Personalizing the Search for Persons: A Recommender-based Approach

Tobias Keim[1], Jochen Malinowski[1], Gregor Heinrich[2], and Oliver Wendt[3]

[1]University of Frankfurt, Mertonstr. 17, 60325 Frankfurt/M., Germany

[2] Fraunhofer IGD, Fraunhoferstr. 5, 64283 Darmstadt, Germany

[3] University of Kaiserslautern, Postfach 3049, 67653 Kaiserslautern, Germany

Abstract. Recommendation systems are widely used on the Internet to assist customers in finding the products or services that best fit their individual preferences. While current implementations successfully reduce information overload by generating personalized suggestions when searching for objects such as books or movies, recommendation systems so far cannot be found in another potential field of application: the personalized search for subjects such as business partners or employees. This is astonishing as (1) the number of CV-, assessment- and social network-data available on the Internet is growing and (2) the complexity and scope of selecting the right partner is much higher than when buying a book. We argue that recommendation systems personalizing the search for people need to be grounded on two pillars: unary attributes on the one hand and relational attributes on the other. We present a framework meeting these requirements together with an outline of a first prototypical implementation.

1 Introduction

Personalization systems such as recommender engines in recent years attracted the interest of many researchers and practitioners. Since Resnick and Varian first established the term "recommender system" in 1997 [26], researchers have been improving recommendation quality and scalability of such systems by various means. While some researchers merged content-based with collaborative filtering in order to overcome sparsity problems and combine the advantages of both approaches [20] [29], others focused on how to reduce the dimensionality of the user-item-matrix underlying collaborative filtering approaches [30] [32]. Today, recommendation systems successfully assist consumers on the Internet in finding products or *objects* based on items similar to the ones the customer himself previously liked or based on items that other customers similar to him liked in the past. However, personalization systems are not yet applied when searching for people or *subjects*. Thus our research question is: What are necessary theoretical enhancements for human recommender systems? We argue that the various bilateral and relational aspects that need to be considered when bringing individuals together imply extending existing approaches

Keim T., Malinowski J., Heinrich G. and Wendt O. (2005).
Personalizing the Search for Persons: A Recommender-based Approach.
In *Proceedings of the 1st International Workshop on Web Personalisation, Recommender Systems and Intelligent User Interfaces*, pages 125-134
DOI: 10.5220/0001421801250134

by relational data. Building on existing theory and own prior research, we derive concrete requirements and present an outline for a recommendation system personalizing the search for individuals.

2 Research Motivation

Information technology in recent years has transformed (1) the ways people find work as well as (2) the ways they effectively work together. With regard to the first aspect, own longitudinal empirical research with the Top-1.000-companies in Germany as well as with over 11.000 job seekers shows that the Internet has replaced print media as the most important recruitment channel [16] [17][1]. With 78% of all vacancies being published within the career section of the corporate website and 49% of open jobs being posted on Internet job-portals, IT-supported channels dominate print media (30%) as a way to attract candidates. Also, over the years the ratio of actual hires generated through job ads on the Internet rises reaching 58% in 2004 [17]. When considering the later stages of the recruitment process such as the treatment of incoming applications and the (pre-) selection of candidates, a diminished importance of IS-support can be observed. However, as digital applications lower application costs, the number of incoming (electronic) applications increases. Thus, companies seek adapted IS-support for the selection stage in order to process the masses of incoming applications efficiently.

