Analyzing Time-decay Effects of Mediating-objects in Creating Trust-links

Hiroki Takahashi and Masahiro Kimura

Department of Electronics and Informatics
Ryukoku University
Otsu 520-2194, Japan
kimura@rins.ryukoku.ac.jp

Abstract. We address the problem of modeling trust network evolution through social communications among users in a social media site. In particular, we focus on a *social trust-link* created between two users having *mediating-objects* such as *mediating-users* and *mediating-items*, and analyze the time-decay effects of mediating-objects on social trust-link creation. To this end, we first introduce the *basic TCM model* that can be regarded as a conventional link prediction method based on mediating-objects, and propose the *TCM model with time-decay* by incorporating an appropriate time-decay function into it. We present an efficient learning method of the proposed model, and apply it to an analysis of social trust-link creation in a product review site. We show that the proposed model significantly outperforms the basic TCM model in terms of prediction performance, and clarify several properties of user behavior for social trust-link creation.

Keywords: Trust-network evolution model, Mediating-user, Mediating-item, Time-decay effect.

1 Introduction

The advancement of Social Media such as eBay, Epinions and Facebook has allowed us to construct large trust networks, where a trust-link (u, v) from a user u to a user v indicates that u trusts v and tends to be influenced by v. Previous studies [7, 6, 13, 14, 8, 11, 18, 3, 17] have already established the importance of trust in social networks for various processes such as information spreading or search.

Modeling human communication behavior in an online world is an underpinning of mining complex social networks, and a central problem in social network analysis. A trust-link established through *mediating-objects* in a social media site can be regarded as the one created through social communications among users. Representative examples of a mediating-object α from a user u to a user v are as follows. The first one is a *mediating-user* α who has both trust-links from user u and to user v. In this case, user u can meet user v by tracing the trust-links via user α. The second one is a *mediating-item* α for which user u found an activity of user v and thereby knew user v. In the case of a product review site, this indicates that user v posted a review for product α, and user u read it with interest. In this paper, we refer a trust-link created between two users having a mediating-object to as a *social trust-link*, and address the problem of modeling the mechanism of how a social trust-link is created in the context of trust network evolution in discrete-time steps.

To confirm the effects of mediating-objects on social trust-link creation, we first introduce a basic model of creating a trust-link under mediating-objects, which is referred to as the *basic TCM model*. Note that the basic TCM model does not reflect when the corresponding mediating-objects were formed, but manages all of them equally. Analyzing to whom a user creates a social trust-link from the viewpoint of mediating-objects is connected with conventional methods for link prediction. In fact, when two representative methods are employed to assess the *mediating-value* of each mediating-object (i.e., its essential value on trust-link creation), the basic TCM model leads to two widely-used methods in the context of link prediction. One is the *naive method* in which all mediating-objects are equally valued regardless of their intrinsic properties. Then, the basic TCM model corresponds to the *Common Neighbor / Feature method* [12, 2] for link prediction. The other is the *A-A method* in which the intrinsic properties of mediating-objects are taken into account and less active mediating-objects (e.g., mediating-users that have a smaller number of trust-links, mediating-items for which a smaller number of users perform activities, etc.) are more highly valued. In this case, the basic TCM model corresponds to the *Adamic-Adar method* [12, 2] for link prediction. As for creation of a social trust-link, it would in general be reasonable to suppose that older mediating-objects are less influential and recent ones are more influential. However, little attention has been given to analyzing the time-decay effects of mediating-objects in creating a trust-link so far.

To analyze the effects of mediating-objects on social trust-link creation in terms of time-decay, we propose the *TCM model with time-decay* by extending the basic TCM model. As time-decay functions that should be incorporated into the proposed TCM model with time-decay, we adopt two typical ones, *exponential decay* and *power-law decay*. We present an efficient learning method of the proposed model, and apply it to an analysis of social trust-link creation in product review site "Epinions"[1]. We show that the proposed model significantly outperforms the basic TCM model in terms of prediction performance, the power-law decay can be more suitable than the exponential decay for the time-decay function, and the A-A method can be more effective than the naive method for determining mediating-values. Moreover, by employing the TCM model with power-law decay under the A-A method, we analyze how an individual user creates social trust-links from the perspectives of mediating-user, mediating-item and time-decay, and clarify several properties of user behavior.

2 Related Work

Social trust-link creation based on mediating-users is closely related to the *triadic closure mechanism* which is derived from the concept that two people with mutual friends have a higher chance to create a link, and which was regarded as a powerful principle to explain link creation in online social networks [10]. By reflecting other social theories as well, this mechanism was extended to predict positive (trust) and negative (distrust)

[1] `http://www.epinions.com`

links in a signed social network [11]. A method of predicting negative links based on positive links and content-centric interactions was also proposed [16]. Weng et al [19] extended the triadic closure mechanism by exploring the role of information diffusion in the evolution of a social network, and showed that shortcuts based on information flow are another key factor in explaining link formation.

