Recommending Collaborators using Social Features and MeSH Terms

Danielle H. Lee & Peter Brusilovsky
School of Information Sciences,
University of Pittsburgh
135 N. Bellefield Ave., Pittsburgh, PA 15260
[hyl12, peterb]@pitt.edu

Titus Schleyer
Center for Dental Informatics, School of Dental Medicine, University of Pittsburgh,
3501 Terrace St., 331, Pittsburgh, PA 15261
titus@pitt.edu

ABSTRACT

Unlike expertise location systems which users query actively when looking for an expert, expert recommender systems suggest individuals without the context of a specific problem. An interesting research question is whether expert recommender systems should consider a users' social context when recommending potential research collaborators. One may argue that it might be easier for scientists to collaborate with colleagues in their social network, because initiating collaboration with socially unconnected researchers is burdensome and fraught with risk, despite potentially relevant expertise. However, many scientists also initiate collaborations outside of their social network when they seek to work with individuals possessing relevant expertise or acknowledged experts. In this paper, we studied how well content-based, social and hybrid recommendation algorithms predicted co-author relationships among a random sample of 17,525 biomedical scientists. To generate recommendations, we used authors' research expertise inferred from publication metadata and their professional social networks derived from their co-authorship history. We used 80% of our data set (articles published before 2007) as our training set, and the remaining data as our test set (articles published in 2007 or later). Our results show that a hybrid algorithm combining expertise and social network information outperformed all other algorithms with regards to Top 10 and Top 20 recommendations. For the Top 2 and Top 5 recommendations, social network-based information alone generated the most useful recommendations. Our study provides evidence that integrating social network information in expert recommendations may outperform a purely expertise-based approach.

Keywords

Expert recommendation, Medical Subject Headings (MeSH), social networks, biomedical research.

INTRODUCTION

The volume of information and resources in biomedical research, such as MEDLINE citations, gene sequences, tools and methods, and funding opportunities, is growing rapidly, often at an exponential rate. Faced with an ever-growing supply of information, researchers must invest increasing effort and time in routine information management, or risk missing relevant material and opportunities to advance their work. This information glut problem also exists when deciding which researchers to collaborate with [25]. For many scientists, collaboration is necessary to solve technical problems, address newly emerging research topics, initiate a research project, engage in inter- and multi-disciplinary work, or develop their professional network. When choosing collaborators, researchers must consider several aspects of potential candidates, such as expertise and skills, social distance, reputation, and personal and demographic traits, such as location, affiliation, ability to collaborate and communication preferences. Social matching systems for professionals or expertise location systems aim to help searchers cope with the information glut problem when trying to find a person to work with [24]. For instance, to help find a software engineer knowledgeable about a certain topic, an expertise location system can search work products or a software source control system to determine potential matches [15].

Most professional social matching systems and expertise location systems are designed to support collaborative problem solving based on a query by the user [20]. Put differently, expertise location is primarily focused on helping a person answer a specific question or solve a specific problem [24]. Expert recommender systems, on the other hand, suggest collaborators in the absence of a specific user request. These systems aim to introduce scientists reciprocally and help them start new collaborative interactions. In a previous study, we developed and evaluated recommendation algorithms solely based on expertise, i. e. "who knows what," inferred from scientists' publication metadata [9]. Subsequent to this study, we began to think about whether and how users' social context, such as "who knows who knows what" and "who is connected to whom," could contribute to making recommendations. We focused on the social context based on the observation that it is often easier for scientists to collaborate with colleagues in their social network, because initiating collaboration with socially unconnected researchers is burdensome and fraught with risk, despite potentially relevant expertise. However, many scientists also initiate collaborations outside of their social network when they seek to work with recognized experts or individuals possessing specific expertise.

Therefore, we compared content-based recommendation algorithms with and without users' social context in this paper. Specifically, we generated four types of recommendations: 1) content-based recommendations using metadata for authors' publications, specifically MeSH terms and author rank; 2) authors' local social network-based recommendation, mimicking traditional collaborative filtering-based approaches; 3) global social network-based recommendation derived from authors' co-authorship networks; and 4) hybrid recommendations combining information about authors' expertise with their social network. We empirically evaluated these recommendation algorithms using a data set of 17,525 biomedical scientists and their 22,542 papers.

RELATED WORK

Due to the collaborative nature in problem-solving, much research and practice in computer-supported cooperative work has focused on expertise finding within organizations [27]. Most expertise location systems help find individuals who are the most knowledgeable about a topic explicitly specified by the user and/or are socially close enough to contact. A sample system is SmallBlue [11], which is an internal IBM system that helps users find experts on a certain topic. It is both content- and social network-based, and visualizes the social network of experts when queried for a specific topic. The system employs private emails and chat logs to determine expertise and social connections. Even though SmallBlue users grant the system explicit access to their personal communications logs, privacy issues may reduce its suitability for other settings, such as academia.

