similarity. Section 4 is devoted to presenting a trust computation model based on interaction and reputation with interest context. Section 5 describes experimental evaluations. Conclusion is presented in Section 6.

2. MODELING SOCIAL NETWORK, ENTRIES AND TOPICS

This section presents briefly the model of social network, a hierarchy structure of users, entry and topic [15–17].

2.1. Model of social network

A social network is defined as a directed graph $\mathcal{S} = (\mathcal{U}, \mathcal{I}, \mathcal{E}, \mathcal{T})$, where

- $\mathcal{U} = \{u_1, \ldots, u_n\}$ is a set of users/peers in a social network.
- $\mathcal{I}$ is a set of all interactions/connections I_{ij} from u_i to u_j, which occurs when u_i dispatches u_j via some "wall" posts, comments, likes, opinions etc. $\|I_{ij}\|$ denotes the number of elements in I_{ij}.

- $\mathcal{E} = \{E_1, \ldots, E_n\}$ is the set of entries dispatched by users in $\mathcal{U}$. $E_i = \{e_{i1}, \ldots, e_{in_i}\}$ are entries delivered by u_i. An *entry* is a brief text piece given by users on items such as papers, books, films, videos, events etc.

- $\mathcal{T} = \{T_1, \ldots, T_p\}$ is a set of *topics* in which each topic is defined as a set of words or terms.

2.2. Hierarchy structure of peers

For each user u_i, we denote L_i^1 to be the set of all users who have direct interaction with u_i, L_i^2 the set of all users having interaction with some user in L_i^1 but not with u_i. Recursively, we can define a sequence of k-level L_i^k of user u_i. We have the following statement (for more detail, see [11]). For every source peer u_i, there exists a number h_i such that $L_i^0, \ldots, L_i^{h_i}$ are subsets of $\mathcal{U}$, called k-neighbors of u_i, and satisfy the following conditions:

1. For every $v \in L_I^k$ $(k = 2, \ldots, h_i)$, v not being interacted with any one in $\cup_{l=0}^{k-1} L_i^l$.

2. $L_i^k \cap (\cup_{l=0}^{k-1} L_i^l) = \emptyset$, for all $k \geq 1$.

The statement permits us to focus on peers on each layer while computing trustworthiness among them.

2.3. Vectorial representation of entries and topics

The vectorial model for representing texts by means of tf-idf has been widely used in various fields of the computer science such as the information retrieval and text mining [1]. Along with work related to extending semantics [2, 3, 9], we utilize the n-gram to extract a text into words and enrich these bags of words into semantics words based on wikipedia (https://vi.wikipedia.org/wiki/). This section reformulates the model in some formal way for our paper. The purpose is to apply the approach to vectorizing entries and topics with word weights in texts. We follow the steps for preprocessing these short texts to obtain bags of words with semantics:

(i) Using the n-gram technique for extracting a text into terms or words;

(ii) Enrich these terms with semantics from wikipedia.

And from now on, in this paper, any document or text is always considered as a set of terms. We make use of the technique tf$-$ idf$(d, D_i) = \text{tf}(d, D_i) \times \text{idf}(d, \mathcal{D})$ for vectorial representation of such entries and topics, where $\text{tf}(d, D_i)$ is the frequency the term d appears in D_i and $\text{idf}(d, \mathcal{D}) = \log(\dfrac{\|\mathcal{D}\|}{1 + \|\{D_i | d \in D_i\}\|})$. The vector representation in the general form is described as follows.

Given a collection of documents $D = \{D_1, \ldots, D_p\}$, each of which is represented as set of terms or words $D_i = \{d_{i1}, \ldots, d_{i_{pi}}\}$. Let $V = \{v_1, \ldots, v_q\}$ be a set of all distinct terms in the whole collection. The weight of term $d \in V$ w.r.t. D_i is defined by the formula $w_d = \text{tf}(d, D_i) \times \text{idf}(d, D)$. And then each D_i is represented as a q-dimension vector $\mathbf{D}_i = (w_1, \ldots, w_q)$, where $w_k = \text{tf}(v_k, D_i) \times \text{idf}(v_k, D)$, $k = 1, \ldots, q$. We utilize the technique to represent entries and topics in vectors, which are described in the rest of this subsection.

2.3.1. Entry vectors

Suppose that $E_i = \{e_{i1}, \ldots, e_{in_i}\}$ and $E_j = \{e_{j1}, \ldots, e_{jn_j}\}$ are two sets of entries dispatched by users u_i, u_j, respectively. Let V_{ij} be a set of distinct terms occurring in both E_i and E_j. Entry vectors e^j_{il}, e^i_{jk} are defined as follows

$$\mathbf{e^j_{il}} = (e^1_{il}, \ldots, e^{\|V_{ij}\|}_{il}),\ l = 1, \ldots, n_i, \tag{1}$$

$$\mathbf{e^i_{jk}} = (e^1_{jk}, \ldots, e^{\|V_{ij}\|}_{jk}),\ k = 1, \ldots, n_j, \tag{2}$$

in which, for each $v_r \in V_{ij}$, $e^r_{il} = \text{tf}(v_r, e_{il}) \times \text{idf}(v_r, E_i)$, $e^r_{jk} = \text{tf}(v_r, e_{jk}) \times \text{idf}(v_r, E_j)$. This representation will be used to estimate the profile similarity of two users which is presented in the next section.

2.3.2. Topic vector and topic entry vector

This subsection describes the vectorial representation of topics $\mathcal{T} = \{T_1, \ldots, T_p\}$ and entries $E_i = \{e_{i1}, \ldots, e_{in_i}\}$ dispatched by user u_i according to topics.

