Engaging the Crowd for better Job Recommendations

Benjamin Kille
Technische Universität Berlin
benjamin.kille@tu-berlin.de

Fabian Abel
XING AG
fabian.abel@xing.com

ABSTRACT

Today, people more and more use business-oriented social networks such as LinkedIn[1] or XING[2] to explore career opportunities and find an interesting jobs. Similarly, companies use those platforms to identify candidates for open positions and post job offers in order to advertise to attract appropriate candidates. Job recommendation systems thus have to solve a reciprocal recommendation problem and have to satisfy the expectations of both users who aim for interesting jobs and companies who aim for appropriate applicants.

In this paper, we discuss challenges of building a reciprocal job recommendation system. Based on an analysis of profile and interaction data, we highlight potential features that such a recommendation system can exploit and discuss opportunities of integrating a user feedback cycle into the recommender algorithm.

Categories and Subject Descriptors

H.3.3 [**Information Systems**]: Information Search and Retrieval—*Information filtering*

Keywords

recommender systems; reciprocal recommendations; social networks; job recommendations, crowd engagement, user-generated content

1. INTRODUCTION AND RELATED WORK

Job advertisements typically comprise characteristics related to the position, candidate, and company. They describe the job role and necessary skills. Job offers paint a picture of the desired candidate. They list references to the educational background and appreciated attributes. In addition, job advertisements introduce the employer. Companies increasingly post job offers on their websites and designated portals. Current as well as future professionals have started to curate profiles and curriculum vitae online. These describe their skills, interests, and the development of their careers. Business-oriented social networks connect professionals and job advertisements. They thereby induce a market place of work. Recruiters get access to a large pool of potential candidates. Professionals can browse large collections of job advertisements. Deciding to apply for a job continues to challenge professionals. They identify with their work which occupies considerable parts of their lives. Advertising jobs spreads wider in the digital era. The set of professionals had reduced to the readers of newspapers before the Internet emerged. Increasing numbers of job advertisements impede professionals' decision processes. Professionals face hundreds of potentially relevant job postings. Recommender systems have established in similar conditions. They support people by suggesting a subset of the overwhelmingly large collection of options. Collaborative Filtering (CF) represents a particularly successful paradigm for this problem (cf. [9]). CF depends on users' preferences being available. Situations in which recommenders lack users' preferences favor content-based filtering (CBF) (cf. [8]). Both techniques and combinations thereof have been used for recommending jobs. Al-Otaibi and Ykhlef [2] as well as Siting et al. [10] surveyed job recommender systems. Concordantly, they report early experiments with CF and CBF followed by hybridization of both techniques. Additionally, both mention more recent approaches with a new paradigm: reciprocal recommendation.

Reciprocal recommendation adheres to slightly adjusted problem description. Traditional recommender systems seek to maximize utility for the recipient. Reciprocal recommender systems seek to maximize the mutual utility of recipient and target (cf. [4]). Much of the research on reciprocal recommender systems concentrates on online dating (cf. [1]). Still, recommending jobs represents an analogous setting. Online dating seeks to match two human beings. Recommending jobs seeks to match an individual to an organization. Both problems comprise two perspectives. In the case of online dating, we either consider the recipient or the recommended user. In the case of job recommendation, we either consider the professional or the employer. We require the system to automatically answer two questions. What is the best job for a given user? Who is the best candidate for a given job? Malinowski et al. [5] decided to tackle both subtasks separately. They define two functions solving one of the tasks each. Subsequently, the combine the results to obtain a more adequate list of recommendations. Mine et al. [7] consider messages exchanged between recruiters and candidates. They transform these interactions into a graph representation. Subsequently, they compute popularity degrees based on the graph. This allows them to more accurately match professionals and re-

[1]`http://linkedin.com`
[2]`http://xing.com`

CrowdRec 2015 19 Sept. 2015, Vienna, Austria
ACM 978-1-4503-3692-5/15/09.
DOI: http://dx.doi.org/10.1145/2792838.2799496.

cruiters. Yu, Liu and Zhang [11] compute the similarity between user profiles and job advertisements. Subsequently, they filter similarities to the mutual benefit of professionals and recruiters.

All proposed methods utilize user profiles, job descriptions, and interaction in between both entities in some way. Users generate these data. Recruiters create job descriptions. Professionals create their profiles and interact with job advertisements. The job recommender depends on the quality of these data. We conclude that business-related social networks depend on contents generated by professionals and recruiters. The remainder of this document introduces shortcomings of existing job recommender systems. Subsequently, we quantify user behavior thus illustrating how we can exploit user-generated content to improve job recommendations.

