2013 IEEE 13th International Conference on Data Mining Workshops

Quantifying and Recommending Expertise When New Skills Emerge

Dongping Fang, Kush R. Varshney, Jun Wang, Karthikeyan Natesan Ramamurthy, Aleksandra Mojsilović, and John H. Bauer
Business Analytics and Mathematical Sciences, IBM Thomas J. Watson Research Center
1101 Kitchawan Road, Yorktown Heights, NY, USA

***Abstract*—In the rapidly changing technological world of today, new technical areas emerge quickly, and new skills related to them garner high demand. In this paper, our goal is to recommend experts for new skills and skill topics. We propose multiple predictive models to utilize data from different enterprise sources: employee assessment data, free-text skill description data, and employee tags from corporate social media. These models include collaborative filtering, content-based, and novel hybrid recommendation approaches. We apply them in an empirical study of real-world corporate data, in which we compare and contrast the models to gain insight on the drivers of performance. The considered data is both structured and unstructured, messy, subjective, and incomplete. The central theme of the paper is to understand how to use data from different sources and what each data source contributes in the expertise management domain.**

***Keywords*-cold-start problem; enterprise social networks; expertise taxonomy; recommendation systems; workforce analytics**

I. Introduction

At no point in history has technological progress been so rapid as today, and at no point has the economy been so service-based. Some of the largest employers in the world are technology-oriented services companies. Employers and employees today have to keep pace with new technology or risk being left behind. In this environment, human capital is the most important asset for a company and it must be cultivated properly. Towards this end, many companies try to quantify and track the skills and expertise of their workers, especially knowledge workers, often through the use of ratings against expertise taxonomies.

Expertise taxonomies are centrally-maintained with dedicated individuals tasked with adding new skills and areas of expertise on a regular basis. An alternative approach for maintaining expertise records is by collaborative creation and management of tags to annotate experts and skills, which is referred to as a *folksonomy* [1]. Central taxonomies are easier to interpret, but are inflexible and are cumbersome to adapt to rapidly evolving skills. Folksonomies adapt organically, but are difficult to interpret since they are not well-structured. Moreover, it is not straightforward to assess employees using either of these two approaches because of subjectivity and variable understanding of skill definitions, subjectivity in level of expertise, and non-compliance, to name a few reasons [2].

In the rapidly changing technological world, new technology topics emerge quickly, and new skills related to them are in high demand. In particular, when new skills/topics emerge, a company would like to know the number and identity of employees in-house that may have the new skill already. However, there are several hurdles in this, ranging from identification of the emerging skill to accurate assessment of these skills by the employees. In this paper, we focus on the problem of assessing the expertise of employees on emerging skills that have not been incorporated into the company's expertise taxonomy.

When new technologies, new business practices, and new job roles are arising at a pace faster than a company can describe these novel aspects of expertise, or evaluate employees against them, the standard approaches for skill management and planning are not effective anymore. Fortunately, modern enterprises are embracing social and collaborative technologies and are beginning to produce social data: data that has not existed previously, but is extremely valuable for understanding the expertise of employees. This new social data can be integrated with traditional enterprise skill assessments, surveys and other expertise data for use in automatically predicting the assignment of new skills. Therefore, in this paper, we develop machine learning methodologies based on combined social and enterprise data that can help companies manage and plan their expertise and organizational knowledge more effectively.

In particular, we start with an existing, incomplete, employee-skill rating matrix and augment this representation with information from skill descriptions and social tags of employees, thereby drawing wisdom from social activity [3]. Recommendation for existing skills under similar settings is reported in [4]. However, the problem of recommending or predicting the employee skill level for new skills cannot be treated on par with the problem of recommending existing skills, since there is no assessment information available in the expertise database for new skills. Therefore, it becomes necessary to leverage key terms associated with the new skill, in order perform reasonable predictions. The problem of predicting expertise among employees on new skills shares mathematical similarities with the cold-start problem seen in the recommendation systems literature [5]. Our approach is predicated on evaluating the similarity between a new skill and the existing skills from the taxonomy, and

978-0-7695-5109-8/13 $31.00 © 2013 IEEE
DOI 10.1109/ICDMW.2013.33

then transferring employee ratings from the existing skills to the new skill.

The remainder of this paper is organized as follows. In Section II, we take a more detailed look into the new skill assessment problem and position it in the recommendation framework. In Section III, we formulate several machine learning and data mining approaches to the new skill recommendation problem, including proposing a novel hybrid approach. In Section IV, we conduct an empirical study of new skills recommendation for a real-world global company, comparing and contrasting results from the several formulations. Finally, in Section V, we provide additional discussions and conclusions.