Suppose that $V_T = \{v_1, \ldots, v_q\}$ is a set of q distinct terms in all $T_i \in \mathcal{T}$. Each topic T_i is defined to be a weighted vector as follows

$$\mathbf{t_i} = (w_{i1}, \ldots, w_{iq}), \tag{3}$$

where $w_{ik} = \text{tf}(v_k, T_i) \times \text{idf}(v_k, \mathcal{T})$, for all $v_k \in V_T$, $k = 1, \ldots, q$. This is a q-dimension vector and called the topic vector.

Each entry $e_{il} \in E_i$ dispatched by u_i is represented in vector w.r.t. topics $T_i \in \mathcal{T}$, which is defined as follows

$$\mathbf{e^t_{il}} = (e^1_{il}, \ldots, e^q_{il}), \tag{4}$$

where $e^k_{il} = \text{tf}(v_k, e_{il}) \times \text{idf}(v_k, E_i)$, all $v_k \in V_T$, $k = 1, \ldots, q$. This is a q-dimension vector and called a topic entry vector.

This representation of vectors is used to estimate the interest measures of users w.r.t. topics and interest similarity of users, which are presented in the next section.

3. DEGREES OF USER'S INTERESTS AND SIMILARITY

This section upgrades formulas of computing user's interest degrees, which have been described in our previous work [16]. We make use of Pearson correlation measure to determine relationship between entries and topics and cosin measure to estimate user's similairty in profiles. Based on these similarity measures, we construct the overall similarity of users which is a combination function of profile and interest similarities.

3.1. Correlation and interest degree

Given two vectors $\mathbf{u} = (u_1, \ldots, u_m)$ and $\mathbf{v} = (v_1, \ldots, v_m)$ with different elements in each vector, the correlation of these two vectors is given by the following formula

$$\text{correl}(\mathbf{u}, \mathbf{v}) = \frac{\sum_i (u_i - \bar{u})(v_i - \bar{v})}{\sqrt{\sum_i (u_i - \bar{u})^2} \times \sqrt{\sum_i (v_i - \bar{v})^2}}, \tag{5}$$

where $\bar{u} = \frac{1}{m}(\sum_{i=1}^{m} u_i)$ and $\bar{v} = \frac{1}{m}(\sum_{i=1}^{m} v_i)$. It is clear that values of correl(x, y) are in $[-1, 1]$. We utilize the function $f(x) = \frac{(x+1)}{2}$ to bound values of correl(x, y) into the unit interval $[0, 1]$. It means that instead of the formula given in (5), the following one (6) will be applied in this paper

$$\mathrm{cor}(\mathbf{u}, \mathbf{v}) = \frac{\frac{\sum_i (u_i - \bar{u})(v_i - \bar{v})}{\sqrt{\sum_i (u_i - \bar{u})^2} \times \sqrt{\sum_i (v_i - \bar{v})^2}} + 1}{2}. \tag{6}$$

Definition 1. Let $\mathcal{P}(E_i)$ be a set of all subsets of entries E_i given by $u_i \in \mathcal{U}$, and $\mathcal{P}(\mathcal{E}) = \bigcup_{u_i \in \mathcal{U}} \mathcal{P}(E_i)$. A function $f : \mathcal{U} \times \mathcal{P}(\mathcal{E}) \times \mathcal{T} \to [0, 1]$ is called an interest measure iff it satisfies the condition $f(u_i, Y_1, t) \leq f(u_i, Y_2, t)$, for all $Y_1, Y_2 \in \mathcal{P}(E_i)$ such that $Y_1 \subseteq Y_2$.

It is easy to prove the following proposition.

Proposition 1. *A function $f_{interest} : \mathcal{U} \times \mathcal{P}(\mathcal{E}) \times \mathcal{T} \to [0, 1]$ is an interest measure if and only if it satisfies the following conditions*

1. *If $cor(\mathbf{e_{i,k}}, \mathbf{t_j}) \geq cor(\mathbf{e_{i,k}}, \mathbf{t_l})$, then $f_{interest}(u_i, e_i, t_j) \geq f_{interest}(u_i, e_i, t_l)$.*
2. *If $cor(\mathbf{e_{i,k}}, \mathbf{t_h}) \geq cor(\mathbf{e_{j,l}}, \mathbf{t_h})$, then $f_{interest}(u_i, e_i, t_h) \geq f_{interest}(u_j, e_i, t_h)$.*

An entry e_{ij} is called θ-entry w.r.t. topic t_k if and only if $\mathrm{cor}(\mathbf{e^t_{ij}}, \mathbf{t_k}) \geq \theta$, where $0 < \theta \leq 1$ is a given threshold. A revised proposition of the statement presented in our previous work [16] is stated as follows.

Proposition 2. *Suppose $\|E_i\|$ is the number of elements in E_i and n_i^t is the number of θ-entries concerned with the topic t given by u_i. The following are interest measures:*

1. $\mathrm{intMax}(u_i, t) = \max_j(\mathrm{cor}(\mathbf{e^t_{ij}}, t)$.
2. $\mathrm{intCor}(u_i, t) = \dfrac{\sum_j \mathrm{cor}(\mathbf{e^t_{ij}}, t)}{\|E_i\|}$.
3. $\mathrm{intSum}(u_i, t) = \dfrac{1}{2}\left(\dfrac{n_i^t}{\sum_{l \in \mathcal{T}} n_i^l} + \dfrac{n_i^t}{\sum_{u_k \in \mathcal{U}, l \in \mathcal{T}} n_k^l}\right)$.

