

Application of social network metrics to a trust-aware collaborative model for generating personalized user recommendations

Iraklis Varlamis¹, Magdalini Eirinaki², and Malamati Louta³

¹ Harokopio University, Department of Informatics and Telematics, Athens, Greece,
varlamis@hua.gr

² San Jose State University, Computer Engineering Department, San Jose, CA, USA
magdalini.eirinaki@sjsu.edu

³ University of Western Macedonia, Department of Informatics and Telecommunications
Engineering, Kozani, Greece
louta@uowm.gr

Abstract. Social network analysis has emerged as a key technique in modern sociology, but has recently gained a lot of interest in Web mining research, because of the advent and the increasing popularity of social media, such as blogs, social networks, micro-blogging, customer review sites etc. Such media often serve as platforms for information dissemination and product placement or promotion. One way to improve the quality of recommendations provided to the members of social networks is to use trustworthy resources. In this environment, community-based reputation can help estimating the trustworthiness of individual users. Consequently, influence and trust are becoming essential qualities among user interactions. In this work, we perform an extensive study of various metrics related to the aforementioned elements, and of their effect in the process of information propagation in social networks. In order to better understand the properties of links and the dynamics of social networks, we distinguish between permanent and transient links and in the latter case, we consider the link freshness. Moreover, we distinguish between the propagation of trust in a local level and the effect of global influence and compare suggestions provided by locally trusted or globally influential users.

Keywords: user recommendations, influence, trust, social networks, centrality, global trust metrics, local trust metrics.

1 Introduction

Social network analysis has been a major area of research for sociologists for many years, but recently has attracted the interest of the Web mining research community. The advent of participatory Web (Web 2.0) and the enormous increase in the use of social networking websites, customer review sites, blogs etc. has turned Web users from information consumers to information producers, thus creating for Web researchers a huge repository of user-provided content. Such content presents features, which are

unique to the Web, in terms of shared authorship, multitude of user-provided tags, inherent connectivity between users and the content they provide, and high update rate. All these characteristics provide a platform that can be exploited in order to mine interesting information about the dynamics of users' interactions.

One common type of analysis is the detection of communities of users with similar interests, and within such communities the identification of the most "influential users". We may define influential users as individuals with many connections within the community, but in general other definitions are possible depending on the type of the community and the social network that interconnects the community members. Influential users act as authorities or hubs within their community and thus play a key role in spreading information. This has obvious implications on "word of mouth" and viral marketing, as indicated in recent studies [1, 2], which in turn makes influential users important for the promotion and endorsement of new products or ideas.

On a slightly different note, another common type of analysis is that of content ranking, in other words finding "influential content", whether this is a product review, a blog or a tweet. Such ranking is becoming increasingly important since online social media expand, in terms of content and users, on a very rapid pace, making navigation cumbersome and time-consuming. This process helps in that the top-ranked items (reviews, blogs, comments, tweets, etc.) can be used as recommendations to the users. Most of existing work in this area generates overall rankings [3–5], and only recently there have been some efforts in personalizing the rankings [6, 7].

Finally, the notion of trust prediction and propagation, used primarily to find trustful (or distrustful) nodes in P2P networks [8] has been adopted in the context of virtual communities [9–12]. It has been shown empirically [13, 14] that in a specific context, it can be assumed that trust may propagate (with appropriate discounting) through the relationship network [9, 14]. Trust has also been shown to be highly correlated to profile similarity and the model is used for recommending items [15–18].

In this work, we bridge these research directions. Our objective is to generate personalized *user recommendations* based on the analysis of implicit or explicit link information between users and user provided content. Depending on the nature of each specific social network, the link information may express trust to the user or simply interest to the content being pointed by the link. Since hyperlinks do not explicitly carry semantic information, our graph analysis model can either discover trustful, influential or interesting nodes depending on the social network.

This paper extends our previous work [7, 19], employing the notion of trust among users as expressed with links, to generate personalized user rankings and recommendations. In this work we incorporate the notion of influence in the ranking algorithm along with the freshness of the trust connections between users and perform an extensive study by integrating several link analysis algorithms in the ranking process in order to get insights of how different influence metrics, such as the degree, closeness, betweenness and centrality or the hub, authority or PageRank scores of a node affect the overall ranking.

The remaining of this paper is organized as follows; in the following section, we present an overview of related research in ranking, influence and trust in social networks. In section 3 we present the background of our work, concerning user recom-

recommendations based on collaborative knowledge from locally trusted or globally influential users. In section 4, we introduce our model, which combines recommendations from trusted (or neighboring) users with those of “influential” users. Section 5 presents the results of the evaluation we performed on a social network which comprises users, links of trust between them and product reviews and ratings. Finally, section 6 summarizes our findings and presents our next steps in this work.

2 Related Work

Studying and analyzing Web 2.0 media, such as social networks, blogs, forums, wikis etc. has gained momentum, resulting in an increase of research in the related fields. Among the several facets of these social media, *trust*, *influence*, and *ranking* are receiving a lot of attention.

The notion of trust propagation within a network was first introduced in the context of P2P networks. The most well-known algorithm in this context is EigenTrust [8]. This reputation management algorithm, is based on the fact that trust is transitive, in that, if a node i trusts another node j , then it also trusts the nodes trusted by j . The algorithm first allows each node to perform a local ranking of the neighboring nodes. These local rankings are then accumulated to give a global “trustfulness” rank to each node in the network. The more “trustworthy” a node is, the more weight their rating of other nodes has.

This idea of trust transitivity, prediction, and propagation, has lately been adopted in the context of virtual communities [14, 9–12]. Golbeck [16] defines the trust rating between two users (called source and sink) as the weighted average of the source’s neighbors’ ratings of the sink. It has been shown empirically [13, 14] that, in a specific context, it can be assumed that trust may propagate (with appropriate discounting) through the relationship network [9, 14]. Massa and Avesani assume that the users state how much they consider trustworthy each other user and his/her ratings and exploit trust propagation over the network in order to infer more trust relationships [18]. A detailed survey of the use of trust in recommender systems is presented in [15].

In our work we introduce the notion of trust in a social network, assuming that trust between a pair of users is already known, either explicitly or implicitly. Our model builds on the idea of integrating local trust and global influence, as suggested in [8]. More specifically, in our model the propagation of trust within a social network is complemented by the more general notion of “influence” within the network and as a result personalized user recommendations are formed considering people that the user trusts and people that have big influence on the whole social network.

Influence in social networks, a topic extensively studied in the pre-WWW era [20], has again emerged as a research topic. One common approach is to model the identification of influencers as a combinatorial optimization problem: given a fixed number of nodes that can be initially activated or infected, find the set of nodes with maximum influence over the entire network - the one that generates the largest cascade of adoptions [1]. Several works build on this Information Cascade (IC) notion proposing various machine learning algorithms [2, 21–24]. Even though such approaches have been shown to improve over traditional social network analysis metrics, they are solely based on

the link structure of social networks, and do not take into consideration other important parameters, such as activity, rate of updates, and trust among users. In the same vein, researchers have investigated the identification of likely influential users through a combination of link analysis techniques [6, 25], as well as user activity-related parameters in order to identify influential users in blogs [26, 27] and social networks [28]. As shown from the analysis above, most of the work in identification of influencers within a social network (real or online) is based on extensions of well known link analysis algorithms and as such, exploit the structural characteristics of the network. In this work we follow a similar approach and incorporate such metrics in our ranking method.

Ranking on the Web is primarily based on the analysis of the Web graph as it is formulated by hyperlinks. In the case of blogs, several ranking algorithms have been suggested that exploit explicit (EigenRumor algorithm [3]) and/or implicit (BlogRank [4, 5]) hyperlinks between blogs. All these algorithms formulate a graph of blogs, based on hyperlinks and then apply PageRank or a variation of it in order to provide an overall ranking of blogs. However, all these algorithms provide a static measure of blog importance that does not reflect the temporal aspects accompanying the evolution of the blogosphere.

