Bayesian Credibility Modeling for Personalized Recommendation in Participatory Media

Aaditeshwar Seth†, Jie Zhang‡, and Robin Cohen

†Department of Computer Science and Engineering, IIT Delhi, India
‡School of Computer Engineering, Nanyang Technological University, Singapore
School of Computer Science, University of Waterloo, Canada
{aseth}@cse.iitd.ernet.in

Abstract. In this paper, we focus on the challenge that users face in processing messages on the web posted in participatory media settings, such as blogs. It is desirable to recommend to users a restricted set of messages that may be most valuable to them. Credibility of a message is an important criteria to judge its value. In our approach, theories developed in sociology, political science and information science are used to design a model for evaluating the credibility of messages that is user-specific and that is sensitive to the social network in which the user resides. To recommend new messages to users, we employ Bayesian learning, built on past user behaviour, integrating new concepts of context and completeness of messages inspired from the strength of weak ties hypothesis, from social network theory. We are able to demonstrate that our method is effective in providing the most credible messages to users and significantly enhances the performance of collaborative filtering recommendation, through a user study on the digg.com dataset.

1 Introduction

In the context of participatory media where web messaging is becoming increasingly prevalent, users are faced with a plethora of messages to view. Current techniques such as RSS feeds are not personalized and users often have to sift their way through hundreds of messages each day. In this paper, we aim to show how artificial intelligence techniques can be effectively introduced in order to assist users in their processing of messages. Our central theme is that fields such as sociology, political science and information science can be instrumental in developing a model for recommending credible messages to users. In particular, the modeling of a user's social network becomes a critical element and the approach of learning about each specific user's messaging preferences is essential in the successful recommendation of messages. We outline the motivating multi-disciplinary research, present our model for determining the credibility of messages to users and then introduce experimental results from a user study on the digg.com dataset (where users view and rate messages), to confirm the value of our proposed approach and its use in recommender systems.

Various researchers have proposed to model credibility as a multi-dimensional construct. Fogg and Tseng [1] reason about credibility criteria used by people to judge the credibility of computerized devices and software, and propose to include the modeling of (a) first-hand experience, (b) bias of a user towards categories of products, and (c) third-party reports about products. A model with similar distinctions is developed in [2] to evaluate the trustworthiness of users in an e-commerce setting. Here, the authors distinguish *witness reputation* (i.e. general public opinion) from *direct reputation* (i.e. opinion from a user's own experience) and include as well *system reputation* (i.e. the reputation from the role of a user, as buyer, seller or broker). These interacting users are modeled as being embedded in a social network of relationships that may be pre-declared or inferred based on the past history of interactions.

From sociology, the *strength-of-weak-ties* hypothesis [3] states that social networks of people consist of clusters with *strong* ties among members of each cluster, and *weak* ties linking people across clusters. Whereas strong ties are typically constituted of close friends, weak ties are constituted of remote acquaintances. The hypothesis claims that weak ties are useful for the diffusion of information and economic mobility, because they connect diverse people with each other. People strongly tied to each other in the same cluster may not be as diverse.

One among many studies based on the *strength-of-weak-ties* hypothesis, [4] traces the changes in political opinion of people before and after the 1996 presidential elections in USA, observed with respect to the social networks of people. It is shown that weak ties (identified as geographically dispersed ties of acquaintances) are primarily responsible for the diffusion of divergent political opinion into localized clusters of people having strong ties between themselves. As indicated by the *strength-of-weak-ties* hypothesis, this reflects that local community clusters of people are often homogeneous in opinion, and these opinions may be different from those of people belonging to other clusters. Furthermore, people have different propensities to respect opinions different from those of their local community members. This reflects that the personal characteristics of people also influence the extent to which they would be comfortable in deviating from the beliefs of their immediate local cluster.

From these studies, we learn that (a) there is value to look at the special case of third-party reporting within a user's cluster or local community, and (b) it is important to allow users to have different weights on the importance of different types of credibilities. Note that this last insight is reinforced by studies in information science [5], which argue that users have different preferences for different types of credibilities discussed so far. Inspired by these studies, we develop and operationalize a multi-dimensional subjective credibility model for participatory media as described next.

2 Bayesian Credibility Model

Knowledge Assumptions: Suppose that we wish to predict whether a message m_k about a topic t and written by user u_j, will be considered credible by user u_i. We consider a scenario where all older messages about topic t written in the past are labeled with the author of each message. In addition, a message may have also been assigned ratings by various recipient users, whenever users would have read the message, based on the credibility of the message for the recipient. The set of credibility ratings of any message are also assumed to be available.