For easy presentation, we denote $intX(u_i, t)$ to be one of the above measures, in which X may be Sum, Cor, Max. The interest vector of users in topics is defined by the following formula

$$\mathbf{u_i^t} = (u_i^1, \ldots, u_i^p), \tag{7}$$

in which $u_i^k = \mathrm{int}X(u_i, t)$ is the interest degree of user u_i in topics $t_k \in \mathcal{T}$ $(k = 1, \ldots, p)$, X may be Sum, Max, Cor. Topic vectors are computed by means of Algorithm 1.

Algorithm 1 Computing topic vector of u_i on topics t

Input: The set of topics $\mathcal{T} = \{t_1, t_2, ..., t_p\}$ and the set of users $\mathcal{U} = \{u_1, u_2, ..., u_n\}$ with entries e_{il}

Output: Topic interest vector of each u_i on topics t, computeTopicVector(u_i, t)

1: $t \leftarrow (w_{i1}, \ldots, w_{iq})$
   $//w_{ik} = \text{tf}(v_k, T_i) \times \text{idf}(v_k, \mathcal{T})$, $v_k \in V_T$.
2: $e_{il}^t \leftarrow (e_{il}^1, \ldots, e_{il}^q)$
   $//\ e_{il}^k = \text{tf}(v_k, e_{il}) \times \text{idf}(v_k, E_i)$, $v_k \in V_T$.
3: **for all** t in $\mathcal{T}$ **do**
4: $\quad u_i^t \leftarrow intX(u_i, t)$
5: **end for**
6: $\mathbf{u}_i^t \leftarrow (u_1^1, \ldots, u_i^p)$
7: **return** $\mathbf{u}_i^t$

3.2. Similarity of users

3.2.1. Similarity of interest

Interest similarity of two peers u_i and u_j in topic t is defined as a cosine similarity of two vectors $\mathbf{u}_i^t$ and $\mathbf{u}_j^t$

$$\text{sim}_{\text{int}}^X(u_i, u_j) = \frac{\langle \mathbf{u}_i^t, \mathbf{u}_j^t \rangle}{\|\mathbf{u}_i^t\| \times \|\mathbf{u}_j^t\|}, \tag{8}$$

in which $\langle u, v \rangle$ is the scalar product, $\times$ is the usual multiple operation and $\|.\|$ is the Euclidean length of a vector; X is Max, Cor or Sum up on the selection of interest degree as defined in Proposition 2.

3.2.2. Profile similarity

Gien two peers u_i and u_j. Profile similarity of two peers u_i and u_j is defined as a cosine similarity of two vectors $\mathbf{e_{ik}^j}$ and $\mathbf{e_{jk}^i}$

$$\text{sim}_{\text{prof}}(u_i, u_j) = \frac{\langle \mathbf{e_{ik}^j}, \mathbf{e_{jk}^i} \rangle}{\|\mathbf{e_{ik}^j}\| \times \|\mathbf{e_{jk}^i}\|}, \tag{9}$$

in which $\langle u, v \rangle$ is the scalar product, $\times$ is the usual multiple operation and $\|.\|$ is the Euclidean length of a vector.

3.2.3. User similarity

Based on the definition of similarity of interest and profile, we have the definition of similarity of users as follows.

Definition 2. The similarity between two users u_i and u_j is defined by the weighted composition of their partial similarities and given by the following formula

$$\text{sim}(u_i, u_j) = \alpha \times \text{sim}_{\text{prof}}(u_i, u_j) + \beta \times \text{sim}_{\text{int}}^X(u_i, u_j) \tag{10}$$

where $\alpha, \beta \geq 0$ and $\alpha + \beta = 1$.

4. TRUST ESTIMATION BASED ON INTERACTION EXPERIENCE AND REPUTATION WITH INTEREST CONTEXT

Trust estimation of a user, called truster, on another user, called trustee, is a function of the following parameters: (i) Interaction experience of truster on trustee; (ii) Degrees of interests by trustee on topics; (iii) Reputation trust which is a reliability degree inferred from some community on some trustee. These stages of computation of overall topic-aware trust are described in this section as follows:

- We present a computational function of experience topic-aware trust which is based on interaction and interest degrees. This is an upgraded version of the studies appeared in our previous work [15–17].
- We describe a formula for estimating reputation topic-aware trust, which is infered from evaluation of some community of trustee. Instead of merely using user's interest similarity [17], in this paper we utilize the novel similarity measure of users which is presented in Section 3.
- We present a model of estimating overall topic-aware trust which is a combination function of experience and reputation topic-aware trust.

Definition 3. A function $\text{trust}_{\text{topic}} : \mathcal{U} \times \mathcal{U} \times \mathcal{T} \to [0, 1]$ is called a topic trust function, in which $[0, 1]$ is an unit interval of the real numbers. Given a source peer u_i, a sink peer u_j and a topic t, the value $\text{trust}_{\text{topic}}(u_i, u_j, t) = u_{ij}^t$ means that u_i (truster) has a confidence on u_j (trustee) of topic t w.r.t. the degree u_{ij}^t.

We will describe in steps for constructing the topic-aware trust function.

4.1. Experience topic-aware trust

Experience trust of user u_i on user u_j, denoted $\text{trust}^{\exp}(i, j)$, is defined by the formula

$$\text{trust}^{\exp}(i, j) = \frac{\|I_{ij}\|}{\sum_{k=1}^{m} \|I_{ik}\|}, \tag{11}$$

where $\|I_{ik}\|$ is the number of interactions of u_i with each $u_k \in \mathcal{U}$.

