A study on social network metrics and their application in trust networks

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

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

Malamati Louta
Harokopio University
Department of Informatics
and Telematics
Athens, Greece
louta@hua.gr

Abstract

Social network analysis has recently gained a lot of interest because of the advent and the increasing popularity of social media, such as blogs, social networks, microblogging, or customer review sites. Such media often serve as platforms for information dissemination and product placement or promotion. In this environment, 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 their effect in the process of information propagation in the virtual world. 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 local and global influence and compare suggestions provided by locally or globally trusted users.

1 Introduction

Social network analysis has been a major area of research for sociologists for many years, but recently it has gained a lot of interest with the advent of Web 2.0 and the enormous increase in the use of social networking websites, customer review sites, blogs etc.

Such media present features unique to the Web, in terms of shared authorship, multitude of user-provided tags, inherent connectivity between users and their posts, 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 identification of communities of users with similar interests, and within such communities the identification of the most "influential" users. A simple notion of influence is the number of 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 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 [3, 8], 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 a very difficult and time-consuming part. 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 [19, 13, 1], and only recently there have been some efforts in personalizing the rankings [25, 23].

In this work, we bridge these two research directions. Our objective is to generate personalized content recommendations based on the analysis of implicit of 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.

We extend our previous work [23] that employed the notion of trust among users as expressed with links along with the freshness of these links to generate personalized rankings, by incorporating the notion of "influence" in the ranking algorithm. We 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 content recommendations based on collaborative knowledge from locally or globally trusted 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 of 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 a big 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.

Several researchers have focused on trust prediction and propagation. Most researchers propose classification models, such as SVM-based methods [16, 18] to assign trust class labels using features such as user profile, user interactions, product reviews, and trust relations. A different approach is that of Lim et. al. [14], that employs the "Trust Antecedent" framework proposed in management science and introduce quantitative - instead of qualitative - features, such as ability, benevolence and integrity in the prediction process. A slightly different line of work focuses on how trust is propagated in a network of people [5, 17]. Whereas in our work we introduce the notion of trust in a social network, we assume that the trust between a pair of users is already known, either explicitly or implicitly. Moreover, trust propagation is thought to be covered by the more general notion of "influence" within such a network.

Influence in social networks, a topic extensively studied in the pre-WWW era [24], 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 [3]. Several works build on this Information Cascade (IC) notion proposing various machine learning algorithms [21, 11, 8, 10, 4]. 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 link analysis techniques [22], as well as user activity-related parameters in order to identify influential users in blogs [2] and social networks [9].

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 [19]) and/or implicit (BlogRank [13, 1]) 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. Recently, some effort has been done to also incorporate the content in the ranking process, when ranking twitterers (TwitterRank [25]).

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.

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. 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 [24]. 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 [15], [23].

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:

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

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:

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

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-adjucent 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:

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

Hubs and Authorities. Both terms were introduced as part of the well-known algorithm HITS [12]. A *hub* is a node with many out-links and an *authority* is a node with many in-links. 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:

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

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

PageRank. PageRank [20] also identifies "authorities" in a graph. 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:

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

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 [23] we presented a personalized recommendation model, which capitalizes on a collaborative rating mechanism that exploits the the edges of a 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. We assume 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 i's intention, called local score (LS).

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

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 second factor can be, for example, the number of links from blog i to blog j, or the number of similar ratings that customers i and j gave for the same products. The definition of weights depends on the type of social network we examine and the importance we give to explicit and implicit expressions of trust or interest.

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. 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 periods t in order to find the score in the current period c.

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

where m stands for the system memory, which means the number of periods back in time that we consider for accumulating the local scores.