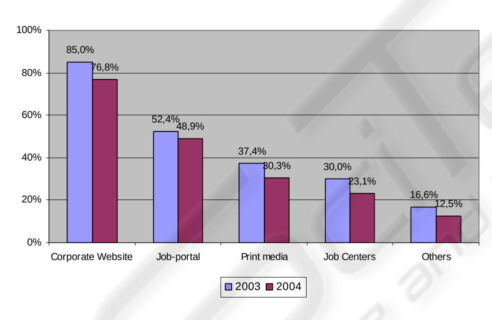

Fig. 1. Ratio of vacancies published

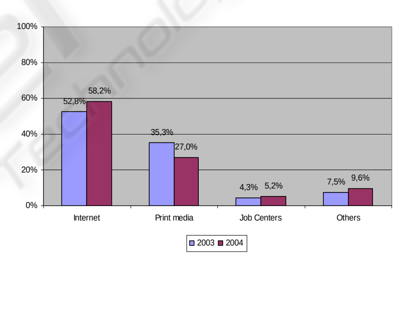

Fig. 2. Ratio of jobs filled

While this empirical research deals with how people find work, other research strands are concerned with how information systems change the ways people effectively work together once the candidate is recruited. Starting from Malone and Laubacher's vision of the "e-lance economy" [21], special attention was paid to the ways communication channels and "discontinuities" of space, time and organizational boundaries characteristic of virtual work influence collaboration patterns [2] [34]. Thus, as work

[1] Companies selected based on revenues; between 151 and 196 companies responding between 2002 and 2004

in changing projects and organizational settings gains importance, individuals are more frequently matched to new colleagues within their lifetime. Beside this, systems for *ad hoc* short-term expert identification streamline the way knowledge is accessed and exchanged between different projects or units beyond document management [1] [9].

3 The Personalized Search for Persons

From these considerations that (1) matching situations within a person's work history will increase and (2) decision support for the matching of collaboration partners will emerge, we started to develop a system for the personalized search for individuals. In the following, we present requirements for such a person-recommender.

3.1 Requirements for Recommending Persons

Team configuration for work contexts has been analyzed by a variety of disciplines. Typically, such problems are considered under the perspective of task-related and social aspects [12], human and social capital or person-job fit and person-team or person-organization fit [31]. Thus, successful team design needs to consider two dimensions:

- The matching of individuals to *tasks* for which the candidate possesses the skills and abilities to carry them out.
- The matching of individuals to *other individuals* with whom the person is able to collaborate successfully.

This latter dimension has major implications for the design of a person-recommender as we cannot consider the selection of candidates as a unilateral decision. While the customer chooses the movie he wishes to watch and not vice versa, this is not the case when recommending people. Selecting a candidate or partner is a bilateral selection decision in which not only the attributes of the item or individual itself need to be considered, but also the relationship between these items or individuals. In separation to the former attributes that can be tied directly to the individual, that we refer to as unary attributes, we denote the latter attributes as relational attributes. Thus, we retain the following key differences when recommending subjects instead of objects:

- Recommending people is a *bilateral* process that needs to take into account the preferences not only of a single person (the active user), but of several persons.
- Even more, recommendations cannot be based on the attributes tied to the items or persons in consideration only, but need to incorporate the underlying relational structure by means of *relational* attributes.
- Finally, as every individual is considered to be unique, we cannot recommend a single item or person several times such as in the case of a

movie or book. As every person can only be selected once, recommendations on the item-level are *not repeatable*. Thus, recommendations cannot be solely based on a user-item matrix but need to incorporate “content”-elements such as the unary and relational attributes mentioned above.

3.2 Towards a Person-Recommender

On our way towards a person-recommender, we implemented two complementary approaches: a CV-recommender and a social network browser that are both going to be presented briefly hereunder. Afterwards we describe how both approaches can be combined leading to a relational recommender system.

The CV-Recommender. In a first step, we built a system recommending CVs that are similar to resumes previously selected by the same recruiter for a specific job-profile considered. The probabilistic hybrid recommendation engine is based on a latent aspect model that understands individual preferences as a convex combination of preference factors [10] [11] [25]. As depicted in figure 3, the recruiter together with the job description is represented in variable x, the preference factors being modeled in variable z. In coherence with our prior considerations, the recruiter by rating a candidate profile or CV with variable v = {"qualified", "not qualified"} does not rate the person itself, but the sum of its attributes. These “content”-elements, taken from the candidate's resume are composed of a quadruple such as a=("mathematical skills", "diploma grade", "1.0", "University of Frankfurt"). Thus, the rating value v depends indirectly on the position considered x and directly on the candidate’s attributes a. With a set of observed values v for an attribute assessed by x and assigned to a, we are able to estimate the model parameters using an Expectation Maximization (EM) algorithm. A detailed description of the approach together with validation results can be found in [4].