In our previous work [7] we introduced a collaborative rating mechanism, which exploits the explicit connections between users and other implicit connections and provides each user with personalized rankings of other users in the network. The rating mechanism employs direct and indirect information from a user's neighborhood. In [29] we presented a global rating model for the blogosphere. The model distinguishes between explicit links between blogs (the links in the blogroll) and implicit links (links between individual posts). The model also captures the time dimension of links. Using links' freshness, we manage to punish blogs that artificially receive a large number of links in a small period of time and are ignored thereafter, we reward blogs that are constantly being referenced by other blogs and successfully distinguish between normal and spam blogs.

To the best of our knowledge, this is the first extensive study of the effect of both overall "influence", as expressed by the analysis of the whole social graph, as well as by personalized aspects of "influence" such as trust, in ranking and recommending other users or content. This paper is an extended version of the paper presented in [19], including more detailed overview of the various algorithms employed, an updated ranking algorithm that incorporates the notion of "freshness" of connections, and a more extensive experimental evaluation.

3 Preliminaries

Social network analysis is the study of social entities (*actors*) and their interactions and relationships. The interaction and relationships are represented as a graph, where each node represents an actor (user), and the edge between two nodes represents their relationship. Several link analysis algorithms have been proposed, that are applied on such graphs in order to identify and analyze the role, position, and influence of each user.

In our work, we employ social network analysis metrics such as centrality and rank prestige, in order to identify the “influential” actors in a social network, in terms of their position in the graph and their connections/interactions with other users [30–32, 20]. *Centrality* identifies as important actors (i.e. users) those that are linked (i.e. involved) extensively with other actors. *Prestige* is a more refined measure since it differentiates between in-links and out-links, focusing on in-links. In other words, the importance of an actor depends on the opinion of other actors, expressed by their ties to her. More specifically, we are interested in *rank prestige*, that also takes into account the prominence of individual actors that participate in this “voting” process. Rank prestige has been the basis upon which both HITS [33] and PageRank [34] were built.

In addition to these global metrics, influence in a local scale is important for all actors. In this context, actors collaborate with the actors they trust and are influenced by their opinions. Moreover, trust and influence are reinforced for certain actors in the circle of trust and decrease for others. In order to model the dynamics of trust and influence in the “neighborhood” of a user, we employ our *collaborative local* scoring mechanism. In what follows, we provide a brief overview of the aforementioned metrics [7, 35].

3.1 Social Network Analysis Metrics

Centrality. The three centrality metrics, namely degree, closeness, and betweenness centrality, identify “key” users of the graph, in terms of information dissemination. Let n denote the size of the graph (i.e. the number of actors/users).

Degree Centrality $Gd(i)$ takes into consideration the node degree $d(i)$ of a user i . The higher the node degree, the more central the user is:

$$Gd(i) = \frac{d(i)}{n - 1} \quad (1)$$

Closeness Centrality $Gc(i)$ of a user i signifies how easily this user interacts with all other users j ($j \in [1..n]$). Let $d(i, j)$ denote the distance of user i from user j , equal to the number of links in a shortest path. Then, according to closeness centrality, the shorter the distance of the user to all other actors, the more central the user is:

$$Gc(i) = \frac{n - 1}{\sum_{j=1}^n d(i, j)} \quad (2)$$

Finally, *Betweenness Centrality* $Gb(i)$ signifies the importance of user i with regards to the flow of information in the social network. If the user is between two non-adjacent users j and k then i has control over their interactions. If i is on the paths of many such interactions (i.e. *between* many users), then this is an important user, having a great amount of *influence* on what happens in the network. Let sp_{jk} be the number of shortest paths between j and k , and $sp_{jk}(i)$, ($j \neq i$ and $k \neq i$) be the number of shortest paths that pass i . Betweenness centrality of a user i is defined as follows:

$$Gb(i) = \sum_{j < k} \frac{sp_{jk}(i)}{sp_{jk}} \quad (3)$$

Hubs and Authorities. Both terms were introduced as part of the well-known algorithm HITS [33]. A *hub* is a node with many out-links and an *authority* is a node with many in-links. The key idea of HITS is that hubs and authorities have a mutual reinforcement relationship, since a good hub points to many good authorities, and a good authority is pointed to by many good hubs. Transferring the algorithm to the social network paradigm, the authority score $Ga(i)$ of user i is the sum of all hub scores $Gh(j)$ of users j that have a (directed) relationship with i (this directed relationship can be, for example, a declaration of trust in a social network or a comment on a blog post). The hub score $Gh(i)$ of user i is defined similarly. Let E be the set of directed edges (i.e. links) in the graph, then the authority $Ga(i)$ and hub $Gh(i)$ scores are iteratively calculated as follows:

$$Ga(i) = \sum_{(j,i) \in E} Gh(j) \quad (4)$$

$$Gh(i) = \sum_{(i,j) \in E} Ga(j) \quad (5)$$

PageRank. PageRank [34] also identifies “authorities” in a graph. The intuition is that the more actors “endorse” or vote for an actor i (i.e. add a link pointing to i), the more important i is. What is more, prominence of the endorsers is crucial, since the vote of important actors is more valuable. Transferring this notion to the social network paradigm, a user i is considered to be influential if a) many other users endorse i (for example by “trusting” i , adding i ’s blog in their blogroll, or becoming i ’s followers), and b) these users are in turn influential. The PageRank score $Gp(i)$ of user i is iteratively computed as follows:

$$Gp(i) = (1 - d) + d \sum_{(j,i) \in E} \frac{Gp(j)}{O_j} \quad (6)$$

where O_j denotes the number of out-links of node j and d is the so-called damping factor.

3.2 Collaborative rating in social networks

In [7] and [19] we presented a personalized recommendation model, which capitalizes on a collaborative rating mechanism that exploits the structure of the social network. The model represents the social network as a directed graph $G = (V, E)$, where users are the nodes V and the implicit or explicit links between users are the edges E of the graph. Explicit links between two users in a social network denote a permanent recommendation and trust, while on the other hand, implicit links represent a temporary interest and thus, a more transient reference to the user being pointed. In any case, it has been assumed that the intention of a user i when adding a link towards user j is to provide a positive recommendation for j to other users in the network.

The model suggests a quantification of user’s i opinion with respect to user j , called local score (LS), which is updated in time, so as to follow the dynamic nature of the social networks and capture the “freshness” of information available.

$$LS_t(i, j) = w_{BR} \cdot BR_t(i, j) + w_{EP} \cdot EP_t(i, j) \quad (7)$$

In Equation 7, the local score for a user j as expressed by another user i , at a certain time period t , is the weighted combination of two factors: a) $BR_t(i, j)$, which corresponds to what i explicitly denotes about j , and b) $EP_t(i, j)$, which corresponds to what i implicitly believes about j . The first factor constitutes of the number of explicit/permanent links, which, for example, in the case of blogs, can be those in the blogroll list of a user, in the case of social networking applications can be the “friend” links or in the case of consumer networks can be the links to the “members of trust”. The second factor considers implicit links and can be, for example, the number of links from blog i to blog j (e.g. if user X , the owner of blog i , adds a link to a post in blog j , owned by user Y , this implicitly means that X likes Y). The definition of weights depends on the type of social network we examine and the relative significance we give to explicit and implicit expressions of trust or interest. Local score estimation takes place at consecutive, equally distributed, time intervals.

Due to the dynamic nature of social networks, users may add new links (i.e. new recommendations) to the same targets, thus, reinforcing their initial recommendations or withdraw some old links, weakening their former positions. In order to capture the links’ “freshness”, we proposed an extension, named *local accumulative score (LAS)*, which aggregates the local scores $LS_t(i, j)$ of previous time periods t in order to find the score in the current period c , attributing higher significance to recent time periods.

$$LAS_c(i, j) = \sum_{\substack{t=c-m+1 \\ t > 0}}^c w_t \cdot LS_t(i, j) \quad (8)$$

Weights are given by the following equation:

$$w_t = \frac{f(t)}{\sum_{\substack{l=c-m+1 \\ l > 0}}^c f(l)} \quad (9)$$

where $f(t) = \begin{cases} m - c + t, & c \geq m \\ t, & c < m \end{cases}$ and $\sum_{\substack{t=c-m+1 \\ t > 0}}^c w_t = 1$.

