Towards a Personalized Blog Site Recommendation System: a Collaborative Rating Approach

Iraklis Varlamis, Malamati Louta (*Member, IEEE*)
Dept. of Informatics and Telematics
Harokopio University of Athens
Athens, Greece
{*varlamis,louta*}*@hua.gr*

***Abstract*—The blogosphere is a part of the Web, enhanced with several characteristics that differentiate blogs from traditional websites. The number of different authors, the multitude of user-provided tags, the inherent connectivity between blogs and bloggers, the high update rate, and the time information attached to each post are some of the features that can be exploited in various information retrieval tasks. Traditional search engines perform poorly on blogs since they do not cover these aspects. In an attempt to exploit these features, this paper proposes a personalized recommendation model, which capitalizes on a collaborative rating mechanism that exploits the hyperlinks between blogs. The model assumes that the intention of a blog owner who creates a link to another blog is to provide a recommendation to the blog readers and quantifies this intention in a local score for the blog being pointed. A set of implicit and explicit links between any two blogs affect the exchanged score. The process is iterative and takes into account the opinion of a set of affiliated blogs and the freshness of links.**

***Keywords*-Blog; recommendation system; collaborative rating; ranking.**

I. Introduction

The idea behind social networking sites is to promote communication and collaboration of web users.Most of the sites in this social networking sphere (e.g. blogs, social bookmarking and multimedia sharing) capitalize on users' feedback in order to provide a trustfulness score for each resource (i.e. user or content). They assume that a user shares a resource with the community (e.g. a blog post, a URL of interest, a photo or her profile) and the community members provide their feedback for this resource. Feedback can be expressed as a comment on a blog post, a thumb-up or thumb-down vote on a photo or a comment on the user profile and can be negative, positive or neutral. The accumulated feedback for all the resources shared by a user defines the user's publicity or trustfulness. Since users provide their resources and feedback frequently, older feedback becomes obsolete and new feedback is used to update the rating of a resource. Traditional ranking models for the web are incapable to support this iterative process. The most successful web ranking models capitalize on explicit hyperlinks, which do not capture the semantics of user feedback.

With these problems in mind, we present a framework for capturing the semantics of social networking and user feedback and a system that exploits collaborative ratings and provides personalized recommendations to the users. The focus of this paper is the blogosphere and as a consequence, the system architecture and the underlying framework are adapted to its structure.

The contributions of this work are: a) it suggests a way for providing suggestions and rating other users' content, b) it describes the mechanism that exploits suggestions and creates personalized recommendations for bloggers.

The following section gives reference to related research works that exploit blogosphere information for providing recommendations. In section 3, we give a synopsis of our view to the blogosphere information, and we present the system architecture and the proposed model for attaching rating information to blogs. Section 4 presents the mathematical formulation of the proposed rating mechanism and explains how personalized recommendations are achieved. Section 5 demonstrates a first application of the model to a social networking application and section 6 concludes this work.

II. Related work

When building a collaborative rating framework for social networks, one should take into account the properties of the network and model all the implicit and explicit rating mechanisms. Yu and Singh [14] propose a reputation mechanism for social networks based on four properties of trust: symmetry, transitivity, self-reinforcement, propagation. In the case of blogosphere, the hyperlinks in the blogroll of a user and the hyperlinks between posts have been frequently employed as positive or negative recommendations [7] or carriers of trust [11], [14], [9]. There is no doubt that symmetry is not valid for hyperlinks: an entry in the blogroll is an indication that of interest to the pointed blog, but the inverse is not necessary. Since self-reinforcement and transitivity can be exploited to bias the final ratings, we should carefully define our model, using trust threshold and transitivity horizon [11], [12]. Finally, the propagation of trust can be modeled using Spreading Activation Networks [16], but this is outside of the scope of this paper.

Another property of hyperlinks as recommendations is their positive or negative nature. Guha et al [5] model the propagation of trust and distrust via positive and negative links, whereas the MoleTrust algorithm in [11] considers that trustful users have a controversiality percentage which is above 0.6. In the case of blogs, Kale et al in [7] introduce the concept of link polarity in order to distinguish

between positive links that point to ”like-minded” blogs, negative and neutral ones.