Modeling network evolution can be connected with link prediction and rating prediction in recommender systems. Various link prediction methods were presented in this context, including supervised trust prediction [13, 14], non-negative matrix factorization based on both link and rating information [17], link prediction with explanations for user recommendation systems [2], link prediction from information diffusion data [5], and link prediction in multiple networks [20]. To represent the temporal change of user preference in a recommender system, Koren [9] incorporated a time decay function, and improved the performance in rating prediction. Tang et al [18] enhanced this framework by combining trust network information, and demonstrated its effectiveness in rating prediction and link prediction. Note that these researches were not intended for directly modeling the dynamics of trust network evolution.

Unlike the previous work such as the approaches mentioned above, we focus on modeling the mechanism of creating trust-links under the presence of mediating-objects, and deal with both mediating-user and mediating-item as mediating-object. Furthermore, we provide a novel model and its efficient learning method in order to analyze the time-decay effects of mediating-objects and the difference of mediating-user and mediating-item in influence strength.

3 Analysis Model

3.1 Problem Formulation

For a social media site offering trust-links and activities, we investigate the evolution of trust network in a given time-period $J = [T_f, T_\ell)$ in discrete-time steps, where T_f and T_ℓ are positive integers with $T_f < T_\ell$. Here, we assume that J is not so long (e.g., around six months) since users' behavioral patterns and preferences in an online world can largely change over a long time frame in general. We focus on a social trust-link (i.e., a trust-link created between two users having a mediating-object), and consider analyzing the effects of mediating-objects in creating social trust-links. For each mediating-object α from a user u to a user v, let $\tau_\alpha(u, v)$ denote the time-step at which α first became a mediating-object from u to v. Here, $\tau_\alpha(u, v)$ is called the *mediating-time* of α from u to v. Since validity period of information cannot in general be so long in an online world, we only treat such trust-links and activities that have been relatively recently generated. Thus, we regard α as a mediating-object from u to v at a time-step t if and only if $t - \Delta t_0 \leq \tau_\alpha(u, v) < t$, where Δt_0 is a positive integer (e.g., around three months) specified in advance, and stands for the validity period of information in this social media site. We also assume that the set of mediating-objects are divided into K (≥ 2) kinds of categories including "Mediating-user", "Mediating-item", etc.

Let U be the set of all users in the site during time-period J. For an arbitrary time-step $t \in J$, let U_t denote the set of all users at time-step t, and let $\bar{E}_t$ $(\subset U \times U)$ denote the set of all trust-links created in the set of users U before time-step t. Then, we have $U = \bigcup_{t \in J} U_t$, and $\bar{E}_{t+1} \setminus \bar{E}_t$ indicates the set of all trust-links created in U at time-step t. For any $u, v \in U_t$ and $k \in \{1, \ldots, K\}$, let $\mathcal{M}_{k,t}(u, v)$ denote the set of all mediating-objects of category k from user u to user v at time step t. We assume that $\mathcal{M}_{k,t}(u, v) \cap \mathcal{M}_{\ell,t}(u, v) = \emptyset$ if $k \neq \ell$. For any $u \in U_t$, we define $V_t(u)$ by

$$V_t(u) = \left\{ v \in U_t \;\middle|\; \bigcup_{k=1}^{K} \mathcal{M}_{k,t}(u, v) \neq \emptyset,\ (u, v) \notin \bar{E}_t \right\}.$$

Note that a social trust-link created at a time-step $t \in J$ is represented by (u, v), where $u \in U_t$ and $v \in V_t(u)$.

Suppose that a user $u \in U_t$ creates a social trust-link to some user belonging to $V_t(u)$ at a time-step $t \in J$. Then, to analyze the effects of mediating-objects on social trust-link creation, we consider modeling the probability $P(v \,|\, u, t)$ that the user u creates a trust-link (u, v) to a user $v \in V_t(u)$ at the time-step t. It is conceivable that the influence of a mediating-object $\alpha \in \mathcal{M}_{k,t}(u, v)$ depends on how close the mediating-time $\tau_\alpha(u, v)$ is to the time-step t. Moreover, it is expected that the larger $t - \tau_\alpha(u, v)$ is, the smaller the influence of α becomes. We can also speculate that how influential a mediating-object $\alpha \in \mathcal{M}_{k,t}(u, v)$ is on the creation of trust-link (u, v) depends on its category k. In this paper, we construct such a model of probability $P(v \,|\, u, t)$ that can analyze the effects of time-decay and category for mediating-objects in creating social trust-links.

3.2 Basic Model

First, we introduce a basic model for evaluating the effects of mediating-objects on social trust-link creation, which is shown to be associated with a conventional method for link prediction when a widely-used simple method is employed to assess the essential value of a mediating-object on the basis of its observed features (see Section 5.2). Here, we fix such an assessment method, and determine the *mediating-value* of each mediating-object $\alpha \in \mathcal{M}_{k,t}(u, v)$ of category k from user u to user v at time-step t. Let $r_{k,t}(\alpha)$ denote the mediating-value of α, where $0 < r_{k,t}(\alpha) \leq 1$.

