***Southeast Asian Journal of Sciences,*** Vol. 08, No 1 (2020), pp. 18-27

# INTEGRATING INFLUENCE INTO TRUST COMPUTATION WITH USER INTERESTS ON SOCIAL NETWORKS

**Dinh Que Tran[1], Phuong Thanh Pham[2]**

[1] *Department of Information Technology*
*Posts and Telecommunications Institute of Technology (PTIT)*
*Hanoi, Vietnam*

[2] *Department of Mathematics and Informatics*
*Thang long University*
*Hanoi, Vietnam*
*E-mail: tdque@yahoo.com, ppthanh216@gmail.com*

.

**Abstract**

The purpose of this paper is to investigate an extension of reputation based topic trust computation to include degrees of user's influence in community. We propose a computation method which is exhibited within two steps: (i) Execute topic trust estimation with influence and interests via interaction among peers; (ii) Perform trust computation from reputation according similarity degrees with trustee peer.

## 1 Introduction

In the real world, the influence of a person on community may change the viewpoint of users in making decisions or selecting items such as goods, books for their purchases. In social networks and on line shopping webs, Modeling and analyzing user's influences in social networks have attracted a great deal

**Key words:** social networks, models of societies, text processing, decision support, distributed systems, artificial intelligence, reliability.
2010 AMS Mathematics classification: 911D30, 91D10, 68U115, 68U35, 68M14, 68M115, 68T99.

of research interests [1] [2] [3] [8] [4] [9] [10] [11]. In this paper, we propose a model of computational influence which is defined by means of user's some interest threshold on topics and the number of feedbacks of users when receiving messages.

From this model, we construct a computation function of trust that is integrated from three factors: experience of user's interaction or experience trust, user's interest degrees on topics and influence weights. We make use of similar measures being constructed from interest and influence degrees to estimate trustworthiness of truster peers on trustees to include reputation in computation. This work is an extension of our previous researches of trust estimation, which is mainly based on interaction experience among partners and their interests on social network [5] [6] [7] [12] [13] [14].

The remainder of this paper is structured as follows. Section 2 presents an updated version of computing degrees of user's interests and entries based on their vectorial representations. Section 3 is devoted to considering computation of user's influence degrees on community. Similarity measures of interests, entries and influences are described in Section 4. Section 5 presents the formulas of reputation based topic trust computation extended for influence of peers. Section 6 is an open problem and conclusions.

# 2 Estimating Degrees of User's Interests based on Vectors of Topics and Entries

This section is first to describe a graphical representation of social network. Then, we present an updated version of computation of the interest degrees based on weighted vectors for topics and user's entries exhibited in our previous work [6].

## 2.1 Social Network

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

- $\mathcal{U} = \{u_1, \ldots, u_m\}$ is a set of users, whose elements are autonomous entities being called *peers*. In this paper, the terms of peer and user are used interchangeably;

- $\mathcal{I}$ is a set of all interactions or connections $I_{ij}$ from $u_i$ to $u_j$. $\|I_{ij}\|$ is denoted to be the number of such interactions. Each interaction between users $u_i$ and $u_j$ is a transaction at an instant time, which occurs when $u_i$ sends to $u_j$ via some "wall" messages such as post, comment, like, opinions etc.

- $\mathcal{E} = \{E_1, \ldots, E_m\}$ is the set of entries dispatched by users $\mathcal{U} = \{u_1, \ldots, u_m\}$, where $E_i = \{e_{i1}, \ldots, e_{im_i}\}$ are entries given by $u_i$. Each *entry* is a brief piece of information dispatched from some user $u_i$ to make a description or post information/idea/opinions on an item such as a paper, a book, a film, a video etc.

## 2.2 Vectorial Representation of Entries and Topics

Suppose that $\mathcal{T} = \{T_1, \ldots, T_p\}$ is a collection of topics, in which each topic is defined as a set of terms or words. The technique $tf - idf(d, D_i) = tf(d, D_i) \times idf(d, \mathcal{D})$ for vectorial representation [7] of such entries and topics are applied, where $tf(d, D_i)$ is the number of times the term $d$ appears in $D_i$ and $idf(d, \mathcal{D}) = log(\frac{\|\mathcal{D}\|}{1+\|\{D_i|d\in D_i\}\|})$. Based on the similarity of vectors, we might classify entries into classes w.r.t. topics and define interest degrees of $u_i$ in topic $t$.