ReferralWeb [8] was an early attempt to locate experts using social networks. This research prototype used a social network graph in order to allow users to find short referral chains to suggested experts quickly. Expertise profiles of users were constructed by mining publicly available Web documents. The system inferred personal expertise from Web pages that mentioned people and topics together. To build social networks, the system perceived pairs of users co-appearing on a Web page as socially connected. Inherent in this approach, however, is a high degree of uncertainty in representing social networks and expertise accurately. It is also unknown how well the approach would work in organizations in which expertise and social connections are represented differently from the Web.

Yang and Chen [28] described an expertise location system built on an educational P2P (peer-to-peer) system at a Taiwanese university. When queried for a topic, the system recommends items posted by users with the highest expertise scores and who are most preferred by the target user. For the system to function properly, human experts must assess each user's expertise and users must rate each other explicitly. As a consequence, the system requires significant ongoing human intervention that is unlikely to be sustainable in all but the most narrow contexts.

The Expertise Oriented Search (EOS) system [1] [10] is designed to help users in identifying expertise and exploring social associations of researchers in computer science. To do so, the system draws on a researcher's 20 most relevant Web pages retrieved through Google and publication list as obtained from the Digital Bibliography and Library Project, and Citeseer, respectively. Topic relevance is propagated through social connections under the assumption that a person's expertise diffuses through interactions in social networks. Both the original topical expertise and propagated relevance values are taken into account when searching for experts.

As opposed to expertise location systems, expert recommender systems respond to users' implicit need to find experts by providing recommendations. McDonald [14] developed and evaluated a system to recommend experts within a software company. The recommendation algorithm integrated two kinds of social networks: work context- and sociability-based. These social networks were constructed partially through user preferences and partially by researchers using various ethnographic methods. An evaluation did not identify one type of network as superior over the other, but suggested that there was a trade-off in recommendations when considering only expertise or social connections, respectively. The social networks in the system were created entirely through manual means, making the approach hard to use in other contexts.

Pavlov and Ichise [21] analyzed the structure of social networks to predict collaborations in a Japanese science institution. They used graph theory to build feature vectors for each expert dyad and applied four machine learning methods, support vector machines, two decision trees and boosting, to predict collaborations. The two decision tree techniques outperformed when precision and recall were combined, and all algorithms were better than the random (control) approach.

Bedrick and Sitting's [2] Facebook application of MEDLINE Publication (MP) is one system described for biomedical research that relies entirely on content for expert recommendations. MP models expertise using MeSH terms

[1] http://www.arnetminer.org/

drawn from publications. It recommends potential collaborators by comparing the angle of small expertise vectors calculated using singular value decomposition. MP did not integrate social network information in its recommendations.

As this brief review shows, many expertise location systems integrate information about users' expertise with social connections. On the other hand, relatively few expert recommender systems have combined expertise and social network information. Previous studies have tended to focus on one or the other, not both. Therefore, in this study, we compared content-based and social approaches to recommendation, both alone and in combination. In the following section, we explain the four algorithms we developed and evaluated.

DESCRIPTION OF RECOMMENDATION ALGORITHMS

We developed our recommendation algorithms using four different perspectives: 1) authors' expertise[2], 2) authors' local social network, 3) authors' global social network, and 4) a hybridized approach. Our first approach to recommendation is purely content-based. We inferred authors' research expertise from metadata about their publications, i.e. MeSH terms and author rank, and used these data to match authors based on the similarity of research topics they worked on. The second, local social network-based approach is a simple social network-based recommendation method which mimics collaborative filtering (CF). We transformed authors' history of collaboration as evidenced by co-authorship to the format of a user-item rating and recommended co-authors of an author's colleagues as potential collaborators. The third, global social network-based recommendation method is more complex than the second approach. We derived the global social network from authors' co-author relationships, but used social network information beyond 1 hop distances. The last method combined both content-based and social approaches. Figure 1 lists the recommendation algorithms we studied.

Metadata-based Recommendations

Our first recommendation method is based on authors' expertise, inferred from authors' publication metadata. In biomedicine, most publications are indexed by a well-defined controlled vocabulary, the Medical Subject Headings (MeSH) [5]. MeSH is the indexing method for MEDLINE, a comprehensive index to the periodical literature in biomedicine maintained by the U.S. National Library of Medicine. As of 2010, MeSH consists of 25,588 distinct terms in an eleven-level hierarchical structure [16]. We assumed that the set of MeSH terms extracted from an author's publications represented the author's research expertise and used them to define a research expertise vector. Then, we compared authors' research expertise vectors pairwise using the Vector Space Model (VSM), one of the most commonly used approaches in information retrieval [13]. We built research expertise vectors in two ways: consisting of 1) MeSH terms only and 2) MeSH terms and author rank, i.e. the position (1^{st}, 2^{nd}, etc.) of the author in the author list.