It can be expected that the probability $P(v \,|\, u, t)$ becomes high if there are many mediating-objects of high mediating-values from user u to user $v \in V_t(u)$ at time-step t. Thus, as a model of probability $P(v \,|\, u, t)$ for $v \in V_t(u)$, we define

$$P^{basic}(v \,|\, u, t) \;\propto\; \sum_{k=1}^{K} \sum_{\alpha \in \mathcal{M}_{k,t}(u,v)} r_{k,t}(\alpha) \qquad (\forall v \in V_t(u)), \tag{1}$$

where it is set that $\sum_{\alpha \in \mathcal{M}_{k,t}(u,v)} r_{k,t}(\alpha) = 0$ if $\mathcal{M}_{k,t}(u, v) = \emptyset$. This model (see Eq. (1)) can be regarded as a basic model of trust-link creation based on mediating-objects, and is referred to as the *basic TCM model*.

3.3 Proposed Model

In order to analyze the effects of mediating-objects on creation of social trust-links in terms of time-decay and category, we consider extending the basic TCM model, and propose modeling probability $P(v \mid u, t)$ as

$$P^{decay}(v \mid u, t;\ \boldsymbol{\lambda}, \boldsymbol{\mu}) \ \propto\ \sum_{k=1}^{K} e^{\mu_k} \sum_{\alpha \in \mathcal{M}_{k,t}(u,v)} r_{k,t}(\alpha)\, f(t - \tau_\alpha(u, v); \lambda_k) \qquad (\forall v \in V_t(u)), \qquad (2)$$

where $\boldsymbol{\lambda} = (\lambda_1, \ldots, \lambda_K)$, $(\lambda_1, \ldots, \lambda_K > 0)$ and $\boldsymbol{\mu} = (\mu_1, \ldots, \mu_K)$, $(\mu_1, \ldots, \mu_K \in \mathbf{R})$ are the model parameters whose values are estimated from the observed data, and it is set that $\sum_{\alpha \in \mathcal{M}_{k,t}(u,v)} r_{k,t}(\alpha)\, f(t - \tau_\alpha(u, v); \lambda_k) = 0$ if $\mathcal{M}_{k,t}(u, v) = \emptyset$. Here, for each $k \in \{1, \ldots, K\}$, λ_k represents the time-decay rate of a mediating-object belonging to category k, and e^{μ_k} represents the relative influence strength of mediating-objects belonging to category k. Moreover, $f(s; \lambda_k)$ is a monotone decreasing function for $s > 0$, and models a time-decay effect. In this paper, we adopt two typical time-decay functions related to human behavior [1, 15, 9]. One is

$$f_{ex}(s; \lambda_k) \ =\ C_0 e^{-\lambda_k s} \qquad (3)$$

for $s > 0$, which is called the *exponential decay function*, and the other is

$$f_{pl}(s; \lambda_k) \ =\ C_1 s^{-\lambda_k} \qquad (4)$$

for $s > 0$, which is called the *power-law decay function*. Here, C_0 and C_1 are the normalization constants with $C_0, C_1 > 0$.

The proposed analysis model (see Eq. (2)) is referred to as the *TCM model with time-decay*. In particular, the TCM model with time-decay function $f_{ex}(s; \lambda_k)$ and the TCM model with time-decay function $f_{pl}(s; \lambda_k)$ are called the *TCM model with exponential decay* and the *TCM model with power-law decay*, respectively.

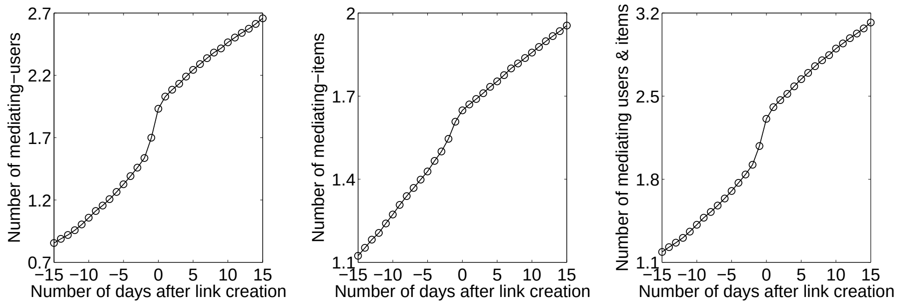
4 Parameter Estimation Method

Let $D_* = \{(u, v, t)\}$ denote the set of all social trust-links created within a time-period $J_* = [T_f, T_*)$, where T_* is a positive integer with $T_f < T_* \leq T_\ell$. Here, $(u, v, t) \in D_*$ indicates that user u created a social trust-link to user v at time-step $t \in J_*$. Then, we consider estimating the parameter values of the TCM model with time-decay from the observed data D_*. Note that to determine the mediating-time $\tau_\alpha(u, v)$ of a mediating-object α from a user u to a user v in the observed time-period J_*, the data in the time-period $J' = [T_f - \Delta t_0, T_f)$ before J_* is also required for the parameter estimation.

