Towards Identifying Teacher TopicInterests and Expertise within an Online Social Networking Site

SEN CAI, SIDDHARTH JAIN, YU-HAN CHANG, JIHIE KIM
University of Southern California, USA

Online social networking tools developed for educators promise better mentorship, professional development and resource sharing among teachers. This paper presents a model of teacher mentoring behaviors within an existing middle school teacher social networking site. We analyze how teachers help others in discussion forums and learn a model that describes their behavior. Our initial results indicate that this mentoring model is relatively consistent with profile answers, suggesting that we can build predictive models of teacher types based on their observed behaviors. We expect that our analysis can be useful in characterizing roles of many teachers who do not provide their profile information.

Key Words: user model, mentor modeling, teacher social network, subject classification, topic distribution

1. INTRODUCTION

Student achievement is boostedwhen teachers are able to share ideas, plan collaboratively, critique and coach each other (O'Hair, McLaughlin, &Reitzug, 2000; Goddard, et al., 2007; McClure 2008); yet, teachers frequently find themselves learning and working in isolation. Fostering collaboration among novice and veteran teachers can improve teacher retention and teacher satisfaction (Kardos and Johnson 2007).One of the key strengths of online communities is their capability to connect the right people with each other. For example, MSP2 (Middle School Portal, http://www.msteacher2.org) provides a good opportunity for teachers to collaborate within and across different school districts. While the site includes forums, blogs, and wikis, it does not pro-actively connect people with each other or with customized resources. Partly as a result, participation rates are fairly low.As a first step towards providing services to recommend potential mentors, we model information seeking and information providing behaviors within the MSP2 social networking site. We categorize the teacher's tasks in which the user appears to possess expertise or desires help, and identify the teacher's subject area. Our methods only rely on each user's forum and blog activity to build these models. To evaluate the results, we compare our predictions about the teachers' areas of expertise to the available data from their responses onthe site registration form.

2. MODELING FORUM AND BLOG INTERACTIONS

We first model the textual exchanges between teachers on the social networking site. Our prior work showed that online forums are often used in exchanging questions and answers about teacher's tasks (Kim et al., 2011), and we provided a summary of the characteristics of the forum posts that led to lengthier, more popular exchanges. Here we focus on predicting a teacher's interests and areas of expertise, based on her forum and blog activity. By predicting interest and expertise, we seek to identify potential mentors who are willing to share their knowledge with other teachers.

We focus on two types of expertise: 1) *Teacher's Tasks,* and 2) *Subject Area*. Teacher's Tasks are teaching activities or issues, such as 'in-class demonstrations', 'integrating technology', and'dealing with student misconceptions'.Teachers vary in their expertise within these task areas. We are also interested in each teacher's subject area. In this paper, we assume a simple topic model, where each forum or blog post islabeled 'math' and/or ‘science’. In addition, we identify the *Speech Act*type of each post as either being a question, answer, or acknowledgement (Kim and Ravi, 2007).

Our corpus containsnearly 2 years of MSP2 forum and blog activity. We manually annotate each forum and blog post with labels for Teacher’s Task, Subject Area, and Speech Act.The tags and scores are given in Kim et al. (2011). This hand-labeled forum and blog activity provides the test data for the methods described in the next section that predict an individual teacher’s subject area and expertise with particular teacher’s tasks.

3. CLASSIFYING TEACHER DISCUSSION TOPICS

Our goal is to create a tool that can match mentors with help-seekers. We need the annotations described in the previous section to assist the system in determining the teachers who possess the needed expertise in particular areas. This is particularly important in fielded social networking sites, such as MSP2, where participants often do not provide detailed profile information when they register to use the site; out of nearly 1200 registered members, more than 60% of them have not yet answered which teacher’s tasks they can help or need help with. Therefore, in order to label such teachers as potential mentors or as help seekers, we need an approach for inferring expertise directly from website participation to fill in the missing information in the user profiles.

We first address the automatic classification of forum and blog posts into the ‘math’ and ‘science’ categories. This classification will be used in Section 4 to determine a teacher’s Subject Area, so that our tool will be able to match teacher mentors and help seekers together based on similar subject area.We use data from NSDL (National Science Digital Library) to build a *word list model* of each Subject Area. An individual post is compared to each word list, and assigned a subject area tag based on whether the post contains a number ofmatching words.