We subsequently extend the local accumulative scores produced for all users in a first step, introducing the concept of *collaborative local score* $CLS_c(i,j)$ at a certain point in time c, for each user j that user i links to (i.e. $(i,j) \in E$). This score aggregates the direct accumulative scores $LAS_c(i,j)$, assigned by i to any user j, with the indirect accumulative scores $LAS_c(k,j)$ assigned to j by all users k that i trusts (i.e. $(i,k) \in E$).

$$CLS_c(i,j)=w_i \cdot LAS_c(i,j)+ \sum_{\substack{(i,j)\in E \\ (k,j)\in E \\ (i,k)\in E}} w_k \cdot LAS_c(k,j) \tag{9}$$

The notion of trust is bound to the permanent link between two users in a social network. For example, in the case of blogs, the permanent links 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".

4 Influence Model

Our objective is to generate personalized recommendations to the users of social media. 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 considered the addition of 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 9 to include a *Global Influence* model GI. This global influence model results in a global ranking of all users in a 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:

$$\begin{aligned} GI(i) = w_d \cdot Gd(i) + w_c \cdot Gc(i) + \\ w_b \cdot Gb(i) + w_h \cdot Gh(i) + \\ w_a \cdot Ga(i) + w_p \cdot Gp(i) \end{aligned} \tag{10}$$

Note that using the aforementioned formula, we may give more importance to one (or more) global influence metrics 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:

$$\begin{aligned} INF_c(i,j) = f(LAS_c(i,j), \sum_{\substack{(i,j)\in E \\ (k,j)\in E \\ (i,k)\in E}} w_k \cdot LAS_c(k,j), \\ \sum_{(m,j)\in E} GI(m) \cdot LAS_c(m,j)) \end{aligned} \tag{11}$$

This function could be, for instance, a weighted sum of the three factors of Equation 11:

$$\begin{aligned} INF_c(i,j) = w_{local} \cdot LAS_c(i,j) + \\ w_{collab} \cdot \sum_{\substack{(i,j)\in E \\ (k,j)\in E \\ (i,k)\in E}} w_k \cdot LAS_c(k,j) + \\ w_{global} \cdot \sum_{(m,j)\in E} GI(m) \cdot LAS_c(m,j) \end{aligned} \tag{12}$$

Equation 12 assumes that the weights are normalized. The weighted sum approach has been used in a related context (identification of influential bloggers) with great success [2]. Alternatively we can produce different rankings using each local and global metric and then merge the rankings in a single ranked list [6], or use an ensemble ranking [7].

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 12 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 ($GI(m)$).

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 will recommend the top-k influential blogs to each user, personalized by the user's personal network of trust and overall influence of blogs. In a microblogging 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.

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 wikipage[1] contains information about product reviews written by the

[1] http://www.trustlet.org/wiki

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 of 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 two distinct subsets: a) one that includes users with a narrow circle-of-trust (set A: users having between 5 and 10 links) and b) one that includes users with an extended circle-of-trust (set B: users having more than 30 links). The two sets which are of equal size (set A contains 5,425 and set B contains 5,405 users) but significantly differ in the connectivity of their nodes.

We first evaluated each model (namely, the local, the collaborative local, and each one of the global influence models) individually (setting the respective weight in Eq. 10 to 1 and the remaining weights to 0). Based on the results of this first experiment, we then combined the collaborative local model with each of the global influence models (setting equal weights for w_{local}, w_{collab} and w_{global} in Eq. 12). Finally, based on our findings, we combined the 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 5.3. We performed each set of experiments for both sets of users (set A and set B).

5.1 Evaluation of the individual models

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 with the trust link 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 U recommended to U_i have matching interests with U_i.

In this first experiment, we generate for each user U (in each of the sets A and B) 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 measure (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 similarity between two users U_i and U_j is defined as the ratio of articles 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 is connected via an explicit trust link ($(U_i, U_t) \in E$).

Figure 1 presents the average similarity values for the top-k matches ($k = 3, 5, 7, 10, 15, 20, 25, 30$) for set A, which comprises of users with few trusted nodes, whereas Figure 2 shows users of set B, who have many trusted nodes in their circle.