The scope of recommendations is another important decision for social networks. Recommending globally accepted resources is a safe approach, but is not useful to most users. As stated in [11], local trust scores are preferred in social networks with many controversial users. The blogosphere is a network that gathers people with different and often contradictory beliefs, and thus, local and decentralized rating schemes for content and users are better than a universal rating mechanism [8], [13], [9],

The models that capture the freshness of references between web pages [1], [2], or scientific papers [3], [15], are based on the fact that ranking algorithms favor old pages and balance this bias using a link (or citation) weighting scheme, which is based on the age of the web page (or paper). In a post-processing step the authority of a pointed node decays, based on the node's age and the incoming links age.

Authors in [10] introduced the notion of trust and collaboration in recommender systems and presented a trust-aware recommender system. In this system, the decision of a user A for an item X is affected by the opinion of a user B in two ways: the explicit trust of A towards B and the implicit similarity between A and B, as denoted by their common interests (i.e. common ratings for the same items). Trust in the blogosphere is expressed through hyperlinks (in blogroll and post level). So, it is reasonable to say that blogroll hyperlinks are definitely above the trust threshold and can be used for recommendations, since what our friendly blogs think about a third resource, strongly affects our opinion for this resource. In this context, we decide to model transitivity of trust in our framework. However, we set a horizon for this transitivity.

In the current work, we consider that ranking in the blogosphere is an iterative process. We exploit two special features of the blogosphere links: a) the difference between blogroll links, which denote a more permanent trust towards the blog being pointed, and post links, which represent a more transient reference to a blog, and b) the timestamp information of a post, which can be employed as a timestamp for a hyperlink. Our aims are to define a framework that allows users to expose their ratings and preferences, to build a model that creates personalized recommendations, and to define the architecture of the system that will assist a blogger in finding recent and trustworthy resources.

III. System Architecture and Semantics

In this study, we propose a personalized blog site recommendation system, exploiting a decentralized and collaborative rating mechanism, which adaptively assigns scores to blogs based on the opinion of other blogs in the system. The designed rating mechanism exploits the hyperlinks between blogs, assuming that creation of a link pointing to a blog indicates an implicit recommendation for the blog being pointed to the rest of the readers.

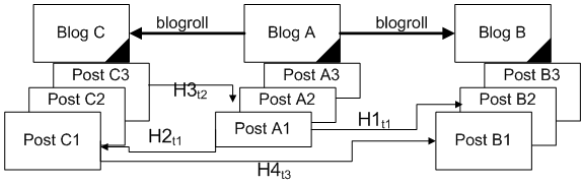

Figure 1. Hyperlink types in the rating model.

A. Link types

Our proposed mechanism distinguishes between blogroll links and hyperlinks at post level as depicted in Figure 1. More specifically, a blogroll hyperlink is a link in the blogroll of blog A pointing to a blog B, which denotes that A gives a permanent recommendation for B. Consequently, A should contribute the score of B. On the contrary, a post hyperlink from an individual post A1 of blog A to a post B2 in blog B denotes a temporary interest of A to the contents of B and consequently should increase the score of B only for a short period of time after the post has been published.

B. Link semantics

The Friend of a Friend (FOAF[1]) and the XHTML Friends Network (XFN[2]) are two metadata schemas for machine-readable pages that describe people, the links between them and the things they create and do. In this work, we suggest a FOAF extension similar to the one presented in [4], which adds properties to the domain of foaf:Person and assigns a level of trust ranging from 0 to 1, to each blog in the roll. For example, the FOAF file for blogger A in Figure 1 is given in Figure 2. This point forward, we will refer to this extended FOAF file with the term FOAFRoll. User details have been omitted for simplicity.

```
<rdf:RDF
 xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
 xmlns:foaf="http://xmlns.com/foaf/0.1/"
 xmlns:foafrate="http://www.foaf.rate.com/foaf/"
 xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#">
<foaf:Person>
 <foaf:name>Blog A</foaf:name>
          ...
 <foaf:weblog rdf:resource="...com/BlogA/"/>
 <foaf:knows>
   <foaf:Person>
    <foaf:name>Blog B</foaf:name>
          ...
    <rdfs:seeAlso rdf:resource="...com/BlogB/"/>
    <foafrate:rate value="0.9" lastupdated="2008-1-1"/>
   </foaf:Person>
   <foaf:Person>
    <foaf:name>Blog C</foaf:name>
          ...
    <rdfs:seeAlso rdf:resource="...com/BlogC/"/>
    <foafrate:rate value="0.8" lastupdated="2009-12-8"/>
   </foaf:Person>
 </foaf:knows>
</foaf:Person>
</rdf:RDF>
```