2. CHALLENGES

XING operates a business-related social network with millions of members. They seek to continuously improve their job recommender. Therefore, they conducted a survey to determine how satisfied their members are with their job recommendations. The survey revealed that a considerable fraction of members were unhappy with their job recommendations. Asked for the reason for their dissatisfaction, $\approx 38.5\,\%$ referred to mismatching field of activity. Inappropriate locations accounted for the dissatisfaction of $\approx 27.5\,\%$ members. Finally, inadequate career levels caused the dissatisfaction of $\approx 16.8\,\%$ of the members. XING's recommender uses a variety of information to suggest the most relevant jobs. A hybrid model combines all aspects covered in the state of the art. It uses interactions to determine preferences necessary for CF. In addition, the system considers user profiles and job advertisements in form of CBF. Still, members deem suggestions frequently inadequate. What causes this disparity?

From professionals' perspectives, the job recommender seeks to determine how well the job advertisement matches their expectations. Consequently, the job recommender models professionals' expectations with available information. Hence, we must ask how well the available information reflect professionals' expectations. Do professionals behave according to the profile they had put up? We frequently observe discrepancies between professionals' profiles and their implicit interests. In other words, professionals tend to click on suggested positions mismatching their profiles. Do profiles reflect the positions professionals strive to acquire? Contrarily, profiles rather constitute experiences and current employments. We hardly find evidence of professionals specifying their expectations. Do professionals actually know what they expect to find? We cannot definitely tell for the complete set of users whether they are aware of their expectations. Still, we suppose that a considerable subset of members is unaware of their expectations. This appears particularly likely for young professionals and job starters. Can professionals express their expectations? As mentioned before, professionals may have a rather blurred picture of their own expectations. When forced to express themselves, they will naturally resort to their own words. This may induce an idiomatic mismatch with recruiters' vocabularies. Thus, a job recommender might fail due to missing links between professionals' profiles and recruiters' advertisements. The above listed phenomena might well cause members' dissatisfaction with the job recommender. A job recommender automatically provides suggestions. It has to deal with the lack of information such as desired field, career level, and location. Section 3 illustrates types of user-generated contents on business-related social networks. Section 4 presents ideas to compensate the lack of information with user-generated content.

3. USER-GENERATED CONTENT

Professionals generate two types of content on business-related social network. First, they create profiles reflecting their experiences, skills, and interests. Second, they interact with elements including job advertisements. The platform offers a variety of actions. These actions include clicking, bookmarking, and replying to job advertisements.

3.1 User Profiles

Reasons to create profiles differ between professionals. Some may seek to attract recruiters to obtain a better position. Some may seek to establish bonds with existing contacts or discover new ones. Others may simply enjoy expressing themselves. Typically, user profiles include demographics, working experiences including the current position, and skills. We find this information in curriculum vitae. Business-related social networks let their members add further information. For instance, they can define what they are looking for or their interests. Additionally, members establish social links as they add contacts. Although, this information typically remains in a restricted area which cannot be publically accessed.

3.2 Interactions

We seek to recognize preferences as we observe professionals interacting with job advertisements. We assume that they will refrain from clicking on irrelevant suggestions. Our observations contain different kinds of feedback. Members can click, bookmark, reply to, and rate job advertisements. Our confidence on the relevance of suggestions depends on the type of action. Clicks can signal interests. They may as well signal curiosity or be due to accident. On the other hand, rating a suggestion highly or even messaging the recruiter represent strong signals for interest. Conversely, we have a large number of clicks available. In contrast, we observe rather few replies, bookmarks, or ratings. Those cause more effort for users compared to clicks.

4. IDEAS

Our ideas to improve the job recommender target the main issues members mentioned in the survey. Members reported that suggested positions mismatched their field of interest. In other words, the system recommended irrelevant positions.

4.1 Improving Relevance Estimation

The systems depends on two information sources as it estimates the match for a pair of professional and job offer. First, the system compares the professional's profile with the job description. Second, the system uses interaction patterns between the sets of professionals and job offers. The former lays the foundation for content-based filtering. The latter supports collaborative filtering. CBF assumes that professionals' profiles and job offers share their contents. We consider professionals' bookmarking behavior. We sample $\approx 40\,000$ professionals. Figure 1 relates skills stated in professionals' profiles and required in bookmarked job offers. We observe that a proportion of $\approx 66.7\,\%$ of professionals' profiles lack any overlap with the job offers' required skills. Further, we note that $\approx 20.7\,\%$ of professionals' profiles share a single skill with bookmarked job offers' requirements.