The word list models are created by fetching NSDL metadata for the top resources returned via a search of math and science keywords, and filtering this metadata for relevant words. We use search terms gathered from: 1) MSP Mathematics Pathway topics hierarchy, 2) MSP Science Pathway topics hierarchy, 3) Topics covered by the Middle School Mathematics test by ETS PRAXIS Series, and 4) Topics covered by the Middle School Science test by ETS PRAXIS Series. Using these keywords as search terms, we then issue a query to NSDL and retrieve the meta-data returned. The top 10 resources from the results are selected, and we extract the ‘description’ metadata from the resources. After removing the stop words (Manning, Raghavan, & Schutze, 2008) from the description content, the remaining words are added to each Subject Area’s word list model. Once all the words are added in the lists, the unrelated words are manually removed from the lists to further rectify them.

For each post and comment in forum and blog, the content is treated as a bag of words. Each bag of wordsis compared with the word lists for Math and Science. The number of matches determines the score for that post or comment. If the score is more than a certain threshold for a subject area, it is classified within that subject area. Thus the same post or comment can be classified within multiple topics. Since the length of each post varies significantly, the threshold is a ratio of the number of matches relative to the overall length of the post. Figures 3.1 and 3.2 show the ROC curves for Math and Science

Classifications with different threshold values. We use the manually annotated blog data as the validation dataset in generating the ROC curves.

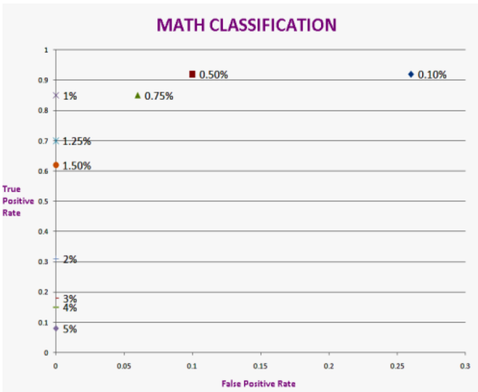

Figure 3.1: ROC Curve for Math Classification

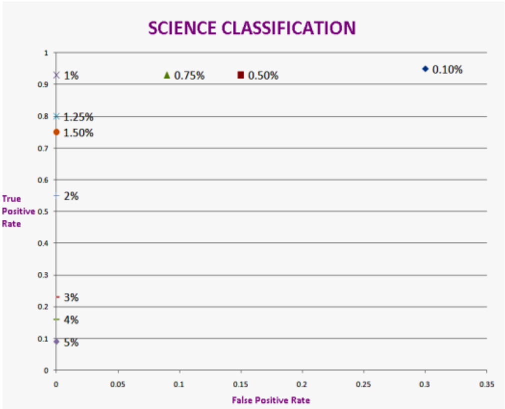

Figure 3.2: ROC Curve for Science Classification

From the graphs we can see that the threshold corresponding to 1% length of the content givesa higher true positive rate, with fewer false positives, compared to others. So this threshold value is used in classifying all the posts and comments, and in subsequent sections to predict teacher expertise and subject area.We have not yet implemented automatic tagging of each post with Teacher's Task labels in addition to the Subject Area labels, but we would expect a similar method to work.

4. IDENTIFICATION OF MENTORS

We next need to identify the teachers who are potential mentors versus the teachers who are seeking help. Due to the lack of profile responses, we use forum participation to infer teacher roles. We then validate our classification results by comparing them to profile answers for those who provided profile information. We will be mainly concerned with precision, which is calculated as follows: If a person was classified as a mentor of a topic and also claims in profile answer, this would count as a correct classification or true positive (TP). If a person was classified as a mentor, but does not claim in profile answer, it will be counted as an incorrect classification or a false positive (FP). As usual, precisionis defined as*Precision = TP / (TP +FP).*Recall is less important, partly because of the sparse profile data. The sparseness makes it difficult to determine whether a person should be in the negative class, since they may have simply been too lazy to respond to the registration profile questions. Furthermore, we are mainly concerned with identifying potential mentors, rather than labeling non-mentors. Thus ourgoal is to optimize precision values, and to obtain a relatively large set of teacher mentors.

Mentor Subject Area. Once the posts and comments in forum and blog are automatically tagged with Math and Science categories, individual members must then be classified based on her contributions that are classified in any of the topics. For each member, the number of posts contributed by him/her for each topic is calculated. If the member has contributed to only one of the topics, then we classify him/her in that topic. If the member has contributed to both topics, then the classification is done in the following manner. First we classify him/her in the topic where he/she contributed to more. For the other topic, if his/her contribution is more than the average among the ones who contributed to the topic, we consider that he/she is interested in the other topic as well. That is, we ignore relatively minor contributions. Once the members are classified for Math and Science topics, the classification is tested based on the profile answers.