For the calculation of the local accumulative score at time period c , only the m most recent local scores formed are considered. Parameter m stands for the “system memory”⁴, which in essence determines the number of periods back in time that we consider for aggregating the local scores. A small value for the parameter m means that the memory of the system is small, whereas large value considers a large memory for the system. Depending on the type of the social network considered, the definition of the value of the parameter m may yield interesting results. Equation 9 in essence

⁴ We assume that the system does not process the complete history of user ratings and friendship links from the very beginning of the community, but only the most recent interactions. Thus we use the term “system memory” to refer to the size (in time units) of this history log.

models the fact that more recent local scores should weigh more in the overall *LAS* evaluation.

We subsequently extend the local accumulative scores produced in a first step, introducing the concept of *collaborative local score (CLS)*. $CLS_c(i, j)$ score aggregates at time period c the direct accumulative scores $LAS_c(i, j)$, assigned by user i to any user j , with the indirect accumulative scores $LAS_c(k, j)$ attributed to user j by all users k belonging in the WS_i set, a subset of the set V , formed by users that user i trusts (hereafter referred to as the “witnesses”). WS_i set comprises the users that are explicitly connected to user i (as determined by user’s i explicit links - depth=1), while adoption of the transitivity property of trust leads to its more general form including the users that are explicitly connected to user’s i trusted users (depth=2), the users that are explicitly connected to the trusted users of user’s i trusted users (depth=3), etc.

$$CLS_c(i, j) = w_i \cdot LAS_c(i, j) + \sum_{k \in WS_i} w_k \cdot LAS_c(k, j) \quad (10)$$

As may be observed, the collaborative local score is a weighted combination of the user’s i direct opinion and the recommendations collected from a number of user’s i witnesses with respect to user j . In general, weight w_k is a measure of the trustworthiness of witness k in the eyes of user i , depends on the transitivity horizon considered and may be a function of the *LAS* attributed to each user in the respective trust chain formed. Considering transitivity horizon equal to n , the weight w_k along a specific trust chain may be given by the following equation:

$$w_k = \frac{w_{u(d=1)} \cdot w_{u(d=2)} \cdot \dots \cdot w_{u(d=n)}}{n} \quad (11)$$

where $w_{u(d=x)}$ denotes the weight of user u in the position/depth x of the trust chain and $w_{u(d=x)} = \frac{LAS(u(d=x-1), u(d=x))}{\sum LAS(u(d=x-1), u(d=x))}$. At this point it should be noted that we have assumed that user i is connected to witness user k only through the specific trust chain. This assumption may be readily relaxed. Besides the multiplicative function considered in equation 11, other functions could be defined as well.

4 Influence Model

Our objective is to generate personalized recommendations to the members of social networks. These recommendations may refer to users, blogs/blog posts, comments, tweets, content reviews, etc. In order for such recommendations to be personalized, a ranking algorithm is needed.

In this work, we propose a model that enhances our previous approach on social networks, by involving both the circle of trust of a user, as well as the overall influential users of a social network in the ranking process. Our objective is to compare and evaluate the importance of different types of users in a social network. Such users might belong to the immediate network of trust of the user, the extended network of trust of the user, or the overall conception of trust among all users in the network. Please note that the same model can be applied to any social medium, for example blogs (where

“trust” is expressed by adding a blog in one’s blogroll), tweets (where “trust” is shown by following a tweeter), or consumer networks (where “trust” is shown explicitly by endorsement or reviews).

To this direction, we extend the collaborative model of Equation 10 to include a *Global Influence* model GI . This global influence model results to a global ranking of all users in the social network, based on their position in the social graph and their connections to all other users. In essence, the global influence $GI(i)$ of user i is an indication of the importance of this user in the whole social graph and is a linear combination of the six models presented in Section 3.1:

$$\begin{aligned}
 GI(i) = & w_d \cdot Gd(i) + w_c \cdot Gc(i) + \\
 & w_b \cdot Gb(i) + w_h \cdot Gh(i) + \\
 & w_a \cdot Ga(i) + w_p \cdot Gp(i)
 \end{aligned} \tag{12}$$

Note that using the aforementioned formula, we may give more importance to one (or more) global influence metric(s) and diminish others.

Our proposed model computes the influence score $INF_c^i(j)$ as a function of the ratings/trust provided for any user j by a) user i , b) the network of trust of user i , and c) the globally influential users:

$$\begin{aligned}
 INF_c(i, j) = & f(LAS_c(i, j), \sum_{k \in WS_i} w_k \cdot LAS_c(k, j), \\
 & \sum_{(m,j) \in E} GI(m) \cdot LAS_c(m, j))
 \end{aligned} \tag{13}$$

This function could be, for instance, a weighted sum of the three factors of Equation 13:

$$\begin{aligned}
 INF_c(i, j) = & w_{local} \cdot LAS_c(i, j) + \\
 & w_{collab} \cdot \sum_{k \in WS_i} w_k \cdot LAS_c(k, j) + \\
 & w_{global} \cdot \sum_{(m,j) \in E} GI(m) \cdot LAS_c(m, j)
 \end{aligned} \tag{14}$$

Equation 14 assumes that the weights are normalized. The weighted sum approach has been used in a related context (identification of influential bloggers) with great success [26]. Alternatively we can produce different rankings using each local and global metric and then merge the rankings in a single ranked list [36], or use an ensemble ranking [37].

The combined model for social networks has a dual meaning: a member of the social network decides upon her own beliefs and upon suggestions of people she trusts and is influenced by the central/powerfull members of the network. The three different weights in Equation 14 represent the balance between the three different types of influence: w_{local} for the user’s own beliefs, w_{collab} for the user’s extended network beliefs and w_{global} for influential users’ beliefs. Moreover, each component weighs differently each participant, with each user k in the network of trust of user i receiving a different weight $w(k)$, and each globally influential user m receiving a weight proportional to her importance in the graph ($GI(m)$). We should point out that the aforementioned model can be easily adjusted to cater for newcomers to a social network. Newcomers don’t have a “circle of trust” yet so the local metrics cannot be applied. In that case the

ranking can be made on the basis of the global influence of each user, and be personalized/adjusted when the user starts making more connections in the network.

The final outcome of our model is a personalized set of influence scores for all other users in a social network. These influence scores can be used to rank the users, and this ranking can be subsequently used to generate recommendations to the current user i . For example, in the blogosphere, the model recommends the personalized top- k list of influential blogs, taking into consideration the user's personal network of trust and overall influence of blogs. In a micro-blogging site such as Twitter, the model will generate a personalized set of trusted and influential "followees", whereas in a social network, the model will generate a personalized set of trusted and influential users.

In the case of decentralized networks (e.g. p2p networks, ad-hoc social networks etc), finding the global influence is difficult or even infeasible, so the local alternative is the only solution. On the other hand, in a web-based social networking application (e.g. Twitter), trust scores and item ratings are all stored in a central node thus permitting the application of both local and global approaches. In the following section we evaluate the local and global approaches and their combinations in a centralized social networking application.

5 Experimental Evaluation

The aim of our study is to compare the performance of local and global models of influence in providing recommendations to the users of social networks and combine them in a single model. For the evaluation of the different models, we employed a dataset which refers to a network of buyers. The *extended Epinions dataset*, which was provided by Epinions and is available through the Trustlet wikipedia⁵ contains information about product reviews written by the members of the Epinions community. It contains approximately 132,000 users who issued 841,372 statements. More specifically, each user provides ratings for users (1 and -1 for trusted and distrusted users respectively) and ratings for the reviews written by other users (ranging from 1 to 6). Finally, the dataset contains information about the author and subject of each review, thus, giving us evidence on the interests of each author.

We model our social graph as follows: The users are the nodes, and the user ratings are the permanent links of the network, used to define the circle-of-trust of each user. The article ratings are considered as the transient expressions of trust or influence. During the preprocessing phase, we kept the 717,667 positive trust ratings and removed self-references, i.e., statements about users trusting themselves.

We divided the dataset in three distinct subsets: a) one that includes users with a narrow circle-of-trust (set A: users having between 5 and 10 outlinks), b) one that includes users with a medium-sized circle-of-trust (set B: users having 15-25 outlinks) and c) one that includes users with an extended circle-of-trust (set C: users having more than 30 outlinks). Sets A and C are of equal size (set A contains 5,425 and set C contains 5,405 users) but significantly differ in the connectivity of their nodes. The size of set B is 2,349.