Definition 4. Suppose that $\text{trust}^{\exp}(i, j)$ is the experience trust of u_i on u_j, $\text{int}X(j, t)$ is the interest degree of u_j on the topic t. Then the experience topic-aware trust of u_i on u_j of topic t is defined by the formula

$$\text{trust}_{\text{topic}}^{\exp}(i, j, t) = \gamma \times \text{trust}^{\exp}(i, j) + \delta \times \text{int}X(j, t), \tag{12}$$

where $\gamma,\ \delta \geq 0,\ \gamma + \delta = 1$.

The parameters $\gamma,\ \delta$ are used to represent the correlation degrees of interest and interaction in social networks. These parameters will be estimated by means of experimental evaluation, which is presented in Section 5. The computation of experience topic-ware trust is given in Algorithm 2.

Algorithm 2 Experience Aware-Topic Trust of u_i on u_j of topic t

Input: The set of topics $\mathcal{T} = \{t_1, t_2, ..., t_p\}$ and the set of users $\mathcal{U} = \{u_1, u_2, ..., u_m\}$ with entries e_{il}

Output: Experience topic-aware trust u_i on u_j of topic t, $\text{computeExpTrust}_{\text{topic}}^{\text{exp}}(i, j, t)$

1: $u_{i,int}^{t} \leftarrow \text{int}X(j, t)$, for t in $\mathcal{T}$
2: $u_{i,exp}^{j} \leftarrow \text{trust}^{\text{exp}}(i, j)$
3: $\text{trust}_{\text{topic}}^{\text{exp}}(i, j, t) \leftarrow \gamma \times u_{i,\text{exp}}^{j} + \delta \times u_{i,\text{int}}^{t}$
4: **return** $\text{trust}_{\text{topic}}^{\text{exp}}(i, j, t)$

4.2. Reputation topic-aware trust

This subsection presents a formula for computation of trust, which is inferred from some community. We restrict consideration of evaluation of community on trustees that have direct interaction with the truster. It means that peers belong to the layer L_i^1 with a truster u_i as presented in Section 2.

Definition 5. Given a source peer u_i and L_i^1 is the $1-$*level* of u_i. The reputation topic-aware trust of u_i on u_j is defined by the formula

$$\text{trust}_{\text{topic}}^{\text{rep}}(i, j, t) = \frac{\sum_{v \in L_i^1} \text{trust}_{\text{topic}}^{\text{exp}}(i, v, t) \times \text{sim}(v, j)}{\|L_i^1\|} \tag{13}$$

in which $\text{sim}(v, j)$ is the similarity of v on u_j being defined in the formula (10).

4.3. Topic-aware trust

The topic-aware trust is a function, which is an integration of the experience topic-aware and the reputation topic-aware trust degrees. It is defined as follows.

Definition 6. Suppose that $\text{trust}_{\text{topic}}^{\text{exp}}(i, j, t)$ and $\text{trust}_{\text{topic}}^{\text{rep}}(i, j, t)$ are the experience trust and reputation trust of u_i on u_j, respectively. Then the topic-aware trust of u_i on u_j of topic t is defined by the formula

$$\text{trust}_{\text{topic}}(i, j, t) = \lambda \times \text{trust}_{\text{topic}}^{\text{exp}}(i, j, t) + \mu \times \text{trust}_{\text{topic}}^{\text{rep}}(i, j, t), \tag{14}$$

where $\lambda,\ \mu \geq 0$, $\lambda + \mu = 1$. The computation of topic-aware trust is executed in steps, which is described in Algorithm 3.

5. EXPERIMENTAL EVALUATIONS

5.1. Problem statement

In Sections 3 and 4, we have described three measures of user interests, the functions of estimating degrees of topic aware trust based on experience and reputation. In this section we present some issues and the corresponding experimental results which are concerned with our model:

Algorithm 3 Topic Trust of u_i on u_j of topic t

Input: The set of topics $\mathcal{T} = \{t_1, t_2, ..., t_p\}$ and the set of users $\mathcal{U} = \{u_1, u_2, ..., u_n\}$ with experience aware topic trust $\mathrm{trust}^{\mathrm{exp}}_{\mathrm{topic}}(i, j, t)$

Output: Topic-aware trust u_i on u_j of topic t, $\mathrm{computeTopicTrust}_{\mathrm{topic}}(i, j, t)$

1: **for all** $v \in L^1_i$ **do**
2: $\quad \mathrm{sum}(i, j, t) \leftarrow \mathrm{sum}(i, j, t) + \mathrm{trust}^{\mathrm{exp}}_{\mathrm{topic}}(i, v, t) \times \mathrm{sim}(v, j)$
3: **end for**
4: $\mathrm{trust}^{\mathrm{rep}}_{\mathrm{topic}}(i, j, t) \leftarrow \frac{\mathrm{sum}(i,j,t)}{\|L^1_i\|}$
5: $\mathrm{trust}_{\mathrm{topic}}(i, j, t) \leftarrow \lambda \times \mathrm{trust}^{\mathrm{exp}}_{\mathrm{topic}}(i, j, t) + \mu \times \mathrm{trust}^{\mathrm{rep}}_{\mathrm{topic}}(i, j, t)$
6: **return** $\mathrm{trust}_{\mathrm{topic}}(i, j, t)$

- The measure of user's interests is defined by one of three functions which are shown in Proposition 2: Max, Cor and Sum. The question is that how those measures affect on user interest in a topic. We utilize the mean deviation to investigate their effects of Max, Cor, Sum on user similarity.
- The experience topic aware trust of a truster u_i on trustee u_j is calculated as a function of degrees of their interaction and trustee's interests give in the formula (12). Our question is that which factor affects more trustworthiness computation. We utilize the mean deviation to define the effects of parameters γ, δ on the estimation.
- The formula (14) represents a computational function of trust estimation of a truster u_i on a trustee u_j by means of community opinion via similarity of interests. Our question is that which factor affects more trustworthiness computation. We utilize the mean deviation to define the effects of parameters λ, μ on the estimation.