II. New Skill Assessment as a Recommendation Problem

Traditional recommender systems are designed to recommend items to users. If we draw a correspondence between items and skills, and between users and employees, then it is possible to approach automatic expertise assessment through recommendation-based solutions. In general these systems can categorized as (a) collaborative filtering, where skill recommendations are based on past assessments by other employees, (b) content-based recommendation, where recommendations are based on skill descriptions or social media attributes/tags of users, and (c) hybrid methods, that combine (a) and (b) [6]. A mathematical framework that incorporates the dual side information for employee/skills as graphs and attempts to perform matrix completion using random walks has been proposed in [7]. In their basic form, such approaches are very effective in recommending existing skills to existing employees.

However, as discussed in the introduction, technologies and skills associated with them are evolving quickly, and thus methods suited for recommending existing skills may not apply. In such dynamic environments, it is critical to develop recommender systems specialized for new skills/topics. Similar situations occur in many other fields, for example when a new movie or a new television program is added to a movie database, or new products are introduced to the market. These can be seen as cold-start problems, because new items have no ratings from customers. Collaborative filtering recommender systems depend exclusively on the history of user ratings and hence they do not work for cold-start problems [8]. In particular, the well-known matrix factorization method, that assumes no additional side information, does not produce reliable results if there is not at least one rating per item and per user [9].

When recommending a new item, content-based recommendation, i.e., recommendation based on the description of the item and a profile of the user's interests is usually used [10]–[12]. However, in [5], the authors propose a method to combine the content information with collaborative filtering for the cold-start problem. A method for incorporating the user profile for the same problem has been proposed in [13]. A different solution has been proposed in [14], where the authors combine association rules and clustering techniques to address this issue. Other approaches for cold-start recommendations are described in [12] and [15]. In this paper, we propose multiple models to combine data from different sources such as existing employee assessments, free-text skill descriptions, and social media tags to address the cold-start problem in skill recommendation. We compare and contrast the relative merits of incorporating data from each of the sources in terms of the recommendation performance.

III. Methodology

Let $\{i\}_{i=1}^{M}$ index a set of employees, $\{j\}_{j=1}^{N}$ index a set of skills, and m_{ij} denote the expertise level of employee i for skill j. The employee-skill assessment matrix $\mathbf{M} \in \mathbb{R}^{M \times N}$ with elements m_{ij} indicates all of the expertise assessments. Given a partial observation of the employee-skill assessment matrix with known elements of $\mathbf{M}$ in the set $\mathcal{O}$, and additional employee and skill features, the goal here is to build a model to predict employees' expertise levels for new skills that are not in the existing set of skills. Using these predictions, we may identify a list of employees who are at the expert level on these new skills. We declare an employee to be an expert on a new skill if his or her predicted expertise level exceeds a threshold value. In the remainder of this section, we describe collaborative filtering, content-based, and hybrid recommendation approaches for new skill recommendation; the hybrid approach is novel.

A. Matrix Factorization for Skill Assessment

Matrix factorization methods have been proven to be very powerful collaborative filtering tools in recommendation systems. We describe both the basic version and a version with side information.

1) Basic Matrix Factorization: The very basic matrix factorization-based approach assumes that the target matrix can be decomposed into two matrices $\mathbf{W} \in \mathbb{R}^{L \times M}$ and $\mathbf{H} \in \mathbb{R}^{L \times N}$ as $\mathbf{M} = \mathbf{W}^\top \mathbf{H}$. Thus, predictions of unknown assessments are

$$\hat{m}_{ij} = \mathbf{w}_i^\top \mathbf{h}_j, \tag{1}$$

where $\mathbf{w}_i$ and $\mathbf{h}_j$, representing latent factors for employee i and skill j respectively, are the ith and jth columns of the matrices $\mathbf{W}$ and $\mathbf{H}$. Including a global bias, employee bias, and skill bias, the factorization may be represented as:

$$\hat{m}_{ij} = \mu + e_i + s_j + \mathbf{w}_i^\top \mathbf{h}_j \tag{2}$$

A basic formulation of matrix factorization can be written using the squared loss function as

$$\arg\min_{\mathbf{W},\mathbf{H}} \sum_{(i,j)\in\mathcal{O}} \|m_{ij} - \hat{m}_{ij}\|^2$$

. We also include additional regularization in the above minimization.

To predict a new skill, we set up a column of all missing values in the assessment matrix and then complete it using the matrix factorization method. We use this basic matrix factorization with bias terms (2) as our baseline model for comparison. We note that this model predicts the expertise level for all new skills as the sum of the global bias and the employee's bias without regard for the semantics of the new skills.