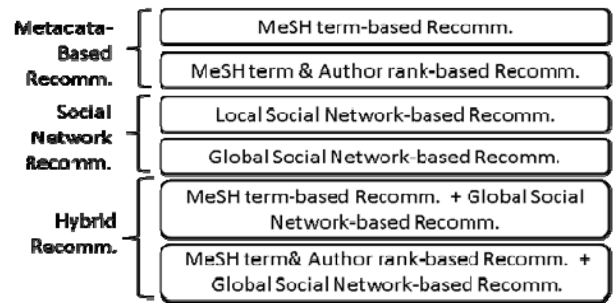

Figure 1. Recommendation Algorithms Under Study

To calculate the expertise similarity of two authors (a_i and a_j), we first computed the Term Frequency and Inverse Document Frequency (TF/IDF) of an author's MeSH term collection (Equation 1). Variable w_{in} denotes the TF/IDF value of a MeSH term n among an author a_i's publications. The TF/IDF is the product of term frequency (tf_{in}) and inverse document frequency (idf_n) [12]. Term frequency tf_{in} measures how many times a term n appears in the author a_i's publications. The higher the term frequency, the higher is the presumed expertise of the author on the subject. However, term frequency alone is insufficient to calculate similarity, because terms occurring frequently *across* authors' collections do not distinguish authors' expertise very well. Therefore, we applied inverse document frequency (idf_n) emphasizing terms which occur less frequently across all authors. Once we had computed the TF/IDF values of authors' MeSH term sets, we used the Cosine similarity of the TF/IDF values (Equation 2) to compare the similarity of research topics between a pair of authors (a_i and a_j). The variable V is the union set of MeSH terms that a_i and a_j have. The Cosine similarity is computed using the TF/IDF values of all terms of both authors.

$$\mathrm{w}_{in} = tf_{in} \times \mathrm{idf}_n \qquad \textbf{eq. 1}$$

$$\mathrm{Cosine}(\mathrm{a_i}, \mathrm{a_j}) = \frac{\sum_{n=1}^{|V|} \mathrm{w_{in}} \times \mathrm{w_{jn}}}{\sqrt{\sum_{n=1}^{|V|} \mathrm{w_{in}^2} \times \sum_{n=1}^{|V|} \mathrm{w_{jn}^2}}} \qquad \textbf{eq. 2}$$

MeSH term-based comparisons are simple because they only consider the collective MeSH terms assigned to each author's publications. However, authors of an article typically have different roles and expertise. Therefore, our second content-based approach takes into account author rank because we hypothesize that it is correlated with expertise. Typically, the first author of a paper is considered to have the highest expertise on the topic of the paper. In this paper, we make the simplified assumption that all authors' expertise on the topic of a paper is proportional to

[2] In this paper, since we mainly focus on users' publications and collaboration history, we use '*user*' and '*author*' interchangeably.

their position in the author list. While this assumption may not hold in all cases (for instance for papers authored by trainees and their advisors), we assume that senior researchers already have enough other publications to demonstrate their research expertise on a specific topic. Therefore, the decrease in expertise for senior researchers due to lower author rank is unlikely to be significant. We compute author rank as shown in Equation 3.

$$o_{im} = \sum_{k=1}^{K} (ta_k - (ao_{ik} - 1))/ta_k \qquad \textbf{eq. 3}$$

$$o_{1A} = \frac{3-(1-1)}{3} + \frac{11-(4-1)}{11} \qquad \textbf{eq. 4}$$

o_{im} is the author a_i's weighted author rank for MeSH term m. Variable ao_{ik} denotes the author a_i's rank for an article k which is indexed by MeSH term m. M is the set of unique MeSH terms assigned to an author's publications ($m \in M$). ta_k is the total number of authors for the article. For example, in Equation 4, Author1 is one of the authors who wrote two papers indexed by the Term *A*. He is the first of three authors and the 4th of 11 authors of these papers, yielding a value of 1.73 for o_{1A}. o_{1A} thus provides the weighted sum of Author1's rank on Term A. In the comparison of two authors (a_i and a_j), we multiply author's term TF/IDF values and authorship rank and then compute Cosine similarity (Equation 5).

$$Cosine(a_i, a_j) = \frac{\sum_{n=1}^{|V|} w_{in}o_{in} \times w_{jn}o_{jn}}{\sqrt{\sum_{n=1}^{|V|}(w_{in}o_{in})^2 \times \sum_{n=1}^{|V|}(w_{jn}o_{jn})^2}} \qquad \textbf{eq. 5}$$

These content-based approaches (MeSH term-based and MeSH term and author rank-based recommendations) could be naïve. As of 2010, MeSH consists of 25,588 distinct terms in an eleven-level hierarchical structure [16]. This hierarchical structure may make it difficult to determine the true semantic similarity of publications. Two closely related papers might be indexed with sibling terms, but would not be considered similar using the first two algorithms we have described. To avoid this problem, we considered exploding source MeSH terms and only using their children at the leaf level as expertise terms. However, according to previous results [9], recommendations based on exploded MeSH terms perform worse than those using original MeSH terms in terms of precision and recall, and hence we did not use the exploded MeSH-based approach in this paper.