Let $V_T = \{v_1, \ldots, v_q\}$ be a set of $q$ distinct terms in all $T_i \in \mathcal{T}$. A topic vector w.r.t. each topic $T_i$ is a weighted one, which is defined as follows

$$\mathbf{t_i} = (w_{i1}, \ldots, w_{iq}) \tag{1}$$

where $w_{ik} = tf(v_k, T_i) \times idf(v_k, \mathcal{T})$, $v_k \in V_T$.

As denoted previously, $e_{il}$ is an entry of terms dispatched by $u_i$. An entry vector w.r.t. topics $\mathcal{T}$, briefly *topic vector*, is a weighted one, which is defined as follows

$$\mathbf{e^t_{il}} = (e^1_{il}, \ldots, e^p_{il}) \tag{2}$$

where $e^r_{il} = tf(v_r, e_{il}) \times idf(v_r, E_i)$, $v_r \in V_T$.

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

$$\mathbf{e^j_{il}} = (e^1_{il}, \ldots, e^{\|V_{ij}\|}_{il}),\ l = 1, \ldots, n_i \tag{3}$$

$$\mathbf{e^i_{jk}} = (e^1_{jk}, \ldots, e^{\|V_{ij}\|}_{jk}),\ k = 1, \ldots, n_j \tag{4}$$

where $e^r_{il} = tf(v_r, e_{il}) \times idf(v_r, E_i)$, $e^r_{jk} = tf(v_r, e_{jk}) \times idf(v_r, E_j)$ $v_r \in V_{ij}$.

Thus, we can define a sequence of topic vectors $\mathbf{e^{t_1}_{il}}, \ldots, \mathbf{e^{t_p}_{il}}$ w.r.t. each entry and a sequence of entry vectors $\mathbf{e^j_{i1}}, \ldots, \mathbf{e^j_{in_i}}$ w.r.t. entries $E_j$. These vectors will be utilized for constructing measures of user's similarity and interests, which are presented in the next subsection.

## 2.3 Interest Degrees

Based on the above definitions of vectors, we can define correlation degrees $cor(\mathbf{e^t_{ij}}, \mathbf{t_k})$ among entries $e_{ij}$ given by $u_i$ w.r.t. topics $t_k$ as follows:

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

where $\bar{u} = \frac{1}{n}(\sum_{i=1}^{n} u_i)$ and $\bar{v} = \frac{1}{n}(\sum_{i=1}^{n} v_i)$. It is clear that values of the function $sim(x, y)$ belong to the interval $[0, 1]$, whereas values of $cor(x, y)$ are in $[-1, 1]$. We may make use of the function $f(x) = \frac{(x+1)}{2}$ to bound values of function $cor(x, y)$ into the unit interval $[0, 1]$.

An entry $e_{ij}$ is called $\epsilon$-*entry* w.r.t. topic $t_k$ if and only if $cor(\mathbf{e^t_{ij}}, \mathbf{t_k}) \geq \epsilon$, where $0 < \epsilon \leq 1$. We consider three interest measures as follows: The interest degree of $u_i$ in topic $t$ is defined by one of the following formulas:

$$intMax(u_i, t) = \max_j(cor(\mathbf{e^t_{ij}}, \mathbf{t})) \tag{6}$$

$$intCor(u_i, t) = \frac{\sum_j cor(\mathbf{e^t_{ij}}, \mathbf{t})}{\|E_i\|} \tag{7}$$

$$intSum(u_i, t) = \frac{1}{2}\left(\frac{n_i^t}{\sum_{l \in \mathcal{T}} n_i^l} + \frac{n_i^t}{\sum_{u_k \in \mathcal{U}, l \in \mathcal{T}} n_k^l}\right) \tag{8}$$

where $n_i^t$ is the number of $\epsilon$-entries concerned about the topic $t$ given by $u_i$.

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

$$\mathbf{u^t_i} = (u_i^1, \ldots, u_i^p) \tag{9}$$

in which $u_i^k = intX(u_i, t)$ 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* as defined in Formulas (6), (7), (8).

The definition of vectors with various degrees is utilized for constructing the similarity of users in their interests which is considered in the next section.

# 3 User's Influence on Community

In this section, we construct an influence degree of a peer based on "backward interaction" by means of dispatching entries on wall. It means that when a peer dispatches a message, feedbacks from the other ones in the forms of "like", "share" etc. are called influences. We utilize Jaccard similarity to measure the degrees of the peer's influences on community. First we observe that:

- The more feedbacks a peer receives, the more he impacts on community
- The more feedbacks a peer receives, the more he attracts

We have the following definition.