To estimate the values of $\boldsymbol{\lambda} = (\lambda_1, \ldots, \lambda_K)$ and $\boldsymbol{\mu} = (\mu_1, \ldots, \mu_K)$ from D_*, we conform to the framework of MAP estimation, and consider maximizing the function

$$\mathcal{L}(\boldsymbol{\lambda}, \boldsymbol{\mu}) \ =\ \sum_{(u,v,t) \in D_*} \log P^{decay}(v \mid u, t; \boldsymbol{\lambda}, \boldsymbol{\mu}) + \sum_{k=1}^{K} \left((b_k - 1) \log \lambda_k - c_k \lambda_k - \frac{{\mu_k}^2}{2{d_k}^2} \right) \qquad (5)$$

with respect to $\boldsymbol{\lambda}$ and $\boldsymbol{\mu}$ (see Eqs. (2), (3) and (4)), where $b_k \geq 1$, $c_k > 0$ and $d_k > 0$ are regularization constants for $k = 1, \ldots, K$. Here, we assume a gamma prior for each λ_k and a Gaussian prior for each μ_k. We consider deriving an iterative algorithm. Let $\bar{\boldsymbol{\lambda}}$ and $\bar{\boldsymbol{\mu}}$ denote the current estimates of $\boldsymbol{\lambda}$ and $\boldsymbol{\mu}$, respectively. By applying Jensen's inequality, we can obtain a convex function $Q(\boldsymbol{\lambda}, \boldsymbol{\mu} \mid \bar{\boldsymbol{\lambda}}, \bar{\boldsymbol{\mu}})$ of $\boldsymbol{\lambda}$ and $\boldsymbol{\mu}$ such that $\mathcal{L}(\boldsymbol{\lambda}, \boldsymbol{\mu}) - \mathcal{L}(\bar{\boldsymbol{\lambda}}, \bar{\boldsymbol{\mu}}) \geq Q(\boldsymbol{\lambda}, \boldsymbol{\mu} \mid \bar{\boldsymbol{\lambda}}, \bar{\boldsymbol{\mu}})$ and $Q(\bar{\boldsymbol{\lambda}}, \bar{\boldsymbol{\mu}} \mid \bar{\boldsymbol{\lambda}}, \bar{\boldsymbol{\mu}}) = 0$. Thus, we can derive an update formula for $\boldsymbol{\lambda}$ and $\boldsymbol{\mu}$ by maximizing $Q(\boldsymbol{\lambda}, \boldsymbol{\mu} \mid \bar{\boldsymbol{\lambda}}, \bar{\boldsymbol{\mu}})$. Due to space constraints, we will elaborate the details of the algorithm in the extended version of our paper.



(a) Mediating-users (b) Mediating-items (c) Mediating users and items

Fig. 1. Relation between the creation of social tust-links and the number of mediating-objects.

5 Experiments

5.1 Epinions Data

We collected real data from social media site "Epinions", where a user can create a trust-link to another user, and post a review for an item in a given set of items. We traced the trust-links by the breadth-first search from the user who was featured as the most popular user in October 2012 until no new users appeared, and collected a set of trust-links and a set of reviews. Then, we aggregated the data into day granularity (i.e., one time-step is set as one day). The collected data included $64,268$ users, $509,293$ trust-links, and $809,517$ reviews for $268,891$ items. We confirmed that all of the indegree distribution (i.e., the fraction of the number of trust-links a user received), the outdegree distribution (i.e., the fraction of the number of trust-links a user created) and the activity distribution (i.e., the fraction of the number of reviews a user posted) exhibit power-law tails, which are known as typical properties of social data in an online world. By taking into consideration the fact that trust-links were constantly generated in 2003, we constructed a dataset from the trust-links and the reviews generated in October 2002 to December 2003. This dataset is referred to as the Epinions data.

5.2 Definition of Mediating-objects

As for mediating-objects to be examined, we focused on two categories (i.e., $K = 2$); category 1 is "Mediating-user" and category 2 is "Mediating-item". For each $t \in J$, let E_t ($\subset U \times U$) denote the set of all trust-links created in the set of users U within time-period $[t - \Delta t_0, t)$. For each $t \in J$ and $u \in U$, let $A_t(u)$ denote the set of items for which user u posted reviews within time-period $[t - \Delta t_0, t)$. Here, by considering the volume of data involved, we simply set the validity period Δt_0 as three months. For each $(u, w) \in E_t$ and $a \in A_t(u)$, let $T_1(u, w)$ denote the time-step at which user u creates trust-link (u, w), and let $T_2(u, a)$ denote the time-step at which user u posts a review for item a. First, we define the category "Mediating-user" as follows: A user α is a *mediating-user* from a user u to a user v at a time-step t when there exist trust-links $(u, \alpha), (\alpha, v) \in E_t$. For a mediating-user α from user u to user v at time-step t, we have $t - \Delta t_0 \leq T_1(u, \alpha), T_1(\alpha, v) < t$, and define the mediating-time $\tau_\alpha(u, v)$ as the maximum of $T_1(u, \alpha)$ and $T_1(\alpha, v)$. Next, we define the category "Mediating-item" as follows: An item α is a *mediating-item* from a user u to a user v at a time-step t when (1) $\alpha \in A_t(u) \cap A_t(v)$ or (2) there exists some user $w \in U_t$ such that $(u, w) \in E_t$ and $\alpha \in A_t(w) \cap A_t(v)$. For a mediating-item α from user u to user v at time-step t, we have $t - \Delta t_0 \leq T_2(w, \alpha), T_2(\alpha, v) < t$, and also define the mediating-time $\tau_\alpha(u, v)$ as the maximum of $T_2(u, \alpha)$ and $T_2(\alpha, v)$. Here, we note that to identify the items for which user v has recently posted reviews and user u can read those reviews with interest, we examine not only the items for which user u has recently posted reviews but also the items for which the users to whom user u has recently created trust-links has recently posted reviews.