The Social Network Browser. As the CV-recommender is focused on what we called unary attributes earlier, we modeled relational attributes in a complementary approach. The network browser shown in figure 4 visualizes trusted social relations that the user then can manually browse, filter the network and search for particular nodes. The social relations used are recommendations between people based on “historic” experience as well as swift trust assessments from candidate interviews via video conferences. A more detailed description of the approaches to swift and historic trust modeled within the system and their elicitation from a user community can be found in [18] and [9]. When navigating the resulting network, by filtering and searching techniques it is possible to identify relevant persons in the graph according to different criteria. This way, important trusted actors in the network can be identified either from an ego-centered position (of the searcher) or globally using graph analysis methods such as shortest-path, relative importance and others also used in social network analysis [33]. In addition, such filtering can initially apply relevant competence criteria, which creates a trust network contextualized on the queried competencies. The motivation of this idea is closely related to research on the relationship between trust, interpersonal cooperation and organizational effectiveness such as [3] or [14].

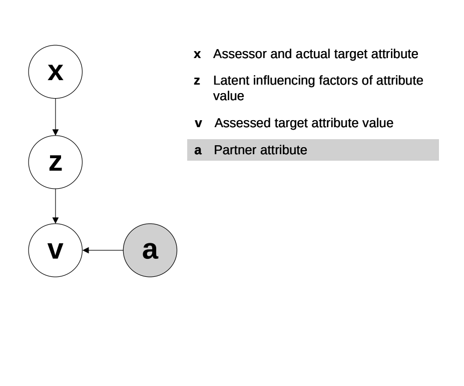

Fig. 3. The probabilistic CV-Recommender

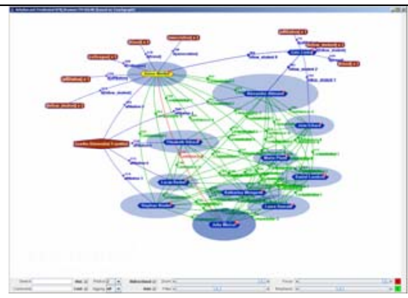

Fig. 4. The network analyzer

4 Towards the Relational Person-Recommender

In order to meet the requirements previously defined, we need to combine the predictive capabilities of the CV-recommender with the descriptive capabilities of the network browser in an automatic setting. This is based on our previously defined requirements where we stated that a person-recommender not only needs to consider individual but also relational attributes. From a theoretical perspective, this is an interesting idea as already Granovetter showed that labor market processes are rooted in social relations [5]. Montgomery argued that the higher quality of information gained from contact networks reduces frictions when entering a new job [22]. Also, the reductions of attraction costs [28] and of screening costs have been mentioned as advantages of partner identification over networks [19].

In order to build such a relational recommender, we developed a trust computational model. Conforming to Richardson, Agrawal and Domingos (2003), we assume that trust can be expressed in a singular value even though it is a complex and multidimensional construct. (In the above network browser, we adopted this scheme already by aggregating the values of the different trust dimension values.) Our current research builds on trust propagation as demonstrated in [6]. Based on findings in the literature and own theoretical considerations we defined three trust propagation and prediction scenarios as depicted in Figure 5(a)-(c).

Figure 5(a) illustrates how the trust level between individuals A and C can be inferred given the trust values t_{AB} and t_{BC} [27]. Figure 5(b) shows a typical collaborative filtering approach to trust propagation that, based on three given relations between four people, infers the missing trusted relation [6]. As a complementary approach to trust propagation, we aim to directly combine individual and relational attributes as depicted in Figure 5(c). Based on the existing individual profiles A, B, C and D as well as a single existing trusted relationship $t_{A,B,}$ we will calculate similarities between user pairs. Dependant on these distances d(x,y) as well as the characteristics of the existing trusted relation $t_{A,B,}$ the system will recommend or not a relationship between people unknown so far. We denote this approach as similarity-based trust propagation. Our next steps include the further development of our existing integrated prototype and its validation with real-life data. Also, we aim to add social network data as an additional variable of the model and extend it by different relation types.