Users may declare a subset of other users as their "friends". We refer to an explicitly declared relationship between two users as a *link* between them, and assume to have knowledge of the social network graph formed by all users and the links between pairs of users. Users may also declare topics of interest to them. We use this information, and the social network graph, to derive the *topic specific social network graph* for topic t, as the induced subgraph of the overall social network graph consisting only of those users and edges between users who are interested in topic t.

For each topic specific social network graph, community identification algorithms such as [6] can identify dense clusters of users and links. We use the definition of *strong* and *weak* ties proposed by [3], and refer to *strong* ties as links between users in the same cluster, and *weak* ties as links between users in different clusters. We use V_{it} to denote the local cluster of users strongly tied to user u_i with respect to topic t.

These assumptions are reasonable in contexts such as the website digg.com, which allows users to construct social networks by declaring some users as their friends. Information about message authorship and ratings given by users to messages is also available. We will show that we can use this knowledge to quantify different types of credibilities for each message with respect to each user. Then, based on ratings given by a particular user to older messages, we can use a Bayesian model to learn preferences of the user towards these different kinds of credibilities of messages. Finally, we can use this learned model to predict whether or not the new message m_k will be considered credible by user u_i.

Bayesian Network: We use the notion of strong and weak ties to develop two characteristics of messages: *context* and *completeness*. We assume that strong ties of a user, ie. close friends in the same social network cluster, share the same context, and hence their opinions contribute to the context of a message. On the other hand, completeness is assumed to be influenced by public opinion and not just the immediate social network cluster Based on this premise, the different types of credibilities that we choose to model are as follows:

- s_{ikt} = *cluster credibility*: This is based on the ratings given by other users in cluster V_{it}, that is, the cluster of user u_i. It denotes the credibility associated by the cluster or local community of u_i to the message m_k written by u_j, based on the belief of the members of the cluster about m_k. We assume that opinions of users in the same cluster will contribute only to adding context to messages; their contribution to completeness is already accounted for through public credibility explained next.
- p_{kt} = *public credibility*: This is based on ratings by all the users, and reflects the public opinion about the credibility for the message m_k written by u_j. Public credibility contributes only to the completeness of messages across all

users, including the users who's opinions have already been accounted in the cluster credibility construct.

- e_{ikt} = *experienced credibility*: This is based only on ratings given by user u_i in the past, and denotes the credibility that u_i associates with the message m_k written by u_j, based on u_i's self belief about u_j. We distinguish between the contributions experienced credibility would make to adding context to the message, or adding completeness.
- l_{ikt} = *role based credibility*: This denotes the credibility that u_i associates with the message m_k written by users having the same role as that of u_j; for example, based on whether the messages' authors are students, or professors, or journalists, etc.

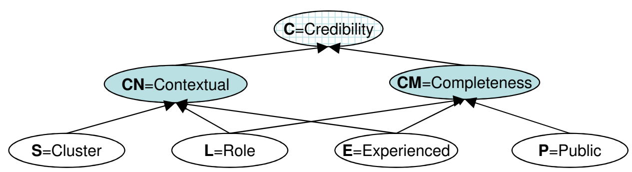

Fig. 1. Bayesian Credibility Model

Each of these credibilities can be expressed as a real number $\in [0, 1]$, and we propose a Bayesian network to combine them into a single credibility score. The model is shown in Fig. 1. Our aim is to learn the distribution for $\mathrm{P}_{it}(\mathbf{C}|\mathbf{E,L,S,P})$ for each user and topic based on ratings given by various users to older messages; here, $\{\mathbf{E,L,S,P}\}$ are evidence variables for the four types of credibilities for a message, and $\mathbf{C}$ is a variable denoting the credibility that u_i associates with the message. Thus, for each topic t, a set of messages M about t will be used during the training phase with samples of $(c_{ik},\ e_{ik},\ l_{ik},\ s_{ik},\ p_k)$ for different messages $m_k \in M$ to learn the topic specific credibility models for u_i. Assuming that a user's behavior with respect to preferences for different kinds of credibilities remains consistent over time, the learned model can now be used to predict c_{ix} for a new message m_x about topic t, that is, $\mathrm{P}_{it}(c_{ix}|e_{ix}, l_{ix}, s_{ix}, p_x)$. We also introduce two hidden variables, to help make the model more tractable to learn, and to capture insights about messages that we developed in prior work [7] – context and completeness, defined as follows:

- *Context* relates to the ease of understanding of the message, based on how well the message content explains the relationship of the message to its recipient. *Simplification* of the meaning of the message [8], can be considered as an outcome of the amount of *context* in the message.
- *Completeness* denotes the depth and breadth of topics covered in the message. The *scope* of the message, or the *opinion diversity* expressed in the message [8], can be considered as outcomes of the degree of *completeness* of the message.