⁵ <http://www.trustlet.org/wiki>

As mentioned in the introduction, depending on the nature of each social network, the proposed model can be appropriately adapted to provide users with personalized recommendations that correspond to trustful or influential users. In the context of the *Epinions* network, user-to-user links express the trust between users, and user-to-item links imply the interest of a user to a specific item, formulating a network of trust among users. A recommendation for a user U_i of this network will be a set of users U that U_i can trust. Since *Epinions* is a buyers’ network, recommending user U_j to user U_i based on the analysis of the graph means that U_i can trust U_j ’s opinion and product reviews. As a result, we expect a big overlap in the lists of items bought (or reviewed) by U_i and U_j . Thus, in order to evaluate our approach, we examine whether the users recommended to U_i (users in U) have matching interests with U_i .

The similarity between two users U_i and U_j is defined as the number of items rated by U_i that have been also rated by U_j . This measure is similar to the bibliographic measure of *coupling* which is based on the number of common references between two users. We compute the average similarity between U_i and the top- k recommended users of each ranking and compare it to the baseline T , which is the average similarity between U_i and the users U_t to whom U_i is connected via an explicit trust link ($(U_i, U_t) \in E$).

The first experiment, examines the effect of the memory size to the performance of the accumulative local model. In the subsequent experiments, we first evaluate each model (namely, the local, the collaborative local, and each one of the global influence models) individually (setting the respective weight in Eq. 12 to 1 and the remaining weights to 0). Based on the results of this experiment, which show that the collaborative local model outperforms the local one, we combine the collaborative local model with each of the global influence models (setting equal weights for w_{collab} and w_{global} and $w_{local}=0$ in Eq. 14). Finally, based on our findings, we combine those global influence models that performed better in the second step with the collaborative local models. The detailed weight values for this experiment are explained in subsection 5.5. We performed each set of experiments for all sets of users.

5.1 The effect of memory in the accumulative local model

In the first experiment, we examine how the memory size affects the performance of the accumulative local model. For this reason, we vary the memory size m between 0 and 15 months, which means that we do not take into account links that have been issued before the last m months in the dataset. For a memory size m and a link issued in period t ($t \in [c - m + 1, c]$) the freshness-related weight w_t for the current period c is given by Equation 9. We employ the users in set B and generate for each user U a ranked list of recommended users, using the local accumulative (L) formation. The average similarity is measured for all users in set-C compared against the top- k users in the recommendation list after removing the current friends of each user. Experimental results on sets A and C have been omitted, since these two sets contain users either with too few trusted users (set A), which results in very few fresh links or with too many trusted users (set C), which results in huge lists of links.

In Figure 1, we present the performance curves for the top-3, top-15 and top-30 users. All intermediate curves (top-5, top-10 etc.) have been omitted, to improve read-

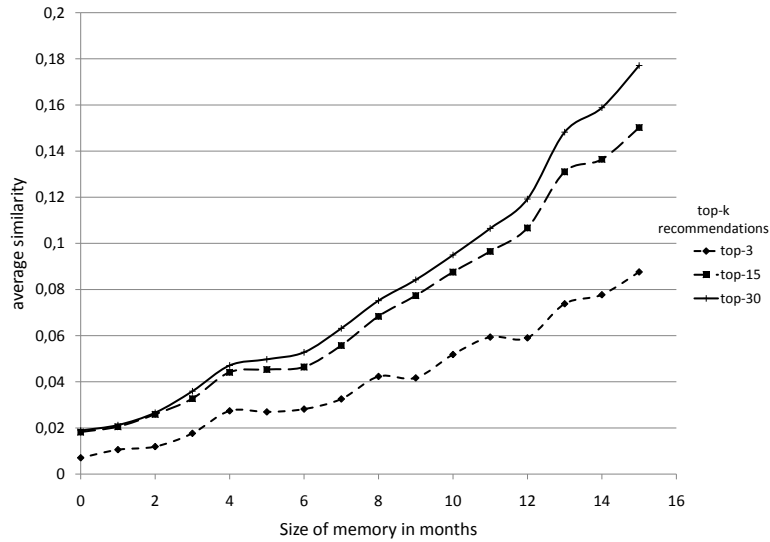


Fig. 1. The effect of links’ freshness and memory size in the performance of the Local Accumulative Score (LAS)

ability of the chart. However, they all fall between the top-3 and top-30 curves depicted in Figure 1. All curves show that an increase in memory size results in an increase to the average similarity. An interpretation of the results is that the use of a user’s rating history can improve the performance of the accumulative local metric. The role of link weighting, according to Equation 9, is crucial, since it prioritizes the fresh links and punishes obsolete ones. For example, recommendations from a friend that has been added a year ago, and whose rating has never been updated again (e.g. with a “like” link), will receive lower values than recommendations from a friend that has been added recently. In this way, we manage to provide recommendations that reflect the recent activity of each user.

In all the following experiments, we decide to use the complete recommendation history for articles and authors. Although it is more expensive, it can significantly boost the system performance as will be demonstrated in the following paragraphs. For example, keeping only the links to authors and articles that have been created in the last 15 months ($m=15$) results in ignoring more than 50% of the link information of the dataset, since more than half of the article links are older than 15 months. As a consequence, the performance of the local accumulative formation measured using average similarity (as depicted in Figure 1) ranges between 0.02 and 0.18, whereas in all the experiments that follow it is constantly above 0.20.

5.2 Evaluation of the individual models

In the second experiment, we generate for each user U (member of sets A, B and C respectively) a ranked list of recommended users, using the local accumulative (L) and

the collaborative local (CL) formation. We also generate the overall (global) rankings of all users using each centrality metric (Gd using degree centrality, Gc using closeness centrality, Gb using betweenness centrality, Gh using hub score, Ga using authority score and Gp using PageRank score). We then select the top- k users from each list. These are the recommended users.

The following Figures (2 -13) show the similarity between a user and the users in the top- k positions of the recommendations' list. As explained previously, similarity is measured in terms of commonalities, i.e. of items that have been rated by both the user and a user in his/her top- k list. When the top- k list contains users with similar interests, the similarity is expected to increase for higher k values, because the number of items that have been rated by both the user and a user in the list is expected to increase. When the top- k list contains users with completely dissimilar interests, the similarity will remain stable. The y-axis in all the following Figures is labeled "average similarity", because it averages the similarity values of all users in the examined set.

A reference ranking is not available so we cannot measure the correlation with our rankings (using Spearman's ρ or Kendall's τ). Golden standards (lists of correct and incorrect recommendations per user) are not available too, so we cannot measure performance using IR metrics in our top- k lists. The use of blind user testing is not feasible in this scale, so we decided to comparatively evaluate our approaches. A standard baseline for comparison could be to recommend to each user k randomly selected users from the opinion dataset. However, this *random* baseline results in very low average similarity values, and is not helpful for comparison. The users who already are in the friend list of a user constitute a more antagonistic baseline for our methods, which is identified as T in figures 2-7 and 11-13. When a method outperforms T in a certain k value, it means that the top- k users in the corresponding recommendation list have more similarities with the target user, than his/her own friends.

Figure 2 presents the average similarity values for the top- k matches ($k=3, 5, 7, 10, 15, 20, 25, 30$) for set A, which comprises users with few trusted nodes. Figure 3 refers to members of set B, who have between 15 and 25 direct neighbors, whereas Figure 4 shows users of set C, who have many trusted nodes in their circle.

From the results presented in Figures 2, 3 and 4 we observe that the collaborative local model (CL) significantly improves the performance of the baseline (T), especially for users with a small circle of trust (set A). This implies that it is useful for a recommendation model to check for suggestions beyond the direct neighbors of a node, in the extended neighborhood of users (in terms of links of trust).

On the contrary, the performance of the global rating models is comparable or even worse than the baseline. In several cases (i.e. when hub, authority or centrality are employed) the performance reaches zero. This means that there are no similarities between the user's likes and those of the top ranked users in the whole network.