	Math	Science
#members classified	37	49
#classified with no profile	12	35
# classified correctly	24	13
# classified incorrectly	1	1
Precision	0.96	0.93

Table 4.1: Classification precision for predicting teacher subject area.

Table 4.1 shows that the precision of the classification is high. Note that a significant number of teachers who are classified as being “Math” or “Science” mentors did not actually select either topic as their area of expertise when they completed their site registration profile. A random sampling of these teachers and manual examination of their contributions verified that our classifications were in fact correct, in the sense that their postings generally did fall in the topic area into which we classified the person. Thus we expect that for the members who have not answered their profile questions about what topic they are interested in, our classification can help fill this gap. That is, we can use the classification results in finding potential mentors who can help with the domain topic in combination with the teacher’s task topics.

Mentor Expertise in Teacher Tasks. We now turn our attention to determining a teacher’s mentorship potential in theTeacher’s Tasks topic areas. We define two criteria for mentorship potential:

- *Expertise*: A mentor must be knowledgeable in a topic area in order to help others. We estimate this factor using the number of knowledge sharing posts the user has written on this topic, and
- *Relative Interest*:A mentor must be willing and interested in sharing their knowledge about a topic. We estimate this factor using the relative number of a user’s posts on this topic, versus their posts on other topics.

More specifically, a mentor must have contributed enough answers to other users’ help-seeking posts to exceed a given threshold. To obtain the answers count, we filtered out self-answers (i.e. answers to one’s own questions), question messages, and simple acknowledgements from the set of all posts by the user. The remaining messages are considered answers. We define this number of answers to be X_{ij}, which isteacher i’s contribution to topic j.

The exact threshold used is a free parameter that we tuned based on our data. Table 4.2 shows four different thresholds we evaluated: In the first method, if individual’s answer count is greater than the average contribution of all members to a specific topic ($x_{ij}>\mu_j$), we classify her as a mentor for that topic. In the remaining three methods, we use the sum of mean and standard deviation, median, and upper quartile as the threshold.

Among those members who provided answers to their profile questions, the average number of topics that they claimed they can help with is 4.45. There are total 17 possible topics of interest. Our *baseline prediction precision*, assuming uniform random selection of topics, is thus 4.45/17 = 0.261. Compared to this baseline precision, we see that using the mean as the classification threshold produces reasonably good accuracy (0.593). It also produces a sufficiently large set of mentors, relative to the total number of active users on the MSP2 site, approximately 80-100 users out of the 1200 registered.

	Classification Threshold	Precision (baseline = 0.261)	# of distinct users	# of Mentorship records considered	# of correct attempts	# Mentors per topic
1	Mean	0.593	12	27	16	5.941
2	Mean+Std. Dev.	0.714	4	14	10	2.411
3	Median	0.434	20	46	20	5.411
4	Upper Quartile	0.485	20	35	17	2.471

Table 4.2: Comparison of identification method with different threshold

Thementorship coverage across teacher task topics is another important factor. In Figure 4.3, we plot the number of mentors in each topic obtained when we use Threshold 1 (Mean) and Threshold2 (Mean + Std. Dev.). Threshold 1 provides better coverage over all the topics with reasonable prediction accuracy that is well above baseline.