Figure 2. A sample FOAFRoll file.

A user's blogroll and consequently the FOAFRoll file are expected to change rarely, however a *lastupdated* attribute will provide evidence on the freshness of each rating.

[1]http://www.foaf-project.org/
[2]http://gmpg.org/xfn/

As far as it concerns hyperlinks in the post level, the latest RSS version supports the Platform for Internet Content Selection (PICS[3]) rating only for channels and not for individual posts. We suggest the enhancement of the item element with a postlink subelement. Each postlink element will contain *id*, *url*, *rating* and *lastupdated* attributes for the targeted post. Consequently, the enhanced RSS file for Blog A in Figure 1, will be as shown in Figure 3.

```
<?xml version="1.0"?>
<rss version="2.0">
 <channel>
  <title>posts of blog A</title>
  <link>...com/BlogA/</link>
          ...
  <pubDate>2009-12-31</pubDate>
          ...
  <item>
   <title>Post A1</title>
   <link>="...com/BlogA/PostA1</link>
   <description>....</description>
   <pubDate>2009-11-8</pubDate>
   <guid>...</guid>
   <postlink id="h1" url=".../BlogB/PostB2"
          rating="0.6" lastupdated="2009-11-8"/>
   <postlink id="h2" url=".../BlogB/PostC1"
          rating="0.7 lastupdated="2009-11-8"/>
  </item>
   <item> <title>Post A2</title>...   </item>
   </channel>
</rss>
```

Figure 3. A sample EnhancedRSS file.

C. The rating mechanism

The proposed rating mechanism adaptively assigns scores to blogs, to be exploited by the designed recommendation system. The mechanism considers both local information (i.e., what the evaluator blog site believes about the blog site under evaluation) and information gathered from affiliated blog sites (i.e., what my friends believe, what the friends of my friends believe, etc.), which are in general trusted blogs which affect the opinion of the evaluator.

The local rating information depicts the image of the evaluator blog site for the part of the blogosphere pointed by it. The local rating assigned by the evaluator blog site A to the target (under evaluation) blog site B is the weighted sum of a) the FOAFroll related rating (i.e., the rating assigned to the target blog if it belongs to the blogroll of the evaluator) and b) the ERSS related rating (i.e., any new posts added daily in the evaluator blog contribute to the rating of the respective blogs they point to). The local rating information is updated during pre-defined time intervals (e.g., per day), while the rating value is changed every time a new hyperlink appears, either in a post or in the blogroll of the evaluator. The contribution of a post hyperlink decreases as days pass and the post becomes old. The aggregation of all the local ratings of a certain evaluator blog site estimated during consecutive time intervals, gives the accumulative local ratings assigned to all blogs pointed by the evaluator.