Social Network-based Recommendations

For social network-based recommendations, we used both local and global views of authors' social network. For local social network-based recommendations, we tried to find the nearest neighbor with a social context similar to that of our target author. We called this social network-based recommendation 'local' because we limited the search to a author's social connections of 1 hop distance, i.e. colleagues of an author's colleagues. To find the nearest neighbors of our target authors, we employed traditional CF-based recommendation and transformed authors' publication activities to rating actions. Traditional CF-based recommendation selects the nearest neighbors based on authors' rating similarities. When users rate a set of items in a manner similar to a target user, recommendation systems define these nearest neighbors as *likely-minded peers* and recommend items that are favored by the peers but not yet discovered by the target user. Analogously, our recommendation recommends persons who have worked with a target author's colleagues but not with the target author himself. Therefore, we considered each author as an item to be rated and writing a paper together as a rating activity in the sense of traditional CF-based recommendations. The frequency of co-authored papers is a rating value. In order to eliminate the effect of variations in co-authorship frequencies among authors, we normalized the frequencies as shown in Equation 6.

$$adj\_freq(a_i \rightarrow ca_j) = \frac{\text{number of coauthored articles by } a_i \text{ and } ca_j}{\text{Total number of } a_i\text{'s articles}} \qquad \textbf{eq. 6}$$

adj_freq is the adjusted frequency of co-authorship between a_i and ca_j. a_j is one of a_i's co-authors ($ca_j \in CA_i$). We calculated the adjusted co-authorship frequencies for all pairs of authors. The example in Figure 2 shows that Author A wrote 4 of his 20 papers with Author C and 10 with Author D, respectively. Author B co-authored 2 papers with Author C and Author F co-authored 5 papers with Author D. Since both Authors B and F have a history of working with Author A's co-authors, they could be considered nearest neighbors of Author A. According to the cosine similarity, Author F is the most similar to our target author, Author A. Therefore, among author F's co-authors, when there is anyone who worked with author F relatively often and didn't work with author A, we recommended the person to author A. More specifically, in order to find the nearest neighbors of local social network-based recommendation, we computed the cosine similarity of adjunct frequencies and predicted Top 5 peers for each target author. Lastly, we suggest the most recommendable people among the Top 5 peers' co-authors according to similarities [7, pp. 19 ~ 20].

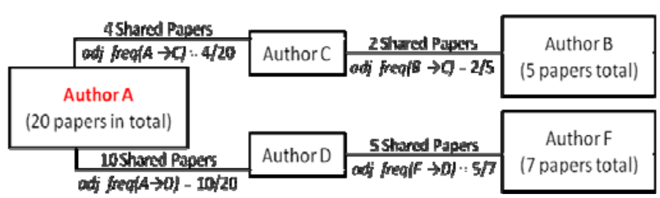

Figure 2. Example for Determining the Nearest Neighbor

Our second social network-based approach attempted to identify authors with strong social connections in the context of their social networks at large. In this approach, we did not limit the scope of the social network to just one hop but navigated the social network regardless of the number of hops. This approach takes the strength of social connections into account. In our paper, we define the strength of social connections through the relative frequency with which individuals have co-authored papers. The strength of a social connection can be propagated throughout a social network, potentially resulting in stronger social connections across multiple hops than for a direct connection [4, 17].

Even though our data set does not represent authors' social networks explicitly, we inferred them from co-author relationships. When researchers co-author a paper they tend to work on the same project, belong to the same institution or have some other kind of relationship, e.g. advisor-advisee. We considered individuals as having a direct social connection if they co-authored at least one paper. Once we had built authors' social networks in this way, we computed the strength of their social connection through two social properties.

The first social property was the '*shortest path*' between two authors. Any sequence of actors (i.e. authors) connected by links (i.e. authorship) in a network is a '*path*', and the number of links traversed along the path is the path distance. Several paths may exist for any given pair of actors. In social network analysis, the shortest path is the quickest way to reach another person. The shortest path is an important property in computing the efficiency of communication and showing the density of networks [18, pp. 138 ~ 139]. To compute the shortest path, we took into account not only the absence/presence of a social connection but also the frequency of interactions (i.e. how many papers two authors wrote together). For example, in Figure 2 Author A wrote a paper with Authors C and D, and therefore, Author A is directly connected to both Authors C and D. However, Author A wrote four papers with Author C and 10 papers with Author D. Thus, Author A seems to be socially closer to Author D than Author C. Therefore, in our computation of shortest path, the more frequently two authors worked together the smaller the path between them. Specifically, as the first step in the shortest path calculation, we compute the distance of a pair of authors (a_i and a_j) who are directly connected to each other as shown in Equation 7. S is a whole social network structure consisting of direct connections, and a pair of authors a_i and author a_j is a member of the social networks, $[a_i, a_j] \in S$.

$$Direct\ Distance(a_i \rightarrow a_j) = \frac{\bar{F}}{freq(a_i, a_j)} \qquad \textbf{eq. 7}$$

Variable $freq(a_i, a_j)$ represents the frequency of co-authorships between author a_i and a_j. The direct distance is then normalized on the center of $\bar{F}$, which is the mean value of frequencies (f_s) of all direct social networks ($s \in S$) [19]. In order to compute the shortest distance of all authors, we applied the Dijkstra algorithm to the direct distance values. The Dijkstra algorithm finds the shortest paths in social networks having non-negative weights of the links by summing the cost of connections [3] [18, pp. 329 ~ 333].