The discrepancy between professionals' profiles and the job advertisements they bookmark explains the job recommender's inability to accurately match both. As professionals deem job offers relevant without any overlap with their profiles, relevance estimation reduces to educated guessing. With insufficient overlap we fail to estimated the relevance of an individual job offer for a

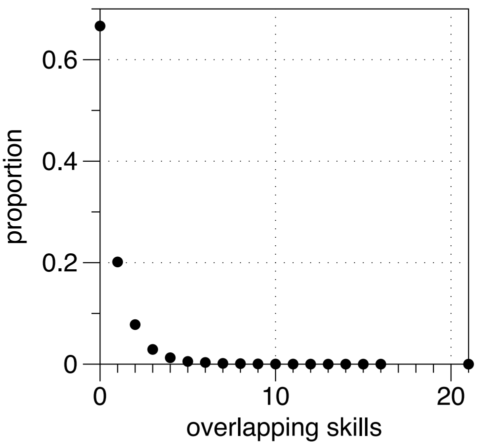

Figure 1: Professionals bookmarked job offers. The figure shows the proportion of users by the overlap between skills in their profiles and skills required in job offers they bookmarked.

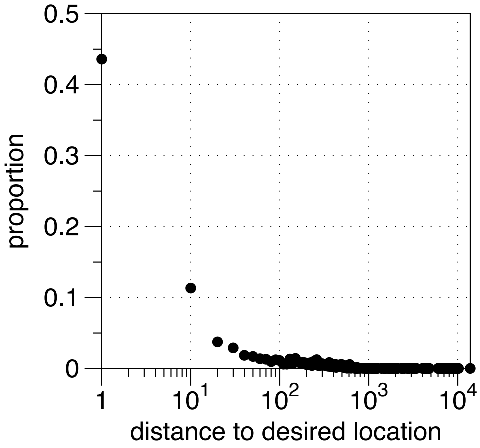

Figure 2: We consider how far professionals are willing to move. The figure relates the proportion of professionals to the distance between their location and their bookmarked job offers' location

specific professional. Users-generated content allows us to improve the matching. As users create expressive profiles and we observe their interactions we construct a latent space. There are several techniques to construct such a latent space. Blei, Ng and Jordan [3] proposed Latent Dirichlet Allocation (LDA). Mikolov et al. [6] describe an architecture featuring neural nets. Using either of these approaches or alternative techniques, we can estimate how relevant a job offer is for a professional. Although, we have to consider the chance that a multitude of job offers obtain similar scores. Conversely, many professionals may be scored similarly with respect to a job offer. Therefore, we consider professionals' behavior as additional information source. Unfortunately, we thereby face considerably scarce signals. We extracted a month of interaction data. Therein, we detect $\approx 0.02\,\%$ of possible pairs of professionals and job postings. We expect a sparse relation. Professionals will find most job advertisements irrelevant. So are only few professionals suited candidates for a specific job offer. Still, collaborative filtering struggles with highly sparse signals. We can improve its performance as we encourage professionals to inrease their activity. The more they interact with job offers, the better the system learns their preferences.

Members noted mismatched locations as another reason for their dissatisfaction with the job recommender. Even though the job advertisement may have been relevant, professionals would refuse the offer due to their location. Professionals may be unwilling to move or commute to the designated location. How shall the job recommender determine whether a location appeals to a professional?

4.2 Score Location

Determining to what degree a professional likes a location challenges the system. Members will hardly explicitly mention their preferences for a collection of locations. Thus, the systems must infer these preferences.

Figure 2 shows the discrepancy the job recommender is facing. We observe that $\approx 43.6\,\%$ of professionals restrict their interest to job offers in their direct surroundings. Still, some professionals bookmark job offers located as far as 10 000 km from their location. We expect a baseline filtering job offers from professionals location to perform well. Although, this will merely cover the set of members unwilling to move or commute longer distances. Alternatively, we may model a professionals willingness to move. Members' working experience includes information on their historic employments. Thereof, we extract the location information. Combined with the chronology of their employments, we obtain sequences of geographic locations. These data enable us to estimate the distance a professional would be willing to move. Alternatively, applying unsupervised learning techniques could reveal irregularly liked locations. For instance, financial trading places such as New York, London, or Frankfurt may be particularly interesting for professionals within the banking sector. Therefore, the location-aware job recommender would score job offers in New York higher than Miami.

Finally, members claimed that the job recommender would not sufficiently consider their career progress. For instance, junior professionals would receive suggestions for director positions. Their lack of experience makes it rather unlikely that they would be considered candidate for such a job offer.