In contrast to the trust aware recommender system architecture presented in [10], which assumes a central repository for ratings, our architecture (see Figure 4), is

[3]http://www.w3.org/PICS/

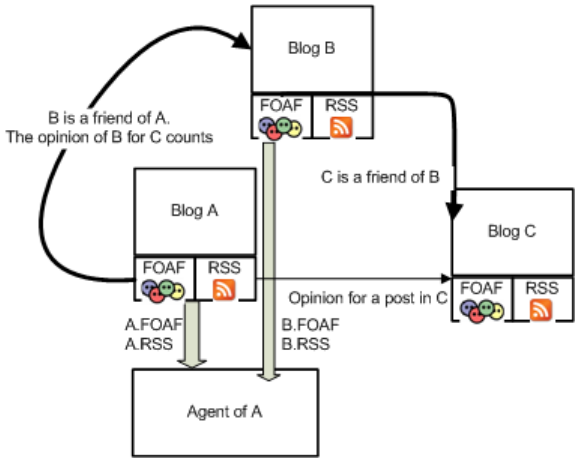

Figure 4. System architecture.

decentralized. Each blog stores safely the ratings estimated during each time period and is assumed that it is willing to share its view in the blogosphere and provide whenever asked for the blog site ratings. Thus, the problem is reduced in finding proper witnesses. In the current version of this study, we consider that the evaluator blog site should take into account the opinion of its friends (e.g., the blogs constituting its blogroll, depth=1), the friends of its friends (e.g., the blogs comprised in the blogroll of the evaluator's friends, depth=2), etc. This way, the evaluator would form a picture on the blogosphere as viewed by its affiliated sites.

The local rating is based on the direct beliefs of the evaluator blog (i.e., information from the local FOAFroll and ERSS files) and is enriched with information gathered from its affiliated blogs (e.g., blog A combines local information with that in the ERSS and FOAFRoll files of blog B). As a result, the collaborative local rating combines the experiences of the evaluator blog A for the target blog C with information regarding C gathered from N affiliated blog sites. The recommendation system will rank the blogs according to the calculated rating and provide a personalized recommendation to the evaluator blog site.

In general, when we aggregate multiple witnesses, we take into account the credibility and freshness of each witness. In an analogous manner, when we combine local ratings from different blogs we must consider the freshness of ratings, which corresponds to a) the freshness of links and b) the freshness of prior rating (i.e., considering the time period during which the rating was estimated) and the credibility of each individual blog, which in context of this study is associated with a weighting factor dynamically updated, reflecting the degree of matching interests between the evaluator blog site and the witness blog sites.

This study is based upon the notion of interacting intelligent agents which participate in activities on behalf of their owners, while exhibiting properties such as autonomy, reactiveness, and proactiveness, in order to

achieve particular objectives and accomplish their goals [6]. Thus, the blog site recommendation agent (BSRA) is introduced and assigned with the role of collecting necessary information from the blog itself and its affiliated blogs, estimating the local, accumulative local and collaborative local blog site ratings, on the basis of which finally provides personalized recommendations to the owner blog site. Our rating mechanism is mathematically formulated in the following section.

IV. Blog site rating system formulation

Let us assume the presence of M Blog Sites BSs falling within the same category with respect to the topics covered and the interests shared. Let $BS = BS_1, BS_2, ..., BS_M$ be the set of Blog Sites in the system. In subsection IV.A, the local blog site rating formation is formally described taking into account only first hand information (i.e. what the evaluator blog site considers about the target blog site), while in subsection IV.B the blog site local rating is collaboratively formed (the evaluator blog site takes into account the opinion of other affiliated blog sites concerning the blog site under evaluation).

A. Local Accumulative Rating

Concerning the local formation of the Blog Site BS_i rating, the Blog Site BS_j may rate BS_i at time period c in accordance with the following formula:

$$LABSR^{BS_j}_{t_p=c}(BS_i) = \sum_{\substack{\text{k=c-n+1} \\ k>0}}^{c} w_{t_p=k} \cdot LSBSR^{BS_j}_{t_p=k}(BS_i) \quad (1)$$

where $LABSR^{BS_j}_{t_p=c}(BS_i)$ is the local accumulative BS_i rating estimated by BS_j at time period $t_p = c$, $LSBSR^{BS_j}_{t_p=k}(BS_i)$ denotes the local rating the evaluator BS_j attributes to the target BS_i at time period $t_p = k$ and weight $w_{t_p=k}$ provides the relative significance of the $LSBSR^{BS_j}_{t_p=k}(BS_i)$ factor estimated at time period k to the overall BS_i rating estimation by the evaluator BS_j.