Note that our modeling method has some interesting design features: the model takes into account personal and contextual opinions of people that may influence their credibility judgements; the model is learned in a personalized manner, and allows accommodating varying degrees of propensities of users to respect opinions of other users; different model instances are learned for different topics, making credibility judgements topic-specific.

3 Credibility Computation

We begin with the following axioms:

– *A-1*: A message is credible if it is rated highly by credible users.
– *A-2*: A user is credible if messages written by her are rated highly by other credible users.
– *A-3*: A user is also credible if ratings given by her are credible, that is, she gives high ratings to messages that appear to be credible to credible users, and low ratings to messages that appear to be non-credible.
– *A-4*: A user is also credible if she is linked to by other credible users.

We henceforth assume that we are operating within some topic t, and drop the subscript for simplicity. We begin with the following information:

– **A[k,n]**: A matrix for k messages and n users, where $a_{ij} \in \{0,1\}$ indicates whether message m_i was written by u_j
– **R[k,n]**: A ratings matrix for k messages and n users, where $r_{ij} \in \{0,1\}$ [2] indicates the rating given to message m_i by user u_j
– **N[n,n]**: A social network matrix where $n_{ij} \in \{0,1\}$ indicates the presence or absence of a link from user u_i to user u_j. We also assume that the clustering algorithm can identify clusters of strong ties among users, connected to other clusters through weak ties.

Our goal is to find a method to compute the evidence variables for the Bayesian model using the axioms given above. The evidence variables can be expressed as the matrices **E[n,k]**, **L[n,k]**, **S[n,k]**, and **P[k]**, containing the credibility values for messages. Here, p_k is the public credibility for message m_k authored by user u_j. e_{ij} and l_{ij} are the experienced and role based credibilities respectively for message m_k according to the self-beliefs of user u_i. Similarly, s_{ij} is the cluster credibility for message m_k according to the beliefs of the users in u_i's cluster V_i. Once these evidence variables are computed for older messages, they are used to learn the Bayesian model for each user. Subsequently, for a new message, the learned model for a user is used to predict the credibility of the new message for the user. We begin with computation of the evidence variable matrix for public credibility **P**; we will explain later how other credibilities can be computed in a similar fashion. Detailed algorithms can be found in [9].

[2] We assume in this paper that the ratings are binary. However, our method can be easily generalized to real-valued ratings as well.

1. Let $\mathbf{P'[n]}$ be a matrix containing the public credibilities of users, and consider the credibility of a message as the mean of the ratings for the message, weighted by the credibility of the raters (*A-1*): $p_k = \sum_i r_{ki}.p'_i/|r_{ki} > 0|$. This is the same as a matrix multiplication $\mathbf{P{=}R}_r\mathbf{.P'}$, where $\mathbf{R}_r$ is the row-stochastic form of $\mathbf{R}$, ie. the sum of elements of each row = 1.
2. The credibility of users is calculated as follows:

2a. Consider the credibility of a user as the mean of the credibilities of her messages (*A-2*): $p'_i = \sum_k p_k/|p_k|$ (or written as $\mathbf{P'{=}A}_c^T\mathbf{.P}$), where $\mathbf{A}_c$ is the column-stochastic form of $\mathbf{A}$; and $\mathbf{A}_c^T$ is the transpose of $\mathbf{A}_c$.

2b. The above formulation indicates a fixed point computation:

$$\mathbf{P'{=}A}_c^T\mathbf{.R}_r\mathbf{.P'} \tag{1}$$

Thus, $\mathbf{P'}$ can be computed as the dominant Eigenvector of $\mathbf{A}_c^T\mathbf{.R}_r$. This formulation models the first two axioms, but not yet the ratings-based credibility (*A-3*) and social network structure of the users (*A-4*). This is done as explained next.