5.2. Experimental data

We collect data from the group of people who love running and share their preferences on website "Dar–DongAnh Runners" (https://racevietnam.com/team/dar-dong-anh-runners/longbien-marathon-2020). Their interests include topics: Fashion in running; Diet as appropriate (health in running); Running tournaments; Running genres such as long-distance running, trail running and technical running, etc. According to the statistics, as of April 30, 2021, the running group consists of 497 members; the number of members participating in posting from 2018 to April 2021 are 89 with 442 posts. There are 218 members who show interactions (e.g, likes, comments) with nearly 10000 comments. The details are given in Figure 1.

Collected data	
Total number of members	**497**
Number of members participating in posting	**89**
Number of posts	**442**
Number of comments	**9970**

Figure 1: Data set

We select six topics to conduct our testing, which are defined by the set of keywords via

Table 1: Topics in running

Running fashion	Health in running	Trail running	Road running	Run tournaments	Running technique
giầy	ăn uống	rừng núi	dẻo dai	giải chạy	khoa học
áo thun	mệt mỏi	hiểm trở	đường trường	checkin	kỹ thuật
thời trang	đau	cây xanh	cự ly	về đích	khởi động
đế mềm	xương khớp	xanh mướt	sức bền	full marathon	marathon
giày thể thao	xương	địa hình	liên tục	haft marathon	hướng dẫn
snecker	thoái mái	leo núi	khoảng cách	thành phố	cần biết
phù hợp	dinh dưỡng	cung đường	rèn luyện	nộp tiền	cơ bản
động tác	nghỉ ngơi	lối mòn	pace	tham gia	chạy bộ
thiết kế	bàn chân	khám phá	giao thông	BIB	cách chạy
đế giày	thể chất	trail	marathon	địa điểm	về đích
màu sắc	Sức khỏe	vách đồi	đua	thể thao	hợp lý
đàn hồi	tinh thần	nguy hiểm	tiếp sức	cung đường	tư thế
giày đi bộ	duy trì	dài ngày	tốc độ	tổ chức	tốc độ
vận động	chất lượng	thời tiết	cung đường	rèn luyện	lịch
đau	cải thiện	xuống sức	chạy	cuộc đua	đầu gối
đồ chạy	thực phẩm	ảnh đẹp	khởi động	hẹn hò	nhịp điệu
đường phẳng	viatamin	checkin	chạy bộ	giao lưu	sải chân
đồi núi	cảm xúc	pace	về đích	sự kiện	năng lượng
garmin	dẻo dai	khoảng cách	bằng phẳng	ủng hộ	cơ thể
nhẹ	bền bỉ	rèn luyện	thành phố	tham dự	chấn thương
patin	ngon	khởi động	thời gian	danh sách	cánh tay
giày	tốt	tốc độ		chào mừng	cơ bắp
bàn chân	phục hồi			hình ảnh	nước rút
tất	tăng cường				bổ sung
đệm	cố gắng				thả lỏng
size	gầy				pace
xỏ	béo				training
giá tiền	tim mạch				
rộp	rèn luyện				
thấm mồ hôi	chấn thương				
xinh đẹp	đầu gối				
hình ảnh					

wikipedia (https://vi.wikipedia.org/wiki/). The table of keywords, illustrated in Table 1, is used to model topics in vectors.

