Personalizing Relevance on the Semantic Web through Trusted Recommendations from a Social Network

Tom Heath, Enrico Motta

Knowledge Media Institute, The Open University,
Walton Hall, Milton Keynes, MK7 6AA, United Kingdom
{t.heath, e.motta}@open.ac.uk

Abstract. Personalization efforts to date have centred on presenting web users with novel items by predicting what they may find relevant. This approach has utility where the user is unsure of exactly what they are looking for, but not where they have a particular information need to satisfy or a particular item to locate. Furthermore, by operating purely on a predefined database of users and items, systems using this approach represent closed worlds and offer poor scalability to new data sets. To address these limitations we propose a technique for personalizing relevance in information seeking activities, based on an understanding of how people seek information and recommendations from their social network. We then describe technical work in progress, based on Semantic Web technologies, that aims to realize this perspective.

1 Introduction

Whilst web sites offering personalization based on conventional web technologies are numerous, the emerging Semantic Web provides an opportunity for richer personalization features to be developed. The availability of structured data adhering to common ontologies enables the integration of user-relevant content from more diverse sources. More importantly however, by allowing users to describe aspects of their context (such as the social networks they are part of) in a standardized way, Semantic Web technologies enable new forms of personalization. In this paper we describe our approach to personalizing relevance in information seeking, through use of Semantic Web technologies and recommendations from social networks.

1.1 Approaches to Personalization on the Web

Personalization on the web has been approached in a number of different ways [1]. Some web sites allow a personalized user experience through changing color schemes or selecting news channels to be displayed on an individual homepage [2]. An alternative approach involves using *recommender systems* [3] [4] [5] that have some knowledge of the user's preferences to select and present content presumed of interest to them, thereby highlighting items of which they may not be aware.

Despite being widely used, these recommender system-based approaches to personalization have a number of limitations. Firstly from a technical point of view, they operate as closed worlds, whereby recommendations can only be given about items that exist within the system, and only the purchase histories of users within the system can contribute to recommendations based on so-called *collaborative filtering* [3]. Secondly, these forms of personalization by speculative recommendation only support the user in carrying out tasks where the solution is unknown or poorly defined, rather than tasks where they have a particular information need to satisfy or a particular item to *locate* [6]. This ignores the role personalization could play in supporting information seeking activities through tailoring relevance to the individual.

1.2 Relevance in Information Seeking

Despite the vast extent of online resources, locating the required piece of information can still present challenges to the user; it may not yet be available on the web, or where a query yields many search results it may be difficult to identify the most appropriate. Resolving these issues requires the identification of sources of additional information not currently available on the web, or means of filtering information based on relevance to the individual's information needs.

Literature on information retrieval has traditionally viewed relevance as a measure of the suitability of a result to the information need of the user *as it is expressed in a query issued to the system*. This relationship between document and query has been referred to as *topical relevance* [7]. However, it would be desirable to measure suitability of the result in relation to the abstract information need of the user [8], whether or not this has been adequately expressed in the query. Building systems that enable such a *personal relevance* requires additional knowledge about the user to be taken into account that may be difficult to express via keyword search.

Attempts to address this issue include [9], where results on a job-seeking site are filtered according to a user profile generated from page view data. However, this approach only enables filtering and personalization to be carried out on items within the closed world of jobs already listed on the site.

1.3 Relevance through Recommendations from a Social Network

Our approach to personalizing relevance is based on identifying the members of their social network the user is most likely to trust as an information source in a given scenario, and using recommendations from these people to personalize search results. Whereas search engines and recommender systems attempt to identify items appropriate to the user, we advocate an approach that identifies the most appropriate sources as a means to identify relevant items.

This source-centric approach using known members of a social network (that we call *Known Person Recommendation*) allows for more complex reasoning to be carried out about the appropriateness of a source than is possible with collaborative filtering systems where other users are unknown. Furthermore, a relevance system driven by social networks is constrained in scope only by the knowledge of the

members of the network, and the ability to infer the source most appropriate to the task. Consequently the approach is not limited to selected domains, as is often the case with existing closed world recommender and personalization systems.

In previous research [10] we identified five factors (*expertise*, *experience*, *impartiality*, *affinity*, and *track record*) that determined from which members of their social network a person would seek recommendations. The criticality and subjectivity of the task were found to influence which factors were most attended to. These findings have informed the technical implementation described below.

2 A System for Personalized Relevance on the Semantic Web

We are currently developing a system to test personalized relevance in locating information about travel resources such as hotels, restaurants, and cultural sights (referred to here as "travel objects"). An architectural overview is shown in Figure 2.