Our second social property measures structural equivalence. We determined the proportion of neighbors that two authors share using the Jaccard similarity (Eq. 8).

$$\text{Jaccard}\,(u_i, u_j) = \frac{\text{Neighbors of } u_i \cap \text{Neighbors of } u_j}{\text{Neighbors of } u_i \cup \text{Neighbors of } u_j} \qquad \textbf{eq. 8}$$

Once we computed the shortest path and structural equivalence for each pair of authors, we added the value of shortest path and Jaccard similarity after min-max normalization [23].

Hybrid Recommendations

Lastly, in our hybrid recommendations, we combined content-based and global social network-based approaches. Specifically, we used a simple version of the *mixed* hybridization strategy. Figure 3, Figure 4 and Equation 9 depict the design in detail.

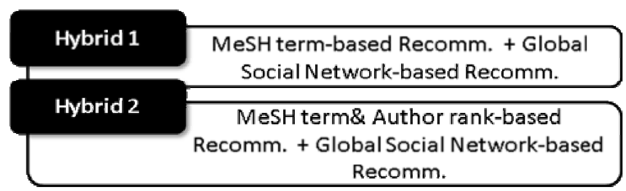

Figure 3. Design of Hybrid Recommendations

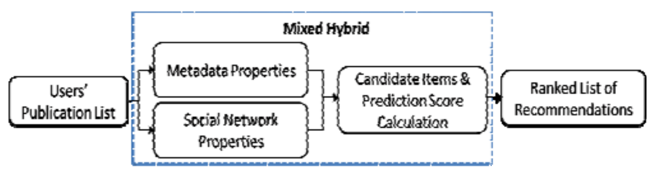

Figure 4. Mixed Hybrid Strategy

$$Hybrid_{a_i} = MP_{a_i} + SP_{a_i} \qquad \textbf{eq. 9}$$

For each of the mixed hybrid recommendations, we combine the score of a content-based recommendation for author a_i, MP_{a_i}, with the score of a global social network-based recommendation, SP_{a_i}. We chose to include the mixed hybrid approach because previous recommendation research has found that exploiting multiple types of information can enhance the quality of recommendations [3].

---

[3] We implemented our Dijkstra algorithm by referring the following Web page. http://www.vogella.de/articles/JavaAlgorithmsDijkstra/article.html#dijkstra (Accessed October, 2010)

Data Set for Experimental Evaluation

We evaluated the quality of our approaches by comparing actual co-author relationships (gold standard) with the predictions generated by our recommendation algorithm. To do so, we constructed a data set of biomedical researchers. We randomly chose 200 researchers in the University of Pittsburgh's Faculty Research Interests Project System [6]. Based on these initial seed authors, we expanded our sample by including all co-authors and the co-authors' co-authors through breadth-first search. This snowball sampling is known to include less isolated pairs [1]. Collexis Holdings, Inc., Columbia, SC, provided a data set containing fully disambiguated authors and their co-author relationships. These relationships were unambiguously defined using an approach similar to that described by Torvik et al. [26]. The number of authors in the data set is 17,525, the number of papers 22,542. The data set included the papers' full citation and co-author relationship information. We imported MeSH terms for each publication directly from PubMed. Table 1 describes the data set in descriptive terms.

Error! Reference source not found. shows paper frequencies across authors. More than half of the authors, 9,650 or 55.1%, have published more than one paper. **Error! Reference source not found.** shows the distribution of the number of co-authors across papers. Most papers, 18,782 or 83.3%, have more than one author. The mean number of MeSH terms per paper is 22.9 (σ = 10.9). These statistics indicate that authors may sufficiently overlap within the data set to calculate author similarity.

Table 1. Experimental Data Set

No. of authors	17,525
No. of publications	22,542
Avg. no. of papers per author	5.4
No. of papers that at least one MeSH term was assigned to	21,806
Avg. no. of MeSH terms per paper	22.9