On the other hand, for users with more neighbors (sets B and C), certain global models (i.e. degree, betweenness and PageRank) perform better than local models when the top- k recommendations are examined. An explanation of this is that users in set B and even more in set C have many direct or indirect neighbors so these users are probably connected to some of the highly connected users of the graph, who also have a high global rating. This is an indication that global rating models in general and "influential"

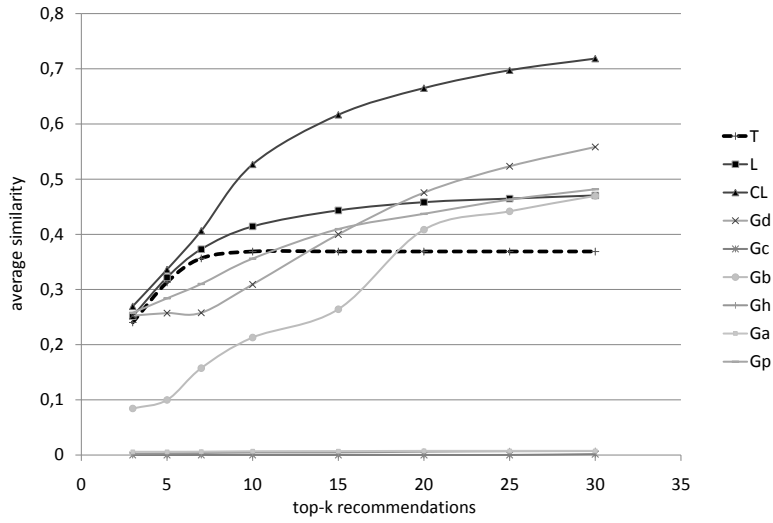


Fig. 2. Comparison of the local, collaborative and individual global rating models for users with few trusted nodes (set A).

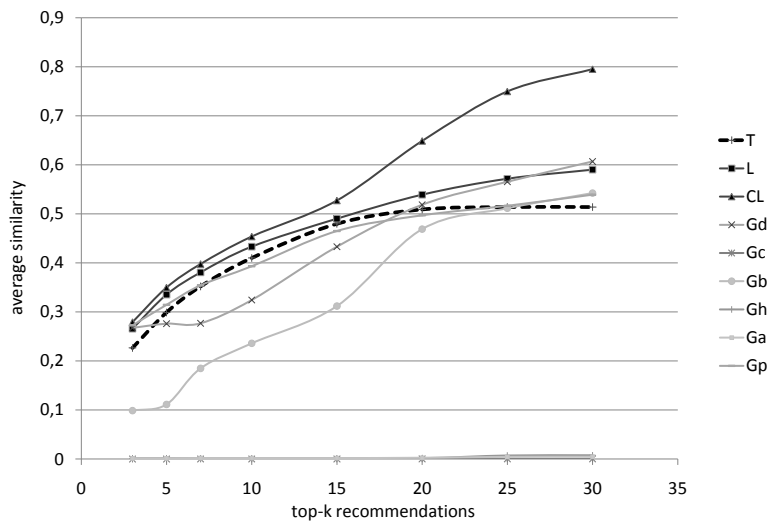


Fig. 3. Comparison of the local, collaborative and individual global rating models for users with many trusted nodes (set B).

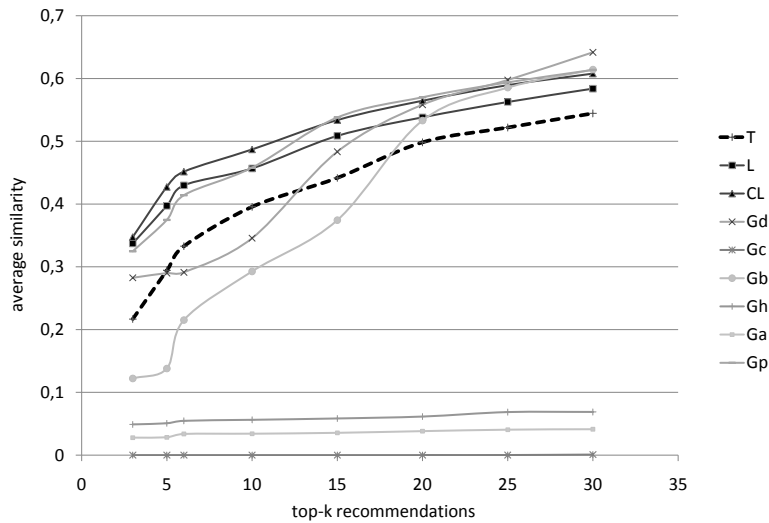


Fig. 4. Comparison of the local, collaborative and individual global rating models for users with too many trusted nodes (set C).

or “central” users can be valuable resources for a recommendation engine, mainly in the absence of local sources of recommendation. Such models can be used to address the “cold-start” problem, in which a user is new and hasn’t yet formed a network of trust. The results also indicate that the three aforementioned global rating models perform better than the remaining global models.

This behavior of the global rating models was anticipated, since, even the top- k users are influential, they do not affect the whole network (especially when a network comprises of thousands of users, as in the *Epinions* case). Thus a recommendation engine might not benefit by looking at such metrics alone, without taking into consideration the direct network of each user. However, some models are able to discover powerful “influentials” and can be combined with collaborative local models. All other global models confuse rather than assist the recommendation engine.

The comparison of results for sets A, B and C shows a higher baseline for set C, where users have many trusted users and thus a lot of recommendations to choose from. The local and collaborative local models manage to further improve performance. The results for users in set B are similar, with the baseline ranging from 0.22 to 0.51 and the collaborative local ranging from 0.28 to 0.79 for $k=3$ and $k=30$ respectively.

Based on the aforementioned results, it is expected that the quality of recommendations is better when they are based on local sources than on globally “influential” nodes. The boost is bigger for smaller values of k , which means that the local models are able to distill the long lists of trusted users and find the most influential users in each circle of trust.

5.3 Combination of collaborative local and global models

Based on the results of our first set of experiments, we decide to combine the collaborative local model with each of the global models. Although the outcome of the experiments showed that only some of the models perform well, we experiment with all combinations of collaborative local with each global model. We assign equal weights to the collaborative local and each of the global scores and evaluate the respective top- k lists. The comparative results are depicted in Figures 5, 6 and 7. They present the average similarity values for all users in each set (figure 5 for set A, figure 6 for set B and figure 7 for set C), comparing each user with the top- k recommended users ($k = 3, 5, 7, 10, 15, 20, 25, 30$) when ranked using the collaborative local model (CLS of Eq. 10) with $w_{CLS}=0.5$ and each one of the global influence measures (Gd, Gc, Gb, Gh, Ga, Gp in Eq. 12) with $w_x=0.5$ (where $w_x \in \{w_d, w_c, w_b, w_h, w_a, w_p\}$). For example the value for CL/Gd and $k = 3$ represents the average similarity between a user i and the top-3 users with the highest weighted combination of degree centrality rating and collaborative local rating using $w_d=0.5$ and $w_{CLS}=0.5$.

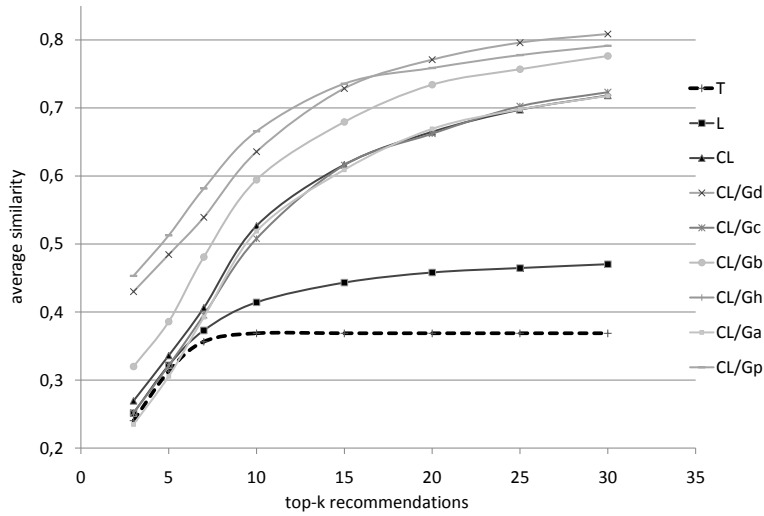


Fig. 5. Comparison of the combination of individual global rating models with the collaborative local model for users with few trusted nodes (set A).