2c. Perform a fixed-point computation to infer the credibilities $\mathbf{G[n]}$ acquired by users from the social network (*A-4*):

$$\mathbf{G{=}(}\beta\mathbf{.N}_r^T + \mathbf{(1\text{-}}\beta\mathbf{).Z}_c\mathbf{.1}^T\mathbf{).G} \tag{2}$$

Here, $\beta \in (0,1)$ denotes a weighting factor to combine the social network matrix $\mathbf{N}$ with the matrix $\mathbf{Z}$ that carries information about ratings given to messages by users. We generate $\mathbf{Z}$ by computing z_i as the mean similarity in credibility ratings of user u_i with all other users. The ratings similarity between a pair of users is computed as the Jacquard's coefficient of common ratings between the users. Thus, z_i will be high for users who give credible ratings, that is, their ratings agree with the ratings of other users (*A-3*). In this way, combining the social-network matrix with ratings-based credibility helps to model the two remaining axioms as well. Note that $\mathbf{Z}_c\mathbf{[n]}$ is a column stochastic matrix and $\mathbf{1[n]}$ is a unit column matrix; augmenting $\mathbf{N}$ with $\mathbf{Z}_c\mathbf{.1}^T$ provides an additional benefit of converting $\mathbf{N}$ into an irreducible matrix so that its Eigenvector can be computed [3]

2d. The ratings and social network based scores are then combined together as:

$$\mathbf{P'{=}(}\alpha\mathbf{.A}_c^T\mathbf{.R}_r + \mathbf{(1\text{-}}\alpha\mathbf{).G}_c\mathbf{.1}^T\mathbf{).P'} \tag{3}$$

Here again $\mathbf{1}$ is a unit column matrix, and $\alpha \in (0,1)$ is a weighting factor. The matrix $\mathbf{P'}$ can now be computed as the dominant Eigenvector using the power method.

3. Once $\mathbf{P'}$ is obtained, $\mathbf{P}$ is calculated in a straightforward manner as $\mathbf{P{=}R}_r\mathbf{.P'}$.

The processes to compute cluster $\mathbf{S[n,k]}$, experienced $\mathbf{E[n,k]}$, and role based $\mathbf{L[n,k]}$ credibilities are identical, except that different cluster credibilities are calculated with respect to each cluster in the social network, and different experienced and role based credibilities are calculated with respect to each user.

[3] This step is similar to the Pagerank computation for the importance of Internet web pages [10].

The cluster credibilities $\mathbf{S[n,k]}$ are computed in the same manner as the public credibilities, but after modifying the ratings matrix $\mathbf{R}$ to contain only the ratings of members of the same cluster. Thus, the above process is repeated for each cluster, modifying $\mathbf{R}$ in every case. For each users u_i belonging to cluster V_i, s_{ik} is then equal to the cluster credibility value for message m_k with respect to u_i. The matrix $\mathbf{Z}$ in the computation on the social network matrix is also modified. When computing the cluster credibilities for cluster V_i, element z_j of $\mathbf{Z}$ is calculated as the mean similarity of user u_j with users in cluster V_i. Thus, z_j will be high for users who are regarded credible by members of cluster V_i because their ratings agree with the ratings of the cluster members.

The experienced credibilities $\mathbf{E[n,k]}$ are computed in the same manner as well, but this time for each user by modifying the ratings matrix $\mathbf{R}$ to contain only the ratings given by the user. The matrix $\mathbf{Z}$ is also modified each time by considering z_j as the similarity between users u_i and u_j, when calculating the experienced credibilities for u_i.

Role based credibility is computed as the mean experienced credibilities of users having the same role. However, we do not use role based credibility in our evaluation because sufficient user profile information was not available in the digg dataset used by us. Henceforth, we ignore $\mathbf{L[n,k]}$ in our computations.

Model Learning: Once the various types of credibilities for messages are calculated with respect to different users, this training data is used to learn the Bayesian model for each user and topic of interest to the user using the Expectation-Maximization (EM) algorithm. The model parameters are learned to predict for user u_i interested in topic t, the probability $\mathrm{P}_{it}(c_{ix}|e_{ix}, s_{ix}, p_x)$ that u_i will find a new message m_x to be credible.

Inference: Now, for a new message m_x, the evidence variables are calculated with respect to a recipient user u_i in one of two ways as described next, and the learned model is used to produce a probabilistic prediction of whether u_i would find m_x to be credible.