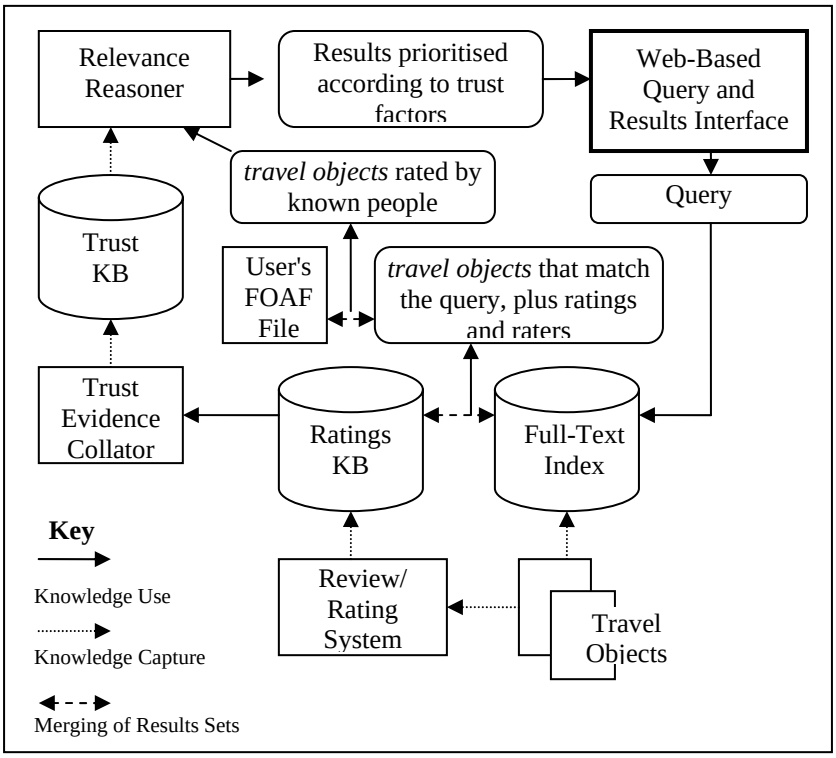

Fig. 1. System Architecture for Personalizing Relevance on the Semantic Web

The system is being piloted in the travel domain to evaluate the feasibility of personalizing relevance in low-criticality high-subjectivity tasks. Consequently, the trust factors being prioritized are **experience** and **affinity**, found in [10] to be most

relied upon in such tasks. If successful, the system will be applied across scenarios with a wider range of characteristics, making use of the full range of trust factors.

Users will provide two types of input to the system: firstly semantic descriptions of their social networks using the FOAF vocabulary [11]. This will give the user freedom and control in how they manage this information, and how much is divulged. Secondly users will use the review/rating system to provide reviews of travel objects. This will serve to populate the knowledge base with information about travel objects that users may wish to locate, and provide data from which trust relationships between a user and members of their social network can be inferred. The extent of these trust relationships will form the basis for providing personalized relevance to users of the system. Whilst it is theoretically possible for users to explicitly provide information about trust relationships the demands this would place on the user are deemed prohibitive.

Use of Semantic Web technologies [12] as a platform for developing such a system helps overcome the limitations of closed world recommender systems in a number of ways. Firstly by acting as a large distributed database, the Semantic Web enables any item to be included within a recommender system, not just those within a centrally administered catalogue. Secondly, providing recommendations are made available in standard formats, users can access the knowledge of members of their social network without all members having to subscribe to the same services, as these can be harvested from multiple locations and easily integrated to form a knowledge base.

2.1 Inferring Trust Relationships from Reviews of Travel Objects

Trust relationships will be inferred based on the affinity between the user and each member of their social network, and the relevant experience of each member of the network.

The experience a member of ones network has of a particular location will be determined by the number of travel objects they have reviewed/rated from that location. Where possible the URI of a travel object will be de-referenced to obtain a machine-readable description of its location. Where such a machine-readable description is not available, tags provided by the user and full-text indexing of the content at the object's URI will offer a best guess as to where it is located.

One potential confound of relying on number of objects rated to infer experience is variation in users' enthusiasm for rating and reviewing objects. The use of evidence about affinity in addition to experience will help to mitigate negative effects of this variation. Whether an affinity exists between a user and a member of their social network will be determined automatically by a combination of the following factors: the extent to which both parties have rated the same tourism objects (i.e. the overlap in rated objects), and the correlation between the ratings given by each party.