Results in Figures 5 - 7 show that highly ranked users (i.e. influential users) may provide additional recommendations which are useful to all authors. Several metrics, such as hub, authority and closeness provide little improvement compared to the collaborative local model. However, degree (CL/Gd), PageRank (CL/Gp) and betweenness (CL/Gb) have further improved the recommendations of the collaborative local model for all the different k values. The average improvement for all the values of k

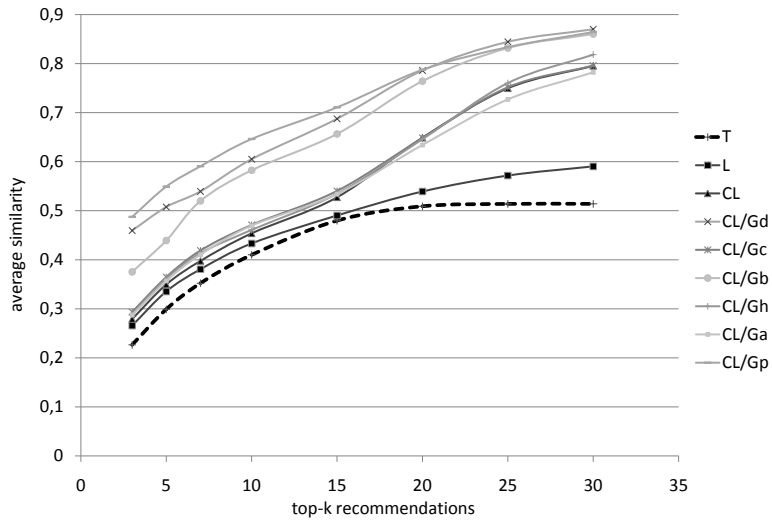


Fig. 6. Comparison of the combination of individual global rating models with the collaborative local model for users with many trusted nodes (set B).

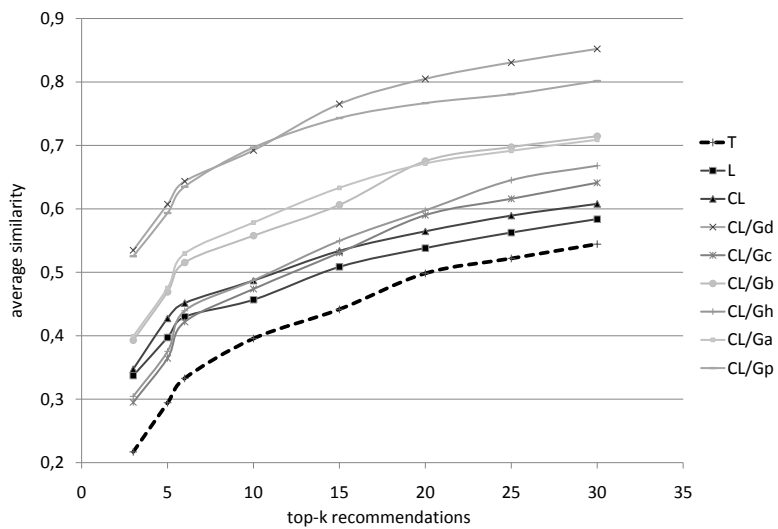


Fig. 7. Comparison of the combination of individual global rating models with the collaborative local model for users with too many trusted nodes (set C).

is in average 0.12, 0.13 and 0.06 for (CL/Gd) , (CL/Gp) and (CL/Gb) respectively. This strengthens our initial belief that global rating models can address the “cold-start” problem, especially when the recommendations coming from globally influential users are combined with those coming from the few people that a user trusts. The results are in accordance to those of the first set of experiments, where PageRank, betweenness and degree centrality outperformed all other global rating models.

More specifically, when we compare the results for the three sets (A, B and C) for the two sets of experiments, we notice that:

- the local methods (local and collaborative local) demonstrate slightly improved results for set C in comparison to set A (average improvement is 0.037)
- the combined methods further increase this improvement (average improvement for PageRank and degree is 0.05)
- the improvement is smaller for users in set B (users with 15 to 25 links) when compared to users of set A (average improvement is 0.035 and 0.018 for the local and combined methods respectively).

5.4 Effect of individual global influence models to the collaborative local model

In an attempt to further improve our results, we combine the local with multiple global models using weighted combinations. However, in order to better understand the influence of each global metric to the collaborative local decisions, we first perform an extended evaluation of the combinations of collaborative local and a single global metric. We decide to evaluate only PageRank, Degree and Betweenness, which are the ones with the stronger influence as shown in figures 5 to 7. We provide results only for set B, which contains users with 15-25 outlinks (trusted friends).

In the previous set of experiments, we combined each individual global influence score with the collaborative local score using equal weights. For example, the combination of the collaborative local metric with PageRank corresponds to $w_{collab}=0.5$ and $w_{global}=0.5$ in Eq. 14 and $w_p = 1, w_d = w_c = w_b = w_h = w_a = 0$ in Eq. 12. In this step, we employ only a percentage of the global influence score (provided by PageRank, Degree or Betweenness) ranging from 10 to 100%. For example, the combination of the collaborative local metric with a 10% of PageRank corresponds to $w_{collab}=0.5$ and $w_{global}=0.5$ in Eq. 14 and $w_p = 0.1, w_d = w_c = w_b = w_h = w_a = 0$ in Eq. 12. The results are depicted in Figures 8 (PageRank), 9 (Degree) and 10 (Betweenness). For readability reasons, we only show the results for combinations of the global and collaborative local scores with $w_i \in \{0.1, 0.4, 0.7, 1\}$.

The results show that PageRank contributes more to the overall performance when its percentage is high (close to 100%). Similarly, Degree reaches its top performance when its percentage is 100%, whereas Betweenness performs better for lower percentage values (e.g. 40%). These are indications that in a combined global influence metric, the weights of the individual global metrics must vary in order to achieve maximum performance. Based on the results of this section, in the following step, we test different combinations of the three (or two) aforementioned global metrics with the collaborative local model using different ratios for each metric.

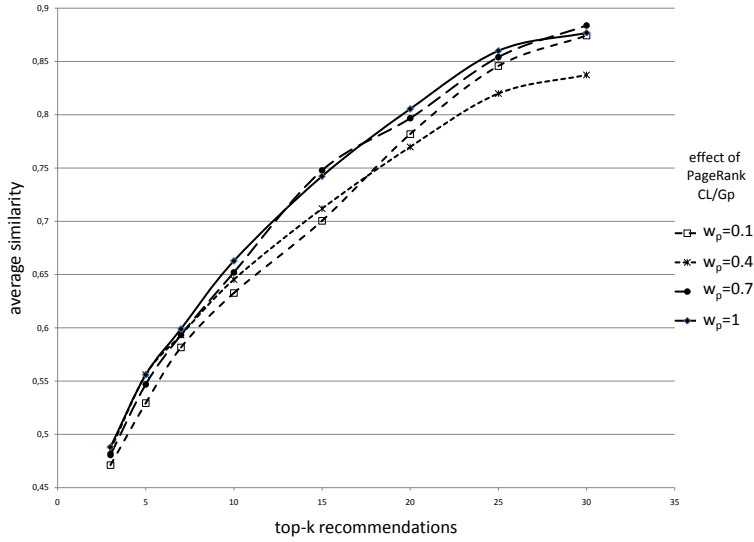


Fig. 8. The effect of PageRank-based global influence on the Collaborative model (set B).

5.5 Combination of multiple collaborative local and global models

In the previous steps, we evaluated each individual global rating model combined with the collaborative local model. The degree, PageRank and betweenness showed the highest performance improvement, so we combine these metrics using Eq. 12 and produce a single combined global rating for each user. We further combine this rating with the collaborative local rating (using $w_{CLS}=0.5$ and $w_{global}=0.5$) and produce the final rating for each user.