2) Incorporating Skill and Employee Features: In the enterprise settings we are working, we have access to extra information about skills and employees beyond just the employee-skill assessment matrix. For example, we have a paragraph describing each skill. However, the basic matrix factorization method has no mechanism to incorporate employee and skill features. Here, we discuss how additional skill information is incorporated into the predictive model. Employee information is incorporated in an analogous way. Also, both additional skill information and employee information can be incorporated simultaneously.

We extract keyword tags from the names of skills and their paragraph descriptions, and incorporate the tags as item features as follows [16]:

$$\hat{m}_{ij} = \mu + e_i + s_j + \sum_{k \in \mathcal{T}_j} v_{jk} t_k + \mathbf{w}_i^\top \left(\mathbf{h}_j + \sum_{k \in \mathcal{T}_j} v_{jk} \mathbf{h}_k^{(t)} \right), \tag{3}$$

where t_k is the coefficient of bias for skill tag k, $\mathcal{T}_j$ is the set of tags for skill j, $\mathbf{h}_k^{(t)}$ is the latent vector for skill tag k, and v_{jk} is the tag weight for kth tag of skill j. A similar formulation to (3) applies for employee tags if we take the weights to reflect employee tag information, index them by employee instead of skill, and add the tag latent vectors to the employee factors $\mathbf{w}_i$ instead of the skill factors $\mathbf{h}_j$.

B. Prediction Based on Skill Similarity

When new skills emerge, they have relationships and correlations with existing skills in the expertise taxonomy. In a content-based recommendation approach, we would like to take advantage of correlations to skills that have been previously assessed by employees. Hence a simple method to predict the assessment of a new skill is by using a weighted sum of known skill assessments. The weights are determined through similarity of the new skill with the existing skills. To be specific, the predicted assessment for employee i on new skill k is:

$$\hat{m}_{ik}^* = \frac{\sum_{j \in \mathcal{J}_{ik}} m_{ij} \operatorname{sim}(j,k)}{\sum_{j \in \mathcal{J}_{ik}} \operatorname{sim}(j,k)}, \tag{4}$$

where $\operatorname{sim}(j,k)$ is the similarity between skill j and skill k, $\mathcal{J}_{ik} = \{j : (i,j) \in \mathcal{O}, \operatorname{sim}(j,k) \geq \operatorname{sim}_{min}\}$, and $\operatorname{sim}_{min} \geq 0$ is a minimum similarity threshold. Similarity between skills can be defined in various ways; one particular choice that we use in the empirical study of Section IV is cosine similarity between weighted vectors of skill tags.

For this content-based approach to be effective, there should be at least some similar existing skills. Imagine that all $\{\operatorname{sim}(j,k)\}_{j \in \mathcal{J}_{ik}}$ are small; then the prediction $\hat{m}_{ik}^*$ will be unreliable because it is based on existing skills that are not very similar. The value $\max_{j \in \mathcal{J}_{ik}} \operatorname{sim}(j,k)$ is a measure of the estimate $\hat{m}_{ik}^*$'s reliability. This approach can be successfully used for the cold-start problem. However, it does not utilize interaction and correlation between employees and predefined skills shown in the data. Also prediction for an employee only uses the assessments of this employee, and does not exploit the assessments of other employees.

C. Hybrid Method

We now propose a hybrid method that combines matrix factorization described in Section III-A and skill similarity-based prediction described in Section III-B. The prediction is expressed as an extension of the matrix factorization with bias terms (2):

$$\hat{m}_{ij} = \mu + \hat{m}_{ij}^* + e_i + s_j + \mathbf{w}_i^\top \mathbf{h}_j, \tag{5}$$

where $\hat{m}_{ij}^*$ is the skill similarity-based prediction from (4). The hope here is to explicitly capture skill similarity in the model while also making use of the collaborative filtering effect.

We can also similarly extend the matrix factorization model with skill and/or employee tags (3) to explicitly include a skill similarity term. We can also extend the hybrid approach with an employee similarity term. To the best of our knowledge, the formulation for bringing in skill similarity and/or user similarity expressed in (5) does not exist in the literature.

Overall in this section, we have described a suite of collaborative filtering, content-based, and hybrid recommendation algorithms for use in predicting assessments of employees on new and emerging skills. The next section compares and contrasts these algorithms on real-world data.

IV. Empirical Study

In this section, we examine expertise assessment in a large multinational Fortune 500 corporation. We conduct our empirical study in two ways. First, since there is no ground truth for truly new skills, we withhold all assessments of some skills from the expertise assessment matrix and treat those withheld skills as new skills. In this way, we can calculate performance metrics for the various recommendation algorithms. Second, we predict assessments for truly new skill topics and compare the results to lists of experts elicited from leaders in the corporation.