Concerning the $LSBSR^{BS_j}_{t_p=k}(BS_i)$ factor estimation, the evaluator BS_j may exploit the following formula:

$$LSBSR^{BS_j}_{t_p=k}(BS_i) = w_{BR} \cdot BR^{BS_j}_{t_p=k}(BS_i) + w_{EP} \cdot EP^{BS_j}_{t_p=k}(BS_i) \quad (2)$$

As may be observed from Equation 2, the local rating of the target BS_i is a weighted combination of two factors. The first factor contributing to the overall BS_i rating value (i.e., $BR^{BS_j}_{t_p=k}(BS_i)$) forms the FOAFroll related factor. This factor is introduced on the basis that the BS_j blogroll provides a list of friendly blog sites frequently accessed/read by the authors of BS_j. It has been assumed that $BR^{BS_j}_{t_p=k}(BS_i)$ lies within the [0,1] range, where a value close to 1 indicates that the target BS_i is a friendly blog site to the evaluator BS_j. In the context of this study, $BR^{BS_j}_{t_p=k}(BS_i)$ is modeled as a variable assuming values that lie within the [0,1] range, taking the value of 0 in case the BS_i does not belong to the blogroll of BS_j at time period k. In essence, BS_j provides a rating of the friendly blog sites in the blogroll, which could be exploited in order to differentiate $BR^{BS_j}_{t_p=k}(BS_i)$ factor for the friendly blog sites comprised in the BS_j blogroll.

The second factor contributing to the overall $LSBSR^{BS_j}_{t_p=k}(BS_i)$ (i.e. $EP^{BS_j}_{t_p=k}(BS_i)$) depends on the fraction of BS_j posts pointing to BS_i at time period k. This factor has been assumed to lie within the [0,1] range and may be given by the following equation:

$$EP^{BS_j}_{t_p=k}(BS_i) = \frac{NoP^{BS_j}_{t_p=k}(BS_i)}{NoP^{BS_j}_{t_p=k}} \quad (3)$$

where $NoP^{BS_j}_{t_p=k}(BS_i)$ number of posts created between time period $t_p = k-1$ and $t_p = k$ pointing to the target blog site BS_i and $NoP^{BS_j}_{t_p=k}$ denotes the total number of the evaluator BS_j posts created in between time period $k-1$ and k. If rating information is available for the post links then $NoP^{BS_j}_{t_p=k}(BS_i)$ can be replaced by the sum of all ratings from BS_i to BS_j and $NoP^{BS_j}_{t_p=k}$ by the sum of all ratings provided from BS_j in the same period.

Weights w_{BR} and w_{EP} provide the relative significance of the anticipated blogroll related part and the posts related factor. It is assumed that weights w_{BR} and w_{EP} are normalized to add up to 1 (i.e., $w_{BR} + w_{EP} = 1$). From the aforementioned analysis, it is obvious that the $LBSR^{BS_j}_{t_p=k}(BS_i)$ factor lies within the [0,1] range.

Weights $w_{t_p=k}$ in equation (1) are normalized to add up to 1 ($\sum_{\substack{\text{k=c-n+1} \\ k>0}}^{c} w_{t_p=k} = 1$) and may be given by equation 4.

$$w_{t_p=k} = \frac{w_k}{\sum_{l=1}^{n} w_l}, \quad where \quad w_k = \begin{Bmatrix} \text{n-c-k, c} \geq \text{n} \\ k, c < n \end{Bmatrix} \quad (4)$$

At this point it should be noted that the authors have assumed that the local rating estimation takes place at consecutive, equally distributed, time intervals. For the formation of the local accumulative BS rating at a time period c, the evaluator considers only the n more recent ratings formed. The value n determines the memory of the system. Small value for the n parameter means that the memory of the system is small, whereas large value considers a large memory for the system. Equation 4 in essence models the fact that more recent local BS ratings should weigh more in the overall BS rating evaluation.