For the Epinions data, we investigated a relationship between the creation of a social trust-link and the number of mediating-objects according to the work of Crandall et al [4]. There were 7,965 social trust-links in 2003. More precisely, there were 6,364 trust-links having mediating-users and 3,577 trust-links having mediating-items. For such a social trust-link (u, v), we examined change in the number of mediating-objects from user u to user v as a function of the number of days after the social trust-link (u, v) was created. Figure 1(a) indicates change in the average number of mediating-users for the social trust-links having mediating-users, Figure 1(b) indicates change in the average number of mediating-items for the social trust-links having mediating-items, and Figure 1(c) indicates change in the average number of mediating users and items for all the social trust-links. We can observe a sharp increase in the numbers of mediating users and items immediately before the social trust-link creation (see Fig. 1(c)). This phenomenon can also be observed for the number of mediating-users (see Fig. 1(a)) and the number of mediating-items (see Fig. 1(b)). These results imply that there exists a correlation between the creation of a social trust-link and the number of mediating-objects, and suggest that incorporating the number of mediating-objects can be a promising approach for modeling social trust-link creation.

5.3 Definition of Mediating-values

As for determining the mediating-value $r_{k,t}(\alpha)$ of a mediating-object $\alpha \in \mathcal{M}_{k,t}(u, v)$ of a category k from a user $u \in U_t$ to a user $v \in V_t(u)$ at a time-step $t \in J$, we employed two representative methods in the experiments.

First, we examined the method of equally assessing all mediating-objects; i.e.,

$$r_{k,t}(\alpha) = 1 \qquad (\forall \alpha \in \mathcal{M}_{k,t}(u, v)).$$

This method is referred to as the *naive method* for mediating-values. In this case, we have $P^{base}(v \,|\, u, t) \propto |\mathcal{M}_{1,t}(u, v)| + |\mathcal{M}_{2,t}(u, v)|$ for $v \in V_t(u)$. Here, note that $\mathcal{M}_{1,t}(u, v) = \mathcal{N}^{out}_{1,t}(u) \cap \mathcal{N}^{in}_{1,t}(v)$ and $\mathcal{M}_{2,t}(u, v) = \left(\bigcup_{w \in \mathcal{N}^{out}_{1,t}(u) \cup \{u\}} A_t(w)\right) \cap A_t(v)$, where $\mathcal{N}^{out}_{1,t}(w) = \{w' \in U_t \,|\, (w, w') \in E_t\}$ and $\mathcal{N}^{in}_{1,t}(w) = \{w' \in U_t \,|\, (w', w) \in E_t\}$ for any user $w \in U_t$. Thus, the basic TCM model can be regarded as a kind of Common Neighbor / Feature method [12, 2] for link prediction.

Next, we examined a method in which 1) mediating-users having a smaller number of trust-links are more highly valued and 2) mediating-items for which a smaller number of users posted reviews are more highly valued. We defined $r_{k,t}(\alpha)$ by

$$r_{k,t}(\alpha) = \frac{C_2}{|\mathcal{N}_{k,t}(\alpha)|} + C_3 \qquad (\forall \alpha \in \mathcal{M}_{k,t}(u,v)),$$

where the constants C_2 (> 0) and C_3 were determined as follows: $r_{k,t}(\alpha) = 1$ if $|\mathcal{N}_{k,t}(\alpha)| = 2$ and $r_{k,t}(\alpha) = 0.2$ if $|\mathcal{N}_{1,t}(\alpha)| = \max\{|\mathcal{N}_{1,t'}(\alpha')|;\ t' \in J_0,\ \alpha' \in U_{t'}\}$ for $k = 1$ and $|\mathcal{N}_{2,t}(\alpha)| = \max\{|\mathcal{N}_{2,t'}(\alpha')|;\ t' \in J_0,\ \alpha' \in A\}$ for $k = 2$. Here, $\mathcal{N}_{1,t}(\alpha) = \mathcal{N}^{out}_{1,t}(\alpha) \cup \mathcal{N}^{in}_{1,t}(\alpha)$, $\mathcal{N}_{2,t}(\alpha) = \{w \in U_t \,|\, \alpha \in A_t(w)\}$ and A is the set of all items. This method is referred to as the *A-A method* for mediating-values. We note that in this case, the basic TCM model can be regarded as a kind of Adamic-Adar method [12, 2] for link prediction (see Eq. (1)).

5.4 Datasets

Using the proposed TCM model with time-decay, we analyzed how social trust-links were created in time-period $J = [T_f, T_\ell)$ for the Epinions data. Let $D = \{(u, v, t)\}$ denote the set of all social trust-links created within time-period J, where $(u, v, t) \in D$ indicates that user u created a social trust-link to user v at time-step t. To evaluate the prediction performance of the proposed model, we divide time-period J into a training time-period $J_0 = [T_f, T_m)$ and a test time-period $J_1 = [T_m, T_\ell)$, and define a training set D_0 and a test set D_1 by $D_0 = \{(u, v, t) \in D \,|\, t \in J_0\}$ and $D_1 = \{(u, v, t) \in D \,|\, t \in J_1\}$, where T_m is a positive integer with $T_f < T_m < T_\ell$. In the experiments, we set $|J_0| = |J_1| = \Delta t_0$; i.e., $T_\ell = T_f + 2\Delta t_0$ and $T_m = T_f + \Delta t_0$. As mentioned before, we also used the data in time-period $J' = [T_f - \Delta t_0, T_f)$ before time-period $J = [T_f, T_\ell)$ in order to determine mediating-time.