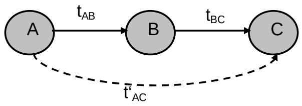

Fig. 5(a). Direct trust propagation

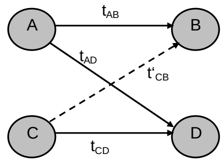

Fig. 5(b). Collaborative trust propagation

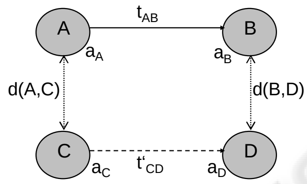

Fig. 5(c). Similarity-based trust propagation

As a basis for the predictive approach, we postulate two work hypotheses: The first hypothesis is that the unary (i.e., propositional) and relational attribute structure latently captures personal qualities that generate degrees of trust, possibly conditioned on specific situations and roles. For instance, looking at a known relational confidence attribute with a source A and a target B (e.g., A assesses B), it is predicted that similar relations (with respect to type and weight) can be measured for sources similar to A and targets similar to B. E.g., if A assesses B positively, C similar to A is predicted to assess D similar to B positively, as well.

The second hypothesis of the approach is that some dimensions of trust are transmissible through a referral network. This means, for example, that looking at such a higher-order trust relation, A trusts B and B trusts C, again possibly conditioned on a situation or role, trust from A to C can be predicted. This is the conceptual basis of referral systems, such as ReferralWeb [15] [36]. The question is what trust dimensions do exhibit this transitive behaviour to which degree.

In particular, the first hypothesis can be mapped to the emergent scientific area of statistical relational learning (SRL), in which graph properties are learned from data and the local graph topology surrounding newly observed nodes are predicted. In this context, we note the work of Jensen, Neville and Wolfe [24] [35] [13] and of Heckerman, Meek and Koller [8], as a basis for a generic social network prediction algorithm. The second hypothesis is related to the friend-of-a-friend principle, which is the basis for transitive trust relations and in fact the basis of the existing system already.

Further, we plan to connect actors with documents to cluster actors by their authorship and roles. This extends the idea of explicit profile creation to implicit methods of profile creation, thus allowing for bootstrapping a real system by connecting it to existing document bases. A scientific basis for work into this direction can be found in [23]. Merging both the content and the social network into a ‘smart’ collaboration network to us seems a promising idea when considering the many real-world knowledge management problems and applications. However, several challenges appear when modelling profiles for the predictive approach to partner matching. These are:

- the modelling of complementarity and compatibility for team building scenarios. This includes the incorporation of research on matching different personal traits with express expertise measures to optimize team staffing.
- the capturing of “inter-rater trust”. Within this functionality, the bias of a rater will be used to remove bias from ratings and will also be incorporated as a specific rater characteristic. This has been partly solved in our existing Opal system via a matrix-based assessment browser as presented in [7].
- the resolution of disreputative scenarios. Situations in which candidates are assessed badly must be resolved in a way that conserves overall integrity and privacy in the community but that still allows marking negative experiences. This is an often-encountered scenario where most rating-based systems capitulate.

5 Validation approach

In order to validate our approach we currently design an experiment as part of a student workshop. We plan to test the aspects of the described recommendation framework in an incremental way. First, students are supposed to enter their CV data into a web-based prototype. The data capturing hereby follows the same rules as it is nowadays done in the various existing job-portals. The CV data together with manually created ratings regarding the match of the students with several job-profiles is then used as input to train the CV-recommender. Based on this training data the system should then be able to predict the match between students and job profiles.

In a second step, Students will be asked to enter relational data into the prototype such as their relations towards fellow students. The relations will be defined based on its direction, duration and intensity. The captured data should then serve as input for the trust computational model. Based on this training data the system should be able to predict previously unknown relations. Finally we aim to combine the separate results into an integrated approach for personalizing the search for persons.