In Figures 11, 12 and 13 we present the results of the baseline T versus the local L , collaborative local CL and six combinations of the collaborative local rating with a combined global rating for user sets A, B and C respectively. We evaluate the following combinations of global metrics: $(CL/GdGbGp)$ with emphasis on the PageRank metric ($w_d = 0.2, w_b = 0.2, w_p = 0.6$), $(CL/GdGbGp(2))$ using equal weights ($w_d = 1/3, w_b = 1/3, w_p = 1/3$), $CL/GdGbGp(3)$ with ($w_d = 0.2, w_b = 0.4, w_p = 0.4$), $(GL/GdGp)$ with ($w_d = 0.5, w_p = 0.5$), $CL/GdGp(2)$ with ($w_d = 1/3, w_p = 2/3$) and $CL/GdGp(3)$ with ($w_d = 2/3, w_p = 1/3$).

The results show that most of the combinations improve the results of the baseline and the collaborative local model with the combinations of PageRank and degree to outperform all other combinations. However only the combinations of the combined PageRank and Degree metrics (i.e. the best global metrics in the previous experiments) manage to further improve the results of the combinations of collaborative local and a single global measure.

Our overall observation based on this experimental evaluation is that the combination of centrality and prestige metrics cannot outperform local metrics in providing recommendations for a specific user. However they can improve the performance of a

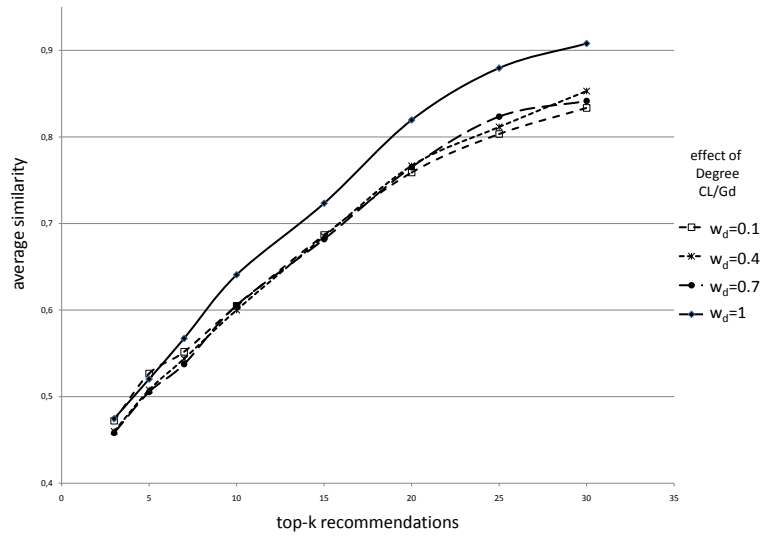


Fig. 9. The effect of Degree-based global influence on the Collaborative model (set B).

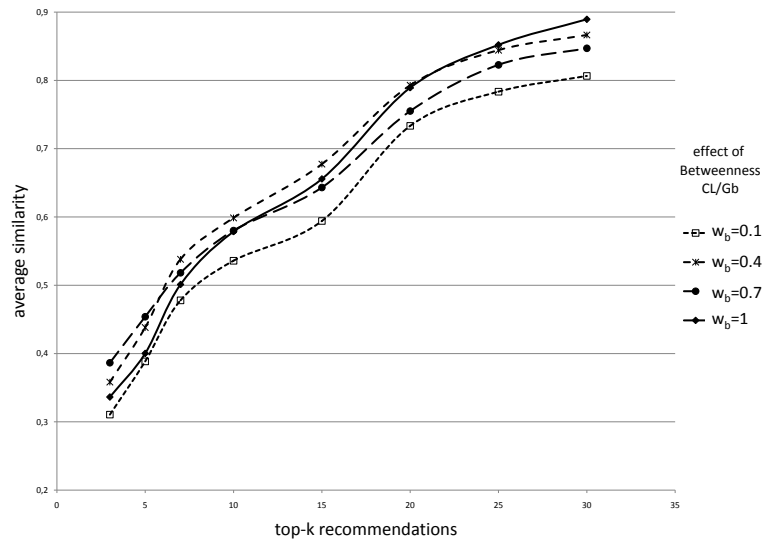


Fig. 10. The effect of Betweenness-based global influence on the Collaborative model (set B).

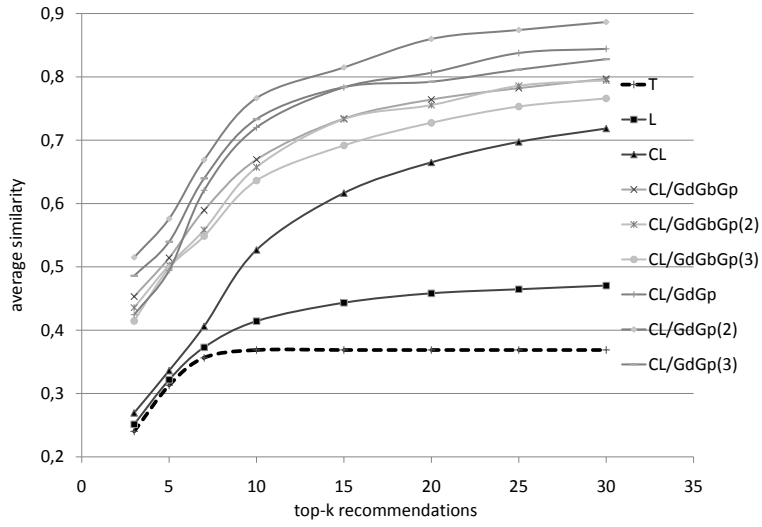


Fig. 11. Collaborative local plus combo of global models (set A).

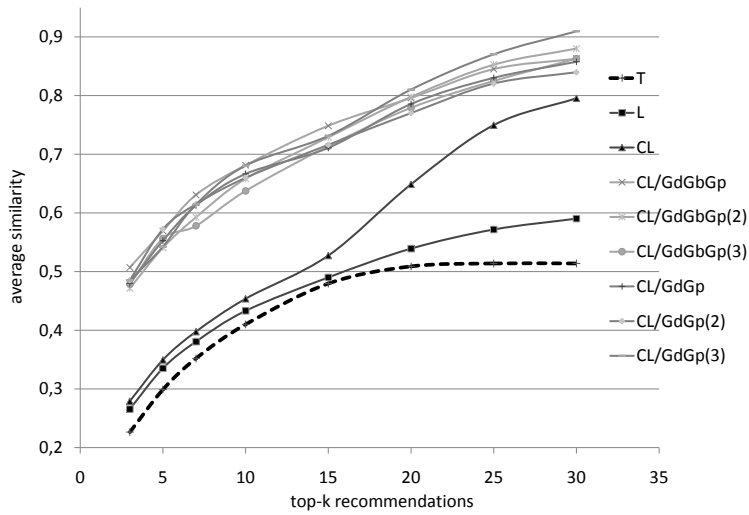


Fig. 12. Collaborative local plus combo of global models (set B).

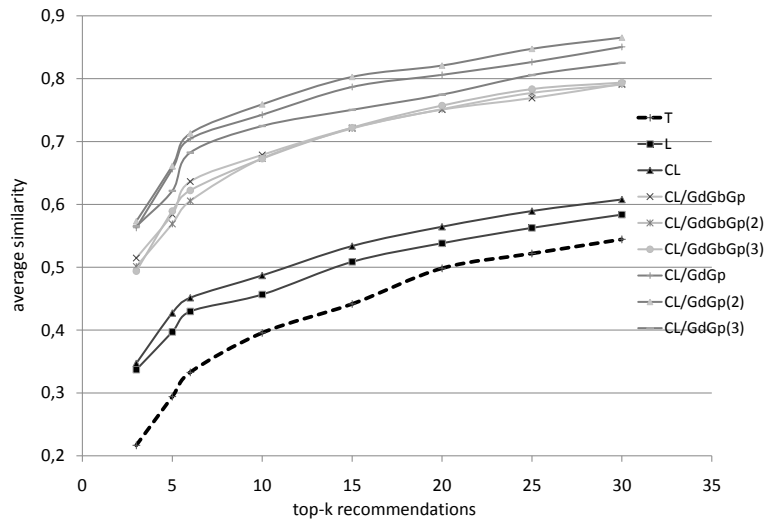


Fig. 13. Collaborative local plus combo of global models (set C).

recommendation engine when combined with collaborative local metrics. Finally, our findings in the performance of combinations of global metrics is that, depending on the nature of the social network and the weights' setup, it is possible to further improve the recommendation engine performance.