A. Expertise Assessment Data

The employee assessment system of the corporation in our study allows employees to rate themselves on more than 16,000 pre-defined skills. They assess themselves on a scale from 1 to 5, with 5 being the highest level of expertise.

Table I
SUMMARY OF ASSESSMENTS IN THE POPULATION

Level	Number of Assessments	Unique Skills	Unique Employees
1	263,445	5,409	10,041
2	181,439	4,459	9,157
3	293,330	5,465	9,632
4	195,785	4,911	8,405
5	20,325	2,522	2,339
Total	954,324	8,529	10,738

Table II
NUMBER OF ASSESSMENTS PER EMPLOYEE OR SKILL

	Mean	Min	Max	Median
Per employee	89	1	546	72
Per skill	112	1	10,197	4

An employee with assessment level 4 or 5 on a skill is considered an expert on that skill. Only a subset of skills is available for an employee to assess depending on the employee's line of business and job role. It is voluntary to assess oneself in the system, so many assessments are missing, and many employees have not assessed any skills at all.

We select one department within the company as our population of study. Our data set consists of employees in that department with at least one assessment. There are 954,324 total assessments with mean level 2.5, involving a total of 10,738 employees and 8,529 skills. Therefore, only about 1% of the elements in the assessment matrix are non-missing. Table I and Table II provide a summary of the data. Table II shows that some skills are assessed by almost everyone yet some are only assessed once. There are 2,124 skills that only have one assessment, i.e., approximately 25%. These skills are practically new skills and hence are difficult to predict.

B. Additional Skill and Employee Information

In addition to the assessments, the corporation maintains a short paragraph description for each skill in the taxonomy giving details on what competencies the skill entails. We extract tags from the skills' names and descriptions, yielding 14,000 tags from all of the skills combined. We then calculate ℓ_2-normalized term frequency-inverse document frequency (TF-IDF) weights for all tags in each skill. We use these as skill features and as input to the cosine similarity measure.

Most modern enterprises are now producing social data that did not exist previously. This data can potentially be extremely valuable, but has not been fully explored for understanding employee expertise. For example, employees with similar technical background in this corporation tend to belong to the same online technical communities and frequently exchange ideas through microblogs and wikis. In particular, tags were extracted from each employee's internal social media and technical community activities. The raw employee tags were very noisy and were cleaned by removing stop words, garbage words, and junk prefixes and suffixes such as "(", but there remains opportunity for further cleaning. We note that the poor quality of some employee tags may undermine their usefulness. In total, 7,038 unique employee tags were extracted. On average there were 26 tags per person for employees with tags, but a good number did not have any tags.

C. Experimental Design

In our empirical study, we compare nine variations of predictive models with different pieces of information as described in Section III. These first four models are the basic matrix factorization approach with biases (2) which we denote MF, matrix factorization incorporating skill tag features (3) which we denote MFs, matrix factorization incorporating employee tag features analogously to (3) which we denote MFe, and matrix factorization incorporating both skill tag and employee tag features which we denote MFse. The fifth model is the content-based recommendation using skill similarity (4) which we denote SS; the minimum similarity sim_{min} is set to 0.05. The final four models are the first four matrix factorization models combined with skill similarity as in (5) which we denote MFSS, MFsSS, MFeSS, and MFseSS, respectively.

Predictive model learning for all matrix factorization methods is implemented via the iterative optimization scheme of [16]. We use 4-fold cross validation to identify the iteration at which the test error is smallest. We then use all data to train the final model, with the number of iterations from cross-validation. This trained model is then used for prediction.

In the first experiment, we randomly choose 10 existing skills with at least 100 assessments as the "new" skills and hold them out when training the predictive models. The ten "new" skills are shown in Table III. Some skills do not have any level 5 assessments. Then we predict these 10 "new" skills and compare the predictions with the true assessments. Finally, since our main goal is to find experts for these new skills, we threshold the predicted expertise level. If the prediction for an employee on a skill is above the threshold, we declare that employee to be an expert on that skill.

In the second experiment, we consider three truly new skill topics: Business Analytics and Optimization (BAO), Cloud, and Smarter Commerce. These are three fast growing areas that are not in the expertise assessment system's existing taxonomy. We asked experts in these areas for a list of keywords describing each new topic. Based on these lists of keywords, we extracted tags and calculated ℓ_2-normalized TF-IDF weights, enabling similarity calculation between existing skills and new skill topics.