From the results presented in Figures 1 and 2 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

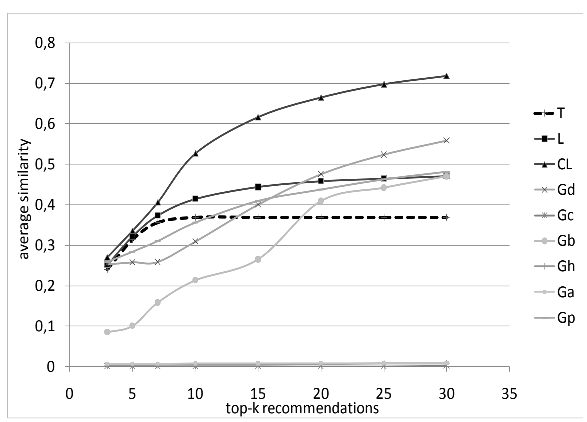

Figure 1. Local vs global models (set A).

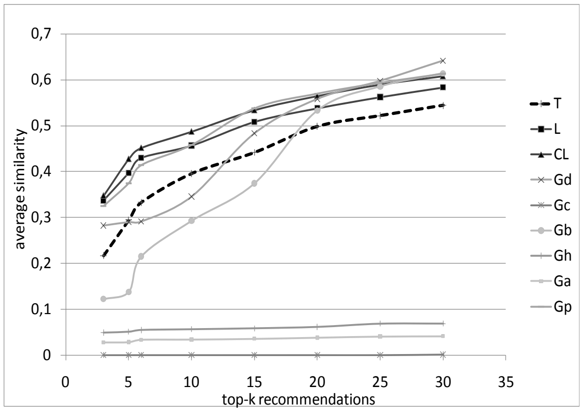

Figure 2. Local vs global models (set B).

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 many neighbors (set B), 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 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. 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 and B shows a higher baseline for the second set, 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 the midpoint (with 15-25 outlinks) 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, thus due to space constraints we don't depict them here.

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.

5.2 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 by merging the two top-k lists and evaluate the results. 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. The comparative results are depicted in Figures 3 and 4. Figure 3 presents the average similarity values for all users in set A, 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 and each one of the global influence measures (Eq. 12 using w_{local}=0, w_{collab}=0.5 and w_{global} =0.5). 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. Figure 4 represents the same experiment for users of set B.

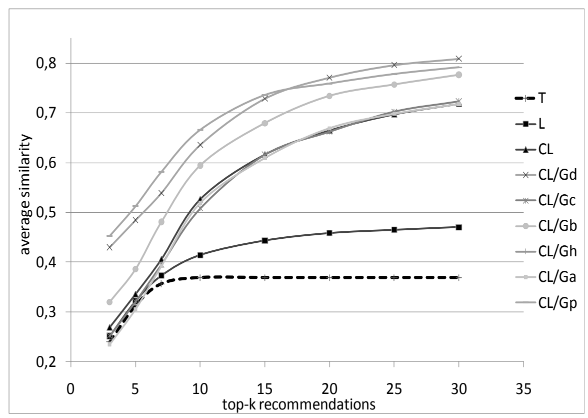

Figure 3. Collaborative local plus Global models (set A).

Results in Figure 3 (set A) 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 is in average 0.12, 0.13 and 0.06 for (CL/Gd), (CL/Gp) and (CL/Gb)respectively. This is an indication that global rating models and "influential" or "central" users can be valuable resources for a recommendation engine, mainly in the absence of local sources of recommendation, or to address

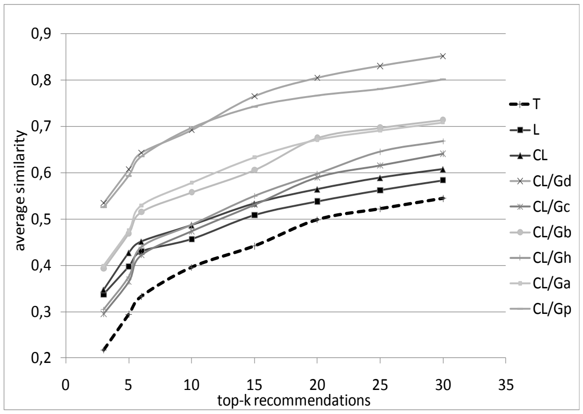

Figure 4. Collaborative local plus Global models (set B).