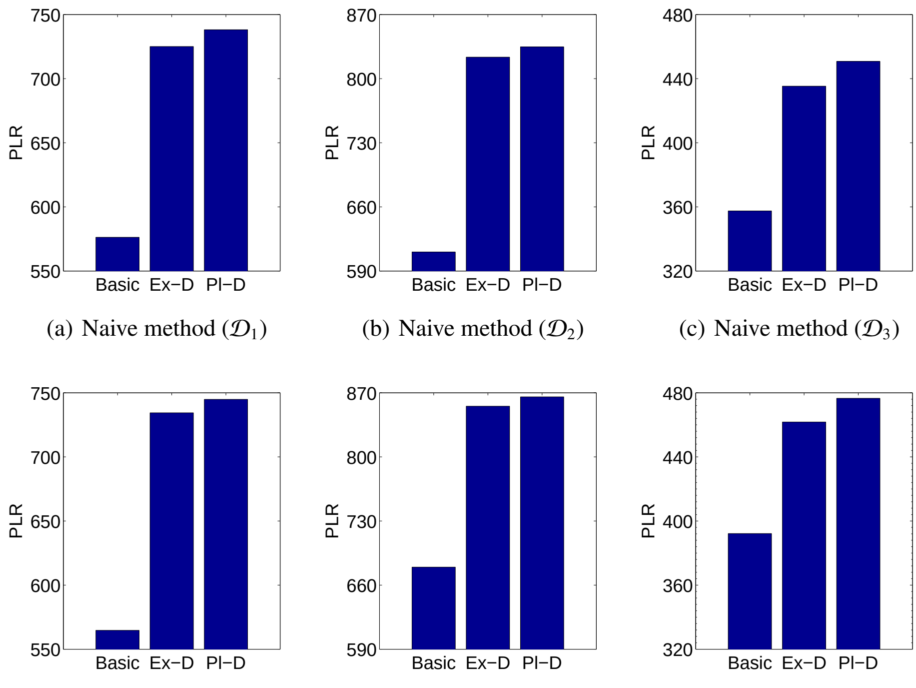
The TCM model with time-decay was learned from training set D_0, and its generalization capability was evaluated using test set D_1, where the regularization constants were set as $b_k = 1$, $c_k = 1$, $d_k = 0.1$ for each category k. We constructed three datasets $\mathcal{D}_1$, $\mathcal{D}_2$ and $\mathcal{D}_3$ by setting J_0 as January to March in 2003 for $\mathcal{D}_1$, April to June in 2003 for $\mathcal{D}_2$ and July to September in 2003 for $\mathcal{D}_3$, respectively. Here, for example, as for $\mathcal{D}_1$, J_1 is April to June in 2003 and J' is October to December in 2002.

5.5 Evaluation Results

For each of the three datasets $\mathcal{D}_1$, $\mathcal{D}_2$ and $\mathcal{D}_3$, we estimated the parameter values of the TCM model with time-decay from training set D_0 (see Section 4), and evaluated the prediction capability of the learned model using test set D_1 in terms of *prediction log-likelihood ratio PLR*. Here, *PLR* is defined by

$$PLR = \sum_{(u,v,t) \in D_1} \left(\log \hat{P}(v \,|\, u, t) - \log \frac{1}{|V_t(u)|} \right) \tag{6}$$

where $\hat{P}(v \,|\, u, t)$ stands for a prediction of probability $P(v \,|\, u, t)$ for $(u, v, t) \in D_1$ by a specified model. Note that $1/|V_t(u)|$ indicates the prediction likelihood of the random guessing for $(u, v, t) \in D_1$. We compared the basic TCM model and the proposed



(a) Naive method ($\mathcal{D}_1$) (b) Naive method ($\mathcal{D}_2$) (c) Naive method ($\mathcal{D}_3$)

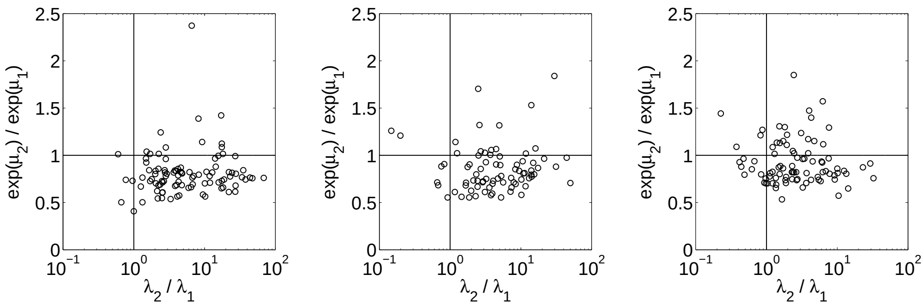
(d) A-A method ($\mathcal{D}_1$) (e) A-A method ($\mathcal{D}_2$) (f) A-A method ($\mathcal{D}_3$)

Fig. 2. Evaluation results of the proposed models.