For evaluation purposes, we elicited lists of experts on the three topics from leaders in the business. This elicited list of business-identified experts is far from complete, containing

Table III
NUMBER OF ASSESSMENTS FOR THE 10 "NEW" SKILLS

Skill	Expertise Level 1	2	3	4	5	Total
Advise on Oracle Appl Technl Foundation Conversion	101	42	32	15	0	190
Apply Knowledge Of Oracle	90	139	246	121	3	599
Apply Knowledge of IBM Offerings/Technology	16	33	55	23	2	129
Design GUI	44	124	248	136	5	557
Implement DB2 UDB for Linux	85	46	32	19	4	186
Implement IBM Enterprise Information Portal (EIP)	106	26	36	14	0	182
Implement WebSphere Appl Server on IBM System z	122	173	156	60	0	511
Perform Configuration Management	132	193	268	102	6	701
Perform DB2 Performance Analysis	36	22	34	11	2	105
Use Business Analysis Work Products	182	76	129	103	11	501
Total	914	874	1,236	604	33	3,661

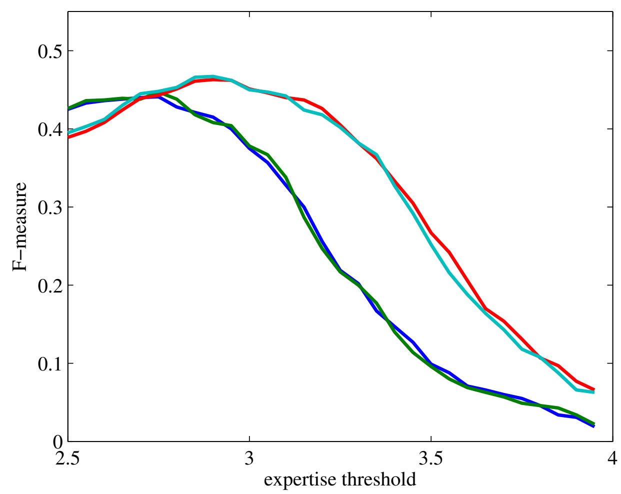


Figure 1. F-measure as a function of expertise threshold for MF (blue line), MFe (green line), MFSS (red line), and MFeSS (cyan line) to illustrate that employee tags alone have little impact on performance.

only 70 experts out of more than 10,000 employees: 42 in BAO, 22 in Cloud, and 6 in Smarter Commerce. Nevertheless, we compare thresholded predictions to these lists to calculate performance metrics. One thing worth mentioning is that BAO, Cloud, and Smarter Commerce are new topics that may consist of many different skills, not just a single skill, and this could make the recommendation process more difficult.

D. Results and Discussion

In this section, we present comparative performance results on the different recommendation models under the two experimental settings.

1) Experiment 1: First, we compare the F-measure (harmonic mean of precision and recall) of the expert prediction as a function of the expert threshold when employee tags are and are not included. As shown in Fig. 1, we see that MF and MFe have approximately the same performance, and that MFSS and MFeSS have approximately the same performance. Compared to the basic matrix factorization, adding skill similarity is clearly beneficial. More to the point of this figure, however, when employee tags are incorporated (alone without skill tags), there is no benefit. Therefore, we drop these two employee tag models (MFe and MFeSS) in the sequel.

In Fig. 2, we plot the precision, recall, false positive rate, and F-measure for the remaining seven models. The models involving skill similarity SS, MFSS, and MFsSS have higher recall and higher false positive rate than the other models. The models incorporating skill tags, MFs, MFse, MFsSS, and MFseSS have higher precision than the others. When both employee tags and skill tags are added to the model, the interaction between them makes a difference. However, the difference is beneficial when skill similarity is not added but

Table IV
ACCURACY AT THE BEST THRESHOLD

Model	Best threshold	F-measure	Recall	Precision	FP rate
MF	2.750	0.441	0.542	0.372	0.192
MFs	2.650	0.495	0.677	0.390	0.223
MFse	2.700	0.505	0.651	0.413	0.195
SS	3.050	0.476	0.516	0.442	0.138
MFSS	2.900	0.463	0.626	0.367	0.227
MFsSS	2.750	0.505	0.684	0.400	0.216
MFseSS	2.650	0.496	0.666	0.395	0.215

is detrimental when skill similarity is added. Incorporating user and skill tags as described in Section III-A introduces too many parameters to be estimated, and hence the results become ambiguous. Therefore we need a better model to directly incorporate the user and skill tag interactions.

The prediction bias and variance are shown in Figure 3. All models have a similar bias pattern: over-prediction for expertise levels 1 and 2 (below mean expertise level), and under-prediction for expertise levels 3 to 5 (above mean). For expertise level 4, the models under-predict by about 1 and for expertise level 5, the under-prediction is even bigger. Some models have larger bias whereas others have larger variance. SS has smaller bias and larger variance than MF, while MFs has slightly smaller bias and smaller variance.