B. Collaborative Local Rating

In order to estimate the rating of a target Blog Site BS_i, the evaluator Blog Site BS_j needs to contact a set WBS of N witness Blog Sites ($WBS \subseteq BS$) in order to get feedback reports on the usability of the BS_i. The set of the N witnesses is a subset of the $BS = BS_1, BS_2, ..., BS_M$ set and can be the blog sites in the FOAFroll of BS_j (i.e., the friends of the evaluator blog site BS_j, depth=1), the blog sites in the FOAFroll of the BS_j friends (i.e., the friends of the friends of BS_j, depth=2), etc. The target BS_i overall collaborative rating $CLBSR^{BS_j}_{t_p=c}(BS_i)$ may be estimated by the evaluator Blog Site BS_i at time period c in accordance with the following formula:

$$CLBSR^{BS_j}_{t_p=c}(BS_i)=w^{BS_j}_{t_p=c}(BS_j)\cdot LABSR^{BS_j}_{t_p=c}(BS_i) + \sum_{k=1,k\neq i}^{N} w^{BS_j}_{t_p=c}(BS_k)\cdot LABSR^{BS_j}_{t_p=c}(BS_i) \quad (5)$$

As may be observed from equation (5), the collaborative rating of the target BS_i is a weighted combination of two factors. The first factor contributing to the rating value is based on the direct experiences of the evaluator blog site BS_j, while the second factor depends on information regarding BS_i past behavior gathered from the N witnesses blog sites.

Weight $w^{BS_j}_{t_p=c}(BS_k)$ provides the relative significance of the rating of the target blog site BS_i as formed by the blog site BS_k to the overall rating estimation by the evaluator BS_j. In general, $w^{BS_j}_{t_p=c}(BS_k)$ is a measure of the credibility of witness BS_k, depends on the transitivity horizon considered and may be a function of the local accumulative blog site rating attributed explicitly or implicitly to each BS_k by the evaluator BS_j. Weight $w^{BS_j}_{t_p=c}(BS_k)$ may be given by the following equation:

$$w^{BS_j}_{t_p=c}(BS_k)=\sum_{BS_{x(d=1\ldots n)}} w^{BS_j}_{t_p=c}(BS_{x(d=1)})\cdot w^{BS_{x(d=1)}}_{t_p=c}(BS_{x(d=2)}) \cdot\ldots\cdot w^{BS_{x(d=n-1)}}_{t_p=c}(BS_{x(d=n)}) \quad (6)$$

where $BS_{x(d=m)}$ denotes the set of blog sites examined at depth=m,

$$w^{BS_j/BS_{x(d=m)}}_{t_p=c}(BS_{x(d=1)}/BS_{x(d=m+1)}) = \frac{LABSR^{BS_j/BS_{x(d=m)}}_{t_p=c}(BS_{x(d=1)}/BS_{x(d=m+1)})}{\sum LABSR^{BS_j/BS_{x(d=m)}}_{t_p=c}(BS_{x(d=1)}/BS_{x(d=m+1)})} \quad (7)$$

and $LABSR^{BS_j/BS_{x(d=m)}}_{t_p=c}(BS_{x(d=1)}/BS_{x(d=m+1)})$ is the local accumulative blog site rating attributed to Blog Site $BS_{x(d=1)}/BS_{x(d=m)}$ by the evaluator $BS_j/BS_{x(d=m+1)}$. One may easily conclude that for the evaluator BS_j it stands $w^{BS_j}_{t_p=c}(BS_j)=1$.

As may be observed from equation 6 the transitivity horizon considered is at most n. Parameter n and the weights w_{BR} and w_{EP} could be defined in the personalized recommendation system in accordance with the specific preferences of the evaluator blog site. In the current version of this study we have limited the transitivity horizon to n=3 (depth =3), as the authors believe that till that point a quite accurate recommendation of interest to the evaluator blog site could be provided. Additionally, according to equation 6 the credibility of a witness is formulated as a multiplicative function. Other functions (e.g., additive) could be defined. Finally, w_{BR} was set to 0.7 and w_{EP} to 0.3, thus giving emphasis to the blogroll information.