6 Conclusions and Future Work

In this work we studied the contribution of various measures in identifying trustworthy or influential actors in a social network in order to recommend them to a specific user. The actors can be users, blogs, or tweets. The measures take into consideration the opinion/trust of the actor for other actors, the opinion/trust of the actor's network of trust, and the overall ranking of all actors, as computed by their position and interconnections in a graph.

Global metrics can be easily manipulated with the creation of fake users and fake links (e.g. Google bombing). On the other hand, local trust metrics are more resistant to attacks, since users get recommendations only from users of trust and ignore the artificially created spam users and their links. A disadvantage of local trust metrics is that they do not apply to the new members of a social network, who have not formed a network of trust yet. The use of global metrics can be beneficial for those users, since it recommends users with raised global influence. Of course, the risk of recommending spam users to the new members of the network exists, but is quickly balanced, as soon as the user starts to add friends. The combination of local and global methods can help us tackle the shortcomings of each separate approach and improve the recommendations' quality.

Our model extends our previous work that generated personalized recommendations based on the network of trust, by incorporating global measures of influence and the notion of time “freshness” of the various user connections. We experimentally compared and evaluated various models, along with several combinations of them. The results showed that global measures are not very useful by themselves in providing recommendations to users, however, when combined with the collaborative local measures, they have a positive impact in the final recommendation set. In the future, we plan to extend our model and study the negative influence as expressed with negative values for trust. Finally, although the aforementioned model is meant to be used offline, it is on our plans to study the complexity of the different measures, a topic of great practical importance in case this model would be adjusted to run for large datasets and/or generate real-time recommendations.

References

1. Domingos, P., Richardson, M.: Mining the network value of customers. In: ACM SIGKDD '01. (2001) 57–66
2. Kempe, D., Kleinberg, J., Tardos, E.: Maximizing the spread of influence through a social network. In: ACM SIGKDD '03. (2003) 137–146
3. Nakajima, S., Tatemura, J., Hino, Y., Hara, Y., Tanaka, K.: Discovering important bloggers based on analyzing blog threads. In: 2nd Annual Workshop on the Blogging Ecosystem: Aggregation, Analysis and Dynamics. (2005)
4. Kritikopoulos, A., Sideri, M., Varlamis, I.: Blogrank: ranking blogs based on connectivity and similarity features. In: 2nd Int. Workshop on Advanced Architectures and Algorithms for internet Delivery and Applications AAA-IDEA'06, ACM (2006)
5. Adar, E., Zhang, L., Adamic, L., Lukose, R.: Implicit structure and the dynamics of blogspace. In: Workshop on the Blogging Ecosystem, WWW. (2004)
6. Weng, J., Lim, E.P., Jiang, J., He, Q.: Twiterrank: Finding topic-sensitive influential twitterers. In: ACM WSDM. (2010)
7. Varlamis, I., Louta, M.: Towards a personalized blog site recommendation system: A collaborative rating approach. the International IEEE Workshop on Semantic Media Adaptation and Personalization, SMAP (2009)
8. Kamvar, S.D., Schlosser, M.T., Garcia-Molina, H.: The eigentrust algorithm for reputation management in p2p networks. In: WWW. (2003) 640–651
9. O'Donovan, J.: Capturing trust in social web applications. In Golbeck, J., ed.: Computing with Social Trust. Springer London (2009)
10. Massa, P., Avesani, P.: Controversial users demand local trust metrics: An experimental study on epinions.com community. In: 20th National Conference on Artificial Intelligence (AAAI-05). (2005)
11. Ziegler, C.N., Lausen, G.: Spreading activation models for trust propagation. In: IEEE Eee'04. (2004)
12. Ziegler, C.N.: On propagating interpersonal trust in social networks. In Golbeck, J., ed.: Computing with Social Trust. Springer London (2009)
13. Golbeck, J.: Generating predictive movie recommendations from trust in social networks. In: iTrust. (2006) 93–104
14. Guha, R., Kumar, R., Raghavan, P., Tomkins, A.: Propagation of trust and distrust. In: WWW. (2004)
15. Golbeck, J.: Trust on the world wide web: a survey. Found. Trends Web Sci. **1**(2) (2006) 131–197

16. Golbeck, J.: Trust and nuanced profile similarity in online social networks. *TWEB* **3**(4) (2009)
17. Massa, P., Avesani, P.: Trust metrics on controversial users: balancing between tyranny of the majority and echo chambers. *Intl. J. on Semantic Web and Information Systems* **3**(2) (2007)
18. Massa, P., Avesani, P.: Trust metrics in recommender systems. In Golbeck, J., ed.: *Computing with Social Trust*. Springer London (2009)
19. Varlamis, I., Eirinaki, M., Louta, M.: A study on social network metrics and their application in trust networks. In: *The international conference on Advances in Social Network Analysis and Mining (ASONAM 2010)*. (2010)
20. Wasserman, S., Faust, K.: *Social Network Analysis*. Cambridge University Press (1994)
21. Richardson, M., Domingos, P.: Mining knowledge-sharing sites for viral marketing. In: *8th ACM SIGKDD*. (2002)
22. Kimura, M., Yamakawa, K., Saito, K., Motoda, H.: Community analysis of influential nodes for information diffusion on a social network. In: *the International Joint Conference on Neural Networks, IJCNN*. (2008)
23. Kimura, M., Saito, K., Nakano, R.: Extracting influential nodes for information diffusion on a social network. In: *22nd AAAI Conference*. (2007)
24. Estévez, P.A., Vera, P.A., Saito, K.: Selecting the most influential nodes in social networks. In: *the International Joint Conference on Neural Networks, IJCNN*. (2007)
25. Song, X., Chi, Y., Hino, K., Tseng, B.: Identifying opinion leaders in the blogosphere. In: *CIKM 2007*. (2007) 971–974
26. Agarwal, N., Liu, H., Tang, L., Yu, P.S.: Identifying the influential bloggers in a community. In: *ACM WSDM*. (2008)
27. Akritidis, L., Katsaros, D., Bozanis, P.: Identifying influential bloggers: Time does matter. In: *Web Intelligence*. (2009) 76–83
28. Kim, E., Han, S.S.: An analytical way to find influencers on social networks and validate their effects in disseminating social games. In: *ASONAM 2009*. (2009)
29. Louta, M., Varlamis, I.: Blog rating as an iterative collaborative process. In Wallace, M., et al., eds.: *Semantics in Adaptive and Personalized Services, Springer Series on “Studies in Computational Intelligence” SCI 279*. Springer-Verlag Berlin Heidelberg (2010) 187–203
30. Freeman, L.: A set of measures of centrality based on betweenness. *Sociometry* (40) (1977) 35–41
31. Freeman, L.: Centrality in social networks: Conceptual clarification. *Social Networks* **3**(1) (1979) 215–239
32. Bonanich, P.: Power and centrality: A family of measures. *The American Journal of Sociology* **5**(92) (1987) 1170–1182
33. Kleinberg, J.: Authoritative sources in a hyperlinked environment. *Journal of the ACM* **46**(5) (1999) 604–632
34. Page, L., Brin, S., Motwani, R., Winograd, T.: The pagerank citation ranking: Bringing order to the web. In: *Technical Report*. Stanford InfoLab. (1998)
35. Liu, B.: *Web Data Mining. Exploring Hyperlinks, Contents, and Usage Data*. Springer (2007)
36. Guha, S., Koudas, N., Marathe, A., Srivastava, D.: Merging the results of approximate match operations. In: *VLDB*. (2004)
37. Hoi, S., Jin, R.: Semi-supervised ensemble ranking. In: *Proceedings of the 23rd national conference on Artificial intelligence-Volume 2, AAAI Press* (2008)