The F-measure balances recall and precision and we use it to decide the expert threshold value. The threshold that maximizes F-measure is different for different models. Since there is a systematic under-prediction for expertise levels 4 and 5, we would like to choose a threshold value below 3.5. The various accuracy measures for the seven models at the F-measure maximizing threshold are summarized in Table IV. The two models with the best F-measure are MFse and MFsSS, where MFse has better precision and MFsSS

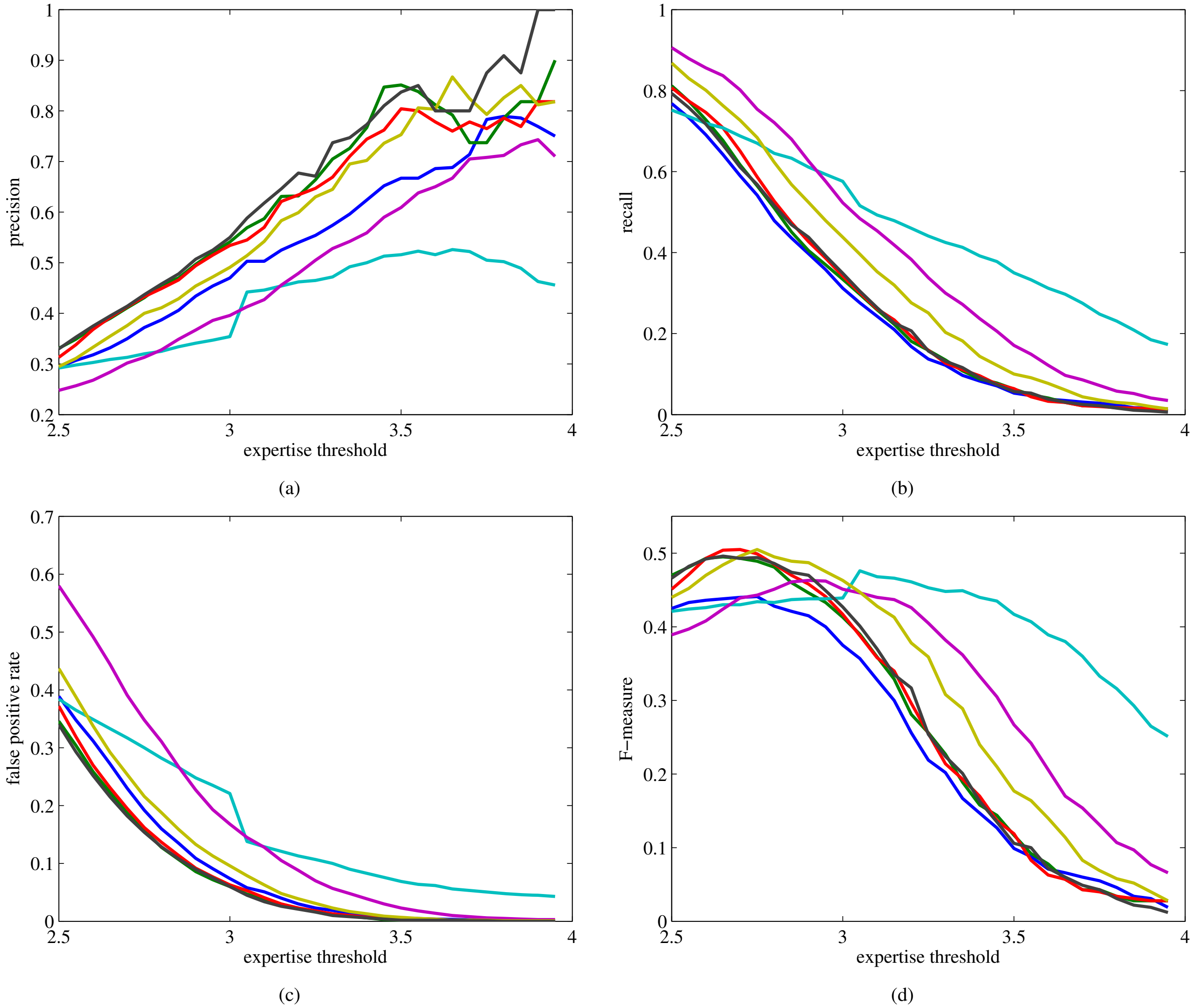


Figure 2. (a) Precision, (b) recall, (c) false positive rate, and (d) F-measure as a function of expertise threshold for MF (blue line), MFs (green line), MFse (red line), SS (cyan line), MFSS (purple line), MFsSS (dark yellow line), and MFseSS (black line).

has better recall.

2) Experiment 2: For identifying experts in the truly new skill topic prediction problem, we use the optimal thresholds for each model from Table IV. In Table V, we present the number of employees predicted to be experts in the three new skill topics at those thresholds. Also in the table, we report how many of the predicted experts were on the business-identified lists.