Our recommendation system exploits, for each evaluator blogsite, collaborative ratings estimated for blogs pointed by the evaluator and for blogs not explicitly connected to the evaluator blog site (e.g., a blog site belonging in the blogroll of a blog site that belongs in the blogroll of the evaluator blog site).

Finally, it should be noted that in the current version of this study, the local accumulative blog site rating estimations of the evaluator and its affiliated blog sites are assumed to be updated during the same time intervals, thus avoiding (due to space limitations and notation complexity) potential time effects (e.g., more recent ratings should weigh more in the overall collaborative blog site estimation).

V. Experimental evaluation

The aim of this work was to present a framework for adding rating information into blog hyperlinks and then build a mechanism that uses hyperlinks as recommendations in a collaborative rating approach. Since a blog dataset that contains ratings about bloggers and blog posts is not available at the moment, we have decided to use a ratings dataset that fits our model as much as possible. For this reason, we selected the *extended epinions dataset*, which was provided by Epinions and is available through the Trustlet wikipage [4].

The dataset 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, it contains user ratings that denote which users are trusted or distrusted (1, -1 respectively) by which users and article ratings (ranging from 1 to 6), which represent how much a certain user rates the review of another user. User ratings can be considered as the blogroll links in our case, whereas article ratings can be perceived as post hyperlinks. Finally, the dataset contains information about the author and subject of each review, thus, giving us evidence on the interests of each author. During the preprocessing phase, we have kept only the 717,667 positive trust ratings, which are supported by our model. We have removed self-references, i.e., statements about users trusting themselves, in order to avoid cycles. From this user set, we pick two subsets that contain user: a) with a short blogroll (setA: between 5 and 10 links) and b) with an extended blogroll (setB: more than 10 links). Both sets are of equal size ($\sim$ 5,000 users).

In order to evaluate our rating model, we generate for each user X (in each set) a list of recommended users, using the local accumulative (L) and the collaborative local (CL) formation. We measure an average similarity between X and her top-K users recommended using (L) and (CL) and compare to the similarity between X and the top-K users in her blogroll (BR). Similarity between two users X and Y is defined as the ratio of articles rated by X that have been also rated by Y. The average similarity values for the top-K matches (k=3,5,7,10,15,20,25,30) for each subset (setA, setB) and method (L, CL, BR) are presented in Figure 5.

The results show that the accumulative local rating (L) may provide better blog recommendations for a user and performance can be further improved if the collaborative local rating (CL) is employed. For users with short blogrolls (set A), the increase on the average similarity

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

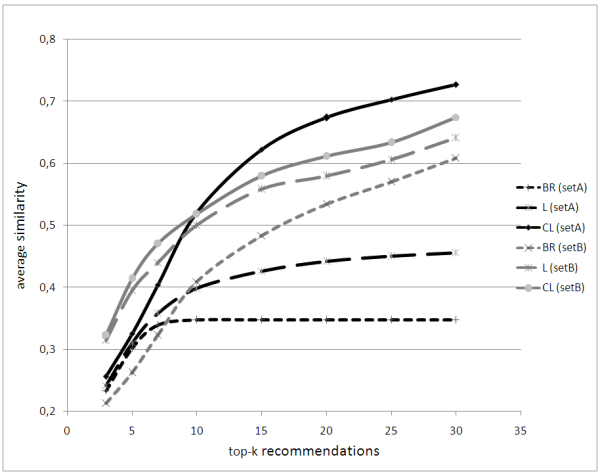

Figure 5. System architecture.

between the user and her recommendations ranges from 4 to 31 % when (L) is employed and raises to 110 % when (CL) is used. For users with long blogrolls the increase is smaller in both (L) and (CL). This indicates that a collaborative rating scheme is preferable for users with small blogrolls since and can assist them in finding resources of interest.