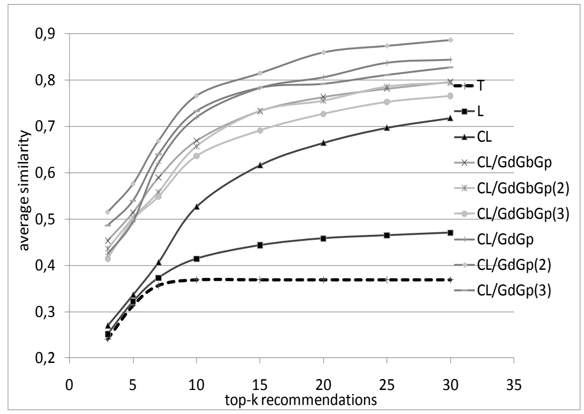

Figure 5. Collaborative local plus combo of global models (set A).

the "cold-start" problem, in which a user is new and hasn't yet formed a network of trust. 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.

Comparing the results with those in Figure 4, that correspond to users with many trusted nodes (set B), we notice that the local methods demonstrate slightly improved results for set B in comparison to set A (average improvement is 0.037) and the combined methods further increase this improvement (average improvement for PageRank and degree is 0.05). The improvement is smaller for the midpoint (users with 15 to 25 links) when compared to users of set A (average improvement is 0.035 and 0.018 for local and global models respectively). This diagram is omitted due to space constraints. Attempting to further improve our results, we combine the local with multiple global models using weighted combinations as explained in the following section.

5.3 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 betweeness showed the highest performance improvement, so we combine these metrics using Eq. 10 and produce a single combined global rating for each user. We further combine this rating with the collaborative local rating (as shown in Eq. 12 using w_{local}=0, w_{collab}=0.5 and w_{global} =0.5) and produce the final rating for each user.

In Figures 5 and 6 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 com-

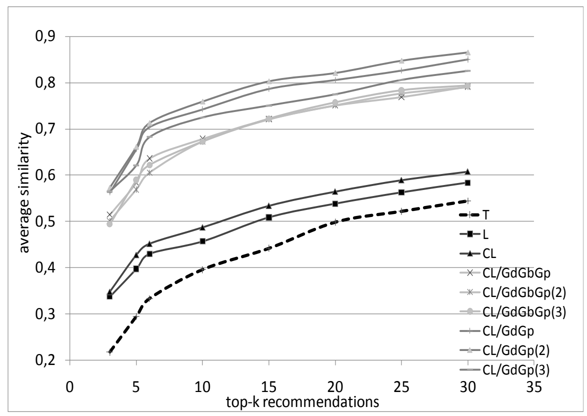

Figure 6. Collaborative local plus combo of global models (set B).

bined global rating for user sets A and B 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 recommendation engine when combined with collaborative local metrics. Finally, the initial findings in the performance of combinations of global metrics are not very promising. However, 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 similar 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. Our model extended an existing model that generated personalized recommendations based on the network of trust, by incorporating global measures of influence. We experimentally compared and evaluated various models, along with several combinations. The results showed that global measures are not very useful by themselves in providing recommendations to users, but, when combined with the collaborative local measures have a positive impact in the final recommendation set. In the future, we plan to perform a more extensive experimental evaluation of the various parameters of our model. Moreover, we intend to extend our model and study the negative influence as expressed with negative values for trust.

References

[1] E. Adar, L. Zhang, L. Adamic, and R. Lukose. Implicit structure and the dynamics of blogspace. In *Workshop on the Blogging Ecosystem, WWW*, 2004.

[2] N. Agarwal, H. Liu, L. Tang, and P. S. Yu. Identifying the influential bloggers in a community. In *ACM WSDM*, 2008.