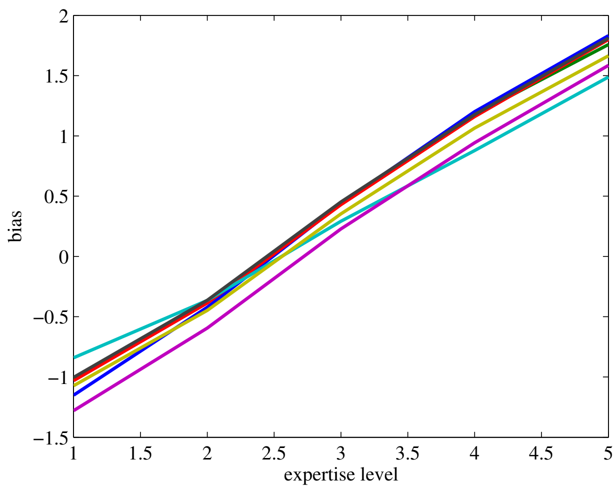
Among different models, MFSS found the most business-identified experts: 37 out of 70. However, this model has a very high number of predicted experts. MFSS performs better in comparison to MF because it predicted fewer experts yet found more employees from the business-identified lists. For the models giving small numbers of predicted experts, the number of business-identified experts that are found is very low. In general, there is much scope for improving the models in terms of finding the business-identified experts. We wish to note here that it is not possible to find some business-identified experts using the available information. For example, there are two business-identified experts who rated themselves at expertise level 1 on all of their assessed skills and particularly on skills related to their business-identified expertise area, and did not have any social tag information.

V. Conclusion

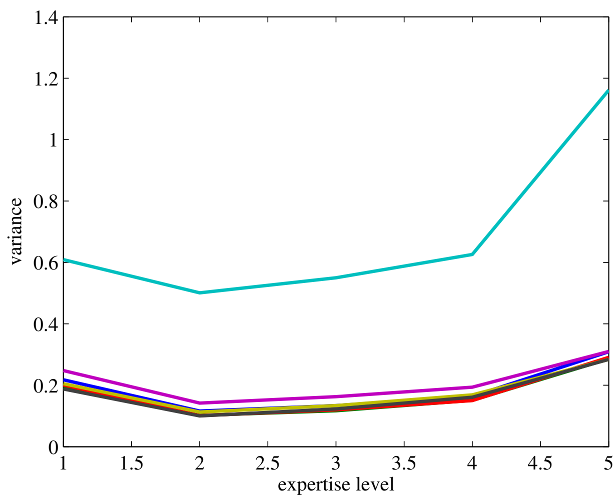
In this paper, our goal was to predict experts for new skills and skill topics. For this challenging cold-start problem, we proposed multiple collaborative filtering, content-based and hybrid recommendation approaches to combine data from

Table V
PREDICTED EXPERTS ON NEW SKILL TOPICS

Model	BAO Business Identified: 42		Cloud Business Identified: 22		Smarter Commerce Business Identified: 6	
	Num. Predicted	Num. Identified	Num. Predicted	Num. Identified	Num. Predicted	Num. Identified
MF	2753	18	2741	11	2771	1
MFs	596	3	497	4	606	0
MFse	505	4	427	3	493	0
SS	739	3	1052	9	729	0
MFSS	2426	20	2493	15	2280	2
MFsSS	359	1	178	2	277	0
MFseSS	354	1	195	2	279	0



(a)



(b)

Figure 3. (a) Bias and (b) variance of predictions as a function of true expertise level for MF (blue line), MFs (green line), MFse (red line), SS (cyan line), MFSS (purple line), MFsSS (dark yellow line), and MFseSS (black line).

different sources: employee assessment data, free-text skill description data, and employee internal social media tags. Effectively combining all different data sources to get better results is challenging. Through an empirical study on real-world corporate data, we found that employee tags alone are not beneficial and need to interact with skill tags to impact expertise recommendation. In the future, we will consider models that deal with user and skill tag interactions more directly.

We also observe that the skill tags are more helpful in prediction, possibly because they model the similarities between skills directly and hence are more aligned to our goal of skill recommendation. All of the methods that we applied to the problem suffer from a similar biased prediction pattern. Prediction using skill similarity has a smaller bias but larger variance. Furthermore, the hybrid model that includes skill similarity for prediction has a higher recall rate. In particular, the novel matrix factorization model that incorporates predictions from skill similarity outperforms basic matrix factorization model in terms of finding business-identified BAO, Cloud, and Smarter Commerce experts.