TCM models in terms of *PLR*, and examined which model is suitable. Figure 2 shows the results for the basic TCM model (Basic), the TCM model with exponential decay (Ex-D) and the TCM model with power-law decay (Pl-D). Here, Figs. 2(a), 2(b) and 2(c) indicate the results for the naive method for mediating-values, and Figs. 2(d), 2(e) and 2(f) indicate the results for the A-A method for mediating-values. Also, the results for datasets $\mathcal{D}_1$, $\mathcal{D}_2$ and $\mathcal{D}_3$ are given in Figs. 2(a) and 2(d), Figs. 2(b) and 2(e), and Figs. 2(c) and 2(f), respectively. *PLR* measures the relative performance versus the random guessing. We observe that the basic TCM model provided much better performance than the random guessing, and reconfirm the importance of exploiting mediating-objects (i.e., the conventional methods for link prediction). Moreover, we see that the TCM model with power-law decay performed the best and the TCM model with exponential decay followed the next. the basic TCM model was always much worse than these two models. These results demonstrate the effectiveness of incorporating time-decay, and also coincide with the observations that many human behaviors follow power-laws [1, 15]. Thus, we focus on the TCM model with power-law decay. Then, in the matter of determining mediating-values, we see that the A-A method can be more effective than the naive method. Hence, for the behavior analysis, we employed the TCM model with power-law decay under the A-A method for mediating-values.

Table 1. The estimated results for the parameters in the TCM model with power-law decay.

	λ_1	λ_2	e^{μ_1}	e^{μ_2}
$\mathcal{D}_1$	0.17	0.47	1.25	0.80
$\mathcal{D}_2$	0.17	0.46	1.15	0.87
$\mathcal{D}_3$	0.25	0.50	1.05	0.96



(a) Results for $\mathcal{D}_1$ (b) Results for $\mathcal{D}_2$ (c) Results for $\mathcal{D}_3$

Fig. 3. Analysis results of user behavior by the TCM model with power-law decay.

Table 1 indicates the estimated results for the parameters in the TCM model with power-law decay under the A-A method. We observe that λ_2 is larger than λ_1, and e^{μ_1} is slightly larger than e^{μ_2}. These results show that from a system-wide point of view, the time-decay rate of a mediating-item is higher than that of a mediating-user, and mediating-users are slightly more influential than mediating-items. This implies that the influence of mediating-items tends to decrease more rapidly than that of mediating-users as time passes.

5.6 Analysis Results

Now, we focus on the behavior of an individual user in creating social trust-links, and analyze it by the TCM model with power-law decay under the A-A method. In the experiments, we investigated the users who created at least 5 social trust-links during the entire time-period J. Such a user is referred to as the *analysis target user*. The number of analysis target users was 96 in dataset $\mathcal{D}_1$, 78 in dataset $\mathcal{D}_2$ and 87 in dataset $\mathcal{D}_3$. For each analysis target user u of each dataset, we estimated the values of parameters λ_1, λ_2, e^{μ_1} and e^{μ_2} in the TCM model with power-law decay by using the data of both the social trust-links created by the user u during the entire time-period J and the corresponding mediating-objects.

Figure 3 shows the analysis results, where it plots e^{μ_2}/e^{μ_1} versus λ_2/λ_1 for all the analysis target users. We observe that the points $(\lambda_2/\lambda_1, e^{\mu_2}/e^{\mu_1})$ plotted on the coordinate plane vary substantially depending on the users. However, the entire tendencies of the results did not depend largely on the datasets. First, most users had the property that λ_2 is larger than λ_1, and e^{μ_2} is slightly smaller than e^{μ_1}, that is, the time-decay rate of a mediating-item tends to be higher than that of a mediating-user, and mediating-items tends to be slightly less influential than mediating-users, which coincides with the system-wide property observed in Section 5.5. Next, the variance for λ_2/λ_1 was much larger than that for e^{μ_2}/e^{μ_1}. This implies that the behavior of each user in creating social trust-links can be characterized by understanding the ratio of the time-decay rate of a mediating-item to that of a mediating-user. Moreover, through our analysis, we can identify several characteristic users, including the users for whom mediating-items were substantially more influential than mediating-users (i.e., $e^{\mu_2} \gg e^{\mu_1}$), and the users for whom the time-decay rate of a mediating-item was substantially lower than that of a mediating-user (i.e., $\lambda_2 \ll \lambda_1$). These results demonstrate the effectiveness of the analysis method based on the TCM model with time-decay.

6 Conclusion

Aiming to construct a better model for trust network evolution through social communications among users in a social media site, we have proposed a novel model that can analyze the effects of time-decay and category with respect to mediating-objects in creating social trust-links. As a basic model of creating social trust-links under mediating-objects, we first introduced the basic TCM model that can be regarded as a conventional link prediction method. In fact, it leads to the Common Neighbor / Feature and the Adamic-Adar methods when the naive and A-A methods are employed to determine mediating-values, respectively. To analyze the effects of mediating-objects in terms of time-decay and category, we proposed the TCM model with time-decay, which derives two models depending on the time-decay functions incorporated; i.e., the TCM model with exponential decay and the TCM model with power-law decay. We presented an efficient method of estimating the values of parameters in the proposed model from the observed data of social trust-links and corresponding mediating-objects. We applied the proposed TCM model with time-decay to real data from product review site “Epinions”, where mediating-users and mediating-items were examined as mediating-objects.