- *Authorship*: The four types of credibilities of the message are considered to be the same as the credibilities of its author with respect to u_i.
- *Ratings*: The cluster and public credibilities are calculated as the weighted mean of ratings for the message given by other users and the credibilities of these users with respect to u_i. The experienced and role based credibilities are the same as the corresponding credibilities of the message author wrt u_i.

As we will show in the evaluation, the ratings method performs better than the authorship method. This allows new users to popularize useful messages written by them because their own credibility does not play a role in the computations. It also allows credible users to make mistakes because the credibility of the author is not taken into account. Given the evidence variables for the new message, and the learned Bayesian model, the probability of u_i finding the message to be credible is computed using standard belief propagation methods such as Markov-Chain-Monte-Carlo (MCMC).

4 Evaluation

We evaluate our method over a dataset of ratings by real users obtained from a popular knowledge sharing website, digg.com [11]. The website allows users to submit links to news articles or blogs, which are called *stories* by the website. Other users can vote for these stories; this is known as *digging* the stories. Stories that are *dugg* by a large number of users are promoted to the front-page of the website. In addition, users are allowed to link to other users in the social network. Thus, the dataset provides us with all the information we need:

- Social network of users: We use this information to construct the social network link matrix between users $\mathbf{N[n,n]}$. The social network is clustered using MCL, a flow-stochastic graph clustering algorithm [6], to produce classifications of ties as strong or weak. The cluster of users strongly connected to user u_i is referred to as V_i.
- Stories submitted by various users: We use this information to construct the authorship matrix $\mathbf{A[k,n]}$. Since all the stories in the dataset were related to technology, we consider all the stories as belonging to a single topic.
- Stories dugg by various users: We use this information to construct the ratings matrix $\mathbf{R[k,n]}$. We consider a vote of 1 as an evidence for credibility of the story, and a vote of 0 as an evidence of non-credibility.

Although the dataset is quite large with over 200 stories, we are able to use only 85 stories which have a sufficiently large number of ratings by a common set of users. This is because we require the same users to rate many stories so that we have enough data to construct training and test datasets for these users. Eventually, we assemble a dataset of 85 stories with ratings by 27 users. A few assumptions we make about the validity of the dataset for our experiments are as follows:

- The submission of a story to Digg may not necessarily be made by the author of the story. However, we regard the submitting user as the message author because it distinguishes this user from other users who only provide further ratings to the messages.
- The ratings provided on the Digg website may not reflect credibility, but rather usefulness ratings given to messages by users. We however consider them to be equivalent to credibility and do not include users who rate more than 65 stories as all credible or all non-credible. We argue that in this pruned dataset, all the users are likely to be interested in the topic and hence all the stories; therefore, the only reason for their not voting for a story would be its credibility.

We use an open-source package, OpenBayes, to program the Bayesian network. We simplify the model by discretizing the evidence variables $\mathbf{E,S,P}$ into 3 states, and a binary classification for the hidden variables $\mathbf{N}$, $\mathbf{M}$, and the credibility variable $\mathbf{C}$. The discretization of the evidence variables into 3 states is performed by observing the Cumulative Distribution Frequency (CDF) and Complementary CDF (CCDF) of each variable with respect to the credibility rating of users. The lower cutoff is chosen such that the product of the CDF for rating=0 and CCDF for rating=1 is maximum, and the upper cutoff is chosen such that the CCDF for rating=0 and CDF for rating=1 is maximum. This gives a high discrimination ability to the classifier because the cutoffs are selected to maximize the pair-wise correlation of each evidence variable with the credibility rating given by the user.

Metrics: We evaluate the performance of the model for each user by dividing the 85 stories into a training set of 67 stories and a test set of 17 stories (80% and 20% of the dataset respectively). We then repeat the process 20 times with different random selections of stories to get confidence bounds for the cross validation. For each evaluation, we use two kinds of performance metrics [12], *Matthew's correlation coefficient* (MCC) and *TPR-FPR*. The MCC gives a single metric for the quality of binary classifications. TPR-FPR plots on an XY-scale the true positive rate (TPR) with the false positive rate (FPR) of a binary classification. The random baseline is TPR=FPR. Points above the random baseline are considered to be good.

All experiments are performed with $\alpha = 0.5$ (eqn. 3) and $\beta = 0.85$ (eqn. 2) which were found to be robust values [9], and also convey our message that all of authorship, ratings, and social networks provide valuable credibility information.