**Definition 1.** *A $\delta$-influence set by $u_i$ on topic $t$ is defined by the following formula*

$$F^{\delta,t}_{i\leftarrow} = \{u_k \| u_k \textit{ issues feedbacks to } e_{il} \in E_i,\ intX(u_k, t) \geq \delta\} \tag{10}$$

*where $0 < \delta \leq 1$ is a given interest threshold.*

**Definition 2.** *The influence degree of $u_i$ on community is defined by the following formula*

$$infDeg(u_i, t) = \frac{\|F^{\delta,t}_{i\leftarrow}\|}{\|\mathcal{I}\|} \tag{11}$$

*where $\mathcal{I}$ is the universe of all users.*

The influence vector of users in various topics is defined by the formulas

$$\mathbf{u_i^{inf}} = (u_i^{inf,1}, \ldots, u_i^{inf,p}) \tag{12}$$

where $u_i^{inf,k} = infDeg(u_i, t_k)$, $k = 1, \ldots, p$.

# 4 Similarities of Influence, Interest and Entries

## 4.1 Similarity of Users

**Definition 3.** *A function $sim : \mathcal{U} \times \mathcal{U} \to [0, 1]$ is a similarity measure iff it satisfies the following conditions:*

*(i)* $sim(u_i, u_i) = 1$, *for all* $u_i \in \mathcal{U}$

*(ii)* $sim(u_i, u_j) = sim(u_j, u_i)$, *for all* $u_i,\ u_j \in \mathcal{U}$

**Definition 4.** *Let $F^{\delta,t}_{i\leftarrow}$ and $F^{\delta,t}_{j\leftarrow}$ be two sets of $\delta$-influence by $u_i$ and $u_j$, respectively. Influence similarity is defined as follows*

$$sim^{\delta,t}_{inf}(u_i, u_j) = \frac{F^{\delta,t}_{i\leftarrow} \cap F^{\delta,t}_{j\leftarrow}}{F^{\delta,t}_{i\leftarrow} \cup F^{\delta,t}_{j\leftarrow}} \tag{13}$$

**Definition 5.** *Influence similarity of two peers $u_i$ and $u_j$ is defined as a cosine similarity of two vectors $\mathbf{u_i^{inf}}$ and $\mathbf{u_j^{inf}}$*

$$sim_{inf}(u_i, u_j) = \frac{< \mathbf{u_i^{inf}}, \mathbf{u_j^{inf}} >}{\|\mathbf{u_i^{inf}}\| \times \|\mathbf{u_j^{inf}}\|} \tag{14}$$

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

Based on this interest vector, we can construct a similar measure in interests as follows:

**Definition 6** ([6])**.** *Interest similarity of two peers $u_i$ and $u_j$ is defined as a cosine similarity of two vectors $\mathbf{u_i}$ and $\mathbf{u_j}$*

$$sim_{int}^{X}(u_i, u_j) = \frac{< \mathbf{u_i^t}, \mathbf{u_j^t} >}{\|\mathbf{u_i^t}\| \times \|\mathbf{u_j^t}\|} \tag{15}$$

*in which $< u, v >$ 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 computation of $u_i^k \in \mathbf{u_i^t}$ and $u_j^k \in \mathbf{u_j^t}$ as defined in Section 2.*

The profile or entries similarity of two users is defined according to entries dispatched by users as follows

**Definition 7.** *Given two users $u_i$, $u_j$ with sets of entries $E_i = \{e_{i1}, \ldots, e_{in_i}\}$ and $E_j = \{e_{j1}, \ldots, e_{jn_j}\}$, respectively. Profile similarity of users is defined by one of the following formulas*

*(i)* $sim_{pro}^{max}(u_i, u_j) = \max_{k,l}(sim(\mathbf{e_{ik}^i}, \mathbf{e_{jl}^j}))$

*(ii)* $sim_{pro}^{sum}(u_i, u_j) = \frac{\sum_{k,l}(sim(\mathbf{e_{ik}^i}, \mathbf{e_{jl}^j}))}{\|E_i\| + \|E_j\|}$

*in which $(sim(\mathbf{e_{ik}^i}, \mathbf{e_{jl}^j}))$ is the usual cosine similarity measure.*

**Definition 8.** *The general similarity, or briefly similarity, between $u_i$ and $u_j$ is defined by the weighted composition of their partial similarities and given by the following formula*

$$sim(u_i, u_j) = \alpha \times sim_{inf}(u_i, u_j) + \beta \times sim_{int}^{X}(u_i, u_j) + \gamma \times sim_{pro}^{Y}(u_i, u_j) \tag{16}$$

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

It is easy to prove the following proposition.