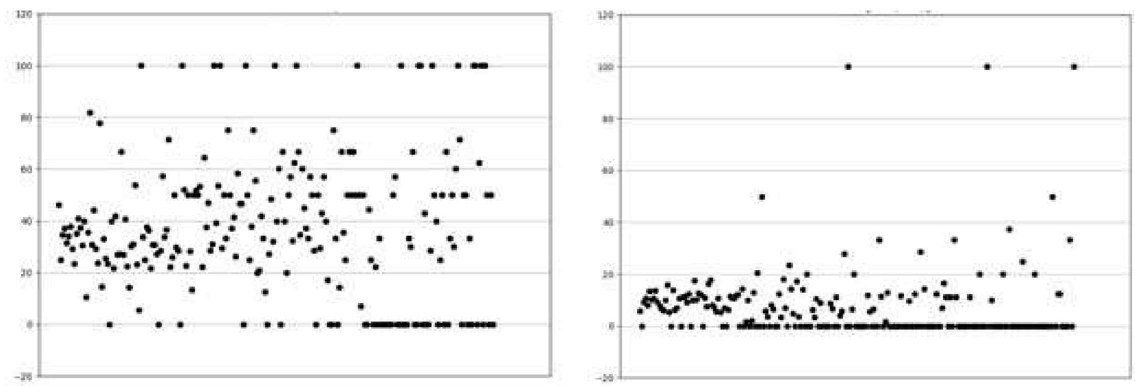
Table 2: Interest of users

Username	Running fashion	Health in Running	Trail running	Road running	Run tournaments	Running technique
Vương Mạnh	0.08	0.13	0.06	0.21	0.46	0.06
Hoàng Dũng	0.5	0.25	0	0	0.25	0
Long Tran Van	0.05	0.19	0.07	0.25	0.35	0.09
Thuan Lam	0.08	0.14	0.08	0.22	0.37	0.11
Minh Ngọc	0.08	0.15	0.06	0.31	0.31	0.08
Dũng Nguyễn	0.08	0.15	0.06	0.24	0.34	0.13
Phùng Thành	0.07	0.2	0.07	0.17	0.38	0.1
Phúc Nguyễn Bá	0.09	0.12	0.1	0.29	0.29	0.11
Nguyễn Đại Dương	0.03	0.17	0.09	0.34	0.23	0.14
Hanh Tum	0.08	0.19	0.06	0.22	0.35	0.09
Phan Toàn	0.08	0.18	0.04	0.21	0.41	0.08
Tran Bac	0.05	0.24	0.03	0.24	0.37	0.07
Nguyễn Hoàng Hải	0.07	0.17	0.03	0.36	0.3	0.06
Vũ Vương	0.06	0.2	0.04	0.2	0.4	0.1
Nguyễn Tiến Tóp	0.21	0.16	0.11	0.26	0.11	0.16
Nguyễn Công Hưng	0.1	0.19	0.07	0.23	0.36	0.05
Hoa Lê	0	0	0	0.18	0.82	0

5.3. Experimental Results

5.3.1. Interests and similarity

A part of user's interests is illustrated in Table 2. The values of interest degrees are ranged from 0.0 to 0.1. The value with 0.0 means that the user is not interested in the respective topic, whereas the value with 0.1 means that the person is interested very much in this topic. The distribution of user's interests by topics is showed in Figure 2. We can see that the user's interest in the topic "Run tournaments" focuses much on 20% to 50%, while the user's interest in the topic "Running technique" will focus on level from 0% to 20%.



(a) Run tournaments (b) Running technique

Figure 2: Distribution in topics

We proceed to calculate the similarity between the user's interest in topics with three degrees Max, Cor, and Sum. The similarity of one user compared with the other users is shown in Figure 3. The testing results with three users Tieu Duong Julia, Nguyen Dac Cu and Duong Minh Nghia, the similarity degrees of Tieu Duong Julia compared with Nguyen Dac Cu and with Duong Minh Nghia are 0.98 and 0.81, respectively. The similarity of two persons in topics is illustrated in Figure 4.

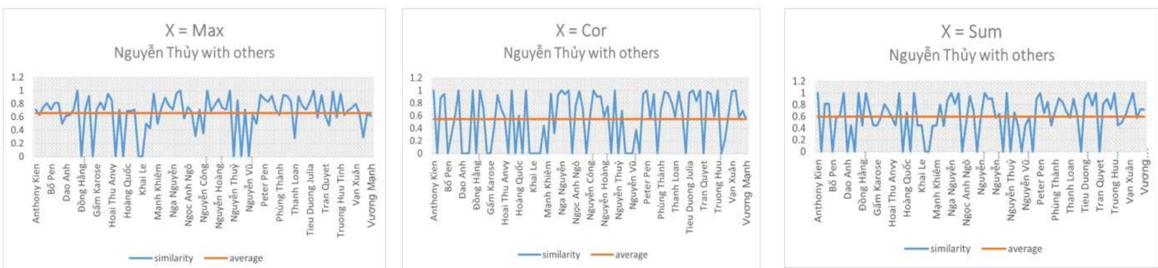


Figure 3: Similarity with different measures

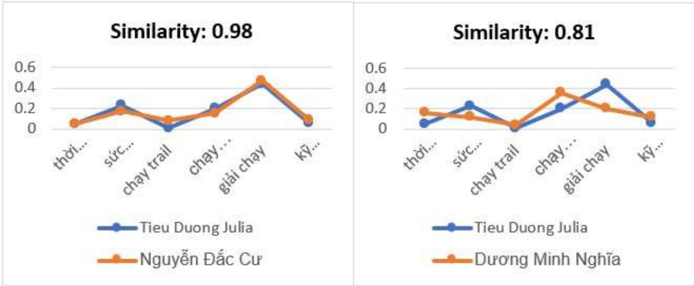


Figure 4: Diagrams with similarity in topics

The testing results show that the Cor measure gives the highest mean deviation value of 0.04411, the Max measure gives the lowest average deviation of 0.02938; and with the Sum measure, the mean deviation is 0.03584. In the next tests, the interest measure Max with the lowest deviation is selected.

5.3.2. Experience and reputation topic-aware trust

This subsection investigates the relationship between experience, reputation trust and user interest degrees. We perform experiments with the corresponding (γ, δ) couples: $(0.9; 0.1)$; $(0.8; 0.2)$; $(0.7; 0.3)$; $(0.6; 0.4)$; $(0.5; 0.5)$; $(0.4; 0.6)$; $(0.3; 0.7)$; $(0.2; 0.8)$; $(0.1; 0.9)$. The experience trust of a truster on a trustee of a certain topic with couples γ, δ is given in Table 3. Figure 5 illustrates the trustworthiness of a user named "Vuong Manh" with 10 other users. In the case of $\gamma = 0.1$ and $\delta = 0.9$, the confidence level is more stable than in the case of $\gamma = 0.9$ and $\delta = 0.1$.