Then, we first showed that the TCM model with time-decay exhibits higher prediction capability than the basic TCM model, demonstrating the effectiveness of incorporating time-decay. We also showed that the TCM model with power-law decay outperforms the TCM model with exponential decay, and as for determining mediating-values, the A-A method is more effective than the naive method. Thus, using the TCM model with power-law decay under the A-A method, we analyzed the behavior of creating social trust-links in the Epinions data. From a system-wide point of view, the time-decay rate of a mediating-item was higher than that of a mediating-user, and mediating-users were slightly more influential than mediating-items. Also, the ratio of the time-decay rate of a mediating-item to that of a mediating-user varied greatly depending on individual users, and enabled us to characterize the behavior of each user in creating social trust-links. Moreover, we identified several characteristic users according to our analysis method. These results demonstrate the effectiveness of the proposed analysis model. Our immediate future work includes extensively evaluating the model for various social media data, investigating various kinds of mediating-objects, and exploring better ways to determine mediating-values.

Acknowledgements. This work was partly supported by JSPS Grant-in-Aid for Scientific Research (C) (No. 26330352), Japan.

References

1. Barabási, A.L.: The origin of bursts and heavy tails in human dynamics. Nature 435, 207–211 (2005)
2. Barbieri, N., Bonchi, F., Manco, G.: Who to follow and why: link prediction with explanations. In: Proceedings of KDD'14. pp. 1266–1275 (2014)
3. Chen, W., Lakshmanan, L., Castillo, C.: Information and influence propagation in social networks. Synthesis Lectures on Data Management 5, 1–177 (2013)
4. Crandall, D., Cosley, D., Huttenlocher, D., Kleinberg, J., Suri, S.: Feedback effects between similarity and social influence in online communities. In: Proceedings of KDD'08. pp. 160–168 (2008)
5. Gomez-Rodriguez, M., Leskovec, J., Krause, A.: Inferring networks of diffusion and influence. In: Proceedings of KDD'10. pp. 1019–1028 (2010)
6. Guha, R., Kumar, R., Raghavan, P., Tomkins, A.: Propagation of trust and distrust. In: Proceedings of WWW'04. pp. 403–412 (2004)
7. Kempe, D., Kleinberg, J., Tardos, E.: Maximizing the spread of influence through a social network. In: Proceedings of KDD'03. pp. 137–146 (2003)
8. Kimura, M., Saito, K., Nakano, R., Motoda, H.: Extracting influential nodes on a social network for information diffusion. Data Mining and Knowledge Discovery 20, 70–97 (2010)
9. Koren, Y.: Collaborative filtering with temporal dynamics. In: Proceedings of KDD'09. pp. 447–456 (2009)
10. Leskovec, J., Backstrom, I., Kumar, K., Tomkins, A.: Microscopic evolution of social networks. In: Proceedings of KDD'08. pp. 462–470 (2008)
11. Leskovec, J., Huttenlocher, D., Kleinberg, J.: Predicting positive and negative links in online social networks. In: Proceedings of WWW'10. pp. 641–650 (2010)
12. Liben-Nowell, D., Kleinberg, J.: The link-prediction problem for social networks. Journal of the American Society for Information Science and Technology 58, 1019–1031 (2007)
13. Liu, H., Lim, E., Lauw, H., Le, M., Sun, A., Srivastava, J., Kim, Y.: Predicting trusts among users of online communities: an epinion case study. In: Proceedings of EC'08. pp. 310–319 (2008)
14. Nguyen, V., Lim, E., Jiang, J., Sun, A.: To trust or not to trust? predicting online trusts using trust antecedent framework. In: Proceedings of ICDM'09. pp. 896–901 (2009)
15. Oliveira, J.G., Barabási, A.L.: Dawin and einstein correspondence patterns. Nature 437, 1251 (2005)
16. Tang, J., Chang, S., Aggarwal, C., Liu, F.: Negative link prediction in social media. In: Proceedings of WSDM'15. pp. 87–96 (2015)
17. Tang, J., Gao, H., Hu, X., Liu, H.: Exploiting homophily effect for trust prediction. In: Proceedings of WSDM'13. pp. 53–62 (2013)
18. Tang, J., Gao, H., Liu, H., Sarma, A.D.: etrust: understanding trust evolution in an online world. In: Proceedings of KDD'12. pp. 253–261 (2012)
19. Weng, L., Ratkiewicz, J., Perra, N., Goncalves, B., Castillo, C., Bonchi, F., Schifanella, R., Menczer, F., Flammini, A.: The role of information diffusion in the evolution of social networks. In: Proceedings of KDD'13. pp. 356–364 (2013)
20. Zhang, J., Yu, P.S., Zhou, Z.: Meta-path based multi-network collective link prediction. In: Proceedings of KDD'14. pp. 1286–1295 (2014)