**Proposition 1.** *For all $u_i, u_j$, $sim_{inf}(u_i, u_j)$, $sim_{pro}^{sum}(u_i, u_j)$, $sim_{pro}^{max}(u_i, u_j)$ and $sim_{int}^{Max}(u_i, u_j)$, $sim_{int}^{Cor}(u_i, u_j)$, $sim_{int}^{Sum}(u_i, u_j)$, $sim(u_i, u_j)$ are similarity measures.*

Thus, for every couple $u_i$ and $u_j$, we can define their similarity degrees in interest, influence, profile and general. The question is that there is any correlation among these measures. The problem will be investigated from the view point of computational trustworthiness and presented in the next section.

# 5 Trustworthiness of Users based on Interaction, Interests and Influences

Based on similarity measures constructed in Section 4, we now develop a method for estimating topic trust among users. Rather than computation merely based on interaction and interest degrees [7], the novel one investigates how the contribution of the community influence in trustworthiness among peers. It means that trust estimation value of a truster peer on a trustee one is a function with the following parameters:

- Interaction experience of truster on trustee
- Interest degrees on topics of trustees
- Influence degree of the trustee peer on community
- Reputation given by similar peers on the trustee in hand

This paper is considered as a complementary work with ones proposed by ourselves [6][7]. We first consider some basic concepts being necessary for constructing such a function.

## 5.1 Levels of Interaction

Given a user $u_i$, we denote $L_i^1$ the set of all users directly interacting 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 $u_i$.

**Definition 9.** *Given $L_i^k$ a k-level of $u_i$. The average similarity threshold of the k-level w.r.t. $u_i$ is defined by the formula*

$$\theta = \frac{\sum_{v \in L_i^k} sim(u_i, v)}{\|L_i^k\|} \tag{17}$$

*where $sim(u_i, v)$ is defined as in Definition 8 .*

From this concept we can define *k-level close friend* as follows:

**Definition 10.** *A peer $v \in L_i^k$ is a k-level close friend of $u_i$ w.r.t. $\theta$ iff its similarity with $u_i$ is greater than similarity threshold $\theta$. Denote $L_i^{k,\theta} = \{v \in L_i^k | sim(u_i, v) \geq \theta\}$*

In this paper, we focus on investigating the class of close friends in 1-level w.r.t. the threshold $\theta$.

**Definition 11.** *An entry $e_{il} \in E_i$ is an acceptable one w.r.t. topic $t$ if $e_{il}^t \geq \delta$, where $\delta$ is a given threshold. Denote $E_i^{t,\delta}$ to be the set of acceptable entries w.r.t. topic $t$ and threshold $\delta$ given by $u_i$.*

### 5.2 Experience based Topic-aware Trust with Interests and Influence

**Definition 12** ([6]). *A function $trust_{topic} : \mathcal{U} \times \mathcal{U} \times \mathcal{T} \to [0,1]$ 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 $trust_{topic}(i,j,t) = u_{ij}^t$ means that $u_i$ (truster) trusts $u_j$ (trustee) of topic $t$ w.r.t. the degree $u_{ij}^t$.*

**Definition 13** ([6]). *Experience trust of user $u_i$ on user $u_j$, denoted $trust^{exp}(i,j)$, is defined by the formula*

$$trust^{exp}(i,j) = \frac{\|I_{ij}\|}{\sum_{k=1,k\neq i}^{m} \|I_{ik}\|} \tag{18}$$

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

Based on the degrees of interaction, user's interest and influence, we can define the *experience topic trust* for sink peers of 1-friends $L_i^1$ of $u_i$. The computation is constructed from the observation: (i) The more a peer interacts with an opponent, the higher it is reliable; (ii) The higher degree of interest a peer owns, the more trust on him it should be assigned; (iii) The higher the degree of influence a peer is, the more reliable it is. We have the following definition.

**Definition 14.** *Suppose that $trust^{exp}(i,j)$ is the experience trust of $u_i$ on $u_j$, $intX(j,t)$ is the interest degree of $u_j$ on the topic $t$ and $infDeg(j,t)$ is the influence degree of $u_j$ on community. Then the experience topic trust of $u_i$ on $u_j$ of topic $t$ is defined by the following formula:*

$$trust_{topic}^{exp}(i,j,t) = \alpha \times trust^{exp}(i,j) + \beta \times intX(j,t) + \gamma \times infDeg(j,t) \tag{19}$$

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

The parameters $\alpha, \beta, \gamma$ are used to represent the correlation degrees of interest, interaction and influence in social networks. These parameters need to be measured by means of experiments.