Measures	$(\gamma;\delta) = (0.1;0.9)$	$(0.2;0.8)$	$(0.3;0.7)$	$(0.4;0.6)$	$(0.5;0.5)$	$(0.6;0.4)$	$(0.7;0.3)$	$(0.8;0.2)$	$(0.9;0.1)$
S.D	0.007596	0.010819	0.014415	0.017615	0.020464	0.023033	0.025389	0.027561	0.02953

Table 3: Standard Deviation (S.D) of Experience topic aware Trust w.r.t. various couples γ, δ

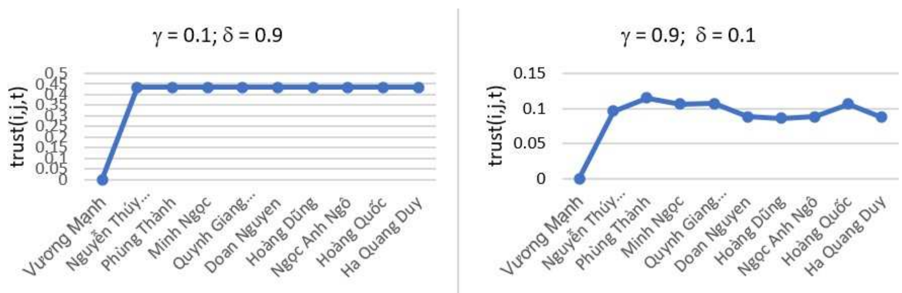


Figure 5: Experience Topic Trust and interests with (γ, δ)

We conduct an experiment to consider which level the reliability given in formula (12) depends on factors of interest and experience trust. Figuer 6 depicts the trustworthiness of user "Vuong Manh" with ten other users in six topics respectively.

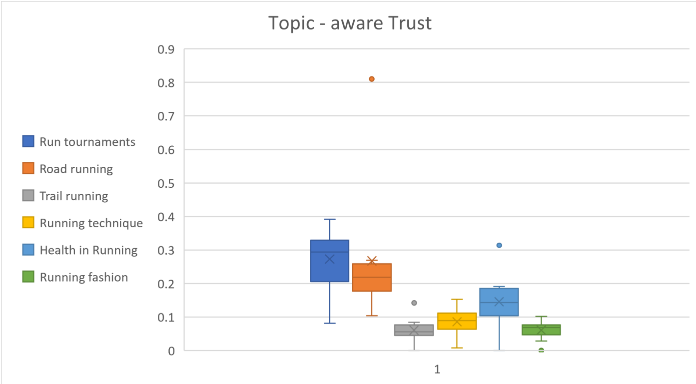


Figure 6: Affect of topics on Trust

In the formula (13), the reliability of u_i on u_j depends on the similarity of users who have direct interaction with u_j. We filtered out 61 users who have a direct link to the user "Ngoc Anh Ngo". Calculating the similarity of those 61 users with the two users "Quynh Giang Doan" and "Bean Nhat Anh" respectively. We get 2 data domains in blue and orange shown in Figure 7. Obviously, the similar values for the user "Quynh Giang Doan" will be distributed mainly in the data domain from 0.8 to 1 while the similar values for the user "Bean Nhat Minh" will only distributed mainly in the range from 0 to 0.5. The reason is that the reliability of the user "Ngoc Anh Ngo" for "Quynh Giang Doan" gives a value of 10.86 while that for "Bean Nhat Anh" is only 4.25.

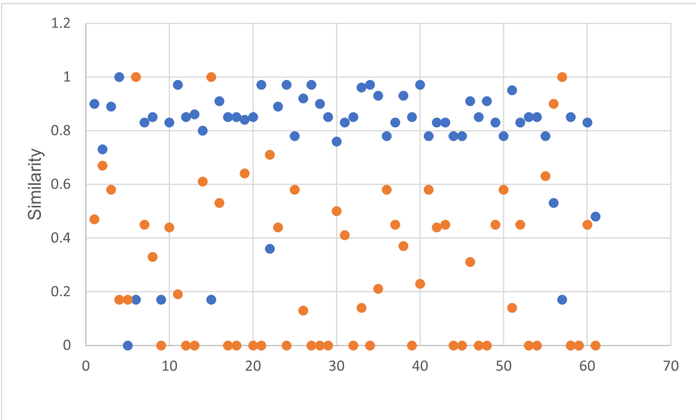


Figure 7: Distribution of data similarity

Similarly, we conduct the experiment with nine couples of (λ, μ) w.r.t. the formula (14) to consider which factor in experience and reputation affects much more on trust estimation. The results are given by Table 4.

Table 4: Standard Deviation (S.D) of topic aware trust values w.r.t. (λ, μ)

Measures	$(\lambda; \mu)$ = (0.1; 0.9)	(0.2; 0.8)	(0.3; 0.7)	(0.4; 0.6)	(0.5; 0.5)	(0.6; 0.4)	(0.7; 0.3)	(0.8; 0.2)	(0.9; 0.1)
S.D	0.018	0.037	0.055	0.073	0.091	0.11	0.128	0.146	0.165

From the result, we choose the couple $(\lambda, \mu) = (0.1; 0.9)$ since it gets the smallest standard deviation. This observation shows that topic aware trust estimation depends on reputation more than user's own experience evaluation.

6. CONCLUSIONS

In this paper, we have proposed a model of topic-aware trust computation, which is a composition of the trust estimation based on experience of direct interaction, degrees of user's interests and reputation based trust. We determine a similarity measure of users which has been constructed by means of the similar ones of profiles and user's interest degrees on topics. Based on the similarity, we proposed the measure of topic aware trust which is inferred from its own experience trust and trust estimation from members of community. Our experimental results showed relationships among types of topic trust and affection of user's interest on trust estimation. We show that the topic aware trust estimation depends on reputation more than user's own experience evaluation. However, there are some limitations in our work. First, in this work, we restrict only consideration of interaction in one direction from truster to trustees. This form needs to be extended to include various forms such as the converses from trustees to truster. Second, how to utilize the propagation to estimate trust among users when there is no direct interaction among them. These issues need to be investigated furthermore. The research results will be presented in our future work.