VI. Conclusions and Future work

This work presented an iterative collaborative process to provide a local rating for a set of blogs using information from the blogroll and post hyperlinks. The rating mechanism is exploited in order to build a personalized recommendation system, assisting bloggers in finding recent and trustworthy resources. The rating model is mathematically formulated, comprising local and local accumulative blog site rating formation (where the accumulative rating is calculated considering the local rating as estimated upon different consecutive time periods) and collaborative local blog site formation (where the evaluator blog site exploits information gathered form other affiliated witnesses blog sites). Our model exploits two special features of the blogosphere: a) the difference between blogroll links, which denote a more permanent trust towards the blog being pointed, and post links, which represent a more transient reference to a blog, b) the timestamp information of a link, which can be assumed from the post's timestamp. Additionally, in conjunction with the system architecture, a suggestion on the semantics that can be attached to each blog is also provided. An initial experimental evaluation shows that the model performs well in providing blog recommendations to users, based on the blogs they have linked to and their affiliated blogs. The next steps of this work is to study how propagation of trust and negative recommendation can be incorporated to the model and how the model can be applied to more social networking applications.

References

[1] E. Amitay, D. Carmel, M. Herscovici, R. Lempel, and A. Soffer. Trend detection through temporal link analysis. *Journal of the American Society for Information Science and Technology*, 55(14):1270–1281, January 2004.

[2] Z. Bar-yossef, R. Kumar, A. Z. Broder, and A. Tomkins. Sic transit gloria telae: Towards an understanding of the web's decay. In *Proceedings of the 13th International Conference on World Wide Web*, pages 328–337, 2004.

[3] K. Berberich, M. Vazirgiannis, and G. Weikum. Time-aware authority ranking. *Internet Mathematics Journal*, 3(2), 2005.

[4] J. Golbeck, B. Parsia, and J. Hendler. Trust networks on the semantic web. In *Proceedings of the 7th International Workshop on Cooperative Intelligent Agents, CIA2003*, 2003.

[5] R. Guha, R. Kumar, P. Raghavan, and A. Tomkins. Propagation of trust and distrust. In *Proceedings of the Thirteenth International World Wide Web Conference*, 2004.

[6] M. He, N. Jennings, and H. Leung. On agent-mediated electronic commerce. *IEEE Transactions on Knowledge and Data Engineering*, 15(4):985–1003, 2003.

[7] A. Kale, P. Kolari, A. Java, T. Finin, and A. Joshi. Modeling trust and influence in the blogosphere using link polarity. In *Proceedings of the International Conference on Weblogs and Social Media ICWSM 2007*, 2007.

[8] S. D. Kamvar, M. T. Schlosser, and H. Garcia-Molina. The eigentrust algorithm for reputation management in p2p networks. In *Proceedings of the Twelvth International World Wide Web Conference*, pages 640–651, 2003.

[9] A. Kritikopoulos, M. Sideri, and I. Varlamis. Blogrank: ranking blogs based on connectivity and similarity features. In *Proceedings of the 2nd international Workshop on Advanced Architectures and Algorithms For internet Delivery and Applications AAA-IDEA'06*. ACM, 2006.

[10] P. Massa and P. Avesani. Trust-aware collaborative filtering for recommender systems. In *Proceedings of the International Conference on Cooperative Information Systems - CoopIS 2004*, 2004.

[11] P. Massa and P. Avesani. Controversial users demand local trust metrics: An experimental study on epinions.com community. In *Proceedings of AAAI-05*, pages 121–126, 2005.

[12] D. Quercia, S. Hailes, and L. Capra. Lightweight distributed trust propagation. In *Proceedings of ICDM'07*, 2007.

[13] L. Raph and A. Aiken. Attack resistant trust metrics for public key certification. In *7th USENIX Security Symposium*, 1998.

[14] B. Yu and M. P. Singh. A social mechanism of reputation management in electronic communities. *Cooperative Information Agents*, pages 154–165, 2000.

[15] P. Yu, X. Li, and B. Liu. On the temporal dimension of search. In *Proceedings of the 13th International World Wide Web Conference on Alternate Track Papers and Posters*, pages 448–449. ACM Press, 2004.

[16] C. Ziegler and G. Lausen. Spreading activation models for trust propagation. In *Proceedings of the 2004 IEEE international Conference on E-Technology, E-Commerce and E-Service (Eee'04), IEEE Computer Society*, 2004.