[3] P. Domingos and M. Richardson. Mining the network value of customers. In *ACM SIGKDD '01*, pages 57–66, 2001.

[4] P. A. Estévez, P. A. Vera, and K. Saito. Selecting the most influential nodes in social networks. In *the International Joint Conference on Neural Networks, IJCNN*, 2007.

[5] R. Guha, R. Kumar, P. Raghavan, and A. Tomkins. Propagation of trust and distrust. In *WWW*, 2004.

[6] S. Guha, N. Koudas, A. Marathe, and D. Srivastava. Merging the results of approximate match operations. In *VLDB*, 2004.

[7] S. Hoi and R. Jin. Semi-supervised ensemble ranking. In *Proceedings of the 23rd national conference on Artificial intelligence-Volume 2*. AAAI Press, 2008.

[8] D. Kempe, J. Kleinberg, and E. Tardos. Maximizing the spread of influence through a social network. In *ACM SIGKDD '03*, pages 137–146, 2003.

[9] E. Kim and S. S. Han. An analytical way to find influencers on social networks and validate their effects in disseminating social games. In *ASONAM 2009*, 2009.

[10] M. Kimura, K. Saito, and R. Nakano. Extracting influential nodes for information diffusion on a social network. In *22nd AAAI Conference*, 2007.

[11] M. Kimura, K. Yamakawa, K. Saito, and H. Motoda. Community analysis of influential nodes for information diffusion on a social network. In *the International Joint Conference on Neural Networks, IJCNN*, 2008.

[12] J. Kleinberg. Authoritative sources in a hyperlinked environment. *Journal of the ACM*, 46(5):604–632, 1999.

[13] A. Kritikopoulos, M. Sideri, and I. Varlamis. 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.

[14] E.-P. Lim, V.-A. Nguyen, A. Sun, J. Jiang, and H. H. Tan. To trust or not to trust? predicting online trusts using trust antecedent framework. In *IEEE ICDM 2009*.

[15] B. Liu. *Web Data Mining. Exploring Hyperlinks, Contents, and Usage Data*. Springer, 2007.

[16] H. Liu, E.-P. Lim, H. W. Lauw, M.-T. Le, A. Sun, J. Srivastava, and Y. A. Kim. Predicting trusts among users of online communities: an epinions case study. In *ACM Conference on Electronic Commerce (EC2008)*, Chicago, 2008.

[17] P. Massa and P. Avesani. Controversial users demand local trust metrics: An experimental study on epinions.com community. In *20th National Conference on Artificial Intelligence (AAAI-05)*, 2005.

[18] Y. Matsuo and H. Yamamoto. Community gravity: measuring bidirectional effects by trust and rating on online social networks. In *WWW*, 2009.

[19] S. Nakajima, J. Tatemura, Y. Hino, Y. Hara, and K. Tanaka. Discovering important bloggers based on analyzing blog threads. In *2nd Annual Workshop on the Blogging Ecosystem: Aggregation, Analysis and Dynamics*, 2005.

[20] L. Page, S. Brin, R. Motwani, and T. Winograd. The pagerank citation ranking: Bringing order to the web. In *Technical Report. Stanford InfoLab*, 1998.

[21] M. Richardson and P. Domingos. Mining knowledge-sharing sites for viral marketing. In *8th ACM SIGKDD*, 2002.

[22] X. Song, Y. Chi, K. Hino, and B. Tseng. Identifying opinion leaders in the blogosphere. In *CIKM 2007*, pages 971–974.

[23] I. Varlamis and M. Louta. Towards a personalized blog site recommendation system: A collaborative rating approach. *the International IEEE Workshop on Semantic Media Adaptation and Personalization, SMAP*, 2009.

[24] S. Wasserman and K. Faust. *Social Network Analysis*. Cambridge University Press, 1994.

[25] J. Weng, E.-P. Lim, J. Jiang, and Q. He. Twitterrank: Finding topic-sensitive influential twitterers. In *ACM WSDM*, 2010.