6 Conclusion

In this paper we argued that recommendation systems so far personalize only the search for objects, but not for subjects. We showed that theoretical extensions such as the integration of relational as well as bilateral aspects into current approaches are necessary in order to build a system personalizing the search for individuals. Based on these requirements and building up on two implementations from previous research, we presented an outline of a first existing prototype integrating both approaches into a single system. Our next steps include the extension of this implementation as well as its validation with real-life data as part of a student workshop to be carried out. The objective is to enhance the matching quality of interpersonal partnership especially for collaboration scenarios by building a bilateral as well as relational recommendation engine personalizing the search for individuals.

Acknowledgements. We gratefully acknowledge the support of the European Union under the Fifth Framework Programme Information Society Technologies (contract number: IST-2000-28295).

References

1. Crowder, R., Hughes, G. and Hall, W. (2002) Approaches to Locating Expertise Using Corporate Knowledge Int. J. Intell. Sys. Acc. Fin. Mgmt., 11 , 185-200.
2. DiTomaso, N. (2001) The loose coupling of jobs: the subcontracting of everyone, in Berg, I. and Kalleberg, A.L. (Eds.) Sourcebook of Labor Markets: Evolving Structures and Processes, Kluwer Academic/Plenum, New York, 247-270.
3. Dunphy, D. and Bryant, B. 1996 Teams: Panaceas or prescriptions for improved performance? Human Relations, 49, 677-699.
4. Färber, F., Keim, T., and Weitzel, T. (2003) An Automated Recommendation Approach to Personnel Selection, Proceedings of the 2003 Americas Conference on Information Systems, Tampa.
5. Granovetter, M.S. (1985) Economic Action and Social Structure: the problem of Embeddedness, American Journal of Sociology, Vol. 91, pp. 481 – 510.
6. Guha, R., Kumar, R., Raghavan, P. and Tomkins, A. (2004) Propagation of Trust and Distrust, Proceedings of the WWW2004-Conference, May 17-22, New York, USA, pp. 403-412.
7. Graham, M. and Kennedy, J. (2004) Exploring and Examining Assessment Data via a Matrix Visualisation, Proceedings of the AVI 2004, Gallipoli, Lecce, Italy, ACM Press. pp. 158-162.
8. Heckerman, D., Meek, C., and Koller, D. (2004) Probabilistic Models for Relational Data, Technical Report MSR-TR-2004-30, Microsoft Research.
9. Heinrich, G. (2004) Teamarbeit nach Mass - Expertisemanagement in Organisationsnetzwerken, Trendkompass Electronic Business - IT-Innovationen & neue Prozesse im Unternehmenseinsatz, IRB-Verlag, Stuttgart.
10. Hofmann, T. (1999) Probabilistic latent semantic analysis, Proceedings of the 15th Conference on Uncertainty in Artificial Intelligence (UAI), July 30-August 1, Stockholm, Sweden, 289-296.
11. Hofmann, T. and Puzicha, J. (1999) Latent class models for collaborative filtering, Proceedings of the 16th International Joint Conference on Artificial Intelligence, July 31 – August 6, Stockhom, Sweden, 688-693.
12. Jackson, S.E. (1996) The consequences of diversity in multidisciplinary work teams, in: West, M.A. Handbook of workgroup psychology, John Wiley & Sons, Sussex.
13. Jensen, D., and Neville, J. (2002) Data Mining in Social Networks, NAS 2002.
14. Jones, G.R. and George, J.M. (1998) The experience and evolution of Trust: Implications for Co-operation and Teamwork, The Academy of Management Review, 23, 3, pp. 531-546.