This study reveals numerous avenues for improvements to the proposed approaches. Nevertheless, the proposed solution has high business value. In an envisioned embodiment, the algorithm will only recommend skills to an employee while allowing the employee to make the final choice. Therefore, some amount of false positives can be tolerated. The approaches proposed in this paper can become a part of the larger skill management system in an enterprise, and can be constantly improved through employee feedback. Such expertise management systems are critical for businesses in today's environment to evaluate knowledge capital and determine strategies for venturing into appropriate new areas. On the other hand, such approaches can also help businesses identify areas where essential skills are lacking and hence impart training to their workforce to bridge the gap accordingly.

ACKNOWLEDGMENT

The authors thank members of the IBM Expertise project team from across the corporation.

References

[1] J. Trant, "Studying social tagging and folksonomy: A review and framework," *J. Digit. Inf.*, vol. 10, no. 1, Jan. 2009.

[2] K. R. Varshney, J. Wang, A. Mojsilović, D. Fang, and J. H. Bauer, "Predicting and recommending skills in the social enterprise," in *Proc. Int. AAAI Conf. Weblogs Soc. Med. Workshops*, Cambridge, MA, Jul. 2013.

[3] C. Chelmis, V. Sorathia, and V. K. Prasanna, "Enterprise wisdom captured socially," in *Proc. IEEE/ACM Int. Conf. Adv. Soc. Netw. Anal. Min.*, Istanbul, Turkey, Aug. 2012, pp. 1228–1235.

[4] J. Wang, K. R. Varshney, A. Mojsilović, D. Fang, and J. H. Bauer, "Expertise assessment with multi-cue semantic information," in *Proc. IEEE Int. Conf. Serv. Oper. Logistics Informatics*, Dongguan, China, Jul. 2013, pp. 534–539.

[5] A. I. Schein, A. Popescul, L. H. Ungar, and D. M. Pennock, "Methods and metrics for cold-start recommendation," in *Proc. ACM SIGIR Conf. Res. Dev. Inf. Retrieval*, Tampere, Finland, Aug. 2002, pp. 253–260.

[6] P. Melville and V. Sindhwani, "Recommender systems," in *Encyclopedia of Machine Learning*, C. Sammut and G. I. Webb, Eds. New York, NY: Springer, 2010.

[7] J. Wang, K. R. Varshney, and A. Mojsilović, "Legislative prediction via random walks over a heterogeneous graph," in *Proc. SIAM Int. Conf. Data Min.*, Anaheim, CA, Apr. 2012, pp. 1095–1106.

[8] K. Zhou, S.-H. Yang, and H. Zha, "Functional matrix factorizations for cold-start recommendation," in *Proc. ACM SIGIR Conf. Res. Dev. Inf. Retrieval*, Beijing, China, Jul. 2011, pp. 315–324.

[9] E. J. Candès and B. Recht, "Exact matrix completion via convex optimization," *Found. Comput. Math.*, vol. 9, no. 6, pp. 717–772, Dec. 2009.

[10] M. K. Jung, M. K. Jang, H. K. Kim, and J. K. Kim, "A new item recommendation procedure based on min-max distance," *Int. J. Soc. Sci. Hum.*, vol. 1, pp. 43–48, May 2011.

[11] P. Cremonesi, R. Turrin, and F. Airoldi, "Hybrid algorithms for recommending new items," in *Proc. Int. Workshop Inf. Heterog. Fusion Rec. Syst.*, Chicago, IL, Oct. 2011, pp. 33–40.

[12] S.-T. Park and W. Chu, "Pairwise preference regression for cold-start recommendation," in *Proc. ACM Conf. Rec. Syst.*, New York, NY, Oct. 2009, pp. 21–28.

[13] X. N. Lam, T. Vu, T. D. Le, and A. D. Duong, "Addressing cold-start problem in recommendation systems," in *Proc. Int. Conf. Ubiq. Inf. Manage. Comm.*, Suwon, Korea, Jan.–Feb. 2008, pp. 208–211.

[14] H. Sobhanam and A. K. Mariappan, "Addressing cold start problem in recommender systems using association rules and clustering technique," in *Proc. Int. Conf. Comput. Comm. Informatics*, Coimbatore, India, Jan. 2013.

[15] A. Gunawardana and C. Meek, "Tied Boltzmann machines for cold start recommendations," in *Proc. ACM Conf. Rec. Syst.*, Lausanne, Switzerland, Oct. 2008, pp. 19–26.

[16] T. Chen, W. Zhang, Q. Lu, K. Chen, Z. Zheng, and Y. Yu, "SVDFeature: A toolkit for feature-based collaborative filtering," *J Mac. Learn. Res.*, vol. 13, pp. 3619–3622, Dec. 2012.