4.3 Consider Career Situation

Professionals state their own career as they enter the working experience part of their profile. The system deduces their current situation therefrom. Other professionals' profiles enable the system to learn the transition probabilities. We randomly sampled 10 000 members. Subsequently, we reduce the collection keeping members with several positions in their profiles. Professionals with a single position lack transitions. We examined the remaining members' careers on how their levels progressed. We distinguish six position levels: student/intern, entry level, professional, manager, executive, and senior executive.

Table 1 illustrates transitions in the careers of 10 000 randomly sampled members. We observe that levels vary in their proportion. The distribution favors professional positions with 39.8 % over executive positions at 5.2 %. Additionally, we consider the chances to move to a specific level given the current level. Entry level positions are most likely to transit to professional positions. For the remaining levels, we observe the tendency to stagnate at the respective levels. For instance, professionals keep their level in 74.2 % of the cases. This emphasizes the dissatisfaction of members with their recommended jobs. Suggesting an executive position to a member currently occupied with an entry level po-

Table 1: We observe the distribution of career levels in a sample of 10 000 professionals. Further, we observe the chances of the next career step. Each row concentrates on an individual level and sums to ≈ 100 %. We note that all levels have the highest relative chance to remain in the level except entry level positions. In those cases, chance are higher to move to a professional position.

	Student	**Entry Level**	**Professional**	**Manager**	**Executive**	**Senior Executive**
Proportion	18.2 %	13.5 %	39.8 %	17.7 %	5.2 %	5.5 %
Student	60.2 %	21.7 %	15.0 %	2.0 %	0.2 %	0.1 %
Entry Level	9.6 %	18.6 %	60.8 %	8.2 %	1.4 %	1.5 %
Professional	3.5 %	3.4 %	74.2 %	14.7 %	1.1 %	3.0 %
Manager	0.8 %	0.6 %	13.2 %	66.8 %	11.4 %	7.1 %
Executive	0.5 %	1.4 %	7.9 %	20.4 %	44.9 %	25.0 %
Senior Executive	2.7 %	1.8 %	12.7 %	15.8 %	20.4 %	46.6 %

sition appears implausible. Merely 1.4 % of the cases we observed a member moving to executive position from an entry level occupation. Consequently, we consider the potential of member profiles to improve the job recommender. We may either estimate general transition probabilities or break them down for certain professional segments or markets.

5. DISCUSSION

Recommending jobs on business-related social networks challenges standard recommendation algorithms. Members tend to scarcely interact with job offers. Thus, collaborative filtering struggles to detect reliable patterns. Professionals and recruiters tend to use different terms to express their offers. Professionals create profiles. Recruiters create job advertisements. We observed that their vocabulary does not necessarily overlap as professionals state their interest in an offer. This limits the applicability of rather simple content-based filtering. Professionals differ in their willingness to move. Unfortunately, they do not necessarily express their preferences clearly. Thus, the systems must infer their preferences. Thereby, the system is prone to errors. Finally, members request suggestions of positions which fit their career levels. Considering transition probabilities may support selecting more appropriate positions. We noted that all such efforts strongly depend on user-generated content. The less the systems knows about either member of job offer, the less likely it will accurately match both. What can we do to encourage members to more actively contribute? We suppose that members might be unaware of the advantages of their increased activity. Therefore, we should encourage them to provide explicit feedback. For instance, members may be asked to rate a set of job offers. Having collected the ratings, we could present them two lists of recommendations. First, the list they initially would have received. Second, the list adjusted with their feedback. As the systems learns, members will receive more and more relevant suggestions. Additionally, we could improve the usability of the user interface. As members enjoy the process of providing valuable feedback, we assume to collect considerable more information.

Another approach of collecting feedback in a more playful fashion is depicted in Figure 3: instead of explicitly asking users to give feedback about their recommendations, the system would provide filtering functionality to the users. As soon as a user selects a tag, the recommendation list will be updated and filtered or tuned with respect to the user's selection. Moreover, the list of *related* tags will be updated so that the user has the chance to broaden or refine their selection. The tags that are listed in the recommendation filtering front-end, may relate to job titles, skills, locations, career levels or industries. By check-marking *"remember for future recommendations"*, users could inform the