Due to the reliance on ratings from users for both population of the system and generation of evidence of trust relationships, the system may suffer from bootstrapping problems. In this case it may be possible to fall back on alternative sources of evidence for factors such as experience. For example, where individuals have shared their photos online and have tagged these with place names or locations, this could be used to infer that they have some experience of this location. Similarly if

an individual has studied or worked in a particular place, as expressed in their FOAF file, this would suggest they have some experience of the locality. Whilst in these cases no specific tourism objects would be available to be recommended by the system as the users had not yet provided ratings, the user could still benefit from identification by the system of potential sources of information.

3 Conclusions and Future Work

Implementation of the system described above is underway. However, a number of challenges remain in both the implementation and adoption of such a system. Whilst we believe the system to be based on sound principles, it is not clear how well the system will scale from a user perspective. A point may be reached where the system contains a volume of knowledge that renders even the relevance mechanisms proposed here unworkable. Such a scenario is analogous to how search engine algorithms have had to be modified as the web has grown in size.

Furthermore, users' views on provision of social network information on the public web will need to be taken into account. Whilst this is fairly widespread among Semantic Web early adopters and users of services such as Tribe.net[1], it is not sufficiently common at present to enable widespread deployment of the proposed system. It also remains to be seen whether the majority of users will be prepared to make this information available for use by the system. Encryption of FOAF files may be required to overcome this issue.

Despite the challenges described, the Semantic Web remains the most appropriate platform for implementing this form of personalized relevance. Achieving the same degree of functionality using conventional web technologies would require the creation of a single system with a vast range of functionalities, and adoption of this one system by all users. It is unlikely that such a closed world approach would gain sufficient uptake by users to reach critical mass.

Following implementation and deployment the system will be evaluated to assess its effectiveness in personalizing relevance for users locating travel information. In addition to collecting quantitative measures such as number of users and number of items rated, users' qualitative experiences of using the system will be recorded. Results provided by the system will also be compared to a baseline set of results from a conventional search engine.

If deemed acceptable and valuable by users, such a system for personalized relevance has the potential to democratize recommendation and personalization on the Semantic Web. At present, personalization is largely limited to closed worlds. Organisations such as e-commerce sites that collect data about user preferences have little incentive to make this data available for use by competing services. However, a system that enables reviews to be made in an open, public world allows many competing services to use this knowledge in providing novel services to the user.

[1] http://www.tribe.net

Acknowledgements

This research was partially supported by the Advanced Knowledge Technologies (AKT) project. AKT is an Interdisciplinary Research Collaboration (IRC), which is sponsored by the UK Engineering and Physical Sciences Research Council under grant number GR/N15764/01. The AKT IRC comprises the Universities of Aberdeen, Edinburgh, Sheffield, Southampton and the Open University.

References

1. Mobasher, B., Cooley, R., Srivastava, J.: Automatic personalization based on Web usage mining. Communications of the ACM 43 (2000) 142-151
2. Manber, U., Patel, A., Robison, J.: Experience with personalization on Yahoo! Communications of the ACM 43 (2000) 35-39
3. Herlocker, J. L., Konstan, J. A., Riedl, J.: Explaining Collaborative Filtering Recommendations. In: Proc. CSCW2000 (2000) 241-250
4. Linden, G., Smith, B., York, J.: Amazon.com Recommendations: Item-to-Item Collaborative Filtering. IEEE Data Engineering Bulletin (2003) 76-80
5. Balabanovic, M., Shoham, Y.: Fab: content-based, collaborative recommendation. Communications of the ACM 40 (1997) 66-72
6. Heath, T., Dzbor, M., Motta, E.: Supporting User Tasks and Context: Challenges for Semantic Web Research. In: Proc. ESWC2005 Workshop on End-User Aspects of the Semantic Web (UserSWeb) (2005)
7. Eisenberg, M., Schamber, L.: Relevance: The Search for a Definition. In: Proc. 51st Annual Meeting of the American Society for Information Science (1988)
8. Park, T. K.: Toward a theory of user-based relevance: a call for a new paradigm of inquiry. Journal of the American Society for Information Science 45 (1994) 135-141
9. Bradley, K., Rafter, R., Smyth, B.: Case-Based User Profiling for Content Personalisation. In: Proc. International Conference on Adaptive Hypermedia and Adaptive Web-Based Systems (AH2000) (2000) 62-72
10. Heath, T., Motta, E., Petre, M.: Person to Person Trust Factors in Word of Mouth Recommendation. In: Proc. CHI2006 Workshop on Reinventing Trust, Collaboration, and Compliance in Social Systems (Reinvent06) (2006). Available from http://www.cs.ucl.ac.uk/reinvent06/
11. Brickley, D., Miller, L.: FOAF Vocabulary Specification.http://xmlns.com/foaf/0.1/ (2005)
12. Berners-Lee, T., Hendler, J., Lassila, O.: The Semantic Web. Scientific American 284 (2001) 34-43