It is easy to see that

**Proposition 2.** *The function $trust_{topic}^{exp}(i,j,t)$ is a topic trust function.*

### 5.3 Reputation based Estimation of Trust for Peers of $L_i^p$ where $p = 1$

**Definition 14** provides a formula for estimating topic trust by truster's experience of interaction with a trustee. However, as previously discussed, the reliability on a peer is also affected by opinions given by reputation about the trustee. Now we consider a method of reputation-based estimation of trust which is resulted from some similarity of peers with the trustee in hand. The topic trust is then called *reputation or reference topic trust* and exhibited in the following definition.

**Definition 15.** *Given a source peer $u_i$. Let $L_i^1$ be the $1-level$ of $u_i$ and $L_i^{1,\theta}$ be the set of 1-level close friends of $u_i$ with the threshold $\theta$. Then, the reputation topic trust is defined by the formula:*

$$trust_{topic}^{ref}(i,j,t) = \frac{\sum_{v \in L_i^{1,\theta}} trust_{topic}^{exp}(i,v,t) \times sim(v,j)}{\|L_i^{1,\theta}\|} \tag{20}$$

It is easy to prove the following proposition

**Proposition 3.** *The function $trust_{topic}^{ref}(i,j,t)$ is a topic trust function.*

## 6 Conclusions

In this paper, we have introduced a method of trust estimation which is constructed from degrees of interaction of peers, interests, influence and reputation. The computation is composed of two stages: (i) First, the experience trust is computed by means of a function of directed interaction, interest and influence; (ii) Second, reference or reputation trust on a trustee is estimated via a function of experience trust of peers which are similar with the trustee. A open problem is that when a peer belongs to $L_i^p$ where $p > 1$, how the estimation of trust on the trustee must be computed via propagation of various levels. We are also currently performing experimental evaluation and comparing with other models on trust computation in social network. The research results will be presented in our future work.

## References


[1] J. Tang et al., *Social Influence Analysis in Large-scale Networks,* Conference KDD09, June 28, 2009

[2] Bingoi et al., *Topic-Based Influence Computation in Social Networks under Resource Constraints,* IEEE Transactions on service computing, Vol. PP, No. 99, 2018.

[3] Ju Fan et al., OCTOPUS: An Online Topic-Aware Influence Analysis System for Social Networks, IEEE 34th International Conference on Data Engineering (ICDE), 2018. Available at: https://ieeexplore.ieee.org/document/8509399

[4] Kan Li et l., Social Influence Analysis: Models, Methods, and Evaluation, Engineering 4 (2018) 40-46.

[5] Dinh Que Tran, *Computational topic trust with user's interests based on propagation and similarity measure in social networks,* Southeast-Asian J. of Sciences, **7**(01) 2019 pp. 18-27.

[6] Dinh Que Tran and Phuong Thanh Pham, *Integrating interaction and similarity threshold of user's interests for topic trust computation,* Southeast-Asian J. of Sciences, **7**(01), 2019, pp. 28-35.

[7] Dinh Que Tran, Phuong Thanh Pham and Thi Hoi Nguyen, *Modeling User's Interests, Similarity and Trustworthiness based on Vectors of Entries on Social Networks,* Southeast Asian Journal of Sciences, **8**(02), 2019, pp. 133-141.

[8] David Crandall, Dan Cosley et al., *Feedback effects between similarity and social influence in online communities,* KDD'08, 2008, USA.

[9] Vedran Podobnik et al., *How to calculate trust between social network users?* In Software, Telecommunications and Computer Networks (SoftCOM), 20th International Conference on, p.1–6. IEEE, 2012.

[10] Chung-Wei Hang et al., *Operators for Propagating Trust and their Evaluation in Social Networks,* Proc. of 8th Int. Conf. on Autonomous Agents and Multiagent Systems (AAMAS), 2009

[11] Wanita Sherchan, Surya Nepal, and Cecile Paris. *A survey of trust in social networks,* ACM Comput. Surv., 45(4):47:1–47:33, August 2013.

[12] Phuong Thanh Pham, Dinh Que Tran, *Incorporation of Experience and Reference-Based Topic Trust with Interests in Social Network,* Advances in Intelligent Systems and Computing 538, Springer, 2017, M. Akagi et al. (eds.).

[13] Dinh Que Tran, Phuong Thanh Pham, *Path Algebra for topic trust computation based on references of users on social networks,* Southeast-Asian J. of Sciences, Vol. 5, No. 1 (2017) pp. 1-8.

[14] Dinh Que Tran. *Classes of functions for topic trust propagation in social network,* Southeast Asian Journal of Sciences, 6(2), 2018.