REFERENCES

[1] D.Manning, P. Raghavan, and H. Schuzte, *Introduction to Information Retrieval.* Cambridge University Press, 2008. https://nlp.stanford.edu/IR-book/pdf/irbookonlinereading.pdf

[2] Bo Jiang and Ying Sha, "Modeling temporal dynamics of user interests in online social networks", *Procedia Computer Science*, vol. 51, 2015, pp. 503–512. https://doi.org/10.1016/j.procs.2015.05.275

[3] Jaeyong Kang and Hyunju Lee, "Modeling user interest in social media using news media and wikipedia", *Information Systems*, vol.65, April 2017, pp. 52–64. https://doi.org/10.1016/j.is.2016.11.003

[4] A. Kalai, A. Wafa, C. Zayami, and I. Amous, "LoTrust: A social Trust Level model based on time-aware social interactions and interests similarity," in *2016 14th Annual Conference on Privacy, Security and Trust (PST)*, 2016, pp. 428–436. Doi: 10.1109/PST.2016.7906967.

[5] J. Golbeck, "Trust on the world wide web: A survey," *Foundations and Trends in Web Science,* vol.2, no. 2, pp.131–197, 2006.

[6] J. Golbeck, "Computing with trust: Definition, Properties and Algorithms", in *2006 Securecomm and Workshops,* 2006, pp. 1-7. Doi: 10.1109/SECCOMW.2006.359579

[7] M. Jamall and M. Ester, "TrustWalker; a random walk model for combining trust-based and item-based recommendation," in *KDD '09: Proceedings of the 15th ACM SIGKDD In-*

ternational Conference on Knowledge Discovery and Data MiningJune, 2009, pp. 397–406. https://doi.org/10.1145/1557019.1557067

[8] W. Jiang and G. Wang, "SWTrust: Generating trusted graph for trust evaluation in Online social networks," in *2011 IEEE 10th International Conference on Trust, Security and Privacy in Computing and Communications,* 2011, pp. 320–327. Doi: 10.1109/TrustCom.2011.251

[9] E. Gabrilovich and S. Markovitch, "Computing semantic relatedness using Wikipedia-based explicit semantic analysis," *IJCAI,* 2007. Available at: https://www.aaai.org/Papers/IJCAI/2007/IJCAI07-259.pdf

[10] W. Sherchan, S. Nepal, and C. Paris, "A survey of trust in social network", *ACM Computing Surveys*, vol.45, no.4, pages 1–33, 2013. https://doi.org/10.1145/2501654.2501661

[11] M.H. Nguyen and D.Q. Tran, "A combination trust model for multi-agent systems", in *International Journal of Innovative Computing, Information and Control ICIC International,* vol. 9, no. 6, June 2013, pp. 2405–2420, 2013.

[12] V. Podobnik, D. Striga, A. Jandras, and I. Lovrek, "How to calculate trust between social network users?," *SoftCOM 2012, 20th International Conference on Software, Telecommunications and Computer Networks,* 2012, pp. 1-6.

[13] M. Richardson, R. Agrawal, and Domingos, "Trust management for the semantic Web", in *The Semantic Web: Proceedings of the 2nd International Semantic Web Conference (ISWC)*, Vol. 2870 of LNCS, pp.351–368, Springer-Verlag, 2003. Available at: http://homes.cs.washington.edu/ pedrod/papers/iswc03.pdf

[14] Chung-Wei Hang et al., "Operators for Propagating Trust and their Evaluation in Social Networks," Department of Computer Science North Carolina State University Raleigh, NC 27695-8206, USA.

[15] D. Q. Tran, "Computational topic trust with user interests based on propagation and similarity measure in social network", *Southeast Asian Journal of Sciences*, vol.7 , no.1, 2019, pp. 18–27.

[16] D. Q. Tran, et al., "Modeling user interests, similarity and worthiness based on vectors of entries in social networks", *Southeast Asian Journal of Sciences,* vol. 7, no. 2, 2019, pp. 133–141.

[17] D. Q. Tran, and P. T. Pham, "Integrating influence into trust computation with user interest on social networks", *Southeast Asian Journal of Sciences,* vol. 8, no 1, 2020, pp. 18–27.

[18] Yang Song, Lu Zhang, and C. Lee Giles, "Automatic tag recommendation algorithms for social recommender systems", *ACM Transactions on the Web*, vol.5, no. ‘, pp. 1–31, February 2011. https://doi.org/10.1145/1921591.1921595

[19] P.T. Pham, M.H. Nguyen, and D.Q. Tran, "Incorporation of experience and reference-based topic trust with interests in social network," *Advances in Information and Communication Technology. ICTA 2016. Advances in Intelligent Systems and Computing,* vol. 538. Springer, Cham. https://doi.org/10.1007/978-3-319-49073-1_31

[20] M.H. Nguyen, D.Q. Tran, "Classes of trust functions for distributed intelligent computing", *Southeast Asian Journal of Sciences*, vol. 1, no. 2, 2012.

[21] Y. Wang, M.P. Singh, "Trust representation and aggregation in a distributed agent system", Department of Computer Science North Carolina State University Raleigh, NC 27695-8206, USA, 2006. Available at: http://www.aaai.org/Papers/AAAI/2006/AAAI06-224.pdf

[22] X. Wang, H. Liu, and W. Fan, "Connecting users with similar interests via tag network inference", in *CIKM '11: Proceedings of the 20th ACM International Conference on Information and Knowledge Management,* October 2011, pp. 1019–1024. https://doi.org/10.1145/2063576.2063723

Received on November 22, 2021
Accepted on May 12, 2022