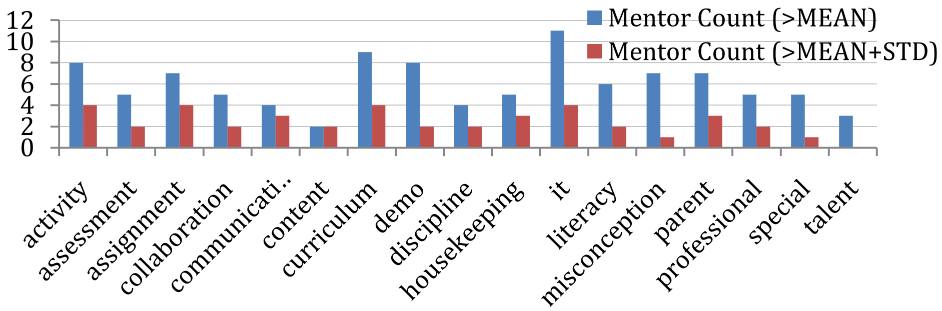

Figure 4.3: Classified mentorship distributions over teacher's task topics

Mentor Interest. Finally, even if a user is determined to possess expertise in a certain area, we also need to gauge their interest in sharing their knowledge on this topic with others. Some users contribute to all topics equally, while others focus on one area. We thusintroducea second factor: the proportion of posts in one topic, relative to total posts in all topics for a specific teacher. A new weighted contribution (WC) metric, which combines a user's topic-answer counts normalized over all users' answers for that topic, together with a user's topic-answer counts normalized over that user's answers over all topics, is defined: $WC_{ij} = \frac{x_{ij}}{\sum_{k\in I} x_{kj}} \frac{x_{ij}}{\sum_{k\in J} x_{ik}}$,wherex_{ij}isteacher i's number of answers for topic j, I is the set of all users, and J is the set of all topics.

	Threshold	Precision (baseline = 0.261)	# of distinct mentors who are specialized in: Only 1 Topic	Only 2 Topics	>= 3 Topics	Total # of distinct mentors
1	$x_{ij} > \mu_j$	0.593	23	8	6	37
2	$WC_{ij} > E[WC_j]$	0.414	40	12	8	60
3	$x_{ij} > \mu_j + \sigma_j$	0.714	6	2	6	14
4	**$WC_{ij} > E[WC_j] + SD[WC_j]$**	**0.636**	**30**	**6**	**0**	**36**

Table 4.5: Analysis of classification using weighted contribution

Using this weighted contribution, we are able to discount generally active teachers as being mentors in topics where they rarely post, relative to their high number of posts to the topics they are truly interested in. As shown in Table 4.5, the total number of distinct mentors shows that methods using weighted contribution (Method 2 and 4) produces a larger set of distinct mentors than using absolute contribution (Method 1 and 3). This will help improve the user interaction on the MSP2 forum, because there will be more distinct individuals involved. As mentioned earlier, the baseline precisionis 0.261. The precision of Method 2 is too low to be adopted. We thus use Method 4: it has a reasonably good precision(0.636), and it classifies a total of30 users as potential mentors who are specialized in one topic and 6 users who are specialized in 2 topics. By determining a user's true interests, we are much more likely to be able to direct help-seeking questions to users who are most likely to respond to those questions.

5. FUTURE WORK

Our future work will extend methods described in previous sections to automatically annotate forum and blog posts with tags corresponding to Teacher's Tasks, and thus replicate the results we presented on automatically determining a user's Subject Area. We are also extending our methods to incorporate search history, click patterns, and other information sources, in order to improve classification precision. Finally, these methods will be used to build a deployable tool within MSP2 that matches potential mentors and help-seekers, promote interactions, and assist teachers in finding the help and resources that they need.

ACKNOWLEDGEMENT

This material is based upon work supported by the National Science Foundation under Grant No. 1044427.

REFERENCES

Adams N. M., & Hand D. J., (1998). *Comparing classifiers when the misallocation costs are uncertain*.

Goddard, Y. L., Goddard, R. D., Tschannen-Moran, M. (2007). A theoretical and empirical investigation of teacher collaboration for school improvement and student achievement in public elementary schools. Teachers College Record, 109(4), 877-896.

Kardos, S. M., & Johnson, S. M. (2007). On Their Own and Presumed Expert: New Teachers' Experience with Their Colleagues . *Teachers College Record Volume 109 Number 9* , p. 2083-2106.

Kim, J., Chang, Y. Cai,S.,Dupar,S. (2011). *Towards Models of Mentoring Dialogue within a Teacher Social* Networking Site, Proceedings of AIED-2011.

Manning, C. D., Raghavan, P., & Schutze, H. (2008). Introduction to Information Retrieval.

McClure C.T., 2008, The Benefits of Teacher Collaboration Essentials on Education Data and Research Analysis, District Administration, 82-83.

McLaren, B., Scheuer, O., De Laat, M., Hever, R., De Groot, R. & Rose, C. 2007. Using Machine Learning Techniques to Analyze and Support Mediation of Student E-Discussions, Proceedings of AIED-2007.

O'Hair, M., McLaughlin,H., & Reitzug, U. (2000). *Foundations of democratic practice.*Harcourt Brace.

Ravi, S., Kim, J.: *Profiling Student Interactions in Threaded Discussions with Speech Act Classifiers*, Proceedings of AIED-2007.