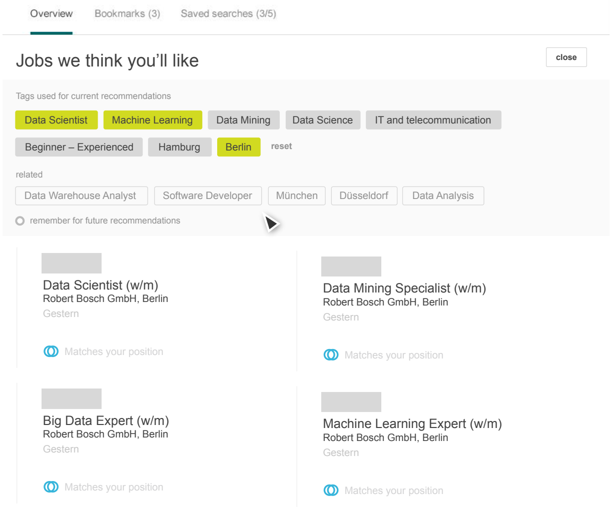

Figure 3: Interactive job recommendations: users are enabled to filter recommendations, get filtering suggestions and can store filtering settings.

recommender system that a given selection should be considered for future recommendations. In addition to giving feedback on recommendations that a user receives, we are also investigating whether there are means to involve users in recommending job recommendations to others. For example, if a user deletes a recommendation and therewith expresses that they are no longer interested in the job recommendation then the recommender system could ask them whether they know people who may be interested into the job or whether they know for what types of users the job may be interesting (e.g. *"This job posting may be interesting to users who work as 'Data Scientist' and are skilled in 'Python' and 'Data Mining'."*).

6. CONCLUSIONS AND FUTURE WORK

Business-related social networks gain popularity among professionals exploring their career opportunities as well as recruiters looking for candidates. They connect both group inducing a digital market. Both groups struggle in their exploratory tasks. The overwhelming amount of professionals and job offers impede finding the subset of most relevant options. Reciprocal recommender systems alleviate the unbalance between relevant and irrelevant contents. Although, they are far from perfect as they face challenges. These challenges mostly come for lack of information.

We have analyzed the problems of mismatched professional fields, locations, and career levels. Each poses a challenge individually. Combined they cause dissatisfaction for members. We proposed to use user-generated content to counteract the lack of information. Encouraging members to contribute represents the crucial part of this process.

Acknowledgments

The work leading to these results has received funding from the EU's Seventh Framework Programme (FP7/2007–2013) under CrowdRec Grant Agreement no. 610594.

7. REFERENCES

[1] J. Akehurst, I. Koprinska, K. Yacef, L. Pizzato, J. Kay, and T. Rej. CCR - A content-collaborative reciprocal recommender for online dating. In *International Joint Conference on Artificial Intelligence*, pages 2199–2204, 2011.

[2] S. T. Al-Otaibi and M. Ykhlef. A survey of job recommender systems. *International Journal of the Physical Sciences*, 7(29):5127–5142, 2012.

[3] D. M. Blei, A. Y. Ng, and M. I. Jordan. Latent dirichlet allocation. *Journal of machine Learning research*, 3:993–1022, 2003.

[4] M. Larson, P. Cremonesi, A. Said, D. Tikk, Y. Shi, and A. Karatzoglou. Activating the crowd: exploiting user-item reciprocity for recommendation. In *Workshop on Crowdsourcing and human computation for recommender systems, ACM RecSys*, 2013.

[5] J. Malinowski, T. Keim, O. Wendt, and T. Weitzel. Matching people and jobs: A bilateral recommendation approach. In *Proceedings of the 39th Annual Hawaii International Conference on System Sciences*, volume 6. IEEE, 2006.

[6] T. Mikolov, K. Chen, G. Corrado, and J. Dean. Efficient estimation of word representations in vector space. *arXiv preprint arXiv:1301.3781*, 2013.

[7] T. Mine, T. Kakuta, and A. Ono. Reciprocal Recommendation for Job Matching with Bidirectional Feedback. (2009):39–44, 2013.

[8] M. J. Pazzani. A framework for collaborative, content-based and demographic filtering. *Artificial Intelligence Review*, 13(5-6):393–408, 1999.

[9] P. Resnick, N. Iacovou, M. Suchak, P. Bergstrom, and J. Riedl. Grouplens: an open architecture for collaborative filtering of netnews. In *Proceedings of the 1994 ACM conference on Computer supported cooperative work*, pages 175–186. ACM, 1994.

[10] Z. Siting, H. Wenxing, Z. Ning, and Y. Fan. Job recommender systems: a survey. In *International Conference on Computer Science & Education (ICCSE)*, pages 920–924. IEEE, 2012.

[11] H. Yu, C. Liu, and F. Zhang. Reciprocal recommendation algorithm for the field of recruitment. *Journal of Information & Computational Science*, 8(16):4061–4068, 2011.