15. Kautz, H., Selman, B. and Shah, M. (1997) Referral Web: Combining Social Networks and Collaborative Filtering, Communications of the ACM, Vol. 40, no. 3, pp. 63-65.
16. Keim, T., König, W. and von Westarp, F. (2004) Bewerbungspraxis 2005 - Eine empirische Untersuchung mit über 11.000 Stellensuchenden im Internet, Research report, University of Frankfurt, 2004.
17. Keim, T., König, W., von Westarp, F. , Weitzel, T. and Wendt, O. "Recruiting Trends 2005 - Eine empirische Untersuchung der Top-1000-Unternehmen in Deutschland und von 1000 Unternehmen aus dem Mittelstand in Deutschland", Research report, University of Frankfurt, 2005.
18. Keim, T., Weitzel, T. (2005) An Integrated Framework for Online Partnership-Building, Proceedings of the 38th Hawaiian International Conference on System Sciences (HICSS-38), Hilton Waikoloa Village, Big Island, Hawaii.
19. Leicht, K. T. and Marx, J. (1997) The Consequences of Informal Job Finding for Men and Women, Academy of Management Journal, 40, pp. 967-987.
20. Melville, P., Mooney, R.J. and Nagarajan, R. (2002) Content-boosted collaborative filtering for improved recommendations, Proceedings of the 18th National Conference on Artificial Intelligence, pp. 187-192.
21. Malone, T. W. and Laubacher, R. J. (1998) The Dawn of the E-Lance Economy, Harvard Business Review, 76 (5), pp. 144-152
22. Montgomery, J. D. (1991) Social Networks and labor-market outcomes: toward an economic analysis, American Economic Review, 81, pp. 1408-1417.
23. McCallum, A., Corrada-Emmanuel, A. and Wang, X. (2004) The Author-Recipient-Topic Model for Topic and Role Discovery in Social Networks: Experiments with Enron and Academic Email, Technical Report, UM-CS-2004-096.
24. Neville, J., Adler, M. and Jensen, D. (2003) Clustering Relational Data Using Attribute and Link Information, Text-Mining & Link-Analysis Workshop, TextLink.
25. Popescul, A., Ungar, L.H., Pennock, D.M. and Lawrence, S. (2001) Probabilistic models for unified collaborative and content-based recommendation in sparse-data environments, Proceedings of the Seventeenth Conference on Uncertainty in Artificial Intelligence, August 2-5, Seattle, USA, pp. 437-444.
26. Resnick, P. and Varian, H. R. (1997) Recommender systems. Communications of the ACM, 40 (3), pp. 56-58.
27. Richardson, M., Agrawal, R. and Domingos, P. (2003) Trust management for the semantic web, Proceedings of the Second International Semantic Web Conference, October 20-23, Sanibel Islands, USA, pp. 351-368.
28. Russo, G., Rietveld, P., Nijkamp, P. and Gorter, C. (2000) Recruitment channel use and applicant arrival: An empirical analysis, Empirical economics, 25, pp. 673-697.
29. Sarwar, B., Karypis, G., Konstan, J. and Riedl, J. (2000) Analysis of recommendation algorithms for e-commerce, Proceedings of the ACM Conference on Electronic Commerce, pp. 158-167.
30. Sarwar, B., Karypsis, G., Konstan, J. A. and Riedl, J. (2000) Application of dimensionality reduction in recommender system – A case study, Proceedings of the ACM WebKDD 2000 Web Mining for E-Commerce Workshop, ACM, New York.

31. Schneider, B., Kristof-Brown, A., Goldstein, H. W. and Smith, D. B. (1997) What is this thing called fit?, in Anderson, N. and Herriot, P. (Eds.): International handbook of Selection and Assessment, John Wiley & Sons, pp. 393-412.
32. Ungar, L. and Foster, D. (1998) Clustering methods for collaborative filtering, Proceedings of the Workshop on Recommendation Systems, AAAI Press, Menlo Park, California.
33. Wasserman, S. and Faust, K. (1994) Social Network Analysis: Methods and Applications, Cambridge University Press.
34. Watson-Manheim, M.B., Crowsten, K. and Chudoba, K.M. (2002) Discontinuities and continuities: a new way to understand virtual work, Information Technology & People, Vol. 15(3), pp. 191-209.
35. Wolfe, A. and Jensen, D. (2004) Playing Multiple Roles, Discovering Overlapping Roles in Social Networks, ICML 2004 Workshop on Statistical Relational Learning and its Connections to Other Fields.
36. Yolum, P. and Singh, M. P. (2003) Emergent properties of referral systems, Proceedings of the 2nd International Joint Conference on Autonomous Agents and MultiAgent Systems (AAMAS), ACM Press.