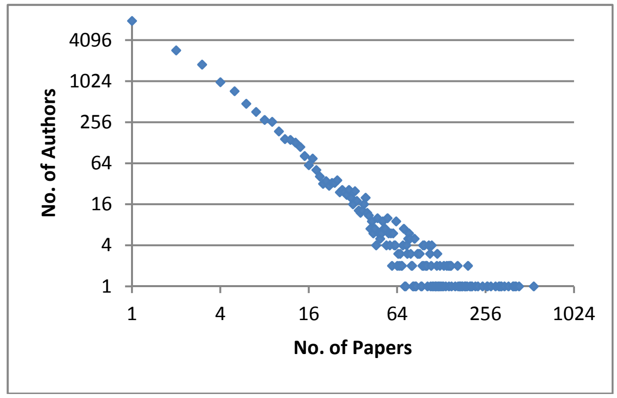

Figure 5. Number of papers per author

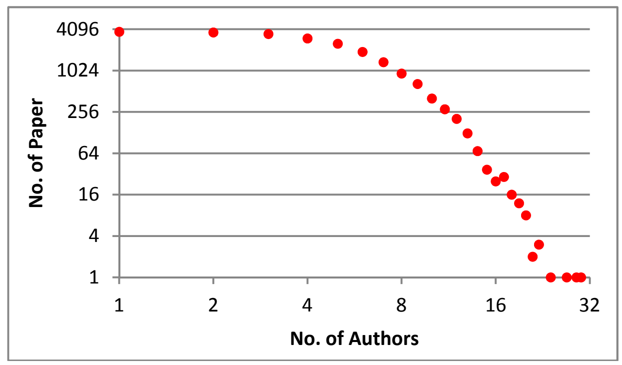

Figure 6. Number of authors per paper

Formal Evaluation

We evaluated the performance of our recommendation algorithms by comparing predicted collaborations to actual collaborations. To do so, we randomly chose 700 authors and divided their publication lists into two sets, articles published before 2007 and published in 2007 or later. We chose the year 2007 as our cut-off point because it separates our data set into a training set of 80% of all publications (articles before 2007) and a test set of the remaining 20%, a common approach in many information retrieval studies. The method for evaluating our algorithms is very practical, since recommendations for new collaborators are based on authors' past interactions. Our evaluation question was whether and to what degree each algorithm could correctly predict co-author relationships in the test set that did not exist in the training set. Among the 700 authors, we excluded 61 who did not publish any papers from 2007 onward, reducing the set of authors to 639. For each of 639 authors, we used our algorithms to predict the Top N individuals who had not co-authored with them before 2007. Then, we checked how many of the Top N authors actually co-authored at least one paper with each author subsequently. We used four evaluation categories – Top 20, 10, 5 and 2 recommendations – because we wanted to evaluate our algorithms with respect to the ranking of recommended items. Authors generally expect higher-ranked results to be more useful than lower-ranked ones. Thus, it was important for us to assess to what degree our algorithms were able to generate top-ranked recommendations.

We assessed final results using two evaluation metrics common in information retrieval: F1-measure and hit rate. The F1 measure is the harmonic value of precision and recall. Precision measures how precise a prediction is and recall measures how complete it is. More specifically, precision at point N (precision@N) is the ratio of the number of correctly predicted items in the Top-N list to N (Eq. 10). Recall at point N (recall@N) is the ratio of the number of correctly predicted items in the Top-N list to the total number of relevant items (Eq. 11). The F1 measure averages the precision and recall values with a bias toward the weaker value (Eq. 12) [22, 29]. We also calculated hit rate, the proportion of authors for whom at least one correct recommendation was generated (Eq. 13).

Last, we calculated the average number of shared papers between target authors and the Top N suggested candidates. We considered a higher number of shared papers as indicative of a closer working relationship, and thus a better recommendation. We compared mean differences of average shared papers using a statistical test.

$$\text{precision@N} = \frac{\text{No. of correct prediction}}{\text{N of top N set}} = \frac{\text{test} \cap \text{top N}}{\text{N}} \quad \textbf{eq. 10}$$

$$\text{recall@N} = \frac{\text{No. of correct prediction}}{\text{size of test set}} = \frac{\text{test} \cap \text{top N}}{\text{test}} \quad \textbf{eq. 11}$$

$$\text{F1 measure} = \frac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \quad \textbf{eq. 12}$$

$$\text{Hit Rate} = \frac{\text{Number of test users whose recommendation is correct at least once}}{\text{Total number of test users}} \quad \textbf{eq. 13}$$

RESULTS

Table 2 shows the results of all Top N evaluation categories (N = 20, 10, 5, and 2) predicted by our algorithms. We checked how the evaluation criteria vary across the proposed recommendations and how many of the Top N authors actually did co-author a paper with each test author. We used a Friedman two-way ANOVA test to test the mean difference of the F1 measure. The difference was considered statistically significant at a p value of 0.01.

In the first Top 20 category, regarding the F1 measure, the Hybrid 2 approach combining authors' expertise information (MeSH terms and author rank) with their global social networks outperformed other approaches. It also produced correct recommendations for a fair number of authors regarding the hit rate. The local social network-based approach generated at least one correct prediction for the largest number of authors, but the recommendation quality as expressed by the F1 value was lower than for other approaches. In the Top 10 category the results were similar to the Top 20. The main difference was that the Hybrid 2 approach performed best in terms of both recommendation quality evaluated by the F1 measure and author coverage evaluated by the hit rate. In both the Top 20 and Top 10 categories, the performance of the Hybrid 2 approach was significantly better than for content-based and local social network-based approaches. However, it was not significantly different from that of global social network-based and Hybrid 1 recommendations.

In the Top 5 and Top 2 categories, the global social network-based recommendation performed the best and the Hybrid 2 recommendation second best. When we compared the mean difference, the performance of content-based and local social network-based recommendations was significantly lower. However, there was no significant difference among global social network, Hybrid 1 and Hybrid 2 approaches in terms of the F1 measure.