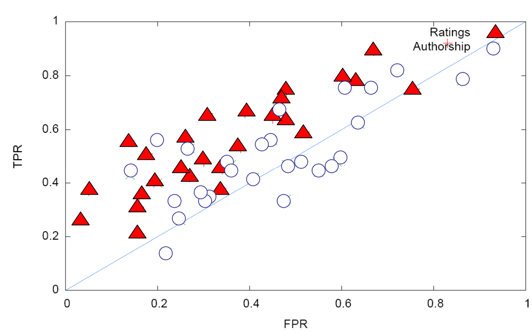

Fig. 2. Performance of Bayesian Credibility Model

Inference Methods: Fig. 2 shows the TPR-FPR plot for ratings and authorship based evidence variable computation when $\alpha = 0.5$ and $\beta = 0.85$. As can be seen visually, the ratings-based method performs better than the authorship-based method. The former gives MCC = 0.156 (σ=0.073), while the latter gives MCC = 0.116 (σ=0.068). However, the authorship performance is still successful for a majority, which is encouraging. This indicates that authorship information may be used to solve the problem of cold-start for new messages that have not acquired a sufficient number of ratings. Similarly, ratings may be used to solve cold-start for new authors who have not acquired sufficient credibility.

Comparison: We next compare our method with other well known methods for trust and reputation computation meant for different applications.

An Eigenvector computation on $\mathbf{A}_c^T.\mathbf{R}_r$ by leaving out the social network part (eqn. 1), is identical to the Eigentrust algorithm [13]. The best choice of parameters could only give a performance of MCC = -0.015 (σ = 0.062). Eigentrust has primarily been shown to work in P2P file sharing scenarios to detect malicious users that inject viruses or corrupted data into the network. The P2P context requires an objective assessment of the trustworthiness of a user, and does not allow for subjective differences, as desired for participatory media.

An Eigenvector computation on the social network matrix (eqn. 2), personalized for each user, is identical to the Pagerank algorithm used to rank Internet web pages [10]. However, this too performs poorly with an MCC = 0.007 (σ = 0.017). This suggests that users are influenced not only by their own experiences, but also by the judgement of other users in their cluster, and by public opinion.

In conclusion, these and other methods we compared perform close to random, even with personalization. We believe this to be due to a fundamental drawback of these methods: they try to form an objective assessment of credibility for users and messages, which is not appropriate for participatory media. Our approach which subjectively model credibility, allowing users to be influenced in different ways by different sources, perform better than objective modeling approaches.

5 Use in Recommender Systems

Our method for credibility computation can be used in two ways to improve recommender systems: (i) Since our method serves to predict the probability of a user finding a message to be credible or non-credible, it can be used as a pre- or post-filtering stage with existing recommendation algorithms. (ii) It can also be adapted to integrate closely with recommendation algorithms; we show how to do this with collaborative filtering (CF) [14] in this section.

A basic CF algorithm works in two steps. First, similarity coefficients are computed between all pairs of users, based on the similarity of message ratings given by each pair. Second, to make a decision whether or not to recommend a new message to a user, the mean of the message ratings given by other similar users is computed, weighted on the coefficients of similarity to these users. If the mean is greater than a threshold, the message is recommended; else it is rejected.

The drawback of the CF method is that it only learns the average user behavior. However, as we have argued, user behavior can be different in different circumstances. We therefore develop an adaptation of our method. Rather than computing a single similarity coefficient between each pair of users, we compute four similarity coefficients based upon whether messages are believed to be highly contextual by both users, or highly complete by both users, or contextual by the first user and complete by the second user, or vice versa. Essentially, we break down the average user behavior into four components based upon the context and completeness of messages to users, as follows:

1. For each user, we run the EM algorithm on training set to learn the model.
2. We use the learned model to infer the probabilities of the hidden variables of context and completeness for each story in the training set: $\mathrm{P}_i(\mathbf{CN}|\mathbf{E,S,P,C})$ and $\mathrm{P}_i(\mathbf{CM}|\mathbf{E,S,P,C})$ shown in Fig. 1. That is, for each story m_j, we infer $\mathrm{P}(cn_{ji}{=}0{,}1|e_{ji}, s_{ji}, p_{ji}, c_{ji})$ and $\mathrm{P}(cm_{ji}{=}0{,}1|e_{ji}, s_{ji}, p_{ji}, c_{ji})$.
3. We then discretize the probabilities for **CN** and **CM** in same way as we did earlier, by finding cutoffs that maximized the product of the CDF for c_{ji}=0 and CCDF for c_{ji}=1. This gives us samples of ($c_{ji} \in \{0, 1\}$, $cn_{ji} \in \{0, 1\}$, $cm_{ji} \in \{0, 1\}$), that is, which stories appear contextual or complete to a user, and the rating given by the user to these stories.
4. For every pair of users, their samples are then compared to produce four similarity coefficients on how similar the users are in their contextual opinion, completeness opinion, and cross opinions between messages that appear contextual to one user and complete to the other, or vice versa.
5. Finally, when evaluating the decision to recommend a test message to a user, the mean of the message ratings is computed over all the four coefficients of similarity, rather than over a single coefficient as in the basic CF algorithm.