These results show that considering both authors' expertise and social connections is important when recommending collaborators. In addition, it appears to be more effective to consider social closeness and structure in wider social networks than to choose candidates from within small networks solely based on the frequency of collaborations.

The last column of Table 2 shows the number of papers written together by co-authors after 2007 in the Top N recommendation list predicted by the six algorithms. For all evaluation criteria (from Top 20 to Top 2), global social network-based recommendations consistently produced the best results. The mean numbers were significantly higher than other recommendations. This result shows that social properties may be more important than authors' expertise in the selection of more productive collaborators.

CONCLUSION AND DISCUSSION

This study explored the importance of the social context in expert recommendation. As opposed to expertise location systems, expert recommender systems to date have mainly focused on either authors' expertise or social networks, rarely on both. Therefore, we examined whether authors tend to work with socially connected people with reasonably compatible expertise or with highly qualified experts outside of their social network.

As our results show, in general, social network-based and hybrid recommendations outperformed the other approaches with respect to the evaluation criteria. These algorithms produced higher F1 scores, and therefore performed more precisely and completely than others. In addition, they predicted more productive collaborations. The general conclusion from these findings is that recommending collaborators solely on the basis of matching expertise may not be as useful as integrating expertise with social network information, or using social network information alone.

When comparing algorithm performance across Top N categories, the global social approaches, in general, outperformed others in the Top 2 and Top 5 categories. This finding is important since higher-ranked, appropriate recommendations have a higher importance for authors than lower-ranked ones, since ranking in search results can influence author perception of relevance. Our interpretation of these results is that when choosing a collaborator, scientists may find the best matches within their social networks rather than based on expertise.

In future work, we plan to refine our algorithms by adding more diverse social perspectives, such as social roles and temporal centrality of authors. We also will examine new algorithms to generate novel and serendipitous recommendations that suggest dissimilar but complementary experts to target authors. In addition, we will work to find ways to reduce the size of the vector space using latent semantic indexing or other clustering methods. Lastly, the results we obtained in this study need to be verified and generalized with larger data sets.

ACKNOWLEDGMENTS

We thank Collexis Holdings, Inc., for providing the data set, and gratefully acknowledge the support of the National Center for Research Resources for this project (grant number UL1 RR024153).

REFERENCES

[1] Ahn, Y.-Y., et al., *Analysis of topological characteristics of huge online social networking services*, in *Proceedings of the 16th international conference on World Wide Web*. 2007, ACM: Banff, Alberta, Canada. p. 835-844.

[2] Bedrick, S. and D. Sittig, *A scientific collaboration tool built on the facebook platform*, in *Proceedings of AMIA 2008 Annual Symposium*. 2008: Washington, DC, USA. p. 41-45.

[3] Burke, R., *Hybrid web recommender systems*, in *The adaptive web: methods and strategies of web personalization*. 2007, Springer-Verlag. p. 377-408.

[4] Cialdini, R.B. and N.J. Goldstein, *Social influence: compliance and conformity.* Annu Rev Psychol, 2004. **55**: p. 591-621.

[5] Coletti, M.H. and H.L. Bleich, *Medical subject headings used to search the biomedical literature.* J Am Med Inform Assoc, 2001. **8**(4): p. 317-23.

[6] Friedman, P., et al., *Development of a MeSH-based index of faculty research interests.*, in *Proceedings of AMIA 2000 Annual Symposium*. 2000.

[7] Jannach, D., et al., *Recommender Systems: An Introduction*. 2010: Cambridge University Press.

[8] Kautz, H., B. Selman, and M. Shah, *Referral Web: combining social networks and collaborative filtering*. Commun. ACM, 1997. **40**(3): p. 63-65.

[9] Lee, D.H. and T. Schleyer, *MeSH term explosion and author rank improve expert recommendations.* AMIA Annu Symp Proc, 2010. **2010**: p. 412-6.

[10] Li, J., et al., *EOS: expertise oriented search using social networks*, in *Proceedings of the 16th international conference on World Wide Web*. 2007, ACM: Banff, Alberta, Canada. p. 1271-1272.

[11] Lin, C.-Y., et al. SmallBlue: Social Network Analysis for Expertise Search and Collective Intelligence. in *IEEE International Conference on Data Engineering*. 2009.

[12] Liu, B., *Informational Retrieval and Web Search*, in *Web Data Mining: Exploring Hyperlinks, Contents and Usage Data*, B. Liu, Editor. 2007, Springer: New York. p. pp. 183 - 236.

[13] Liu, B., *Web Data Mining: Exploring Hyperlinks, Contents, and Usage Data (Data-Centric Systems and Applications)*. 2009: Springer.

[14] McDonald, D.W., *Recommending collaboration with social networks: a comparative evaluation*, in *Proceedings of the SIGCHI conference on Human factors in computing systems*. 2003, ACM: Ft. Lauderdale, Florida, USA. p. 593-600.