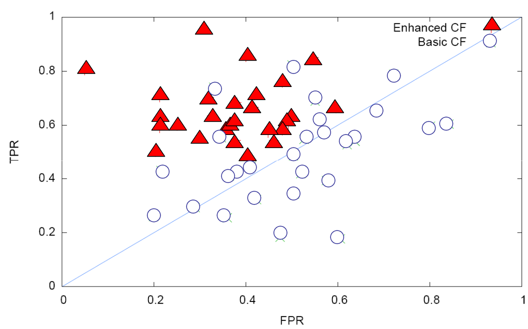

Fig. 3. Enhancement of Collaborative Filtering

Fig. 3 shows the performance of the basic CF scheme and our enhanced version. The basic scheme performs worse than random for many users, but when enhanced with breaking up the average user behavior into contextual and completeness components, the performance improves considerably. The mean MCC for the basic scheme is 0.017 ($\sigma = 0.086$), and for the enhanced scheme is 0.278 ($\sigma = 0.077$), a sixteen-fold improvement. We consider this to be a huge improvement over the existing methodologies for recommendation algorithms, especially to build applications related to participatory media.

6 Conclusions

In this paper, we made use of insights from sociology, political and information science, and HCI, to propose a subjective credibility model for participatory media content. We formulated the model as a Bayesian network that can be learned in a personalized manner for each user, making use of information about the social network of users and ratings given by the users. We showed that our method works better than existing methods on trust and reputation computation. In addition, an adaptation of our method to recommendation algorithms such as collaborative filtering (CF) was able to improve CF on our dataset. This encourages the use of sociological insights in recommender system research.

References

1. Fogg, B., Tseng, H.: The elements of computer credibility. In: Proceedings of SIGCHI. (1999)
2. Sabater, J., Sierra, C.: Regret: A reputation model for gregarious societies. In: Proceedings of the Fifth International Conference on Autonomous Agents Workshop on Deception, Fraud and Trust in Agent Societies. (2001) 61–69
3. Granovetter, M.: The strength of weak ties. American Journal of Sociology **78**(6) (1973)
4. Baybeck, B., Huckfeldt, R.: Urban contexts, spatially dispersed networks, and the diffusion of political information. Political Geography **21** (2002)
5. Rieh, S.: Judgement of information quality and cognitive authority on the web. Information Science and Technology **53**(2) (2002)
6. Dongen, S.: MCL: A Cluster Algorithm for Graphs. PhD Thesis, University of Utrecht (2000)
7. Seth, A., Zhang, J.: A social network based approach to personalized recommendation of participatory media content. In: Proceedings of ICWSM. (2008)
8. Bryant, J., Zillman, D.: Media Effects: Advances in Theory and Research. Lawrence Erlbaum Associates (2002)
9. Seth, A.: An infrastructure for participatory media. PhD thesis, University of Waterloo (2009)
10. Brin, S., Page, L.: The pagerank citation ranking: Bringing order to the web. Technical Report http://dbpubs.stanford.edu:8090/pub/1999-66 (2001)
11. Lerman, K.: Social information processing in news aggregation. IEEE Internet Computing **11**(6) (2007)
12. Davis, J., Goadrich, M.: The relationship between precision-recall and roc curves. In: Proceedings of ICML,. (2006)
13. Kamvar, S., Scholsser, M., Garcia-Molina, H.: The eigentrust algorithm for reputation management in p2p networks. In: Proceedings of WWW. (2003)
14. Adomavicius, G., Tuzhilin, A.: Toward the next generation of recommender systems: A survey of the state-of-the-art and possible extensions. IEEE Trans. Knowledge and Data Engineering **17**(6) (2005)