[15] McDonald, D.W. and M.S. Ackerman, *Expertise recommender: a flexible recommendation system and architecture*, in *Proceedings of the 2000 ACM conference on Computer supported cooperative work*. 2000, ACM: Philadelphia, Pennsylvania, United States. p. 231-240.

[16] Medicine, U.N.L.o. *Fact Sheet: Medical Subject Headings (MeSH)*. May 24, 2010]; Available from: http://www.nlm.nih.gov/pubs/factsheets/mesh.html.

[17] Mohsen, J. Modeling and Comparing the Influence of Neighbors on the Behavior of Users in Social and Similarity Networks. 2010.

[18] Newman, M., *Networks: An Introduction*. 2010, USA: Oxford University Press.

[19] Opsahl, T. *Average shortest distance in weighted networks*. [cited 2010; Available from: http://toreopsahl.com/2009/01/09/average-shortest-distance-in-weighted-networks/.

[20] Pavlov, M. and R. Ichise, *Finding Experts by Link Prediction in Co-authorship Networks*. Proceedings of the Workshop on Finding Experts on the Web with Semantics (FEWS2007) at ISWC/ASWC2007, ed. V.Z. Anna, et al. 2007, Busan, South Korea.

[21] Pavlov, M. and R. Ichise, *Finding Experts by Link Prediction in Co-authorship Networks*, in *Proceedings of the Workshop on Finding Experts on the Web with Semantics (FEWS2007) at ISWC/ASWC2007, Busan, South Korea*, V.Z. Anna, et al., Editors. 2007.

[22] Sarwar, B., et al. Application of dimensionality reduction in recommender systems a case study. in *In ACM WebKDD Workshop*. 2000.

[23] Shalabi, L.A., Z. Shaaban, and B. Kasasbeh, *Data Mining: A Preprocessing Engine.* Journal of Computer Science. **2**(9): p. 735-739.

[24] Terveen, L. and D.W. McDonald, *Social matching: A framework and research agenda.* ACM Trans. Comput.-Hum. Interact., 2005. **12**(3): p. 401-434.

[25] TK, S., et al., *Conceptualizing and advancing research networking systems.* ACM transactions on Computer Human Interaction, In Press.

[26] Torvik, V.I., et al., *A probabilistic similarity metric for Medline records: A model for author name disambiguation: Research Articles.* J. Am. Soc. Inf. Sci. Technol., 2005. **56**(2): p. 140-158.

[27] Wellman, B., *Computer Networks As Social Networks.* Science, 2001. **293**(5537): p. 2031-2034.

[28] Yang, S.J.H. and I.Y.L. Chen, *A social network-based system for supporting interactive collaboration in knowledge sharing over peer-to-peer network.* Int. J. Hum.-Comput. Stud., 2008. **66**(1): p. 36-50.

[29] Yuan, Q., et al. Augmenting Collaborative Recommender by Fusing Explicit Social Relationships. in *Workshop on Recommender Systems and the Social Web, Recsys 2009*. 2009. New York, NY, USA.

Table 2. Results of Top N Recommendations

		F1-measure	Hit-Rate	Avg. No. of shared papers
Top 20 Results	**MeSH-term**	8.06%	21.25%	0.46
	MeSH-term + Author Rank	8.93%	24.69%	0.53
	Local Social Network	9.30%	**34.69%**	0.85
	Global Social Network	9.86%	31.72%	**0.94**
	Hybrid 1 (MeSH-term + Social Network)	9.94%	30.63%	0.76
	Hybrid 2 (MeSH-term + Author Rank + Social Network)	**10.58**%	31.72%	0.82
Top 10 Results	**MeSH-term**	10.71%	14.38%	0.29
	MeSH-term + Author Rank	11.95%	17.66%	0.36
	Local Social Network	11.58%	23.13%	0.50
	Global Social Network	12.47%	26.25%	**0.64**
	Hybrid 1 (MeSH-term + Social Network)	12.69%	26.09%	0.59
	Hybrid 2 (MeSH-term + Author Rank + Social Network)	**12.83**%	**26.88%**	0.61
Top 5 Results	**MeSH-term**	11.89%	9.84%	0.19
	MeSH-term + Author Rank	13.75%	11.72%	0.21
	Local Social Network	13.99%	17.19%	0.33
	Global Social Network	**14.86**%	19.06%	**0.40**
	Hybrid 1 (MeSH-term + Social Network)	**14.86**%	19.06%	0.36
	Hybrid 2 (MeSH-term + Author Rank + Social Network)	14.79%	**19.22%**	0.36
Top 2 Results	**MeSH-term**	11.41%	5%	0.08
	MeSH-term + Author Rank	13.86%	7.03%	0.11
	Local Social Network	16.16%	9.06%	0.16
	Global Social Network	**16.61**%	11.56%	**0.21**
	Hybrid 1 (MeSH-term + Social Network)	16.24%	11.56%	0.20
	Hybrid 2 (MeSH-term + Author Rank + Social Network)	16.42%	**11.72%**	0.20